Eric Budish
Paul G. McDermott Professor of Economics and Entrepreneurship
The University of Chicago Booth School of Business
5807 S. Woodlawn Ave., Chicago, IL 60637
Web: www.ericbudish.org | Email: eric.budish@chicagobooth.edu

January 18th, 2024

Ms. Vanessa A. Countryman, Secretary
U.S. Securities and Exchange Commission
100 F Street NE
Washington, DC 20549-1090

Re: Regulation NMS: Minimum Pricing Increments, Access Fees, and Transparency of Better Priced Orders (Release No. 34-96494; File No. S7-30-22)

Dear Ms. Countryman,

I appreciate the opportunity to comment on the Commission's proposed modifications to Regulation National Market System ("Reg NMS"). I am an economics professor at the University of Chicago Booth School of Business who researches market design – designing the "rules of the game" for markets, to borrow Milton Friedman's phrasing – with a specific focus on the design of financial exchanges. Market design research assumes that participants in a market act optimally in their rational self-interest with respect to market rules, but takes seriously the possibility that the market rules themselves may be sub-optimal. I believe that this approach brings a useful perspective to debates over equity market structure. I write independently and have no financial conflicts of interest. My research and policy writing on this topic is collected at: https://ericbudish.org/research/financial-markets/.

I support the proposed reforms to Reg NMS. In my view, this set of rules changes – primarily, a finer tick-size for tick-constrained stocks, a lower access fee cap, and harmonization of pricing increments for on-exchange and off-exchange trading – will reduce both investors' costs and the overall complexity of U.S. equity markets. I come to this view for three main reasons.

Basis for Conceptual Support of the Proposed Reforms

1. Reducing the tick-size constraint for tick-constrained stocks will reduce excess rents from artificially constrained prices. These excess rents lead to a speed race to the top of the book, which increases complexity, and the rents come at the expense of investors via a higher cost of liquidity.

There are two main sources of excess rents from trading speed in modern financial markets: sniping stale quotes (aka latency arbitrage), and the race to the top of the book. My research has primarily focused on the sniping / latency arbitrage issue, which requires a change to market design to fix.[1] Fixing the race to the top of the book is conceptually much simpler – just reduce the artificial price constraint!

[1] See Eric Budish, Peter Cramton and John Shim, 2015, "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response", *Quarterly Journal of Economics,* for the conceptual definition of sniping stale quotes and latency arbitrage, analysis of why the problem is inherent to continuous limit order book markets,

I think it is terrific that the SEC is considering this change and recommend that the SEC use a methodology along the lines of that in Aquilina et al (2022) and Brugler and Hendershott (2023), using exchange message data, to quantify the rents at the top of the book before and after the policy change.[2]

2. Putting on-exchange and off-exchange trading on a more level playing field, by reducing tick sizes and harmonizing pricing increments, will reduce artificial reasons to trade off exchange to "hack the penny." This both reduces complexity and improves the incentives to provide on-exchange displayed liquidity, which will improve price discovery and reduce costs.

Chair Gensler and many market observers have noted with concern the decreasing share of trading volume that takes place on exchanges as opposed to in off-exchange trading venues. The combination of the reduced tick-size and the harmonization of price increments on exchange and off exchange will eliminate one of the incentives to trade off exchange, which my collaborators and I colloquially like to call "hacking the penny." That is, trading off exchange simply to be allowed to trade at prices between tick increments that are not allowable on exchange.

I applaud the SEC for eliminating this artificial incentive to trade off exchange as opposed to on exchange.

There will still remain two substantive, economic incentives to trade off exchange as opposed to on exchange: (i) off-exchange trading can siphon off less-adversely-selected order flow (e.g., order flow from retail investors), and (ii) off-exchange liquidity provision isn't exposed to sniping, whereas on-exchange liquidity provision has to pay the sniping tax.

Both of these factors are quantitatively significant. On (i): The SEC's analysis in its proposal of the Order Competition Rule found that the adverse selection of retail investors trading off-exchange (i.e., the price impact of retail orders processed by wholesalers) is more than 70% lower than the adverse selection of anonymous trades on exchange. I summarized this evidence in my comment letter on the Order Competition Rule.[3] On (ii): Aquilina et al (2022) estimates that the sniping tax is 17-33% of the cost of on-exchange liquidity depending on the measure used.

