Eric Budish
Professor of Economics
Centel Foundation/Robert P. Reuss Faculty Scholar
The University of Chicago Booth School of Business
5807 S. Woodlawn Ave., Chicago, IL 60637
Phone: 773.702.8453 | Email: eric.budish@chicagobooth.edu

January 11, 2019

European Securities and Markets Authority
Submitted electronically via www.esma.europa.eu

Response to ESMA's Call for Evidence: "Periodic Auctions for Equity Instruments" (ESMA70-156-785)

Dear ESMA,

I appreciate the opportunity to comment in response to your "Call for evidence: Periodic Auctions for Equity Instruments". I am an economics professor at the University of Chicago Booth School of Business who researches market design, 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 given market rules, but takes seriously the possibility that the rules themselves may be sub-optimal. In particular, since 2010, I have been researching periodic auctions — my research uses the term "frequent batch auctions" — as a potential alternative or competitor to the continuous-time limit order book that is prevalent around the world, including in Europe.

My research shows that the limit order book market design has a simple mathematical flaw — the combination of (i) treating time as continuous, and (ii) processing requests to trade serially (i.e., one-at-a-time), causes "sniping", or arbitrage rents from symmetric public information, to be a built-in equilibrium feature of the market. Sniping harms liquidity provision, hurts investors, leads to a socially wasteful arms race for speed, and offends common economic intuitions about what constitutes an efficient market. My research shows that frequent batch auctions — (i) putting time into a discrete unit (e.g., 1 millisecond), and (ii) batch processing requests to trade that arrive at the same time — directly addresses the problems with the continuous limit order book. FBAs eliminate sniping, enhance liquidity provision, benefit investors, and stop the arms race for speed. They also computationally simplify the market and lead to an easier-to-interpret paper trail.

My responses to your specific questions about periodic auctions follow below. I hope that these responses are useful to ESMA and will be pleased to discuss further as helpful.

Key References

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* 130(4): 1547-1621.

Eric Budish, Peter Cramton, and John Shim. 2014. "Implementation Details for Frequent Batch Auctions: Slowing Down Markets to the Blink of an Eye." *American Economic Review Papers and Proceedings* 104(5): 418-424.

Eric Budish, 2017. "Will the Market Fix the Market." American Economic Association and American Finance Association joint luncheon keynote address. Video available online at https://www.aeaweb.org/webcasts/2017/luncheon

Q1. Do you agree with the two main differences identified to distinguish conventional periodic auctions from frequent batch auctions? If not, please explain why.

ESMA's call for evidence describes two "main differences" between frequent batch auctions and conventional periodic auctions:

1. "The duration of frequent batch auctions is very short and lasts only some milliseconds"
2. "Whereas conventional periodic auctions are scheduled by the trading venue, for frequent batch auctions two different models for triggering an auction currently exist … [first is] trigger a 'call period' every time a pair of opposing orders can be matched … [second is] trigger an auction as soon as one order has been submitted

I agree about the first main difference. FBAs as currently implemented in Europe have short durations measured in milliseconds. In fact, FBAs could be run as frequently as sub-millisecond and still have significant economic benefits, as I will describe in detail in my response to Q9.

I somewhat disagree about the second main difference. The Call is certainly correct insofar as these are the main scheduling/triggering models *currently used* for FBAs in Europe. However I would like to emphasize that FBAs as I and my coauthors have envisioned them in our research (Budish, Cramton and Shim, 2014, 2015) would ideally be run at frequent, scheduled intervals. For instance, the interval could be 1 millisecond, or potentially sub-millisecond, or potentially a bit slower.

FBAs as introduced and analyzed in my research, with frequent, scheduled auctions, may be viewed as more of an alternative to, or direct competitor to, continuous-time limit order book trading. Indeed, my research argues that the limit order book has a simple mathematical flaw — the combination of (i) treating time as continuous, and (ii) processing requests to trade serially (i.e., one-at-a-time), causes "sniping", or arbitrage rents from symmetric public information, to be a built-in equilibrium feature of the market. Sniping harms liquidity and leads to a never-ending arms race for speed. FBAs, as we introduce and analyze, directly correct this flaw. The combination of (i) putting time into a discrete unit (e.g., 1 millisecond), and (ii) batch processing requests to trade that arrive at the same time, transforms competition on speed (who can snipe the stale quote the fastest) to competition on price (who offers the best price), and in so doing eliminates sniping.

Overall, I would encourage ESMA to think of frequent batch auctions — at least as Budish et al. (2015) propose and analyze — as just like a centralized limit order book market, but with time put into discrete units and orders that are received at the "same time" processed in batch. It's not some completely different animal.

Q2. Do you agree with the observation of a rising market share for equity trading on frequent batch auctions?

