Journal of Financial Economics 147 (2023) 317–337

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Refusing the best price?☆

Sida Li[a], Mao Ye[b,*], Miles Zheng[c]

[a] Brandeis University, United States
[b] Cornell University and NBER, 365 Sage Hall, 106 E Ave, Ithaca, NY 14853, United States
[c] University of Illinois at Urbana-Champaign, United States

ARTICLE INFO

Article history:
Received 12 October 2021
Revised 10 November 2022
Accepted 11 November 2022
Available online 28 November 2022

Keywords:
Regulation NMS
Order types
Routing
Make/take fees
High-frequency trading

ABSTRACT

The Regulation National Market System (Reg NMS) links fragmented stock exchanges by routing orders to the National Best Bid and Offer (NBBO). As the NBBO ignores exchange fees, 62% of routings lead to worse net prices. An increase in fee differences increases the market share captured by orders that refuse Reg NMS routings, particularly for stocks whose fees account for a large portion of transaction costs. Heterogeneous opportunity costs rationalize routing choices: non-routable orders entail lower non-execution costs than routable orders. Our results indicate that fees and clientele segmentation drive the proliferation of order types in the Reg NMS era.

1. Introduction

To promote competition, most developed countries now allow a stock to be traded outside its listing exchange (O'Hara and Ye 2011; Foucault et al., 2013). One central question for policymakers is whether and how they should link these fragmented markets. U.S. regulators, for example, chose to connect fragmented markets through the Regulation National Market System (Reg NMS) in 2005. Reg NMS establishes the national best bid and offer (NBBO) across competing stock exchanges and promulgates rules designed to route orders submitted to any exchange to the NBBO. Under Reg NMS rules, exchanges are interconnected and each exchange serves only as a point of entry for the same destination—the NBBO.

Surprisingly, we find that orders that refuse Reg NMS routing comprise 57% of trading volume on the New York Stock Exchange (NYSE), and the incentive to circumvent Reg NMS is so strong that it has led to a proliferation of order types.[1] Why do most orders refuse Reg NMS routing? Why do some other orders accept Reg NMS routing? The answers to these two questions provide the key to understanding and improving market design, as creating the NBBO and routing to the NBBO are two cornerstones that differentiate the market structures of U.S. exchanges from

☆ Nikolai Roussanov was the editor for this article. We thank Jim Angel, Dan Bernhardt, Ekkehart Boehmer, Eric Bryce, Colin Clark, Carole Comerton-Forde, Nicolas Crouzet, Ian Dew-Becker, Thomas Ernst, Harry Feng, Thierry Foucault, Joel Hasbrouck, Edwin Hu, Pankaj Jain, Phil Mackintosh, Albert Menkveld, Bruce Mizrach, Dermot Murphy, Maureen O'Hara, Marios Panayides, Steven Poser, Andriy Shkilko, Jeffery Smith, Elvira Sojli, Chester Spatt, Jeremy Stein, Kumar Venkataraman, Sunil Wahal, Chen Yao, and seminar participants at Louisiana State University, the New York Stock Exchange, Stockholm Business School, the University of Illinois at Urbana-Champaign, Rutgers University, the Citadel, Two Sigma Investments, AFA 2021, the 3rd Future of Financial Information Conference, CICF 2021, EFA 2021, NFA 2021, SFS 2022, and Microstructure Online Seminars Asia Pacific for helpful discussions and suggestions. We thank the New York Stock Exchange for providing us with their proprietary data for this paper. Ye acknowledges support from National Science Foundation grant 1,838,183 and the Extreme Science and Engineering Discovery Environment (XSEDE).

* Corresponding author.
E-mail address: my87@cornell.edu (M. Ye).

[1] Mackintosh (2014) identified 133 order types across U.S. stock exchanges in 2014, the majority of which were designed after the implementation of Reg NMS for the purpose of refusing Reg NMS routing.

https://doi.org/10.1016/j.jfineco.2022.11.004
0304-405X/

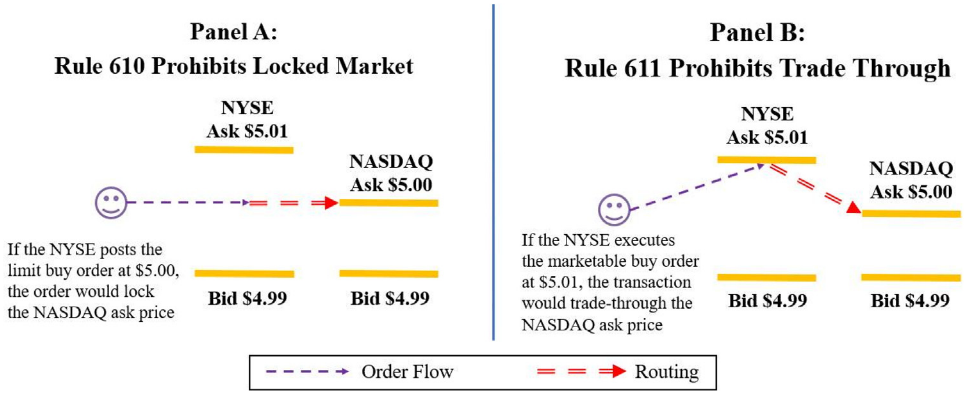

Fig. 1. Routing to unlock the market and routing to prevent trade-through, This figure illustrates two types of routing led by Reg NMS. The purple dotted lines represent the order flow sent to the NYSE, while the red double-dashed lines represent routing performed by the NYSE. Panel A shows routing designed to unlock the market. The $5.00 limit order would improve the best bid on the NYSE, but Rule 610 prohibits one exchange from establishing a bid (ask) price that is the same as the ask (bid) price on another exchange. The NYSE unlocks the market by routing the limit buy order to take liquidity from NASDAQ. Routing to unlock the market leads to the same gross price but a worse net price after paying the 0.3-cent routing fee per share. Panel B shows the routing to avoid a trade-through. The buy order is now marketable on the NYSE, but the NYSE routes the order to avoid a trade-through—the execution of marketable orders at prices that are inferior to displayed quotes on other exchanges. Rule 611 routing (the no trade-through rule) leads to a greater price improvement than the 0.3-cent routing fee.

those of exchanges in other countries.[2] We find that the answers to these two questions come from how Reg NMS defines the best price and, more broadly, the best execution. (1) Most orders refuse Reg NMS routing because it defines the NBBO based on a stock's gross price but not its net price. Consequently, we find that Reg NMS routes 62% of orders to worse net prices. (2) Reg NMS focuses solely on the best price, whereas the best execution includes other dimensions such as fill rates and the opportunity costs of non-fills. We find that the opportunity costs of non-fills for orders that accept routings are significantly higher than the fees. That is, if these orders refused routing to worse net prices, they would be filled with even worse net prices in the future. On the other hand, orders that refuse routing incur almost zero opportunity costs. Therefore, heterogeneous opportunity costs rationalize heterogenous routing choices and thereby the proliferation of order types that serve heterogeneous clienteles.

We first show that the difference between best gross price and best net price explains why some order types refuse the NBBO. Reg NMS defines the NBBO based on gross prices, but traders' true transaction costs also include routing fees. Therefore, routing improves net prices only when it strictly improves gross prices. Unfortunately, only 38% of routings improve gross prices, with the rest aiming to unlock the market, which leads to the same gross prices and worse net prices. Panel A of Fig. 1 visualizes routings that unlock the market and explains why they lead to worse net prices. Suppose that the NYSE's best ask price is $5.01 and the NYSE's best bid price is $4.99. If a buyer submits a limit order at $5.00 to the NYSE, the order improves the NYSE best bid to $5.00. In January 2010, NYSE provided a rebate of 0.10 cents per share for a liquidity-providing orders upon execution. Yet Rule 610 of Reg NMS prohibits the order from improving the NYSE bid if it locks the market, that is, if the improved bid price is equal to the ask price on another exchange.

The default solution to unlocking the market is to route the limit order to another exchange and take liquidity at its limit price. Such routing does not, however, improve the gross price but instead leads to a worse net price because the order needs to pay a routing fee of 0.30 cents. The 0.40-cent difference between the rebate and the routing fee seems small, but it is larger than the profits earned by making liquidity. Worse net prices explain why most non-marketable orders add "do-not-ship" (DNS) instructions, which ask the NYSE to cancel an order if it needs to be routed. We find that DNS limit orders earn a small profit of 1.57 bps after collecting rebates but would lose 0.89 bps if they had paid routing fees.

We conduct difference-in-differences tests to provide further evidence that fees and rebates cause routing refusals. As the NYSE charges flat fees for all stocks in our sample, routing to worse net prices reduces liquidity-providing revenue to a greater extent for stocks with narrower bid–ask spreads. In addition, narrower bid–ask spreads make it more difficult to improve spreads without locking quotes from other exchanges. Therefore, we find that DNS limit orders capture a 10-percentage-point higher market share for stocks whose bid–ask spreads fall into the lowest tercile than for stocks whose bid–ask spreads fall into the highest tercile. We then exploit two exogenous shocks to exchange fees. On May 1, 2020, the NYSE increased make rebates from 0.10 cents per share to 0.13 cents per share and then to 0.15 cents per share on January 3, 2011. As the routing fee was always set at the 0.30-cent level, the cost difference between making liquidity on the NYSE and taking liquidity outside increased from 0.40 cents to 0.43 cents and then to 0.45 cents. We find that use of the DNS tag increases by 2.28 (1.22) percentage points after the first (second) fee change. Finally, the difference-

[2] For example, MiFID in Europe neither establishes an NBBO nor requires an exchange to route orders to other exchanges (European Securities and Markets Authority 2019). MiFID delegates the responsibility for ensuring best execution to investment firms, not to stock exchanges.

in-differences comparison shows the heterogeneous treatment effects of fee changes. As fee changes are relatively more important for stocks with narrower spreads, we find that, after the first (second) fee change, the market share captured by DNS limit orders increases 2.62 (2.55) percentage points higher for stocks whose bid–ask spreads fall into the lowest tercile than for stocks whose bid–ask spreads fall into the highest tercile.

Although routing to worse net prices can explain why some orders refuse routing, it gives rise to another puzzle regarding why some other orders accept worse net prices. We find that the answer comes from the difference between the best price and the best execution. Reg NMS makes price the sole criterion for determining how and where orders will be executed (Glassman and Atkins 2005; Spatt 2014), but the best execution for investors also includes other dimensions such as the probability that execution occurs.[3] As DNS instructions cancel an order if it needs to be routed, the order submitter may need to resubmit the order later. The order submitter incurs positive opportunity costs if the order will be filled at a worse future price (Perold 1988). We find that the opportunity costs for routed limit orders are much higher than the fee costs if routed limit orders reject routings. That is, if these orders had refused to be routed to worse net prices, they would have been filled at even worse prices in the future. On the other hand, the opportunity costs for rejecting routing are not significantly positive for DNS orders. We find that DNS limit orders are more likely to come from liquidity-providing high-frequency-traders (HFTs). DNS limit orders win 72% of races to establish time priority in liquidity provision. DNS limit orders also successfully cancel stale quotes 43% of the time, but routable limit orders cancel only 9% of stale quotes before they are adversely selected. Taken together, our results indicate the following clientele segmentation based on heterogeneous opportunity costs. HFTs use sophisticated DNS orders to reduce fee costs for providing liquidity, while non-HFTs accept routing to reduce the opportunity costs of non-fills.

Our paper provides the first analysis of exchange routing. We find that the most well-known reason for routing—Rule 611 (the no trade-through rule)—accounts for only 31% of routing.[4] The remaining routings aim to comply with Rule 610, whose name, “access to quotations,” seems unrelated to routing. Indeed, most textbook and policy discussions of routings focus only on Rule 611, while discussions of Rule 610 focus only on its clauses (a)–(c), which grant fair access to quotes and cap access fees at 0.30 cents per share (e.g., Hasbrouck 2007, Foucault et al. 2013). Surprisingly, we find that Rule 610, through its overlooked clause (d), strictly dominates Rule 611 in terms of the market share captured by routing. Rule 610 (d) prohibits an exchange from displaying a bid (ask) price that locks or crosses an existing ask (bid) price on other exchanges. Among the 69% of routings led by Rule 610 (d), 7% still improve prices because they resolve crossed markets—that is, if an order displays a bid (ask) price that is higher (lower) than the ask (bid) price on another exchange. The remaining 62% of routings result in the same gross prices and worse net prices by resolving locked markets.

More interestingly, the risk of being routed to worse net prices depends crucially on order types. Intuitively, an order type that does not display liquidity cannot lock the displayed market and is thereby subject to Rule 611 but not Rule 610. We find that routings always improve prices for market orders and Immediate-Or-Cancel (IOC) orders because they refuse to provide or display liquidity. In contrast, routing worsens net prices for 78% of routable limit orders because they can display quotes that lock markets. Limit orders could avoid routings to worse net prices by using DNS instructions to cancel orders once they lock markets. Therefore, although one goal of Reg NMS is to encourage traders to compete by displaying more aggressive quotes (SEC 2005), Rule 610 creates an incentive to cancel such quotes. Another way to circumvent Rule 610 is to hide quotes completely. We find that routing improves prices for 99.95% of fully hidden orders. In comparison, routings lead to worse net prices for 81.50% of partially hidden orders (i.e., iceberg orders) because their displayed portions can still lock markets. Therefore, although Reg NMS was designed in part to encourage traders to compete by displaying quotes, Rule 610 creates an incentive for traders to hide their quotes. In Section 7 we briefly discuss cross-exchange variations and time-series evolutions of order types. Most order-type innovations aim to remodel order types that are subject to Rule 610. These innovations unlock a market along three product lines: canceling, hiding, or automatically repricing orders to less aggressive prices. Along with the proliferation of order types, we see further increases in the market shares captured by non-routable orders. 57% of executed volume in 2010–2011 comes from non-routable orders, while the proportion was as high as 79% in June 2021.[5]

Our paper contributes to the literature on competition between stock exchanges (Yao and Ye, 2018; Budish et al., 2021). We provide one explanation of why exchanges compete directly on speed. Existing studies of speed competition focus on traders' speed.[6] Within this framework, exchange speed is mostly irrelevant if an exchange creates the same latency for all traders. Yet all major U.S. exchanges compete aggressively on speed. They invest in reducing exchange latency and adopt price/time priority to reward traders who establish fast quotes, even though the SEC allows alternative priority rules such as price/size pri-

[3] See, for example, FINRA Rule 5310, which is available at https://www.finra.org/rules-guidance/rulebooks/finra-rules/5310, and FINRA Regulatory Notice 15–46 (Guidance on Best Execution Obligations in Equity, Options and Fixed Income Markets), available at https://www.finra.org/rules-guidance/notices/15-46.

