A Theory of Stock Exchange Competition and Innovation: Will the Market Fix the Market?

Eric Budish (U. Chicago & NBER)
Robin S. Lee (Harvard & NBER)
John Shim (Notre Dame)

Econometric Society Meetings
January 2022
A Design Flaw in Modern Stock Exchanges

- As of 2021, all 16 stock exchanges in the US use a market design called the continuous limit order book.
A Design Flaw in Modern Stock Exchanges

- As of 2021, all 16 stock exchanges in the US use a market design called the continuous limit order book.

- Academic research has shown that this market design gives rise to "latency arbitrage" (Budish, Cramton and Shim 2015):
  - Rents from symmetric public information
  - As distinct from asymmetric private information (Kyle 1985; Glosten Milgrom 1985)

- Latency arbitrage rents, in turn, lead to an arms race for trading speed. Socially wasteful, harms liquidity.

- Empirical evidence using exchange "message data" (full back and forth) (Aquilina, Budish, O'Neill 2022):
  - Modal race 10-15 microseconds
  - Races 20%+ of volume
  - Races 33% of price impact, effective spread
  - Potential for 17% reduction in market cost of liquidity

- This paper studies the incentives of exchanges to innovate to address latency arbitrage and the arms race for speed.
A Design Flaw in Modern Stock Exchanges

- As of 2021, all 16 stock exchanges in the US use a market design called the continuous limit order book
- Academic research has shown that this market design gives rise to “latency arbitrage” (Budish, Cramton and Shim 2015)
  - Rents from symmetric public information
  - As distinct from asymmetric private information (Kyle 1985; Glosten Milgrom 1985)
- Latency arbitrage rents, in turn, lead to an arms race for trading speed. Socially wasteful, harms liquidity.
A Design Flaw in Modern Stock Exchanges

- As of 2021, all 16 stock exchanges in the US use a market design called the continuous limit order book

- Academic research has shown that this market design gives rise to “latency arbitrage” (Budish, Cramton and Shim 2015)
  - Rents from symmetric public information
  - As distinct from asymmetric private information (Kyle 1985; Glosten Milgrom 1985)

- Latency arbitrage rents, in turn, lead to an arms race for trading speed. Socially wasteful, harms liquidity.

- Empirical evidence using exchange “message data” (full back and forth) (Aquilina, Budish, O’Neill 2022):
  - Modal race 10-15 microseconds
  - Races 20%+ of volume
  - Races 33% of price impact, effective spread
  - Potential for 17% reduction in market cost of liquidity
A Design Flaw in Modern Stock Exchanges

- As of 2021, all 16 stock exchanges in the US use a market design called the continuous limit order book

- Academic research has shown that this market design gives rise to “latency arbitrage” (Budish, Cramton and Shim 2015)
  - Rents from symmetric public information
  - As distinct from asymmetric private information (Kyle 1985; Glosten Milgrom 1985)

- Latency arbitrage rents, in turn, lead to an arms race for trading speed. Socially wasteful, harms liquidity.

- Empirical evidence using exchange “message data” (full back and forth) (Aquilina, Budish, O’Neill 2022):
  - Modal race 10-15 microseconds
  - Races 20%+ of volume
  - Races 33% of price impact, effective spread
  - Potential for 17% reduction in market cost of liquidity

- This paper studies the incentives of exchanges to innovate to address latency arbitrage and the arms race for speed
We must consider...whether the increasingly expensive search for speed has passed the point of diminishing returns. I am personally wary of prescriptive regulation that attempts to identify an optimal trading speed, but I am receptive to more flexible, competitive solutions that could be adopted by trading venues. These could include frequent batch auctions or other mechanisms designed to minimize speed advantages.

...

A key question is whether trading venues have sufficient opportunity and flexibility to innovate successfully with initiatives that seek to deemphasize speed as a key to trading success in order to further serve the interests of investors. If not, we must reconsider the SEC rules and market practices that stand in the way.
We must consider...whether the increasingly expensive search for speed has passed the point of diminishing returns. I am personally wary of prescriptive regulation that attempts to identify an optimal trading speed, but I am receptive to more flexible, competitive solutions that could be adopted by trading venues. These could include frequent batch auctions or other mechanisms designed to minimize speed advantages.

...  