I do not have any special knowledge to add to ESMA's understanding of FBA market shares. The only view I would like to express is enthusiasm that FBAs have gone from being negligible to being several percentage points of European trading. That strikes me as a very promising development for the market.

Q3. What are in your view the main factors driving this development?

Conceptually, I see two potential factors for why European market participants would find FBAs attractive as currently implemented in Europe:

1. To avoid the "latency arbitrage tax". That is, to avoid trading in a venue in which liquidity is unnecessarily expensive to access or expensive to provide, because of the risk of being "sniped" (see Budish et al., 2015) in a race to react to new public information.
2. To trade at prices within the BBO, for example at the midpoint.

Factor #1 is specific to frequent batch auctions as opposed to limit order books. This "sniping tax" is the subject of my academic research (Budish et al., 2015) and is discussed in detail in my response to Q1.

Factor #2 is not specific to frequent batch auctions in that there are many other kinds of off-exchange trading venue designs that try to facilitate mid-point trading (e.g., many different dark pools in the United States).

Economically, I see Factor #1 as a response to a flaw in the continuous-time limit order book market design, whereas I see Factor #2 as a response, primarily, to tick-size constraints. Factor #2 also reflects market participants' ability to segment themselves based on adverse selection risk, for instance, retail traders who statistically are unlikely to possess an informational advantage may be able to trade at a price closer to the midpoint than a likely-to-be-informed institutional investor, even in the absence of tick-size constraints.

As market regulators, I urge you not to "throw the baby out with the bath water". Market regulators have good reason to be wary of off-exchange trading that is motivated by circumventing tick-size constraints. But, at the same time, regulators should applaud and encourage market design innovations that allow market participants to trade in a way that avoids the negative aspects of high-frequency trading, and in particular the tax imposed by sniping / latency arbitrage.

Q4. Do you agree with the four characteristics identified by ESMA? Please explain.

ESMA describes four important characteristics of frequent batch auctions as:

1. The application of pre-trade transparency
2. Short auction duration
3. Price determination within the best bid and offer price
4. Self-matching features

My views on (1) the application of pre-trade transparency, are nuanced; indeed, I think the concept of pre-trade transparency is misunderstood both for frequent batch auction markets and for continuous limit order book markets. (In a world with latency – that is, the real world – the best you can hope for is to know where the market was a latency ago, and hope that the market will still be there a latency from now!). Please see my response to Q6 for details.

Re (2), short auction duration: I agree with ESMA.

Re (3), price determination within the BBO: I agree with ESMA, but emphasize that in an ideal world, FBAs would also be able to discover prices outside the BBO (potentially sweeping in liquidity from the continuous market in this case, see Budish et al., 2014). See my response to Q12 for details.

Re (4), self-matching features: I have no special information to add to ESMA's understanding.

Q5. Do you consider that other characteristics of frequent batch auctions may explain their success and/or raise questions in terms of compatibility with the MiFID II transparency provisions? Please explain.

Yes. In addition to the four characteristics ESMA emphasizes in paragraph 24 and question 4, I would also add "avoiding sniping, or latency arbitrage" as a central characteristic of frequent batch auctions that may explain some of their success.

As I emphasize in my research, sniping / latency arbitrage is (1) caused by the continuous-time limit order book market design, (2) is like a tax on liquidity. Liquidity is more expensive to provide because liquidity providers have to incorporate the potential expense of getting sniped, and this in turn makes liquidity more expensive for investors, because in a competitive market the cost of getting sniped will be passed on to demanders of liquidity.

FBAs avoid this tax on liquidity! That is a key reason why they should be successful in the market.

Q6. What is your view on the level of pre-trade transparency applied by systems that initiate auctions upon the receipt of a first order? In particular, should pre-trade transparency already be applied as of the start of an auction, irrespectively of whether there is a potential match or not? Please explain.

I believe the public discussion about pre-trade transparency and frequent batch auctions is somewhat confused and confusing. In my response to Q6 let me attempt to clarify.

To start, let me make clear what I view as the ideal information policy in a frequent batch auction market. This is FBA as defined and analyzed in Budish, Cramton and Shim (2015).

1. Orders remain outstanding until either executed or canceled. If an order is not executed in one batch auction, it remains outstanding for the next, and the next, etc., until either executed or canceled. Orders can be canceled at any moment in time.
2. At the conclusion of each auction, the market reports:

a. All trades that occurred in the auction (prices and quantities)
b. All orders that remain outstanding (prices and quantities)

What does this information policy accomplish? Imagine that the FBAs occur every 0.001 second or even faster. Then, every 0.001 second, market participants (realistically, algorithms), see all outstanding bids and asks for each security, and they also see if any trade occurred. This means that there is a direct analogue to the "bid" and "ask" in an FBA market. Traders can see the best outstanding order to buy (the bid) and the best outstanding order to sell (the ask). They can also see full depth-of-book on both sides of the market – kind of like demand and supply curves.