[4] Panel B of Fig. 1 shows how Rule 611 routing differs from Rule 610 routing. Again, the NYSE's best ask price is $5.01 and the NASDAQ best ask price is $5.00, but the trader submits a marketable order. The NYSE cannot execute the order at $5.01 because doing so leads to a trade-through (the execution of marketable orders at prices that are inferior to displayed quotes on other exchanges). One approach to resolve trade-throughs would be to route marketable orders to national best ask prices on NASDAQ.

[5] https://www.nyse.com/markets/nyse/trading-info#equities-order-types, accessed on August 12, 2021.

[6] Menkveld (2016) provides an excellent survey for the literature on speed competition between traders.

ority and frequent batch auctions (SEC 2015).[7] Our paper indicates that Rule 610 is one driver of speed competition between exchanges, as a slow exchange tends to lock quotes from a fast exchange. The refills of partially displayed reserve orders provide an ideal laboratory for studying the impact of exchange latency because refills involve exchanges' but not traders' latency. We find that a partially displayed reserve order can be routed to another exchange because it locks the market during the brief latency in refilling its displayed portion.[8] Exchange latency costs order submitters because most routings led by exchange latency execute at worse net prices. Exchange latency also costs stock exchanges because slow exchanges lose trading volume and pay fees to fast exchanges. Therefore, Rule 610 explains why no major U.S. exchanges have incentives to delay liquidity-making orders.[9]

Combined with studies that examine other rules under Reg NMS, our paper helps explain the post-Reg NMS U.S. market structure. Rule 612 imposes a minimum price variation (tick size) of one cent. Tick sizes and fees explain why gross quotes frequently lock. When the gross quotes of two exchanges lock, their net quotes do not lock. Instead, the true bid–ask spread equals the sum of rebates on these two exchanges. Rule 610 and Rule 612 then implicitly impose a minimum net spread of 1 tick plus the sum of rebates. The large market share that DNS orders capture, particularly for stocks with narrower bid–ask spreads, indicates that the true bid–ask spread is often narrower than the mandated minimum. The differences between the true and imposed minimum bid–ask spreads then lead to speed races to capture rents (Yao and Ye 2018), and we find that DNS limit orders win such speed races.

Chao et al. (2019) show that Rule 612 leads to fragmented stock exchanges that charge varying fees. In this paper, we show that exchange fees and Rule 610 generate order types that are designed to avoid fees. Spatt et al. (2018) and the Equity Market Structure Advisory Committee (EMSAC 2017) argue that Reg NMS is highly prescriptive in mandating how executions must occur. We find that the proliferation of order types creates flexibility for such prescription. For example, Rule 610 forbids locked markets, but traders who lock markets may have heterogeneous intentions. Heterogeneous order types give traders customized keys to unlock the market. Traders who aim to execute their orders can choose routable limit orders to resolve the locked market by execution. In contrast, traders who aim to provide liquidity can choose non-routable orders to secure the rebates. Taken together, our paper and previous studies help to explain how Reg NMS generates four features of the current market: high-frequency trading, market fragmentation, exchanges with heterogeneous make/take fees, and the proliferation of order types.

Our proprietary data provide detailed information on order types that are excluded from existing datasets. For example, exchange direct-feed data, such as NASDAQ ITCH, or consolidated data such as NYSE Trade and Quote Data (TAQ), can show that an order displays 100 shares at $10, but these datasets do not contain information indicating whether the liquidity comes from a routable limit order, a non-routable limit order, or a refill of a partially displayed order. Existing datasets decompose these complex orders into a simple binary world: an order is called a limit order when it makes liquidity and it is called a market order when it takes liquidity. This data limitation imposes a constraint on the literature on order types, which focuses on the binary choice between limit orders and market orders.[10] We provide the first anatomy of complex order types and rationalize their evolution and proliferation.

2. Data

Our sample period runs from January 1, 2010 through March 1, 2011, and we select a stratified sample of 109 stocks. Table 1 Panel A provides information on the sample selection process. Starting with all securities listed on the NYSE in December 2009, we apply standard filters (Boehmer 2005; O'Hara and Ye 2011) to remove non-common equities, dual-class shares, real estate investment trusts (REITs), and common stocks of non-U.S. companies. We also exclude stocks that, at the end of the year 2009, are priced below $5.00. We obtain our final sample by ranking the remaining 1086 stocks based on their trading volumes and then pick every tenth stock, starting with the stock with the highest trading volume. Our final list of 109 sample stock tickers is presented in Table 1 Panel B. In Table 1 Panel C, we provide summary statistics for the characteristics of these 109 stocks.

We study Reg NMS routing and stock-exchange order types using proprietary message-level data from the NYSE: System Order Database (SOD).[11] The granularity of timestamps is 0.01 s for order submissions and 0.001 s for or-

[7] The NASDAQ PSX exchange tried price/size priority in 2010, but it attracted little volume and reverted to price/time priority in 2012. The NASDAQ PSX revived price/size priority in 2014 but once again attracted little volume. The NYSE always rewards the trader who sets the quote with higher priority over other traders, no matter whether the other trader is a designated market maker, floor broker or a participant in the limit-order book.

[8] A routable reserve order will be evaluated for routing both on arrival and each time the display quantity is replenished (NYSE rule 7.31.d.1.D). Consider a reserve order of 5000 shares that displays only 100 shares at an ask price of $10. Once the 100-share trade is executed, the NYSE needs to fill another 100 shares. During the small latency of the refill, another exchange may display a bid price of $10. The NYSE then needs to route the reserve order out to take liquidity because the refilled ask price of $10 locks the market.

[9] The Chicago Stock Exchange, NASDAQ OMX PSX, and CBOE EDGA all proposed an asymmetric speed bump that delays only liquidity-taking orders but not liquidity-making orders. The IEX is the only U.S. stock exchange that implements a symmetric speed bump that delays all orders but has not attracted much displayed trading volume. In the period of January–June 2020, only 13.4% of the IEX's trading volume came from displayed orders while 86.6% of its trading volume came from non-displayed orders (Aquilina et al., 2021). The NYSE American implemented a symmetric speed bump in 2017 but removed it in 2019.

[10] See Seppi and Parlour (2009) for a survey of the limit-order-book literature.

[11] Hasbrouck (1992) introduced an earlier version of SOD data, which is one of the four files in his TORQ database. To accommodate the proliferation of new order types, our SOD data contain many more columns than the TORQ data in the early 1990s. Most importantly, our data come with special-order instructions such as routing and reserve decisions. The SOD data are observed at the child-order level. For studies that analyze data at

Table 1

Sample selection, tickers of sample stocks, and summary statistics for sample stocks, In Panel A, we report our sample selection criteria. In Panel B, we list the tickers of our 109 sample stocks. In Panel C, we present summary statistics (at the beginning of the year 2010). *Market Cap* is market capitalization ($billion). *Share Volume* is the annual shares of a stock traded (in billions). *Share Price* is the nominal share price.

Panel A: Sample selection criteria

All NYSE securities	2413
Non-common stock equities (American Depository Receipts, units, certificates, and shares of beneficial interest)	−565
Common stocks of non-U.S. companies, closed-end funds, Real Estate Investment Trusts, and Americus Trust components, and exchange-traded funds	−517
Dual-class stock	−152
Price (December 31, 2009) < 5	−93
Universe sample	1086
Final sample by selecting every tenth stock from our universe sample (December 2009)	109

Panel B: Tickers of sample stocks

ACI, AGL, AGP, AIG, AMD, AME, AOS, ARM, ATI, B, BAC, BBT, BGG, BW, BYI, CGA, CMN, CNS, CPN, CPO, CPX, CSS, CSX, CYH, DGI, DPL, DVN, ELY, EMR, ENZ, ES, ESE, EXP, FCN, FMR, FOR, GAS, GCI, GD, GEO, HL, HLX, HNZ, HRS, HW, HXL, IEX, IFF, IPI, IWA, JBL, JLL, JMP, JWN, KCI, KSU, MCD, MCS, MCY, MDC, MHP, MMC, MOS, MTG, MTZ, N, NOV, NPK, NWN, OKE, OSK, PCP, PPD, PVR, RAI, RGA, RGR, RGS, ROK, RRC, RRI, SBX, SHW, SMG, SON, SUG, SUR, SY, TC, TE, THC, TIN, TLB, TRC, TUP, USB, VMI, VQ, VSH, VVI, WFR, WGL, WGO, WMS, WPP, WWW, XCO, XEC, Y

Panel C: Summary statistics for sample stocks

	N	Mean	SD	5%	25%	50%	75%	95%
Market Cap	109	6.3	15.28	0.23	0.79	2.17	3.73	26.29
Share Volume	109	1.36	8	0.02	0.09	0.23	0.61	2.41
Share Price	109	31.63	31.19	7.54	13.62	26.54	41.14	68.17

der executions and cancellations. The message-level data include all messages sent by market participants to the NYSE, including two types of information not found in other U.S. datasets. First, existing data on U.S. stock markets provide information only on actions that change the status of the LOB or lead to trades (Aquilina et al., 2021), whereas message-level data also include actions that neither change the status of the LOB nor lead to trades. Therefore, existing U.S. datasets truncate the number of observations in message-level data. Second, existing data on U.S. stock markets usually provide a simple dichotomy regarding whether an order provides or demands liquidity, whereas message-level data provide detailed information that reveals order types. Therefore, existing data on U.S. stock markets truncate the message-level data in terms of variables that apply to each observation. Our conversation with the NYSE suggests that no traders have access to U.S. message-level data, but sophisticated traders deploy algorithms to obtain, from public data, noisy estimates of this important information indicating order types.[12]

2.1. Recovering truncations in observations

Existing data from U.S. stock markets include observations only when they affect the displayed limit order book (LOB) or lead to a trade. Aquilina et al. (2021) use message-level data from the London Stock Exchange to address one type of truncation: failed attempts to take or cancel liquidity. Our SOD data come from the U.S., and we analyze three other types of observation truncations.

First, SOD data include failed attempts to add liquidity. For example, a DNS limit order is canceled if it locks or crosses quotes on another exchange. Second, SOD data include detailed information pertaining to reserve orders. Because the purpose of such orders is to hide trading interests, they leave no or limited traces in other datasets. SOD data include order prices, total sizes, displayed sizes, and refilled sizes if displayed portions are consumed. Third, and most importantly, SOD data track orders routed outside the NYSE. These routing-out orders do not leave footprints in local markets, but they are essential for studying connections between and competition across exchanges.[13]

Table 2 summarizes the market shares of routing volume and executed volume for each destination exchange. The results reported show that NYSE Arca receives 53.86% of routed-out volumes.[14] This number is almost double the NYSE Arca's market share of trading volume (27.74%). Two economic forces explain the NYSE's preference for routing to NYSE Arca. First, NYSE Arca has the same owner as the NYSE. Second, NYSE Arca and the NYSE use the same data center, which is also the same location for the official NBBO consolidator (known as the "SIP") of NYSE-listed stocks ("Tape A" stocks). Routing to NYSE Arca, therefore, involves the least latency.

2.2. Recovering truncations in variables

SOD data include detailed information regarding order types. In contrast, existing data from U.S. stock markets

the institutional-parent-order level, see, for example, Frazzini et al. (2018), Anand et al. (2021), and Beason and Wahal (2021).

[12] For example, Bacidore (2020) details algorithms that are designed to infer hidden liquidity.

[13] When the NYSE routes an order to NASDAQ, the trade will be reported by NASDAQ to the SIP. NASDAQ will not re-route the order to another exchange.

[14] When multiple exchanges offer the best price, the routing initiator can choose the routing destination among them.

Table 2
Market share of routing volume and market share of executed volume In this table, we report routing from the NYSE to various exchanges (sorted by the number of shares routed), the number of routed trades, the number of routed shares, the exchanges' corresponding operators in 2010, their market shares of routing volume over total routed volume, and their market shares of trading volume over the consolidated volume of all 13 exchanges.