---

the definition of frequent batch auctions (FBAs), and the analysis of how FBAs address the sniping issue. See Matteo Aquilina, Eric Budish and Peter O'Neill, 2022, "Quantifying the High-Frequency Trading Arms Race", *Quarterly Journal of Economics,* for quantification of the issue using exchange message data, including especially the failed-take and failed-cancel messages that are the empirical signature of a trading race.

[2] Specifically, define a race to provide activity at a new price level analogously to how Aquilina et al (2022) define a race to trade at a stale quote, looking for multiple firms that submit liquidity at the same price level sufficiently close together in time that it is physically impossible that they are reacting to each other. Then measure the dollarized profits from being first in this race, accounting for fees and rebates, and add it all up. This is conceptually simple but requires comprehensive message data (i.e., the full back-and-forth traffic between market participants and the exchange) and accurate time stamps.

[3] See Topics 1-2 of Eric Budish, 2023, "Comment Letter Re: Order Competition Rule Proposal (Release No. 34-96495; File No. S7-31-22).

For these reasons, there is a significant portion of off-exchange trading that these proposed reforms will not effect, perhaps the large majority, but directionally the proposed reforms create a more level playing field between on-exchange and off-exchange trading, which is to be commended. On-exchange, displayed, liquidity provision is vitally important and if anything ought to be structurally advantaged, not structurally disadvantaged as it is currently.

3. Reducing the tick-size and access fees together will reduce artificial reasons to fragment on-exchange trading with fee schedules that "hack the penny." This will reduce costs and complexity.

At the current access fee cap of $0.0030 per share (30 "mills"), most on-exchange trading volume is on exchanges with a "maker-taker pricing model" in which the taker of liquidity pays a take fee of $0.0030 per share, the regulatory maximum, and the provider of liquidity receives a rebate that averages around $0.0028 per share. The exchange's net fee is thus $0.0030 - $0.0028 = $0.0002 per share, or $0.0001 per-share-per-side.[4]

Before continuing, it is worth pausing to reflect on just how competitive this net fee is – this is about as close to Bertrand price competition as one sees.[5] Exchanges make money from other revenue sources that are not as structurally competitive, such as proprietary data and co-location/connectivity services, but trading fees themselves are really low (see Budish, Lee and Shim, forth.). This is a signal achievement of Reg NMS.

However, the market has converged on a pretty bizarre way to charge such a competitive net fee – charging one side of the trade 15 times the net fee, while rebating the other side of the trade 14 times the net fee. There are a few reasons why the market's current equilibrium way of setting a small net fee is sub-optimal.

First, it creates a paper trail where the price paid or received gross of fee is economically quite different from the price paid or received net of fee. For example, even though gross prices are quoted in whole pennies such as $0.01, $0.02, $0.03, etc., the net prices one can actually buy at on-exchange are really $0.013, $0.023, $0.033, etc., while the net prices one can actually sell at on-exchange are really $0.007, $0.017, $0.027, etc. The smaller the access fee is, the more reflective the paper trail is guaranteed to be of the prices actually paid or received by investors. (See Implementation Details Remark #2 below).

Second, the combination of the coarse tick size of one penny, and the high-access-fee cap of $0.0030, creates both the incentive and the means to use multiple exchange fee schedules to price at different points between tick amounts (see Chao, Yao and Ye, 2019). The "inverted" exchanges charge a fee to the provider of liquidity, and provide a rebate to the taker of liquidity, so that the net price charged lands at different amounts within the penny. For example, if the take rebate is $0.0016 and the maker

[4] See Eric Budish, Robin Lee, and John Shim, forthcoming, "A Theory of Stock Exchange Competition and Innovation: Will the Market Fix the Market?", *Journal of Political Economy* for this and related data.