For most stocks, in most milliseconds, NOTHING HAPPENS. Thus, if a trader wants to "buy at the ask" or "sell at the bid", they can do so, just as in a continuous limit order book market.

In the rare milliseconds where there is lots of activity, this burst of activity gets batch processed via an auction, to determine the price, quantity traded, etc. This is different from in a limit order book, in that priority is determined by who offers the best price, not who is fastest. This difference is the whole point, and is what eliminates latency arbitrage / sniping.

Now let us contrast information policy in an FBA to information policy in a conventional limit order book market. In both markets, participants can "see the best bid and ask". In both markets, participants can "usually" trade at the best bid or best ask, if they want to, assuming nobody else is trying to do so at around the same time. In both markets, if there is a burst of activity (e.g., in response to some public pricing signal, like a jump in a related asset), traders who try to buy at the ask, or sell at the bid, might fail to do so. In a continuous market they fail if they are not first, in the FBA market they fail if their price offered is not the best.

In sum … information policy is very similar between a limit order book and the Budish et al. (2014, 2015) definition of FBA. The difference being that the same information that is disseminated in continuous time in a limit order book market, is disseminated in extremely frequent discrete time in an FBA market.

I do not want to vouch for the information policy of all current embodiments of frequent batch auctions (or periodic auctions). Rather, what I wish to highlight for ESMA is that, in the version of frequent batch auctions that I have offered in my research (and that has been peer reviewed, received academic awards, etc.), the information policy is very closely analogous to that in a continuous-time limit order book market.

There is no such thing as perfect pre-trade transparency in a limit order book market, because we live in a world with latency – you can only see where the market was a latency ago, and hope that it will still be that way a latency from now. This same imperfect version of pre-trade transparency is achievable in the Budish et al. (2015) design of a frequent batch auction market.

Q7. What is your view on the level of pre-trade transparency applied by systems that initiate auctions upon the identification of a possible match? In particular, do you consider that systems locking in prices at the beginning and/or allowing the submission of orders pegged to the midpoint meet the pre-trade transparency requirements? Please explain.

Please see my response to Q6 for my detailed views on FBAs and pre-trade transparency.

Q8. Would you see benefit in frequent batch auction systems providing information on market/order imbalance? Please explain.

Please see my response to Q6 for my detailed views on FBAs and pre-trade transparency.

Q9. Do you consider the auction length of frequent batch auctions as appropriate? In particular, how does the short auction length contribute to fair and orderly trading? Please explain.

I would encourage ESMA to think about the auction length as follows.

First, there is a significant economic difference between continuous-time serial-process trading (i.e., a limit order book) and discrete-time batch-process trading that arises even with extremely fast batch intervals, such as 1 millisecond or potentially even finer. The difference is that the combination of discrete-time and batch processing eliminates sniping, or latency arbitrage. Mathematically, "symmetric public information" does not lead to arbitrage profits; to earn profits you have to know something the rest of the market does not know (i.e., have asymmetric information). Whereas, in a continuous market, even symmetric public information creates arbitrage rents; these rents in turn harm liquidity and come at the expense of investors.

Second, there are then additional costs and benefits that come from further lengthening the batch interval. The longer is the batch interval, the more trading interest can be collected, but the more difficult it might be for algorithmic trading firms to provide liquidity the way they do in continuous markets, and the longer market participants have to wait to get certainty about their trade.

Another way to express my view is that I think of moving from continuous-time to very frequent discrete-time as having benefits without easily identifiable costs[1], but the move from very frequent discrete time to longer time intervals as having both additional potential benefits and easy-to-identify potential costs. What counts as "very frequent discrete time" is an engineering question as opposed to an economic one; based on my understanding of the relevant engineering details, 1 millisecond is comfortably enough to batch process among a large set of sophisticated market participants, and I have heard many market participants express the view that even 100 microseconds or finer might be sufficient in some circumstances. Economically, the key is that time is discrete and orders are batched.

Q10. Would you see benefits in having a longer auction duration? Do you consider that the auction duration should take into account the liquidity and/or type of instruments traded (e.g. a longer auction duration for less liquid instruments)? Please explain.

[1] Social costs, that is. Specific market participants with a comparative advantage in speed would lose their advantage in a very frequent discrete-time market. These participants may have a vested interest in preserving continuous-time trading and thwarting discrete-time trading.

Please see my detailed response to Q9 on the costs and benefits of substantially lengthening the auction duration. In essence, I think of moving from continuous-time to very frequent discrete-time as having significant benefits without easily identifiable social costs (though some market participants with a comparative advantage in speed will lose rents), but the move from very frequent discrete time to longer time intervals as having both additional potential benefits and easy-to-identify potential costs.