Destination	Number of routed trades	Number of routed shares	Market share (routed volume / total routed volume)	Operator (2010)	Market share (exchange trading volume / consolidated volume)
NYSE Arca	4408,234	1015,980,000	53.86%	NYSE	27.74%
NASDAQ	2877,660	493,862,000	26.18%	NASDAQ OMX	31.97%
BZX	1426,802	254,230,000	13.48%	Bats Global	16.97%
EDGX	225,823	49,951,000	2.65%	Bats Global	5.52%
National Stock Exchange	281,237	49,503,100	2.62%	NSX	1.42%
Boston Stock Exchange	73,749	11,315,900	0.60%	NASDAQ OMX	7.83%
Chicago Stock Exchange	15,414	4634,200	0.25%	CHX Holdings	0.59%
EDGA	30,654	3769,500	0.20%	Bats Global	6.17%
Philadelphia Stock Exchange	8900	1706,100	0.09%	NASDAQ OMX	0.58%
BYX	7296	992,100	0.05%	Bats Global	0.99%
CBOE	3350	546,000	0.03%	CBOE	0.24%
AMEX	0	0	0.00%	NYSE	0.00%

provide only binary information indicating whether an order provides or takes liquidity. Under this simplification, we know the displayed price of a liquidity-providing order but not the price function underlying the order; we know the displayed size but not the actual size of the order; we know when an order enters but not when the order expires; we know whether an order is willing to provide or take liquidity but not whether it refuses to provide or take liquidity. The SOD data allow us to recover this important information from two variables: time-in-force (TIF) and special-order instructions (SOIs). TIF provides information indicating whether an order can provide liquidity or not, and SOIs refine an order type by stipulating the conditions under which the order provides or demands liquidity. TIF and SOI combinations define an order type, and we provide the taxonomy of order types in the next section.

3. Taxonomy of order types and summary statistics

We categorize order types in SOD data along three lines. In Section 3.1, we introduce four order types that only take liquidity, which are refinements of market orders. In Section 3.2, we introduce four order types that can make liquidity at their limit prices, which are refinements of limit orders. We focus our main analysis on these eight order types, which account for more than 99% of the continuous trading volume. In Section 3.3, we introduce order types that are excluded from our main analysis either because they fall outside continuous trading hours or because they capture very low market shares.

The best way to understand an order type is to examine the conditions under which an order of that type is executed. An order can be executed in three ways: by providing liquidity on the NYSE, by taking liquidity from the NYSE, or by being routed to away exchanges to take liquidity. Table 3 shows the percentages of the executed volumes that take local liquidity, take liquidity from an away market, or make liquidity locally for each order type.

3.1. Order types that take liquidity only

Panel A of Table 3 summarizes the results for four order types that can only take liquidity. As liquidity-taking orders can never lock or cross the market, they are subject to Rule 611 but not Rule 610. The simplest and most famous order type that takes liquidity is the market order, which buys or sells at the best price obtainable. Market orders do not specify their limit prices, and we do not find any SOIs for market orders. The other three types of liquidity-taking orders contain limit prices, but their TIFs are IOC. Therefore, they refuse to provide liquidity at their limit prices. A plain IOC order does not add further instructions. Like market orders, a plain IOC order executes by taking liquidity from the NYSE or from other exchanges. Nevertheless, plain IOC orders are much less likely than market orders to take liquidity from outside, as routing accounts for 33.64% of executed volume for market orders but only 1.93% of executed volume for plain IOC orders.

Traders can add further instructions to IOC orders to refuse Rule 611 routing. Rule 611 prohibits the *execution of trades* at one venue at prices that are inferior to quotes displayed at another venue. A trade can bypass Rule 611 routing in either of two ways: canceling or sweeping. An NMS IOC order refuses routing by canceling the order. An IOC ISO complies with Rule 611 by sweeping protected quotes on all other exchanges that are priced better than the limit price of the IOC ISO. An exchange can then directly execute an IOC ISO without checking protected quotes on other exchanges. As NMS IOC and IOC ISOs refuse to make liquidity and refuse to take liquidity outside of the NYSE, 100% of their executions take liquidity on the NYSE.

We find that non-routable orders dominate share volume across order types that only take liquidity. IOC ISOs account for 15.66% of total share volume, while NMS IOC orders account for 11.35% of share volume. Market and plain IOC orders account for only 4.48% and 2.84% of share volume.

3.2. Order types that can make liquidity at their limit prices

We identify an order as a limit order if its TIF occurs at the end of the trading day (a Day order) and if it does not impose further restrictions on prices other than its limit price. Panel B of Table 3 provides summary statistics for four types of limit orders. A plain limit order includes no

Table 3
Summary statistics for order types, In this table, we report summary statistics for each order type, including the number of orders, number of trades, executed share volume, the market share of the executed share volume, average order size, and fill rates. The fill rates are calculated as the executed share volume divided by the submitted share volume. Conditional on order execution, we also present the percentage of the executed volume that takes local liquidity, routes to away markets, or makes liquidity locally. Multiplying the numbers reported in Column (8) by those reported in Column (10) will lead to the routed orders as a percentage of submitted share volume. We leave the fill rates for Good-Till-Cancel and Stop orders blank because some of them do not have expiration times. The average order size includes the unfilled portion of an order.

Classification (1)	Order Type (2)	Number of Orders (3)	Trades (4)	Executed Volume (5)	Average Order Size (6)	Market Share (7)	Fill Rate (8)	Take Local Liquidity (9)	Route (10)	Make Liquidity (11)
Panel A: Order types that take liquidity only										
Liquidity-taking	Plain market	6364,008	6364,008	1775,278,400	278.96	4.48%	100.00%	66.34%	33.64%	0.00%
	Plain IOC	9117,590	4196,808	1127,373,900	268.63	2.84%	19.48%	98.07%	1.93%	0.00%
	IOC ISO	84,834,619	25,410,600	6207,890,700	244.30	15.66%	22.48%	100.00%	0.00%	0.00%
	NMS IOC	28,223,979	14,652,500	4499,703,400	307.09	11.35%	31.86%	100.00%	0.00%	0.00%
Panel B: Order types that make liquidity at their limit price										
Orders that make liquidity	Plain limit	723,538,058	43,360,100	9560,971,500	220.50	24.12%	2.71%	29.78%	14.28%	55.94%
	DNS limit	951,053,287	44,243,300	8069,151,300	182.38	20.36%	3.72%	11.43%	0.00%	88.57%
	Reserve	110,936,921	6660,529	1917,804,300	287.94	4.84%	2.09%	31.09%	16.30%	52.62%
	DNS Reserve	61,385,431	4077,848	732,919,700	179.73	1.85%	1.56%	5.00%	0.00%	95.00%
Panel C: Other order types										
Auction	Market-on-close	897,083	897,083	3890,610,000	4336.96	9.82%	100.00%	100.00%	0.00%	0.00%
	Market-on-open	346,417	346,417	798,674,000	2305.53	2.02%	100.00%	100.00%	0.00%	0.00%
	Limit-on-open	24,351,806	373,525	289,195,000	774.23	0.73%	1.96%	0.00%	0.00%	100.00%
	Limit-on-close	1420,217	212,533	448,966,000	2112.45	1.13%	15.30%	0.00%	0.00%	100.00%
Low Market Share	Good-Till-Cancel	–	36,612	20,495,800	559.81	0.05%	–	11.30%	4.28%	84.42%
	Stop	–	150,566	50,090,400	332.68	0.13%	–	68.28%	31.65%	0.07%
	Primary Pegged	1483,135	1256,004	237,375,000	188.99	0.60%	78.12%	0.71%	0.01%	99.28%
	Buy Minus Zero Plus	20,598	13,263	2330,600	175.72	0.01%	45.88%	7.71%	12.91%	79.38%

SOIs and is subject to Rule 610 if it displays a quote that locks or crosses an away market and is subject to Rule 611 if it is marketable upon arrival. A trader can refuse these routings by adding a DNS tag to create a DNS limit order.

A reserve limit order adds a "non-display" instruction to a plain limit order. The trader must specify a published quantity ("PUBQTY") of the order, and the order will not be displayed for more than PUBQTY shares. Although they are not officially defined by the NYSE, orders with PUBQTY $= 0$ are usually called "hidden" orders and those with PUBQTY > 0 are called "iceberg" orders. Hidden orders are not subject to Rule 610 because they cannot *display* a quote that locks or crosses the market. Iceberg orders are subject to both Rule 610 and 611 because their displayed portions can still lock or cross the market. A DNS reserve limit order adds both DNS and non-display instructions to a plain limit order.

Although plain limit orders and reserve limit orders are willing to provide liquidity at their limit prices, they often take liquidity upon arrival. Plain limit orders take liquidity from the NYSE for 29.78% of their executed volume and take liquidity outside the NYSE for 14.28% of their executed volume. Consequently, plain limit orders provide liquidity for only 55.94% of their executed volume. Reserve limit orders make liquidity for 52.62% of their executed volume, take NYSE liquidity on the NYSE for 31.09% of their executed volume, and take outside liquidity for 16.30% of their executed volume.

DNS limit orders and DNS reserve limit orders are much more likely than routable limit orders to make liquidity, not only because they refuse to take liquidity from outside but also because they take a lower volume of liquidity on the NYSE. DNS limit orders take liquidity for only 11.43% of their executed volume, and the percentage is as low as 5.00% for DNS reserve limit orders. These results suggest that one factor that drives refusal of Reg NMS routing by limit orders is the preference for making but not taking liquidity.

Non-routable orders also dominate market shares for orders that are designed to provide liquidity. DNS limit orders capture a market share similar to that of plain limit orders (20.36% versus 24.12%). DNS reserve orders capture a market share of 1.85%, while reserve orders capture a market share of 4.84%. In total, non-routable orders account for 57% ($\frac{15.66+11.35+20.36+1.85}{100-9.82-2.02-0.73-1.13}$) of continuous (non-auction) trading volume.

3.3. Other order types

Panel C of Table 3 provides summary statistics for other order types in our sample. Traders can choose the TIF as AUC (auction) such that the order participates only in open/reopen/close auctions. As auction orders cannot participate in continuous trading, they are not subject to either Rule 610 or Rule 611. As we focus on evaluating the performance of exchange routing and explaining the incentive to refuse exchange routing, we exclude auction orders for the sake of brevity. We also exclude orders with Good-Till-Cancel TIFs because they are extremely rare (0.05%) and because we do not know the limit price of a Good-Till-Cancel order if the order is submitted before the first day of our sample period.

We exclude the three other order types listed in Panel C from further analysis because their combined volume represents less than 1% of the market share and because these orders are neither marketable nor have well-defined limit prices that enable us to calculate price improvements.

A stop order, also known as a stop-loss order, buys or sells once the price of a stock becomes worse than the stop price. Investors generally use a stop order to limit a loss or to protect profits on stocks that they own. Stop orders do not affect displayed limit order books unless their stop prices are triggered. Stop orders experience a route-out rate (31.65%) that is similar to that of market orders because most stop orders, once triggered, become marketable orders.

Primary Pegged orders track the bid or ask prices, and traders may modify the reference price by choosing a specific number of ticks away from the bid and ask price, such as buying at the bid price minus one tick or buying at the bid price plus one tick. A Primary Pegged order may also include a limit price so that the order no longer pegs the reference price once it moves outside the price limit. Because Primary Pegged orders in our sample are designed to peg the best bid or ask price, it is rare (0.01%) for a pegged order to be routed out.

"Buy-Minus-Zero-Plus" (BMZP) orders peg to a reference price that is the higher of the current bid price and the last trade price. This reference price dovetails with the requirements of SEC Rule 10b-18, which is designed to prevent price manipulation by discouraging firms from repurchasing their shares at prices that are higher than the highest independent bids or the most recent transaction prices, whichever is higher. In its SEC filing, the NYSE states: "*The BMZP instruction is designed to assist member organizations in their compliance with the 'safe harbor' provisions of Rule 10b-18 under the Act ("Rule 10b-18") for issuer repurchases.*" A BMZP order may also include a limit price as the upper bound of the purchase price, but its limit price is much less relevant than its reference price. We find that limit prices of BMZP orders are 18 bps above the NYSE ask price upon entry, which indicates they would be immediately marketable without the constraints of reference prices. Therefore, issuers are willing to pay higher prices than the bid price and the last transaction price, but the implicit price ceiling imposed by Rule 10b-18 constrains issuers' ability to repurchase shares (Li et al., 2021). BMZP orders include all three types of executions, although most executions come from making liquidity on the NYSE. A BMZP order realizes a 30-day return of 706 bps, consistent with the fact that repurchasing firms are informed in their stocks (Dittmar and Field, 2015). All but two BMZP orders accept Reg NMS routing, indicating that their users care more about filling their orders than paying exchange fees or achieving execution speeds measured in milliseconds.

4. Do Reg NMS routings improve prices for routable orders?

In this section, we examine whether Reg NMS routings improve prices for routable orders. The analysis provides the first evaluation of exchange linkages created by Reg NMS, which are designed to provide the best prices for liquidity-taking orders and encourage displays of liquidity-making orders on multiple markets (SEC 2005). Studying routable orders helps us understand why non-routable orders refuse Reg NMS routing.

We define a price improvement as the difference between an order execution price and its limit price, adjusted for the order direction. As market orders do not have limit prices, we use the best available NYSE price at the time of routing as the benchmark for a price improvement. We find that 62% of Reg NMS routings lead to the same gross prices and worse net prices. In Table 4 Panel A we itemize routing performances for routable order types.

We find that Rule 610 serves as the main driver of routings to worse net prices, as routing always improves prices for order types that are exempted from Rule 610. As plain IOC and market orders cannot lock a market, the NYSE routes them to comply with Rule 611 but not Rule 610. The results reported in Table 4 show that Reg NMS routings improve gross prices by at least one cent for plain IOC and market orders. As the fee difference is less than one cent, Reg NMS also routes plain IOC and market orders to better net prices.

Routing for plain limit orders and reserve limit orders are subject to both Rule 611 and Rule 610. The Rule 611 routing can be differentiated from Rule 610 routing by examining whether plain limit orders and reserve orders are marketable upon arrival. If they are marketable on the NYSE upon arrival, Rule 611 prohibits the NYSE from executing the order at a price that is worse than the NBBO. Rule 611 routings account for 13.57% (13.10%) of routings for plain (reserve) limit orders. Again, Rule 611 routings improve prices by at least one tick.