A key question is whether trading venues have sufficient opportunity and flexibility to innovate successfully with initiatives that seek to deemphasize speed as a key to trading success in order to further serve the interests of investors. If not, we must reconsider the SEC rules and market practices that stand in the way.
We must consider...whether the increasingly expensive search for speed has passed the point of diminishing returns. I am personally wary of prescriptive regulation that attempts to identify an optimal trading speed, but I am receptive to more flexible, competitive solutions that could be adopted by trading venues. These could include frequent batch auctions or other mechanisms designed to minimize speed advantages.

A key question is whether trading venues have sufficient opportunity and flexibility to innovate successfully with initiatives that seek to deemphasize speed as a key to trading success in order to further serve the interests of investors. If not, we must reconsider the SEC rules and market practices that stand in the way.
We must consider...whether the increasingly expensive search for speed has passed the point of diminishing returns. I am personally wary of prescriptive regulation that attempts to identify an optimal trading speed, but I am receptive to more flexible, competitive solutions that could be adopted by trading venues. These could include frequent batch auctions or other mechanisms designed to minimize speed advantages.

... 

A key question is whether trading venues have sufficient opportunity and flexibility to innovate successfully with initiatives that seek to deemphasize speed as a key to trading success in order to further serve the interests of investors. If not, we must reconsider the SEC rules and market practices that stand in the way.
Private vs. Social Returns to Market Design Innovation

- Implicit presumption: the market will fix the market.
Implicit presumption: *the market will fix the market.*

- Natural instinct: if the current design is sufficiently inefficient, then there will be a private incentive to fix the inefficiency.
- Standard case when private and social returns to innovation are aligned (e.g., Griliches 1959)
Private vs. Social Returns to Market Design Innovation

- Implicit presumption: the market will fix the market.

  - Natural instinct: if the current design is sufficiently inefficient, then there will be a private incentive to fix the inefficiency

  - Standard case when private and social returns to innovation are aligned (e.g., Griliches 1959)

- But private incentives don’t always align with social

  - (e.g., Arrow 1962, Nordhaus 1969, Hirshleifer 1971, Mankiw Whinston 1986, ...)
Private vs. Social Returns to Market Design Innovation

▸ Implicit presumption: *the market will fix the market.*
  ▸ Natural instinct: if the current design is sufficiently inefficient, then there will be a private incentive to fix the inefficiency
  ▸ Standard case when private and social returns to innovation are aligned (e.g., Griliches 1959)

▸ But private incentives don’t always align with social
  ▸ (e.g., Arrow 1962, Nordhaus 1969, Hirshleifer 1971, Mankiw Whinston 1986, ...)

▸ This paper’s insight: incumbents’ private innovation incentives misaligned with social *precisely because they earn rents from the arms race.*
1. A Theory of Status Quo Competition Among Exchanges
   - Synthesized single exchange
   - Competitive trading fees
   - Economic profits from speed technology that are not dissipated

2. Empirical Validation of the Theory

3. Incentives for Market Design Innovation
Institutional Background: Key Regulations

Two key regulations shape modern electronic stock exchange competition in the US:

► Unlisted Trading Privileges (UTP)
  ▶ Any stock can be bought or sold on any exchange
  ▶ Model: same asset trades on all exchanges, *perfectly fungible*

► Regulation National Market System (Reg NMS)
  ▶ Order Protection Rule: roughly, on an order-by-order basis, transaction must execute at exchange(s) with best quote
  ▶ Dissemination and Access Rules: exchanges must make quotes easily electronically accessible.
  ▶ Model: *frictionless search and access*

► Contrast with
  ▶ Futures exchanges (no fungibility), Uber/Lyft (no frictionless multi-homing), Pagano-style models (single-homing)
Model of the Status Quo: Plan

- Point of departure: Budish, Cramton and Shim (2015) model

- Add 3 things:
  1. Informed traders, Glosten-Milgrom style
  2. Multiple exchanges
     - Exchanges strategic
     - Trading across the multiple exchanges shaped by UTP, Reg NMS
  3. New solution concept, Order Book Equilibrium (see paper)

- (Note: model is of regular-hours trading. Not opening/closing auctions, nor role of listings.)
Model of the Status Quo: Game Timing

- **STAGE 1**: Exchanges set prices. Trading fees $f_j$, exchange-specific speed technology fees $F_j$.

- **STAGE 2**: Speed technology adoption. TFs choose where to buy ESST.