[5] For context, if you add up all U.S. regular-hours on-exchange trading fees, for the whole of the U.S. stock market, the sum comes to less than StubHub's fee revenue for secondary market event ticket transactions. Even though the secondary market for event tickets is tiny in volume terms relative to the stock market (on the order of $10-$20bn per year versus on the order of $100 trillion per year), fees for secondary market event tickets transactions are about 30% of resale value, which is about 100,000 times higher on a percentage basis.

fee is $0.0018 (a realistic example), the net fee is the same $0.0002 per share, but now the prices one can actually buy at on-exchange are really $0.0084, $0.0184, $0.0284, etc., and the prices one can actually sell at on-exchange are really $0.0116, $0.0216, $0.0316, etc.

So, different exchanges are optimal to use for different prices within the penny, which is a recipe for artificial fragmentation, a confusing paper trail, and overall excess complexity. Excess complexity, in turn, is a recipe for excess rents, agency conflict and distrust. I think it is terrific that the SEC is proposing reforms that will mitigate some of this excessive complexity.

*

Remarks on Implementation Details

These are the three conceptual reasons I like the proposals. I would like to make three remarks about the specific implementation details.

Implementation Details Remark #1: I do not have a strong view on the precise magnitudes for the narrower tick sizes or the precise regulatory mechanism used to determine the narrower tick sizes. I find the SEC's analysis and approach reasonable overall. I think it is easier to be confident about the bounds on the "what is too coarse of a tick" side of the problem than on the "what is too fine of a tick" side of the problem.

I read the SEC's rule proposal with a strong prior that reducing the tick size for tick-constrained stocks is a good idea (for the conceptual reasons outlined above) but without a strong prior on how precisely to accomplish this goal, both in terms of the overall magnitudes and the precise regulatory mechanism.

I found the SEC's overall approach reasonable. Specifically, I think it is sensible to use the Time-Weighted Average Quoted Spread as an indication for what tick size is appropriate for a particular stock. I also thought it was good that the SEC synthesized evidence from past academic literature with its own additional analysis of the tick-size pilot to put some bounds on the problem.

Let me discuss the two sides of the problem in turn.

Too coarse. Economically, the tick size is "too coarse" if there are significant economic rents from providing liquidity at the top of the book that cannot be competed away by improving the price level by one increment, because the increment is too coarse. This is a classic case of a regulatory price constraint generating excess economic rents. The SEC concludes with confidence that the tick increment is too coarse if the Time-Weighted Average Quoted Spread is less than 2 times the tick increment (i.e., there are "less than 2 ticks within the spread"). For example, if a stock has a $0.01 increment and mostly trades with a bid-ask spread of $0.01 or $0.02, it would likely benefit investors to go to a narrower spread.

I found this evidence reasonable and also economically intuitive given priors from theory.

Too fine. The bound in the "too fine" direction is conceptually harder. The reason is that what makes a tick size too fine is special to the market design – in a continuous limit order book market, a too-fine tick

size leads to an explosion of message traffic in which market participants constantly improve the price by economically tiny amounts. This is sometimes called "pennying," though the idea is easiest to see if one imagines the tick size is orders of magnitude smaller than a penny, for instance what my collaborators and I call a "micro-dollar", or $0.000001.[6]

There are a few reasons why pennying is negative for the market. First, the volume of message traffic per se can be hard for exchanges to manage. Even with a reasonable tick size, Aquilina et al (2022) found that about 4% of latency-sensitive trading races were won by the "wrong" message as measured by the time the message arrived at the exchange's computer systems. If time is measured finely enough, and messages arrive fast and furious enough, exchange computer systems have an impossible task.

Second, and more importantly, the risk of being pennied reduces the incentive to provide stable, deep liquidity on exchange. In effect, there is no longer such a thing as time priority because priority can always be superceded for an economically trivial amount of money.

The SEC finds in its analysis that a Time-Weighted Average Quoted Spread more than 15 times the increment (e.g., an average spread of $0.15 or more in the current market with $0.01 increments) is associated with harm to market liquidity. I find this conclusion intuitively plausible – again, keeping in mind that this issue is specific to the continuous market.