If plain limit orders and reserve limit orders are not marketable on the NYSE upon arrival, their routings are subject to Rule 610, which prohibits them from establishing quotes that lock or cross a quote on another exchange. Rule 610 accounts for 86.43% (86.90%) of routing of plain limit (reserve limit) orders. Among these Rule 610 routings, 8.95% (7.40%) experience price improvements of one cent or above because they resolve crossed markets. In sum, routing improves prices for only 22.51% (20.49%) of plain (reserve) limit orders. The other 77.49% (79.51%) of plain (reserve) limit orders are routed to the same gross prices to resolve locked markets.

Reserve orders of *zero* display size and reserve orders of *positive* display size enable us to further compare routings led by Rule 610 and Rule 611. We report the results in Panel B of Table 4. Because Rule 610 applies only to displayed quotes, reserve orders of *zero* display size cannot lock the market, and we find that they receive price improvements 99.95% (1 – 0.05%) of the time. Once the displayed size (PUBQTY) is non-zero, the case for zero price improvement rises dramatically, from 0.05% to 81.50%. Therefore, although one goal of Reg NMS is to encourage traders to display their quotes through Rule 611, Rule 610 creates incentives for traders to fully hide their quotes to avoid routing to worse net prices. Bessembinder and Venkataraman (2004), Bessembinder et al. (2009), and Chakrabarty et al. (2020) study the trade-off between dis-

Table 4
Price improvements by routing, In this table we report price improvements for routable orders. Panel A presents realized price improvements by order type. For all order types except market orders, price improvement is defined as the difference between the execution price and the limit price. Market orders do not have limit prices. For them, price improvement is defined as the difference between the execution price and the best available price on the NYSE. In Panel B, we report price improvements for reserve orders categorized by whether they can lock the market and whether they are routed out upon arrival. Reserve orders that are partially displayed (PUBQTY > 0, known as "iceberg orders") may lock the displayed market, so their routing is subject to both Rule 610 and Rule 611. Reserve orders with no displayed quantities (PUBQTY = 0, known as "hidden orders") do not lock the displayed market, so their routing is subject only to Rule 611.

Panel A: Price improvements for routed-out orders

	Plain market	Plain IOC	Plain limit	Reserve		
Improvement	Route by 611	Route by 611	Route by 610	Route by 611	Route by 610	Route by 611
0	0.00%	0.00%	77.49%	0.00%	79.51%	0.00%
0.01	62.28%	70.49%	4.41%	8.52%	3.58%	5.85%
0.02	15.84%	12.00%	1.02%	1.72%	1.46%	2.06%
0.03	6.24%	4.52%	0.60%	0.75%	0.71%	0.97%
0.04	2.86%	2.46%	0.48%	0.48%	0.50%	0.82%
0.05	1.52%	1.59%	0.53%	0.42%	0.46%	0.87%
0.05+	11.27%	8.94%	1.91%	1.68%	0.69%	2.53%
Total	100%	100%	86.43%	13.57%	86.90%	13.10%
#Obs	597,203,600	21,758,300	1365,306,700	312,544,900		

Panel B: Price improvements for hidden and iceberg orders

Improvement	Reserve orders		
	PUBQTY = 0	PUBQTY > 0 AND Route Out Upon Arrival	PUBQTY > 0 AND Route Out After Posted
0	0.05%	81.50%	76.88%
0.01	20.10%	9.16%	11.61%
0.02	6.63%	3.45%	3.37%
0.03	3.79%	1.63%	0.89%
0.04	17.77%	0.90%	0.60%
0.05	28.23%	0.66%	0.10%
0.05+	23.42%	2.71%	6.55%
Total	100%	100%	100%
#Obs	7562,200	304,804,200	178,500

playing and hiding quotes on one exchange. Reg NMS routing and fees create a trade-off between displaying and hiding quotes on linked exchanges. Hidden orders may fail to interact with liquidity demanders on other exchanges through Rule 611, but hidden orders may also reduce the risk of locking the quotes of liquidity providers on other exchanges.

Surprisingly, we find that a reserve order with a positive display size can also lock the market during the refill process. For example, on July 1, 2010, a reserve sell limit order of 10,000 shares for ELY (Callaway Golf Co.) established, at 9:43:10.100, an ask price at $6.04 using its display component of PUBQTY = 100 shares. The order was executed and refilled on the NYSE until the displayed 100 shares were consumed at 9:50:47.752. Three milliseconds later, the NYSE routed the order to NASDAQ and the Boston Stock Exchange to take liquidity at $6.04 instead of refilling the order. Note that the NASDAQ and Boston Stock Exchange bids occur only within the three-millisecond window; otherwise, the NASDAQ and Boston Stock Exchange bids would have locked the NYSE ask and have been routed to the NYSE. Therefore, the reserve sell limit order is routed out as a result of the latency in refilling the shares. We find that 76.88% of the routing orders led by refilling experience no price improvement, as shown in the last column of Panel B of Table 4.

Routings during refills provide a clean environment in which to examine the impact of exchange latency. Holding all else equal, Rule 610 implies that a slow exchange tends to lock a fast exchange. It is extremely difficult to test this hypothesis because (1) it is difficult to separate trader latency from exchange latency, and (2) it is difficult to obtain data from two exchanges with perfectly synchronized timestamps. Routings during refills help us surmount these two challenges because (1) refills involve only exchange latency but not trader latency, and (2) we know that exchange latency causes these routings because a refilled order cannot be routed out if exchange latency is zero. Although routings during refills are rare events, affecting only 0.052% of the volume, the intuition that a slow exchange tends to lock a fast exchange should hold more generally. Although Reg NMS does not impose time priority across exchanges on the same side of the book (Foucault and Menkveld 2008), our results indicate that Rule 610 imposes time priority across exchanges on the opposite side of the book. Under Rule 610, exchange latency costs order submitters, because their orders are routed to worse net prices; exchange latency also costs slow exchanges, because they need to pay take fees to fast exchanges and also lose trading volume. Therefore, Rule 610 incentivizes stock exchanges to be fast.

We proceed with our analyses as follows. In the main text below, we show that fees provide an incentive for traders to refuse Rule 610 routing. We defer the discussion of Rule 611 routing refusal to the Appendix A, for two reasons. First, because Rule 611 routing always improves the net price, the incentive to refuse Rule 611 routing arises not from price but from technical details involved in reducing the latency under specific market conditions. We defer the discussion of these technical details to the Appendix a to focus our paper on the economic driver of the fees. Second, and more importantly, because Rule 611 always improves prices, it creates no incentive for exchanges to further refine order types that refuse 611 routings. Most new order types aim to provide innovative ways to comply with Rule 610, and we summarize these innovations in Section 7.

5. An incentive to refuse Rule 610 routing: fees

In the previous section, we show that Rule 610 tends to route orders to the same gross prices or worse net prices after fees. Therefore, fees provide a natural incentive for traders to refuse Rule 610 routing. DNS limit orders earn a small profit of 1.57 bps after collecting rebates but would have lost 0.89 bps if they had paid routing fees. The negative profit led by routing fees gives DNS limit orders incentives to cancel orders if they need to take liquidity from another exchange. Next, we test the hypothesis that DNS orders refuse Reg NMS routing to avoid fees.

We test our hypothesis in the cross-section, in time series, and through difference-in-differences. In the cross-section, the market share captured by DNS orders should be greater with low-spread stocks because the difference between a making rebate and a routing fee is a fixed dollar amount per share. When the bid–ask spread is narrow, the fee difference represents a larger portion of the spread, and the incentive to collect rebates should be more pronounced for those stocks. Also, narrow bid–ask spreads make it easier for traders to lock or cross a market when they improve quotes. In time series, when the NYSE increases maker rebates, the incentive to collect rebates strengthens and we should expect to observe an overall increase in DNS usage. Finally, the difference-in-differences test should show that an increase in maker rebates increase DNS usage to a greater extent for narrow bid–ask spread stocks than for wide bid–ask spread stocks. We discuss the institutional details associated with exchange fees in Section 5.1 and then report our empirical tests of the three predictions in Section 5.2.

5.1. Shocks to fee differences

U.S. stock exchanges charge heterogeneous fees for orders that make liquidity, take liquidity locally, and route to other exchanges. Most exchanges charge liquidity takers and offer rebates to liquidity makers. At the beginning of our sample period, the NYSE charges a take fee of 0.18 cents per share and it provides rebates of 0.10 cents per share for the basic tier of liquidity makers.[15] The NYSE earns 0.18 − 0.10 = 0.08 cents per share. When the NYSE routes an order to another exchange, it takes liquidity from that exchange and pays a take fee to that exchange. Reg NMS mandates a cap for a take fee of 0.30 cents per share for all exchanges. In our sample period, some exchanges, such as NASDAQ, NYSE Arca, and BYX, charge a take fee of 0.30 cents per share, which equals the fee cap. The NYSE charges a routing fee of 0.30 cents per share so that it will not lose money by routing orders out. For a trader, the fee difference between making liquidity on the NYSE and taking liquidity from another exchange is 0.40 (= 0.10 + 0.30) cents per share.

We test our empirical predictions with two shocks to the NYSE rebates in our sample period. On May 1, 2010, the NYSE increased the makers' basic rebate from 0.10 cents per share to 0.13 cents per share and further increased it to 0.15 cents per share on January 3, 2011.[16] These two shocks increased the difference between the make rebates and the routing fees from 0.40 cents to 0.43 cents and 0.45 cents per share, in turn.

5.2. Empirical tests of dns order usage

To assess the shock's heterogeneous impact on stocks, we partition our sample into three groups based on each stock's nominal bid–ask spread in January 2010. The narrow-spread group consists of stocks with a time-weighted average bid–ask spread that is narrower than 1.39 ticks (the 33rd percentile). The wide-spread group consists of stocks with a spread wider than 2.37 ticks (the 67th percentile). There are roughly 36 stocks in each group. We pool DNS limit orders and DNS reserve orders together and call them DNS orders. We define DNS market share as executed shares from DNS orders divided by the total volume of all four types of orders that can provide liquidity. The sample period spans one month before and after the fee change.

Table 5 shows the test results pertaining to the three empirical predictions. First, the results reported in the first and the second columns of Panels A and B show higher DNS market shares in narrow-spread stocks than in wide-spread stocks. The difference is about 10 percentage points, and this pattern holds for all the months in our sample period. Second, the results reported in the first rows of Panels A and B show that the increases in makers' rebates lead to larger DNS market shares. After the rebate increase of May 2010 (Jan 2011) of 0.03 (0.02) cents per share, DNS market shares increase by 2.28 (1.22) percentage points. Third, the results reported in the last row of Panels A and

[15] Higher-volume tiers earn higher levels of rebates. For example, at the beginning of our sample period, the rebate was 0.17 cents per share for the highest-volume tier.

[16] In general, the rebate changes are similar for all volume tiers. For example, for the highest volume tier, the rebate increased from 0.17 cents per share to 0.20 cents per share on May 1, 2010, and further increased to 0.22 cents per share on January 3, 2011. The only exception is that rebates for Designated Market Makers (DMMs) did not change. Nevertheless, the market share of DMMs is small in normal market conditions (Clark-Joseph et al., 2017). Also, the changes do not reduce rebates for any traders. The detailed fee changes are documented in NYSE filings 34-62082 and 34-63642, in turn. January 1– 2, 2011 are not trading days, so the after period is the entire trading days of January 2011. Around our event windows, there were no concurrent fee changes in major competing exchanges, i.e., NASDAQ, BZX, and NYSE Arca.

Table 5

DNS limit-order usage and sensitivity to fee changes, This table presents estimates of changes in average daily DNS order market shares around two NYSE make fee changes. Panel A presents changes in DNS market shares from April 2010 to May 2010. Panel B presents changes in DNS market shares from December 2010 to January 2011. The DNS market share of a stock month is the sum of liquidity-making orders' executed shares with DNS tags divided by the total volume of all liquidity-making orders. The DNS market share is calculated for low-, medium-, and high-spread stocks as well as the difference between the low and high groups of stocks over the two-month period. The partitions are based on each stock's average nominal bid–ask spread in the baseline month of January 2010. Heteroskedasticity-consistent standard errors are reported in parentheses. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

DNS limit orders' average market share			
Panel A: NYSE make fee change on May 1, 2010			
	Before (April 2010) (21 days)	After (May 2010) (20 days)	After – Before
Full Sample	33.05	35.33	2.28***
(109 stocks)			(0.51)
Low Spread	38.11	40.98	2.86***
(36 stocks)			(0.77)
Medium Spread	35.53	38.84	3.31***
(36 stocks)			(0.60)
High Spread	28.09	28.32	0.24
(37 stocks)			(0.87)
Low - High	10.03***	12.65***	2.62**
	(1.51)	(1.48)	(1.17)
Panel B: NYSE make fee change on Jan 3, 2011			
	Before (Dec 2010) (22 days)	After (Jan 2011) (20 days)	After – Before
Full Sample	35.65	36.87	1.22**
(106 stocks)			(0.50)
Low Spread	40.07	43.26	3.20***
(35 stocks)			(0.71)
Medium Spread	38.30	39.81	1.51***
(35 stocks)			(0.57)
High Spread	30.26	30.91	0.65
(36 stocks)			(1.02)
Low - High	9.81***	12.35***	2.55**
	(1.52)	(1.49)	(1.24)

B show that the increases in makers' rebates lead to heterogeneous treatment effects. Narrow-spread stocks experience an increase in DNS market shares of 2.86 (3.20) percentage points, while wide-spread stocks experience only an increase of 0.24 (0.65) percentage points. The market share of DNS limit orders increases 2.62 (2.55) percentage points more for stocks whose bid–ask spreads fall into the lowest tercile than for stocks whose bid–ask spreads fall into the highest tercile.