Q11. In your experience, how often do frequent batch auctions result in a match, and how many transactions are executed per frequent batch auction on average?

The only point I would add to ESMA's understanding is that in a frequent batch auction as envisioned in Budish et al. (2015), it would be quite common for there to be batch intervals with zero trade. In the current market, for most stocks, in most milliseconds, there is zero trade. Thus the same would be true in most frequent batch auction intervals.

Q12. Do you consider frequent batch auction systems as non-price forming systems? Please explain. Should a characteristic of any trading system be that it is always price forming in order to operate without a waiver? Please explain.

Ideally a frequent batch auction should contribute to price discovery … after all that is the whole point of an auction! A frequent batch auction as envisioned in Budish et al. (2015) would do just that.

If, on the other hand, a so-called "auction" is being used to facilitate trade ONLY at the BBO midpoint, then it is not contributing to price discovery, but rather free-riding off of price discovery elsewhere. This, in my view, would be an "auction in name only".

Ideally, a frequent batch auction operating in the context of an otherwise continuous market should sometimes facilitate trade at the best bid, sometimes at the best ask, sometimes strictly within the BBO … and, if allowed, sometimes discover prices /outside/ the BBO. After all, the whole point of an auction is to discover the right price. There is some discussion of the mechanics of how a frequent batch auction could operate alongside a continuous market (e.g., sweeping in liquidity from the continuous market when the auction discovers a price outside the BBO) in Budish et al. (2014).

I do not have an opinion about the "operate without a waiver" question that ESMA poses.

Q13. Do you consider that these functionalities resemble reference price systems (in particular when matching transaction at mid-point)? Please explain.

Please see my detailed response to Q12.

Q14. How do frequent batch auctions ensure multilaterality and interactions of trading interests in the price formation process (e.g. diversity of participating members, average number of participants, distribution of orders involved per transaction)?

What I would emphasize in response to your question about multilaterality is that, in a frequent batch auction market, it would be common for institutional investors to interact BOTH with other investors and with professional trading firms. There is a critical role in financial markets for professional liquidity providers. I sometimes encounter the fantasy that investors can somehow "disintermediate" liquidity providers and only trade with other investors, all without ever paying a bid-ask spread. This is bad economics! Instead, the goal should be to facilitate trade amongst investors and professional trading firms, but without the tax imposed by sniping.

Q15. Do you consider that the possibility of pegged orders might weaken the price determination logic? If yes, which measures would you recommend?

Pegged orders can be appropriate for market participants who do not have a high-frequency view on price. They contribute to price discovery because they have a direction (e.g., a pegged order to buy versus a pegged order to sell have different effects on the overall demand/supply for an asset).

Q16. How frequently are mechanisms used to prevent an auction uncross at a price outside the EBBO or PBBO (e.g. patterns and occurrences)?

No special information to add to ESMA's understanding of this issue.

Q17. What are your views on self-matching functionalities, and in particular member preferencing, in the context of frequent batch auction systems taking into account their short auction length? Do self-matching functionalities, and in particular member preferencing, coupled with other features of frequent batch auctions (short duration, locked-in prices) contribute to fair and orderly trading?

No special information to add to ESMA's understanding of this issue.

Q18. Do you consider that self-matching functionalities, and in particular member preferencing, on frequent batch auction systems may be used to formalise privately negotiated transactions?

No special information to add to ESMA's understanding of this issue.

Q19. In your opinion, is the feature of member preferencing indispensable for the success observed in frequent batch auction systems since the application of MiFID II?

No special information to add to ESMA's understanding of this issue.

Q20. How do you determine on which execution venues to conclude transactions. Please explain.

Not applicable.

Q21. Which execution venues attracted the most trading volume following the suspension of dark trading venues under the DVC and why? Please substantiate your answer by quantitative data where available.

No special information to add to ESMA's understanding of this issue.

Q22. Should trading under frequent batch auctions become subject to stricter requirements in the future, to which type of execution venues do you expect the current trading volume under frequent batch auctions to migrate to?

My hope would be that frequent batch auctions, as envisioned in Budish et al. (2015), become a larger part of the market. In their ideal form, frequent batch auctions transform competition on speed into competition on price, eliminate sniping from the market, enhance liquidity, stop the high-frequency trading arms race, and lead to a computationally simpler and more transparent market.

I strongly urge ESMA not to "throw the baby out with the bath water". I see the potential for legitimate regulatory concerns if there are implementations of FBA that are really "auction in name only" and just facilitate off-exchange midpoint trading. But ESMA should preserve, and indeed nurture, FBAs that address the negative aspects of high-frequency trading.

Kind regards,

Eric Budish