Overall Magnitudes. So where does this leave us on the magnitudes? In the too coarse direction, I think we can be confident from the data and from theoretical intuition that a tick size that is more than half the Time-Weighted Average Quoted Spread (TWAQS), i.e., <2 ticks within the spread, is too coarse. In the too fine direction, I think we can be conceptually confident that there is such a thing as too fine a tick in the continuous market, and the SEC's empirical analysis suggests that a tick size that is less than $1/15^{th}$ of the AQS is too fine, i.e., >15 ticks within the spread is too fine.

That leaves a fairly wide range, from 1/2 to $1/15^{th}$ of TWAQS, or from 2 to 15 ticks within the average spread. Where and how exactly to draw the lines within these bounds seems open to multiple reasonable points of view. I found the SEC's main proposal reasonable, which targets 4-8 ticks intra-spread and would initially have tick sizes of $0.002, $0.005 and $0.01, with scope for some stocks to have tick sizes of $0.001 if their TWAQS is small enough post implementation. I also found the SEC's Alternative Tick Threshold proposal on pg. 281 reasonable. This alternative targets 2-4 ticks intra-spread and would initially have tick sizes of $0.005 and $0.01, with scope for some stocks to have tick sizes of $0.0025 if their TWAQS is small enough post implementation. The case outlined on pg. 285 for allowing ticks wider than $0.01, for stocks with high-enough TWAQS, is reasonable too. There are surely multiple other reasonable approaches. For this reason, measurement of the effects of the implemented policy and explicit provision to iterate on the details if called for by the data will be essential (see Implementation Details Remark #3 below).

Note: Discrete Time would Simplify the Issue Significantly. With apologies for being a broken record, I want to at least mention that these issues with tick sizes and pennying basically go away in the discrete-

[6] There is beautiful empirical evidence of the pennying phenomenon in the context of internet advertising auctions (Edelman and Ostrovksy, 2007) and a very nice equilibrium characterization of the phenomenon, also called "flickering quotes", due to Baruch and Glosten (2013).

time batch auction markets that I have envisioned in my research. There isn't the same explosion of message traffic, and the best quote is best for an economically meaningful amount of time, a full discrete-time increment. Priority is still price then time, but, with time discrete, there isn't a race to the top of the book – all orders that wish to provide liquidity at the same price level that arrive at the same time are treated equally (e.g., with pro-rata tie breaking for orders that arrive at the same discrete time, while there is priority to orders that have been resting for longer).

Indeed, in "Flow Trading" my collaborators and I envision a market in which time is discrete while prices and quantities can be significantly more continuous than they are currently. It would be feasible to trade in millionths of dollars (micro-dollars) and billionths of shares (nano-shares) without any issue.[7]

Implementation Details Remark #2. If the access fee cap is more than half the tick size there can be a ranking violation where trade A at a strictly higher price gross of fee than trade B has a strictly lower price net of fee, and vice versa. If the access fee cap is more than the full tick size there can be a ranking violation between protected quotes, even without rebates. These observations favor an access fee cap that is small relative to the tick size, all else equal, as envisioned in the proposal.

The current access fee cap of $0.0030 is 30% of the current tick increment of $0.01. This effectively means that the price net of fee can be 30% of a tick different from the price gross of fee – for example, a taker that buys at $10.01 is really paying $10.013 net of fee, and a taker that sells at $10.01 is really receiving $10.007 net of fee. This in turn means that the difference in price net of fee between a trade in which the taker is a buyer versus a trade in which the taker is a seller can be 60% of a price increment – the difference between the $10.013 for the aggressive buyer and the $10.007 for the aggressive seller is $0.0060 or 60% of a tick.

As discussed above, this already leads to a fairly confusing paper trail. I want to caution that if the access fee cap were to exceed half of the tick size, the paper trail can be not only confusing but can literally misrank trades.

For example, suppose that the access fee cap were $0.0055 not $0.0030 as it is presently and the tick size is $0.01 as it is presently. Then a taker that buys at $10.01 could really be paying $10.0155, and a taker that sells at $10.02 could really be receiving $10.0145 – so it is possible that a trade in the tape at $10.02 is for a lower price net of fee than a trade in the tape at $10.01. It might look to some market observers like the price just went up from $10.01 to $10.02, while other market observers with better data – specifically, with direct-feed data that they can use to maintain the state of exchanges' limit order books on a message-by-message basis – would know that actually the price just went down (i.e., from $10.0155 to $10.0145).