6. Clientele segmentation in routing decisions

The next question is why some orders accept while others refuse routings. We find that client segmentation explains heterogeneous routing choices. In Section 6.1, we show that HFTs are likely to be dominant users of DNS orders. In Section 6.2, we show that DNS orders are more likely to improve the bid–ask spread than non-routable orders. In Section 6.3, we show that routable orders incur much larger opportunity costs for non-fills than DNS orders. Finally, we summarize our results in Section 6.4. Taken together, our results indicate the following clientele segmentation of stock exchange order types: HFTs refuse routings to reduce fee costs in providing liquidity, whereas non-HFTs accept routing to increase their fill rate.

6.1. DNS orders win speed races to provide liquidity

The literature shows that two types of agents aim to collect rebates and avoid fees. The first type comprises HFTs that aim to profit from market-making (Brogaard et al., 2014). The second type comprises brokers who aim to maximize rebates but not execution quality for their customers (Battalio et al., 2016). Although our data do not include the I.D.s of order submitters, we can still provide some evidence pertaining to the main clientele of non-routable orders. Even if we cannot "see" the driver of a vehicle, we can obtain a noisy signal of the driver's information by observing the car. In Section 6.1.1., we show that DNS orders win speed races to secure front-of-the-queue positions. In Section 6.1.2., we show that DNS orders win speed races to cancel stale quotes. Taken together, what we report in this section indicates that the main clientele of DNS orders consists of HFTs.

6.1.1. Speed races to establish front queue position

Li et al. (2021) model speed races to provide liquidity. Because Rule 612 of Reg NMS imposes a one-cent tick size (a minimal price variation) on any stock traded at a price above $1, the discrete price constrains price competition and creates rents for liquidity provision. Speed becomes essential in capturing such rents because orders at

Table 6

Speed races to establish front-of-the-queue positions and cancel stale quotes, This table presents speed races between liquidity-making orders. In Panel A, we present speed races that are run to establish front-of-the-queue positions. The first responders are non-marketable limit orders that refill liquidity at the same price as the price where the liquidity is consumed. To ensure that a responding limit order is triggered by a liquidity-taking order, we require that the refill order be submitted within 0.1 s. The average number of first responders per day is reported in the second column, and in the following columns we report the average numbers of first responders per day by order types. The market shares of responding orders are reported under the counts. In Panel B, we report the outcomes of speed races between those that are run to snipe and those that are run to cancel stale quotes. We track the outcomes of all outstanding limit orders 0.1 s before the first IOC snipers arrive. The escape rate is the proportion of limit orders that are canceled before sniping. All numbers are averaged at the stock-day level.

Panel A: Establishing front-of-the-queue positions

Average	First responder count per day				
trades per day	Total races	Plain limit	DNS limit	Reserve	DNS Reserve
1652	523 (31.7%)	121 23.1%	374 71.5%	11 2.1%	17 3.3%

Panel B: Canceling stale quotes

	Plain limit	DNS limit	Reserve	DNS Reserve	Total
Number of outstanding limit orders 0.1 s before races	106.00	274.37	11.91	20.46	412.74
Sniped in races	96.17	157.08	11.77	17.23	282.25
Escape from sniping	9.83	117.29	0.14	3.23	130.49
Escape rate	9.27%	42.75%	1.18%	15.79%	31.61%

the front-of-the-queue position enjoy higher execution priority (Yao and Ye (2018)).

We measure such speed races following Li et al. (2021). The model shows that traders continue to add shares to the queue until the marginal profit equals zero. A new profitable queue position opens when a marketable order executes with existing orders in the queue. Therefore, we examine which order type is more likely to react after a marketable order moves the queue forward. We define the first responder to a buyer-initiated (seller-initiated) trade as the sell (buy) limit order that (1) is submitted within 0.1 s after a transaction, (2) is not marketable, (3) whose limit price is identical to the previous trade price, and (4) is such that no race to take liquidity occurs in the following 0.1 s.[17] If multiple orders satisfy conditions (1)–(4), we take the first arriving order as the first responder. For example, on January 22, 2010, a marketable limit sell order of 100 shares for XCO arrived at the NYSE at 12:27:18.51, executed at the current bid price of $20.00. The next message the NYSE received was a DNS limit buy order of 100 shares at $20.00, and it refilled the best bid price at 12:27:18.53. The next message was a plain limit buy order of 100 shares at $20.00, which arrived at 12:27:18.56. It arrived 0.03 s later than the DNS order, so it was placed later in the queue at the NYSE best bid price. No race to take liquidity occurred in the following 0.1 s. We consider the DNS limit buy order in this example to be the first responder.

The results reported in Panel A of Table 6 show that first responders are more likely to be DNS limit orders. On average, $\frac{523}{1652}$ =31.7% of liquidity-taking orders are followed by orders to refill at the same prices within 0.1 s. DNS limit orders win races $\frac{374}{523}$ =71.5% of the time, while Plain limit orders win races $\frac{121}{523}$ =23.1% of the time, suggesting that DNS limit orders are more likely to achieve front-of-the-queue positions. We also find that DNS reserve limit orders are more likely to achieve front-of-the-queue positions than reserve limit orders (3.3% versus 2.1%), although both carry a small fraction of first responders. Indeed, because hidden orders are assigned lower execution priority than displayed orders, neither type of reserve limit order is likely to be the first responder to a favorable queue position.

6.1.2. Speed races to cancel stale quotes

In this section, we show that DNS orders are more likely than routable limit orders to cancel stale quotes. Liquidity-providing orders become “stale” when the fundamental value changes (Copeland and Galai 1983; Foucault et al., 2003; Budish et al., 2015). Thus, liquidity providers race to cancel stale quotes while other traders race to snipe stale quotes. We define speed races aimed to snipe or cancel stale quotes following Aquilina et al. (2021). We require every race to include the arrival of two or more IOC orders at the same stock, side, and price level within 0.1 s. At least one IOC order fails to execute (the “loser”). We then require some liquidity-providing orders to display liquidity 0.1 s before the first IOC order arrives, and these liquidity-providing orders are the targets of sniping. These orders would execute if they did not cancel. A limit order wins the speed race if it successfully cancels.[18] The results reported in Panel B of Table 6 show that 412.74 orders are targets of sniping on an average stock day. Among these orders, 282.25 are sniped, while 130.49 (31.61%) orders escape from sniping during the races.

[17] We add this requirement because Li et al. (2021) show that investors may use aggressive limit orders to stimulate HFTs to demand liquidity. Investors lose money by providing liquidity, but they choose to stimulate HFTs if the cost is lower than paying the bid–ask spread.

[18] We require at least two IOC orders to participate in a race, which helps rule out other drivers of cancellation. For example, HFTs can cancel their limit orders because they want to back-run (Yang and Zhu (2020)) or front-run (Baldauf and Moller (2020)) IOC orders from institutional traders. It is unlikely, however, that two institutional traders arrive at the same time and submit IOC orders at the same price purely by chance (Menkveld 2018).

Table 7
Price aggressiveness in the limit order book, In this table we report price aggressiveness of plain limit, DNS limit, reserve, and DNS reserve limit orders. These orders can either be marketable or add liquidity to the limit order book. For marketable orders, the most aggressive orders walk the book and take liquidity at multiple prices, followed by orders that take liquidity at one price. For non-marketable orders that add liquidity to the limit order book, the most aggressive orders aim to improve the NYSE BBO, followed by orders at the BBO, and then orders away from the BBO (1 tick from the BBO or > 1 tick from the BBO). An order can lock the market only when it aims to improve the BBO on the NYSE. That is, it can lock the market when it is more aggressive than existing limit orders on the NYSE but not aggressive enough to take liquidity from the NYSE.

Category	Plain limit	DNS limit	Reserve	DNS Reserve
Trade at multiple prices	0.04%	0.00%	0.06%	0.00%
Trade at one price	4.69%	4.30%	2.10%	3.38%
Improving BBO by > 1 tick	0.93%	2.08%	0.98%	3.73%
Improving BBO by 1 tick	1.76%	9.79%	1.79%	5.19%
Order Placement at BBO	6.86%	23.95%	5.49%	18.12%
Order 1 tick from BBO	4.71%	10.44%	2.81%	6.79%
Order > 1 tick from BBO	81.02%	49.46%	86.78%	62.80%

The results reported in Panel B of Table 6 show that DNS limit orders are more likely to be targets of stale-quote sniping: when the race starts, 274.37 DNS limit orders are on the limit-order book, but 117.29 DNS limit orders successfully cancel, leading to an escape rate of 42.75%. The escape rate for plain limit orders is only 9.27%, and the number drops as low as 1.18% for reserve limit orders. DNS limit orders successfully cancel 42.75% of orders before sniping, indicating that they are more likely to be placed by HFTs. Their decisions to cancel also indicate that DNS limit orders are less likely to be placed by brokers who aim to maximize rebate revenues because canceling orders reduces the rebate revenue for brokers.

6.2. DNS orders are more likely to improve liquidity than routable limit orders

In this section, we find evidence that DNS orders are more likely to improve the bid–ask spread than their routable counterparties. This result has two implications. First, DNS orders are more likely to be placed by traders who aim to provide liquidity. Second, because DNS orders aim to improve the local bid-ask spread, the risk that they lock the market is higher than their routable counterparties.

In Table 7, we show the price aggressiveness of non-routable orders relative to their routable counterparties. The most aggressive orders walk the book and take liquidity at multiple prices, followed by orders that take liquidity at one price. The most aggressive non-marketable orders aim to improve the NYSE BBO, followed by orders at the BBO and then orders away from the BBO. An order can lock the market only when it aims to improve the BBO on the NYSE. That is, it can lock the market when it is more aggressive than existing limit orders on the NYSE but not aggressive enough to take liquidity from the NYSE.

The results we report in Table 7 show that DNS limit orders are more likely to improve the best bid and ask on the NYSE than plain limit orders (9.79% + 2.08% vs. 1.76% + 0.93%), and DNS reserve limit orders are more likely to improve the quotes than reserve limit orders (5.19% + 3.73% vs. 1.79% + 0.98%). Also, DNS limit orders are more likely to be at the BBO on the NYSE than plain limit orders (23.95% vs. 6.86%), and DNS reserve limit orders are more likely to be at the BBO on the NYSE than reserve limit orders (18.12% vs. 5.49%). These comparisons indicate that non-routable orders are more likely than routable orders to provide liquidity.

Among DNS limit orders that improve prices, we find that 82% ($\frac{9.79}{2.08+9.79}$) improve the price by only one tick. This choice is natural not only because one tick is the least costly way to improve the quotes but also because the average spread in our sample is two ticks, leaving little room for improving quotes by two ticks. Finally, the following economic reasoning shows that routing cannot improve the price for an order that improves a local quote by one tick. We illustrate its economic mechanism in Fig. 2.

Suppose that the NYSE best bid is \$5.00 and the ask price is equal to or higher than \$5.02 (for the sake of brevity, the ask price is not included in the figure). Consider a buy limit order that improves the NYSE best bid to \$5.01. Notice that the order is not subject to Rule 611 because the NYSE cannot execute the order. The order is subject to Rule 610 if it locks or crosses the best ask price on another exchange. The buy limit order can receive price improvement only if the ask price on other exchanges is \$5.00 or lower. Such an ask price should not exist because Rule 610 would require the ask price of \$5.00 or lower to transact with the NYSE bid of \$5.00 to resolve the crossed or locked market. Therefore, when the buy limit order is routed outside, it will take liquidity at \$5.01, experience no price improvement, and yet pay the routing fee and lose potential rebates for making liquidity on the NYSE.

6.3. Routable limit orders incur higher opportunity costs for non-fills

In the previous two sections we show that DNS orders are more likely to be placed by HFTs, for whom routing refusal may help them to reduce the fee costs for providing liquidity, than to be placed by other traders. The next question is why not all traders choose to refuse routings. One

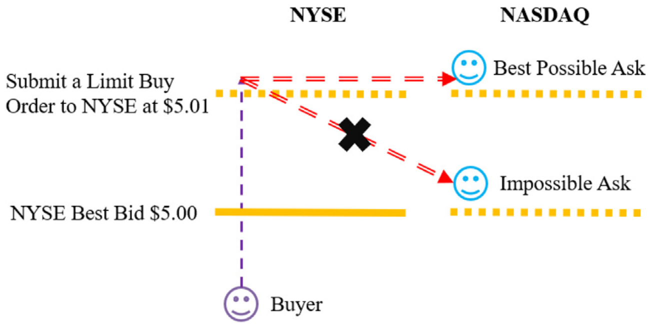

Fig. 2. Routing cannot improve prices for limit orders that improve local quotes by one tick, This figure illustrates why most limit orders do not obtain price improvements when routed. When an order improves the local price by only one tick, it is impossible to obtain any price improvements when being routed. For example, suppose the current NYSE best bid is $5.00 and that a limit buy order attempts to improve on the NYSE bid price by one tick, i.e., the order has a limit price of $5.01. If the NYSE routes the order to NASDAQ, it must be the case that NASDAQ has an ask price equal to or lower than $5.01. However, the NASDAQ cannot have an ask price that is equal to or lower than $5.00. Otherwise, the NASDAQ would have violated Rule 610(d): it locked/crossed the NYSE bid of $5.00. Therefore, the best possible ask price on the NASDAQ is $5.01. If the NYSE $5.01 limit buy order has been routed out, it would not receive any price improvements. Solid horizontal lines represent existing limit orders, and dashed horizontal lines represent limit prices of incoming orders or hypothetical orders. Single-dashed arrows represent order flow and double-dashed arrows represent routing.

Table 8

Opportunity costs for non-fills, In this table, we present the opportunity costs for non-fills of routed limit orders and Reg NMS Rejected DNS orders as well as the associated statistical inferences. For routed limit orders, the opportunity cost for non-fills is the difference between the fill price and a future benchmark price. For Reg NMS Rejected DNS orders, the opportunity cost for non-fills is the difference between the limit price (capped with the NYSE local BBO) and a future benchmark price. In Columns (1)–(7), the benchmark prices are the midpoint prices 1 second (3 s; 5 s; 10 s; 30 s; 1 min; 5 min) following the order submission. In Columns (8)–(14), the benchmark prices are closing prices from the end-of-day to 30 days following the order submission. Numbers in bold are statistically significantly different from zero at the 5% level. Standard errors are clustered at day and ticker level.

	#Obs	1sec (1)	3sec (2)	5sec (3)	10sec (4)	30sec (5)	1min (6)	5min (7)	EOD (8)	2days (9)	3days (10)	5days (11)	10days (12)	20days (13)	30days (14)
Routed Plain Limit	1,040,912,900	**1.69**	**1.97**	**2.12**	**2.27**	**2.48**	**2.46**	**2.44**	**4.13**	**6.81**	**7.42**	**9.1**	**4.22**	**8.89**	**15.35**
Reg NMS Rejected DNS Limit	7,872,641,300	**−0.27**	−0.05	0.03	0.09	**0.17**	**0.13**	**−0.22**	**0.95**	0.88	−1.37	**−4.3**	**−7.45**	**−7.11**	**−8.38**
Difference	–	**1.96**	**2.02**	**2.09**	**2.18**	**2.31**	**2.33**	**2.66**	**3.19**	**5.93**	**8.79**	**13.40**	**11.67**	**16.00**	**23.73**
Routed Reserve Limit	158,639,500	**1.99**	**2.27**	**2.46**	**2.73**	**3.12**	**3.26**	**2.71**	**5.71**	**8.27**	6.66	0.79	10.28	**24.92**	**56.48**
Reg NMS Rejected DNS Reserve Limit	963,796,900	**−5.63**	**−5.52**	**−5.48**	**−5.43**	**−5.27**	**−5.13**	**−5.38**	**−6.98**	−6.45	−5.16	−3.01	2.59	2.99	−4.11
Difference	–	**7.62**	**7.79**	**7.94**	**8.16**	**8.39**	**8.39**	**8.09**	**12.69**	**14.72**	**11.82**	3.8	7.69	**21.93**	**60.59**

possibility is that users of routable limit orders are naïve when making order choices and their brokers do a disservice by accepting routing. Although we cannot completely rule out this hypothesis, we find that heterogeneous opportunity costs for non-fills provide one rationale for the heterogeneous routing choices.

Consider an order that is filled on another exchange; the order would be rejected if its submitter chose the DNS option. To fill the order, the order submitter then needs to resubmit the order later. Therefore, the opportunity cost of refusing routings is the difference between the fill price and a future benchmark price.[19] Table 8 presents the results based on the benchmark prices at various horizons. In Columns (1)–(7), the benchmark prices are the midpoint prices 1 s (3 s; 5 s; 10 s; 30 s; 1 min; 5 min) following the order submission. In Columns (8)–(14), the benchmark prices are closing prices from the end-of-day to 30 days following the order submission. Intuitively, a high opportunity cost implies that the order submitter is "correct" in accepting Reg NMS routing because the order submitter would fill the order at a higher cost if she had rejected routing.

The results reported in Table 8 show that routed orders incur opportunity costs that are significantly higher than the fees. For example, at the end of the day, the opportunity cost for routed limit orders is 4.13 bps, which is statistically significantly different from 0 at the 5% level and much higher than the average fees (2.11 bps). The results for routed reserve orders are even higher. For example, the opportunity cost is as high as 5.71 bps if we use the end-of-day closing price as the benchmark. Therefore, accepting routing is rational because these routed orders would cost more if their submitters refused routing and resubmitted orders later.

Interestingly, we find that it is also rational for DNS orders to refuse routing because they incur low opportunity costs for non-execution. To calculate the opportunity costs to DNS orders for routing refusals, we collect all DNS orders rejected by Reg NMS. These orders would be filled if they were to accept routing. We assume that these orders would be filled at their limit price or current BBO,

[19] Alternatively, our measure is the implementation shortfall in Perold (1988). Because an order gets no fill if it rejects routing, our implementation shortfall includes opportunity costs but not execution costs.

whichever is better.[20] The Table 8 results indicate that the opportunity cost for DNS limit orders is close to 0. For example, the opportunity cost is only 0.95 bps before the fee and is statistically insignificant. The opportunity costs for DNS reserve limit orders are actually negative. This result reflects in part the fact that DNS reserve limit orders tend to improve the price more aggressively when improving the bid-ask spread. Among all price-improving orders, only 17.5% of DNS limit orders ($\frac{2.08}{2.08+9.79}$, Table 8) improve the price by more than one tick, a percentage that runs as high as 41.8% ($\frac{3.73}{3.73+5.19}$) for DNS reserve limit orders. The result that reserve orders improve prices more aggressively is consistent with Bessembinder et al. (2009).

The Table 8 results also reveal that the difference in opportunity costs between routed limit orders and rejected DNS orders increases over time. For example, the opportunity costs for routed limit orders are 1.96 bps higher than those for rejected DNS orders at the one-second horizon, a difference that increases to 2.66 bps at the 5-minute horizon and 5.71 bps at the end-of-day horizon. The time-series pattern again indicates that submitters of routed limit orders should fill their orders quickly by accepting routings. Therefore, heterogeneity in the opportunity costs for non-execution rationalizes routing choices.

In summary, we do not find evidence that routed orders are naïve. Instead, their high opportunity costs indicate that they are more likely to be placed by informed traders.[21] For example, the opportunity costs for non-fills of reserve limit orders are as high as 56.48 bps at the 30-day horizon, indicating large price movements in their direction.

6.4. Summary: clientele segmentation

In summary, our results provide interrelated evidence that clientele segmentation drives routing decisions. DNS orders are more likely to be placed by HFTs who aim to provide liquidity than to be placed by slow traders. These DNS orders entail low opportunity costs, probably because liquidity-providing orders are less likely to involve directional bets. The low opportunity costs make the fee costs more salient and incentivize HFTs to refuse routing. Users of routable orders entail high opportunity costs for non-execution, which incentivizes them to increase their fill rates and execute their orders by accepting routing.

7. Evolution of order types and policy implications

Fees and clientele segmentation not only rationalize the use of non-routable orders but also rationalize the invention of new order types. Our paper covers order types on the NYSE from January 1, 2010 through March 1, 2011, but its economic intuition helps to explain order types on other exchanges as well as the evolution of order types. In Section 7.1, we show that one main driver of order-type innovations is the effort to prevent orders that aim to provide liquidity from taking liquidity and paying the fees. One particular focus of these innovations involves bypassing Rule 610, which often routes orders to worse net prices.[22] Therefore, Rule 610 and the continual effort to bypass Rule 610 provide insights into the design of market structure in fragmented markets. We summarize the policy implications of our paper in Section 7.2.

7.1. Evolution of order types

In our sample period, the only way to refuse Rule 610 routing is to use DNS orders. DNS limit orders suffer from two limitations. First, these limit orders refuse to pay routing fees, but they are still subject to take fees, and we find that 11.43% of DNS limit orders take liquidity from the NYSE. In 2014, the NYSE introduced add-liquidity-only (ALO) orders, which refuse to take liquidity on any exchange, including the NYSE. In Section 7.1.1 we introduce two generations of ALOs. Second, DNS limit orders comply with Rule 610 by canceling orders. To avoid unnecessary cancellations, the NYSE and other U.S. exchanges invented a series of new order types to comply with Rule 610. In Section 7.1.2, we summarize these new order types with respect to four product lines: slide orders, Hide Not Slide orders, DAY ISOs, and the complicated but nowadays widely used variant—ALO DAY ISOs. The names of order types on other exchanges may vary, but their economic functions are similar. The combination of order types discussed in Sections 7.1.1 and 7.1.2 address unnecessary liquidity-taking and unnecessary canceling, and they superseded DNS limit orders on the NYSE in 2016.

7.1.1. Two generations of ALO orders that avoid unnecessary liquidity-taking

ALO orders. The difference between the take fee and the make rebate is smaller than the difference between the routing fee and the make rebate. Still, traders who want to avoid routing fees may also want to avoid take fees.

[20] We use hypothetical fill prices for rejected DNS orders because our data do not include real-time prices on away markets. We expect very few DNS orders to obtain price improvements if they were routed outside. Table 4 results indicate that 77.49% of plain limit orders and 79.51% of reserve limit orders fail to obtain price improvements. That number should be much higher for DNS orders, for two reasons. First, Table 7 results show that DNS orders are more likely to improve prices by only one tick, and Fig. 2 shows that routing would not improve prices for orders that improve the price by one tick. Second, Table 7 shows that DNS orders almost never execute at multiple prices, indicating that their submitters excel in monitoring the market and controlling the prices and sizes of their orders. Therefore, it is very unlikely that DNS submitters reject routing if there are true price improvements. Nevertheless, the result that DNS orders entail lower opportunity cost still holds even if they are equally likely to obtain price improvements over their routable counterparts. For an apples-to-apples comparison, we assume that all routed orders are filled without price improvements. The opportunity cost for routed limit (reserve) orders then reduces to 3.07 (4.78) bps at the end-of-day horizon. Therefore, assuming a worse fill price would reduce the opportunity costs, but not by much. The underestimation for DNS orders should be even smaller because it is even less likely that they obtain price improvements than it is that routable orders obtain such improvements.

[21] Certainly, some orders need to lose money if routable limit orders are informed. Li et al. (2021) show that market orders and good-till-cancel limit orders are more likely to uninformed traders.

[22] In contrast, there are almost no efforts to further refine order types that bypass Rule 611, as the rule routes orders to better net prices. We show in the Appendix A that order types that refuse Rule 611 routings aim to solve a technical issue in latency.

Therefore, in 2014, the NYSE introduced ALOs. The first-generation ALO orders refuse to take liquidity at any price. The NYSE later updated ALO orders so that they can take liquidity or be routed out if they receive a price improvement of at least one tick.[23] In other words, although the name of the order type ("add-liquidity-only") suggests that it is designed only to add liquidity, this second-generation ALO order would take liquidity when a price improvement is higher than the fee difference.

The introduction of ALO orders and the later update of such orders follow two economic mechanisms revealed by our paper. First, liquidity-making orders refuse to take liquidity at their limit prices because of higher cum-fee costs. Second, a 1-cent price improvement overwhelms the fee differences. The updated ALO accepts Rule 611 routings but still refuse Rule 610 routings that worsens the net price.

7.1.2. Four new order types that avoid unnecessary cancelation

Rule 610 prohibits an exchange from locking quotes on other exchanges, yet it does not instruct exchanges regarding how to unlock markets. DNS orders unlock the market through cancelation. However, cancelation is costly to HFTs because they need to resubmit orders and may lose front-of-the-queue positions. Later, the exchanges invented order types that provide innovative ways to unlock markets without canceling orders.

Slide orders. A slide instruction reprices (i.e., slides) a quote if it locks the quotes on another exchange. For example, an NYSE limit buy order that aims to improve the local bid to $5.01 would lock the market if the NBO is also $5.01. As a DNS instruction cancels the order, its submitter needs to resubmit the order at a less aggressive price or to wait until the NBO moves up. The slide instruction implements these two functions automatically by first repricing the limit order at $5.00 so that it does not lock the market. If the market ticks up to an NBO of $5.02, the slide instruction would reprice the order to buy at $5.01. The NYSE introduced slide orders in 2016, and NASDAQ and the CBOE also offer slide orders.[24]

Hide Not Slide orders. In Section 6.1, we show that DNS limit orders win the majority of speed races to secure front-of-the-queue positions. In Section 4, we show that Reg NMS allows completely hidden orders to lock a market because they are not displayed quotations. The Direct Edge Exchange introduced Hide Not Slide orders that enable their submitters to both hide orders and win front-of-the-queue positions. A Hide Not Slide instruction hides an order if it locks a market. Once the market unlocks, this order lights up. These orders enjoy time priority over slide orders because the timestamps for Hide Not Slide orders are order initial-entry times while the timestamps for slide orders are order reprice (slide) times. A 2012 *Wall Street Journal* article claims that Hide Not Slide orders can jump ahead of other orders in the queue.[25] Direct Edge describes a single "price sliding" process, but they actually "*offered three variations of 'price sliding' order types. The exchanges' rules did not completely and accurately describe the prices at which those orders would be ranked and executable in certain circumstances, and they also failed to describe the execution priority of the three order types relative to each other and other order types.*"[26] On January 12, 2015, the SEC announced that "*two exchanges formerly owned by Direct Edge Holdings and since acquired by BATS Global Markets have agreed to pay a $14 million penalty to settle charges that their rules failed to accurately describe the order types being used on the exchanges. The penalty is the* SEC's *largest against a national securities exchange.*"[27] This penalty provides a unique case that summarizes three implications of our paper. First, two orders that differ by one instruction alone can have dramatically different economic functions and outcomes. Second, Rule 610 drives innovation in order types. Third, queue position is crucial in liquidity provision.