If the access fee cap were to exceed the full tick size, there can additionally be a ranking violation between quotes, which can interact in a confusing manner with Reg NMS's order protection rule. Specifically, it would be possible that there are two quotes, one on exchange A and one on exchange B, such that A's price is lower gross of fee while B's price is lower net of fee. One of the reasons the access fee cap was originally set to be $0.0030 in Reg NMS was to ensure that this could not happen: "The

[7] Eric Budish, Peter Cramton, Albert Kyle, Mina Lee and David Malec, 2023, "Flow Trading," Becker Friedman Institute Working Paper.

adopted rule thereby assures order routers that displayed prices are, within a limited range, true prices" (Regulation NMS Final Rules Release, pg. 27).

As the SEC contemplates significant reductions in both the tick size and the access fee cap, it should be mindful of this issue of the ratio of the access fee cap to the tick size. The proposal language seems appropriately mindful of this issue as one factor among several (proposed rule pgs. 97-98).

Implementation Details Remark #3. Especially if the SEC hard-codes numbers into regulation, it should do so in a way that explicitly provides for measuring the effects of the policy and the possibility of revising the policy and its hard-coded numbers in response to the measured effects.

I like that the SEC's proposal varies the tick size based on data on the Time-Weighted Average Quoted Spread. But even that is a hard-coded formula – i.e., the tick size for a particular stock is flexible but in a rigid manner.

I encourage the SEC to explicitly provide for both the empirical measurement of the effects of the rule – ideally using not only quoted spreads but also depth measures, actual trading costs for large investors, and direct analysis of message traffic – and the possibility of revising the rule in response to the data. It is not a mathematical constant that >15 ticks is too fine or <2 ticks is too coarse, and there is a lot of room even within these bounds for multiple reasonable approaches.

*

Conclusion

To conclude, I am broadly supportive of these proposals to reform Reg NMS. The combination of a lower tick size, lower access fee, and harmonized price increments between on-exchange and off-exchange trading will reduce excess rents, improve liquidity, reduce complexity, and yield a more level playing field between on-exchange and off-exchange trading. On-exchange trading will still be hindered by the cost of latency arbitrage, and off-exchange trading will still have the structural advantage of being able to segment order flow and provide liquidity without risk of getting sniped, but these are good steps to take and I commend the SEC for its efforts to improve our capital markets.

Sincerely,

Eric Budish

Bibliographical Details of Academic Papers Referenced

Aquilina, Matteo, Eric Budish and Peter O'Neill (2022). "Quantifying the High-Frequency Trading 'Arms Race'." *The Quarterly Journal of Economics*, 137(1), 493-564. Data and Code Appendix at https://github.com/ericbudish/HFT-Races.

Baruch, Shmuel and Lawrence R. Glosten (2013). "Flickering Quotes". Working Paper.

Brugler, James and Terry Hendershott (2023). "Public and Private Information in Quotes and Trades." Working Paper.

Budish, Eric, Peter Cramton and John Shim (2015). "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response." *The Quarterly Journal of Economics*, 130(4), 1547-1621.

Budish, Eric, Robin Lee and John Shim (forthcoming). "A Theory of Stock Exchange Competition and Innovation: Will the Market Fix the Market?" *Journal of Political Economy.*

Budish, Eric, Peter Cramton, Albert S. Kyle, Jeognmin Lee, and David Malec (2023). "Flow Trading." Becker Friedman Institute Working Paper.

Chao, Yong, Chen Yao and Mao Ye (2019). "Why Discrete Price Fragments U.S. Stock Exchanges and Disperses their Fee Structures." *The Review of Financial Studies*, 32(3), 1068-1101.

Edelman, Benjamin and Michael Ostrovsky (2007). "Strategic Bidder Behavior in Sponsored Search Auctions." *Decision Support Systems*, 43(1):192-198.

Related Comment Letter Referenced

Budish, Eric (2023). "Comment Letter Re: Order Competition Rule Proposal (Release No. 34-96495; File No. S7-31-22)