DAY ISOs. DAY ISOs make liquidity first by taking liquidity. A DAY ISO, if marketable on arrival, will be immediately traded with a contra-side interest in the NYSE book up to its full size and limit price. Any untraded quantity of a DAY ISO will be displayed at its limit price. The sender of the DAY ISO complies with Rule 610 by sending ISOs to other exchanges to clear the locked or crossed quotes. DAY ISOs represent only 0.25% of trading volume as of April 2021, yet its ALO variant is used more frequently by a wide margin (4.08%), as we elaborate below.[28]

ALO DAY ISOs. In Section 4, we show that an order may lock a market because of exchange latency. In Section 6.1, we describe speed races to the head of the queue. Exchange latency also imposes a cost on an order that races to be the first to establish a new price. For example, to obtain snapshots of quotes on other exchanges, exchanges use cables, which are slower than the microwave towers used by fast liquidity providers. Therefore, an exchange may think a newly established local quote locks quotes on other exchanges even when those quotes have already disappeared (MacKenzie 2021, p.179). Fast traders may send in ALO orders repeatedly until an exchange accepts them, but exchanges later invented a more advanced solution that helps them overcome exchange latency: ALO DAY ISOs. With a DAY ISO tag, an order can establish a quote on the NYSE regardless of the NYSE's own view of the NBBO.

[23] NYSE Rule 7.31.e.2.B.ii. The same is true for NASDAQ and CBOE "Post Only" orders. In other words, ALO orders refuse to be routed to the "best" gross prices, but they do not refuse better net prices.

[24] The official NYSE name of the order type is "non-routable limit orders" (NYSE rule 7.31.e.1.A). We have renamed this order type based on its economic function, which is to slide the quote if the quote locks a market. A slide order on NASDAQ is called a "Price to Display" order. Price to Display orders are available solely to participants that are market makers (NASDAQ rule 4702.b.2.A). The CBOE version is named "Price Sliding" orders (see, e.g., BZX rule 11.9.g.).

[25] *The Wall Street Journal*, "For superfast stock traders, a way to jump ahead in line," September 19, 2012. Available at https://www.wsj.com/articles/SB10000872396390443989204577599243693561670. As Hide Not Slide orders can "jump ahead" of other orders in the queue, traders who use other orders in the queue may find it harder to fill their orders. The increased non-execution probability may incentivize non-HFTs to execute their orders by accepting routing and paying the fees.

[26] Ibid.

[27] "The SEC charges Direct Edge Exchanges with failing to properly describe order types," January 12, 2015. Available at: https://www.sec.gov/news/pressrelease/2015-2.html.

[28] See https://www.nyse.com/publicdocs/nyse/NYSE_Group_Executed_Order_Type_Usage.xlsx.

Therefore, the ALO DAY ISO can post on the NYSE book earlier than regular ALO orders. With an ALO tag, the order refuses to take liquidity (or pay take fees) on the NYSE.

7.2. Policy implications

Our paper indicates that fees and clientele segmentation drive innovation in and the proliferation of order types to bypass Rule 610. In this section, we discuss our paper's policy implications. One possibility is maintaining the status quo. Although Rule 610 can route orders to worse net prices, the invention of order types that refuse Reg NMS routings provides a market solution to the problem. We also find that high opportunity costs for routed limit orders rationalize why they accept routings. If policymakers are concerned with the complexity of order types and market liquidity, we offer several policy solutions.

As fees distort the best price, such price distortion could be addressed by defining the NBBO based on net prices rather than gross prices. In addition to preventing routing to worse net prices, the net NBBO also reduces the incidence of locked markets, as the true bid and ask prices on two exchanges do not lock when gross prices lock. Therefore, the net NBBO can weaken the incentive to design complex order types. Such a policy would however face the challenge that exchanges often charge different tiers of fees for different traders (Spatt 2020). Therefore, the prerequisite for a net NBBO is first to remove fee tiers.

The other possible way to reduce price distortion is to reduce the fee cap (SEC 2018) or remove rebates (Harris 2018). Both policies reduce gaps between gross and net prices and weaken the incentive to design order types to collect rebates and avoid fees. One side effect of this solution might be that fee-cap reduction makes prices more discrete. Yao, and Ye (2018) show that fees and rebates are devices that enable stock exchanges to create more continuous net prices when the gross price is constrained by a discrete tick size. Imposing fee caps would eliminate the set of feasible net transaction prices and could also reduce liquidity.[29]

The other solution that might be effective involves allowing the gross prices on two exchanges to lock. In addition to weakening the incentive for traders to use and for exchanges to design complex order types, such a policy might also improve liquidity. When the gross quotes on two exchanges lock, the net bid–ask spread is not zero: it is the sum of the make rebate on the two exchanges. As 62% of routing aims to unlock markets and 43.40% of liquidity-making orders are designed to remove the risk of locking markets, our results indicate that the net bid–ask spread is often lower than one tick plus the sum of rebates. Therefore, allowing gross prices on two exchanges to lock may significantly improve liquidity. One limitation of this solution is that it replaces the existing prescriptive policy (Rule 610) with a new prescriptive policy. Rule 610 forbids locked markets, and the default solution to unlocking markets is to execute the order that locks the market. This solution implicitly assumes that all orders that lock the market are open to taking liquidity at their limit prices even if they need to pay the fee. If we allow the locked market, we go to another extreme by assuming that no orders that lock the market want to take liquidity at their limit prices. The results we report in Section 6 imply, however, that orders that entail higher opportunity costs may prefer execution to fee-saving. If we allow the locked market in general, exchanges may choose to withhold orders that are marketable on another exchange and thereby reduce their fill rates.

One less extreme and prescriptive approach is to allow an order to lock the market only when their submitters specify that they aim to improve liquidity. We have shown in this paper that the existing methods for complying with Rule 610 ask traders to cancel liquidity, hide liquidity, or even worsen liquidity by sliding orders to worse prices. In summary, Rule 610 discourages orders from improving displayed liquidity, so perhaps orders that aim to improve liquidity should be allowed to lock the market. We call such liquidity-improving orders "Non-Routable Lock" (NRL) limit orders. We illustrate NRL orders in Fig. 3. Suppose a stock's break-even ask price is \$5.001, the break-even bid price is \$4.999, and the rebate for providing liquidity is \$0.002. It would then be profitable after rebates to post limit orders at \$5.00 to buy and sell. Suppose that NASDAQ first establishes the best ask at \$5.00. In the current market environment (Panel A), a liquidity provider on the NYSE can display a bid price only at \$4.99 or below (Panel A1). If a plain limit buy order arrives at \$5.00, it will be routed to NASDAQ by Rule 610 (Panel A2). Therefore, the minimum bid–ask spread is 1 cent and liquidity takers can receive at most \$4.99 to sell. Panel B illustrates the environment with the addition of NRL limit orders, in which the liquidity provider can display a bid price at \$5.00 in NYSE, pushing the nominal bid–ask spread to zero (Panel B1). Suppose then that an NYSE trader arrives and submits a routable limit order to buy at \$5.00. The order is then routed to NASDAQ and executed at \$5.00 because the order does not refuse routing (Panel B2). Therefore, unless a trader refuses routing, her order is protected by Rule 610, the "Access Rule," because the order still has access to quotes on other exchanges. Certainly, exempting NRL from Rule 610 may lead to a zero bid–ask spread, and regulators are concerned that a zero bid–ask spread may confuse the market (SEC 2005). Our results indicate, however, that mandating a minimum 1-cent displayed spread can also confuse the market through order types that cancel or worsen liquidity to unlock the market. Therefore, policymakers should weigh the benefits and costs of allowing zero gross bid–ask spreads.

One final possible solution is to reduce the tick size such that the minimum bid–ask spread is no longer constrained by one tick plus the sum of rebates. A reduction in the tick size would also reduce the incidence of locked markets and routings to worse net prices. Moreover,

[29] The following extreme example provides the intuition. Suppose that regulators set the fee cap at 0. The rebate must then be 0 because exchanges do not want to lose money. Then all net transaction prices occur in multiples of ticks and bid–ask spreads are also in multiples of ticks. Yao and Ye (2018) show that, when the fee cap is not 0, competing exchanges would choose heterogenous fee structures and thereby subpenny transaction prices.

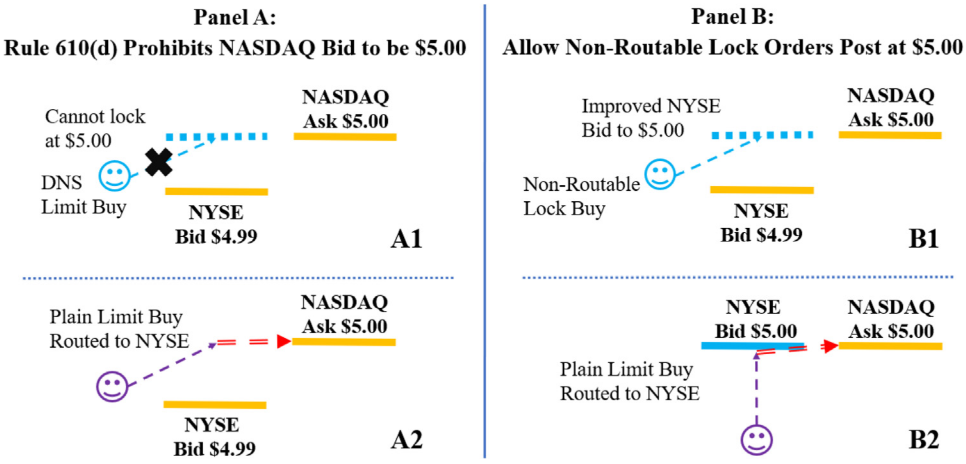

Fig. 3. Illustration of "Non-Routable Lock" (NRL) limit orders, This figure illustrates the proposed "Non-Routable Lock" (NRL) limit orders. Suppose a stock's break-even ask price is $5.001, the break-even bid price is $4.999, and the rebate for providing liquidity is $0.002. Posting limit orders at $5.00 to buy or sell are then both profitable after rebates. Solid horizontal lines represent existing limit orders, and dashed horizontal lines represent limit prices of incoming orders. Single-dashed arrows represent order submissions, and double-dashed arrows represent exchange routings. Panel A illustrates the current market environment, in which no displayed orders can lock the market. In this environment, non-routable orders cannot display liquidity at $5.00 bid (Panel A1), and plain buy limit orders at $5.00 are routed to NASDAQ (Panel A2). Panel B illustrates the market environment if NRL limit orders can lock the market. In this environment, an NYSE NRL can provide liquidity and reduce the nominal bid–ask spread to 0 (Panel B1), and plain limit orders are still protected by Rule 610 "Access Rule" and are routed to NASDAQ (Panel B2).

Chao et al. (2019) show that reducing the tick size would reduce the importance of fees and rebates. In addition, as we show that locked markets and fees drive complex order types, a reduction in the tick size may reduce the complexity of order types. Finally, Yao and Ye (2018) show that decreasing the tick size discourages speed competition to capture rents created by discrete pricing while also improving liquidity. Certainly, a smaller tick size may increase the number of price levels for traders to monitor, and regulators may need to consider the costs of price-level complexity when they reduce the tick size.

8. Conclusion

Reg NMS creates two cornerstones of the U.S. equity market structure. First, it consolidates quotes from fragmented exchanges and establishes national best bid and ask prices. Second, it routes orders to the exchange that displays the best price. Surprisingly, 62% of Reg NMS routings lead to worse net prices. Although the well-known reason for routing orders outside, Rule 611 (the no trade-through rule), always improves prices, Rule 611 accounts for only 31% of routings. The remaining 69% of routings aim to comply with the much less well-known Rule 610 (d). About 90% of Rule 610 routings lead to the same gross prices and worse net prices after routing fees. Routing to worse net prices particularly affects orders that aim to improve prices. They can collect rebates if they successfully provide liquidity but need to pay routing fees once they are routed outside.

In our sample period, traders can avoid routing to worse net prices by using DNS orders, which cancel orders if they need to be routed outside. As fees are uniform across stocks, we find that the market share that DNS orders capture is larger for stocks whose bid–ask spreads are narrower because fee differences become economically more significant for these stocks. We also find that two exogenous shocks that raise the rebates increase the market share captured by DNS orders, and our difference-in-differences tests show that these two shocks lead to a greater increase in the DNS market share for low bid–ask spread stocks than for high bid–ask spread stocks.

Clientele segmentation provides an explanation of heterogenous routing choices. DNS orders are more likely to come from liquidity-providing HFTs, and they entail low opportunity costs for non-execution. Therefore, it is rational for DNS orders to refuse routings as fees are economically significant relative to other costs. We find that routable orders are more likely to be placed by non-HFTs than by HFTs. As we find that routed orders would incur high opportunity costs for non-execution, it is rational for their submitters to accept routing because their opportunity costs outweigh the fee costs. In summary, our results indicate that order-type innovations provide customized means of complying with Reg NMS and that Reg NMS is one driver of the proliferation of order types.

In light of the fact that a large fraction of trading volume originates in orders that are designed to avoid small fees, we provide insights into a proposed New Jersey State Government transaction tax of 0.25 cents per share. We find that stock trading is highly sensitive to small differences in fees at the magnitude of the proposed transaction tax. Therefore, our results justify the aggressive response of the NYSE and NASDAQ to the proposed transaction tax. Both exchanges activated their backup sites in Chicago to demonstrate their capacity to pull their business out of New Jersey.[30] This indicates that a seemingly

[30] Matthew Leising, Bloomberg, 2020, "Leaving N.J. for Chicago gives no easy tax fix to Nasdaq, NYSE." Available at: https://www.bloomberg.com/news/articles/2020-09-24/leaving-n-j-for-chicago-gives-no-easy-tax-fix-to-nasdaq-nyse.

Table A.1
Price aggressiveness for liquidity-taking orders, In this table we report results indicating the price aggressiveness of liquidity-taking orders. Large liquidity-taking orders may result in multiple execution records either at one price level or at multiple price levels ("walk up the book"). We count the number of orders that result in trades at multiple prices and one price and then divide them by the total number of orders.

Category	Plain Market	Plain IOC	IOC ISO	NMS IOC
Trade at multiple prices	5.46%	1.32%	0.53%	0.25%
Trade at one price	94.54%	98.68%	99.47%	99.75%

small transaction tax would fundamentally change the U.S. trading landscape.

Disclosure statement

The author declares that he has no relevant or material financial interests that relate to the research described in this paper.

Data Availability

The authors do not have permission to share data.

Appendix A. Refuse Rule 611 routing: latency

This paper's main results focus on non-routable orders that can provide liquidity. Here in this Appendix A, we discuss non-routable orders that cannot provide liquidity. As IOC orders refuse to provide liquidity, they cannot lock the market. The results reported in Table 4 show that IOC orders would never be routed to worse net prices because their routings are driven by Rule 611 but not Rule 610.[31] We find that latency drives refusals of Rule 611 routing.

The motivation to design IOC ISOs, according to the SEC, is the desire to provide institutional investors with immediate access to liquidity at multiple price levels, in multiple markets, to fill large block trades with parallel order submissions (SEC 2005; Chakravarty et al., 2012). The results reported in Table A.1 show, however, that 99.47% of IOC ISOs do not sweep multiple price levels. Therefore, the main driver of IOC ISO usage diverges from the SEC's purpose in designing ISOs.

Next, we show that exchange latency helps to explain why liquidity-taking orders refuse Rule 611 routing. Exchanges are subject to geographic, consolidation, and transmission latencies (SEC 2018). Therefore, fast traders may prefer routing orders themselves, especially when considering strategies that are sensitive to speed. We analyze one such strategy: sniping stale quotes.

In Section 6, we consider cases in which liquidity-providing orders win speed races to snipe stale quotes. Here, we consider cases in which liquidity-providing orders fail to cancel and in which IOC orders win the race. If $q > 1$ IOC orders successfully snipe stale quotes, each order counts as winning $1/q$ of the race. We next consider which type of IOC order wins the race.

The results reported in Table A.2 show that there are 258.22 cases where, on an average stock-day, at least one IOC order wins. This number is lower than the total number of orders being sniped (282.25), as documented in Section 6, because some races involve sniping more than one order.

We find that non-routable orders are more likely than plain IOC orders to win speed races. For example, IOC ISOs win 7.31 races when they compete with plain IOC orders while IOC ISOs lose 4.65 races. The odds of winning that IOC ISOs face are $\frac{7.31}{7.31+4.65} = 61\%$. NMS IOC orders win 4.62 cases over plain IOCs and lose 3.16 cases to plain IOCs. The odds of winning are 59.38%. IOC ISOs win slightly more races in head-to-head competition with NMS IOCs: $(\frac{49.60}{49.60+45.02}) = 52.4\%$ versus $(\frac{45.02}{49.60+45.02}) = 47.6\%$.

We next explain why IOC ISOs win the majority of speed races over other IOC orders and why other IOC orders still win some of the speed races over IOC ISOs. IOC ISOs differ from other non-routable orders, such as NMS IOC orders and DNS limit orders, along one dimension. No other non-routable orders create speed advantages by themselves, because the NYSE processes orders in the order received, even if there is more latency inherent to processing one type of order than another. In Section 6, we show that DNS limit orders win speed races to provide liquidity, but that is not because DNS limit orders enjoy a speed advantage; it is because their submitters are faster. IOC ISOs can create a speed advantage by themselves, but in some very specific scenarios. For example, suppose that the ask on the NYSE is $5.01 and the best ask on NASDAQ is $5. The value then jumps such that the new ask price should be $5.03 and the new bid price should be $5.02. The ask quotes on both the NYSE and NASDAQ then become stale. Suppose that the NASDAQ ask has already been sniped (potentially by the ISO sender herself), but the NYSE thinks that the $5 ask still exists because of exchange latency. In this case, the NYSE would route an IOC order to NASDAQ and reject an NMS IOC order. IOC ISOs, however, can execute directly at the $5.01 ask.

The benefit enjoyed by IOC ISOs in this specific scenario comes with two costs related to regulatory compliance. First, IOC ISO submitters need to have fast connections to all exchanges, including the smallest ones. Second, on a trade-by-trade basis, IOC ISO submitters might be asked to show regulators that they have not violated Rule 611 at the time of execution. Traders need to weigh these compliance costs against the benefits of IOC ISOs, which explains why other types of IOC orders can still win some speed races.

Traders that lack speed advantages will lose money in "sniping" events. Next, we briefly note that stock exchanges invented order types that mitigate sniping risks for traders.

M-ELOs. NASDAQ introduced midpoint extended life orders (M-ELOs) in 2018.[32] Trades of this order type refuse

[31] Still, fees may play a role because the routing fee is usually higher than the take fee. Therefore, it is cheaper for traders to take liquidity directly from another exchange than to let the NYSE take liquidity from another exchange on its behalf. Nevertheless, we find that market shares captured by non-routable IOC orders, particularly IOC ISOs, are much less sensitive to fee changes (untabulated for the sake of brevity).

[32] See https://www.nasdaq.com/solutions/midpoint-extended-life-order-m-elo.

Table A.2
Speed races to snipe stale quotes, In this table we report the pairwise speed race counts for sniping stale quotes across three types of IOC orders. We define the applicable speed races in Section 6.1.2 and this Appendix A. The order types displayed in the rows are race winners while the order types displayed in the columns are race losers. If $q > 1$ orders win (lose) a race, each order is counted as winning (losing) $1/q$ of the race. All numbers are averaged at the stock-day level.

Winner\Loser	Plain IOC	IOC ISO	NMS IOC	Total winning races	Winning rate
Plain IOC	6.16	4.65	3.16	13.98	5.41%
IOC ISO	7.31	91.16	49.60	148.06	57.34%
NMS IOC	4.62	45.02	46.55	96.18	37.25%
Total Losing races	18.09	140.83	99.31	258.22	

to trade with IOC orders and are non-displayed, although they welcome trades with resting limit orders as well as other M-ELOs. M-ELOs thus aim to avoid being sniped by restricting the counterparties to a trade. This order type resonates with one intuition conveyed by our paper. Even though M-ELO submitters cannot screen counterparties based on the counterparties' identities, the order types used by their counterparties signal trader types.

D-limit orders. Another order type that is designed to avoid sniping risk is the Discretionary limit (D-limit) order introduced by the Investors Exchange (IEX) in 2020.[33] The IEX runs an algorithm that is designed to dynamically forecast sniping risk. When sniping risk is low, a D-limit order behaves like a regular limit order. When sniping risk is high, the IEX automatically reprices a D-limit order to one tick worse than the NBBO. D-limit orders are thus designed to reduce sniping risk by dynamically adjusting their limit prices. As U.S. regulators do not encourage stock exchanges to slow liquidity-taking orders to a greater extent than they slow liquidity making orders (asymmetry speed bumps), the IEX implemented only a symmetric speed bump. D-limit orders help the IEX effectively implement asymmetric speed bumps because it can update liquidity-making orders based on its adverse-selection signals ("crumbling quote indicators") such that these liquidity-making orders may escape before sniping orders arrive.

[33] See https://www.sec.gov/rules/sro/iex/2020/34-89686.pdf. In official order-type descriptions, IEX call sniping risk "adverse selection by latency-arbitrage trading strategies."

References

Anand, A., Samadi, M., Sokobin, J., Venkataraman, K., 2021. Institutional order handling and broker-affiliated trading venues. Rev. Financ. Stud. 34, 3364–3402.
Aquilina, M., Budish, E., O'Neill, P., 2021. Quantifying the high-frequency trading "arms race": a simple new methodology and estimates. Q. J. Econ. 137 (1), 493–564.
Baldauf, M., Mollner, J., 2020. High-frequency trading and market performance. J. Financ. 75, 1495–1526.
Battalio, R., Corwin, S.A., Jennings, R., 2016. Can brokers have it all? on the relation between make-take fees and limit order execution quality. J. Financ. 71, 2193–2238.
Beason, T., and Wahal, S., 2021, The anatomy of trading algorithms. Working paper.
Bessembinder, H., Panayides, M., Venkataraman, K., 2009. Hidden liquidity: an analysis of order exposure strategies in electronic stock markets. J. Financ. Econ. 94, 361–383.
Bessembinder, H., Venkataraman, K., 2004. Does an electronic stock exchange need an upstairs market? J. Financ. Econ. 73, 3–36.
Boehmer, E., 2005. Dimensions of execution quality: recent evidence for U.S. equity markets. J. Financ. Econ. 78, 553–582.
Brogaard, J., Hendershott, T., Riordan, R., 2014. High-frequency trading and price discovery. Rev. Financ. Stud. 27, 2267–2306.
Budish, E., Cramton, P., Shim, J., 2015. The high-frequency trading arms race: frequent batch auctions as a market design response. Q. J. Econ. 130, 1547–1621.
Budish, E., Lee, R.S. and Shim, J., 2021. A theory of stock exchange competition and innovation: will the market fix the market?. Working paper.
Chakrabarty, B., Hendershott, T., Nawn, S., and Pascual, R., 2020. Order exposure in high frequency markets. Working paper.
Chakravarty, S., Jain, P., Upson, J., Wood, R., 2012. Clean sweep: informed trading through intermarket sweep orders. J. Financ. Quant. Anal. 47, 415–435.
Chao, Y., Yao, C., Ye, M., 2019. Why discrete price fragments U.S. stock exchanges and disperses their fee structures. Rev. Financ. Stud. 32, 1068–1101.
Equity Market Structure Advisory Committee, 2017, Framework for rule 611 and 610 discussion.
European Securities and Markets Authority, 2019, MiFID II/MiFIR Review Report No. 1.
Dittmar, A., Field, L., 2015. Can managers time the market? Evidence using repurchase price data. J. Financ. Econ. 115 (2), 261–282.
Foucault, T., Menkveld, A.J., 2008. Competition for order flow and smart order routing systems. J. Financ. 63, 119–158.
Foucault, T., Pagano, M., Röell, A., 2013. Market Liquidity: Theory, Evidence, and Policy. Oxford University Press.
Foucault, T., Röell, A., Sandås, P., 2003. Market making with costly monitoring: an analysis of the SOES controversy. Rev. Financ. Stud. 16, 345–384.
Frazzini, A., Israel, R., and Moskowitz, T., 2018, Trading costs, Working paper.
Glassman, A.C., and Atkins, P., 2005, Dissent of commissioners cynthia A. Glassman and Paul S. Atkins to the adoption of regulation NMS. Available at https://www.sec.gov/rules/final/34-51808-dissent.pdf.
Hasbrouck, J., 1992, Using the TORQ database, Working paper.
Hasbrouck, J., 2007. Empirical Market microstructure: The Institutions, Economics, and Econometrics of Securities Trading. Oxford University Press.
Li, S., Wang, X., Ye, M., 2021. Who provides liquidity and when? J. Financ. Econ. 141 (3), 968–980.
Li, S., Ye, M., and Zheng M., 2021, Financial regulation, clientele segmentation, and stock exchange order types, Working paper.
Li, X., Ye, M., and Zheng, M., 2021, Price ceiling, market structure, and payout policies, Working paper.
MacKenzie, D., 2021. Trading At the Speed of Light Trading at the speed of light. Princeton University Press.
Mackintosh, P., 2014. Demystifying order types. KCG Market Insights. Available at http://www.smallake.kr/wp-content/uploads/2016/02/KCG_Demystifying-Order_Types_092414.pdf .
Menkveld, A.J., 2016. The economics of high-frequency trading: taking stock. Annu. Rev. Financ. Econ. 8, 1–24.
Menkveld, A.J., 2018. High-frequency trading as viewed through an electron microscope. Financ. Anal. J. 74, 24–31.
O'Hara, M., Ye, M., 2011. Is market fragmentation harming market quality? J. Financ. Econ. 100, 459–474.
Perold, A., 1988. The implementation shortfall: paper versus reality. J. Portf. Manag. 14, 4–9.
Spatt, C., 2014, Statement for house subcommittee on capital markets and government sponsored enterprises (GSEs) hearing on "Equity market structure: a review of sec regulation nms. Available at https://docs.house.gov/meetings/BA/BA16/20140228/101818/HHRG-113-BA16-Wstate-SpattC-20140228.pdf.
Spatt, C., 2020, Is equity market exchange structure anti-competitive?, Working paper.

Spatt, C., 2018, The new realities of market structure and liquidity: where have we been? where are we going? in Public Policy and Financial Economics: Essays in Honor of Professor George G. Kaufman for His Lifelong Contributions to the Profession, edited by Douglas Evanoff, Tassos Malliaris, and George Kaufman, 247–263.
U.S. Securities and Exchange Commission, 2005. Final rule: Regulation NMS. U.S. Securities and Exchange Commission File No. S7-10-04.
U.S. Securities and Exchange Commission, 2018. Transaction Fee Pilot for NMS stock. U.S. Securities and Exchange Commission File No. S7-05-18.
Yang, L., Zhu, H., 2020. Back-running: seeking and hiding fundamental information in order flows. Rev. Financ. Stud. 33, 1484–1533.
Yao, C., Ye, M., 2018. Why trading speed matters: a tale of queue rationing under price controls. Rev. Financ. Stud. 31, 2157–2183.