- **STAGE 3**: Infinitely repeated Trading Game:
  - Publicly observed state at start of trading game: value $y$, state of order books $\omega$.
  - Period 1: Trading firms submit limit orders to exchanges
    - E.g., offer to buy $q$ shares at $y - \frac{s}{2}$, sell $q$ at $y + \frac{s}{2}$
    - Serial processing at each exchange. Updated $\omega$ public.
  - Period 2: Nature selects one of four events:
    1. $\lambda_{invest}$: an investor arrives and can trade on all exchanges
    2. $\lambda_{private}$: an informed trader privately observes a jump in $y$, and can trade on all exchanges; information about $y$ then revealed
    3. $\lambda_{public}$: there is a publicly observed jump in $y$, and TFs engage in “sniping race” on all exchanges
    4. $1 - \lambda_{invest} - \lambda_{private} - \lambda_{public}$: nothing happens
Equilibrium Properties I: Single Synthesized Exchange

- In Stage 3, the seemingly fragmented exchanges “aggregate up” into a *single synthesized exchange*
  - Frictionless search and access -> all liquidity competed down to the same bid-ask spread across all exchanges
  - One-for-one relationship between depth and volume ensures the marginal unit of liquidity is equally attractive on all exchanges. (See Glosten 1994; Ellison Fudenberg 2003)

- Key economic point: frictionless search very different economically from “standard” platform competition (Rochet and Tirole 03, Armstrong 06)
  - E.g., can show strictly positive frictions leads to tipping
  - Analogy: Diamond (1971) search model. Large economic difference between 0 and $\epsilon$
Equilibrium Properties II: Competitive Trading Fees

- Frictionless search and access → Bertrand competition on trading fees
  - Investors are perfectly elastic w.r.t. spreads and trading fees.
  - Hence, if \( f_j > f_k \), then any trading firm providing liquidity on \( j \) can more profitably provide on \( k \).

- Contrast: models with single-homing → network effects → supra-competitive trading fees (e.g., Pagano 1989, Cantillon-Yin 2008)
If exchanges make positive profits from selling ESST, will they dissipate these rents in competition for share by setting negative trading fees \((f < 0)\)?

In standard add-on pricing models firms that anticipate monopoly rents in an add-on good dissipate these rents in competition to sell the pre-add-on good

- Ex: printers and cartridges, hotels and minibars
- (see Ellison 2005, Gabaix-Laibson 2006)
If exchanges make positive profits from selling ESST, will they dissipate these rents in competition for share by setting negative trading fees ($f < 0$)?

In standard add-on pricing models firms that anticipate monopoly rents in an add-on good dissipate these rents in competition to sell the pre-add-on good

- Ex: printers and cartridges, hotels and minibars
- (see Ellison 2005, Gabaix-Laibson 2006)

- No.
If exchanges make positive profits from selling ESST, will they dissipate these rents in competition for share by setting negative trading fees ($f < 0$)?

In standard add-on pricing models firms that anticipate monopoly rents in an add-on good dissipate these rents in competition to sell the pre-add-on good

- Ex: printers and cartridges, hotels and minibars

(see Ellison 2005, Gabaix-Laibson 2006)

No. Difference here: *binding money pump constraint*.

Marginal costs to execute a trade are extremely low (modeled as zero). To dissipate rents would require negative prices.

But a negative price creates an obvious money pump: TFs will execute $\infty$ trades
Exchanges capture rents via positive ESST fees $F^*$. These are not dissipated.

- Intuition: even though stocks are fungible across exchanges, latency sensitive trading opportunities are not
Equilibrium Properties IV: Division of Latency Arb Rents

- Exchanges capture rents via positive ESST fees $F^*$. These are not dissipated.
  - Intuition: even though stocks are fungible across exchanges, *latency sensitive trading opportunities are not*

- Yet, even as price-setters in Stage 1, exchanges cannot extract *all* sniping rents:
Equilibrium Properties IV: Division of Latency Arb Rents

- Exchanges capture rents via positive ESST fees $F^*$. These are not dissipated.
  - Intuition: even though stocks are fungible across exchanges, *latency sensitive trading opportunities are not*

- Yet, even as price-setters in Stage 1, exchanges cannot extract *all* sniping rents:
  - Intuition: TFs can steer volume by providing good liquidity
  - Formally: “lone-wolf” deviation. We show that this is the most attractive deviation in Stage 2, so ruling it out is sufficient.

- Upper bound on total ESST revenues: $\frac{M}{(M-1)(N-1)} \Pi(s^*)$. ("30%")

- Model delivers *strictly interior split of latency arbitrage rents*. 
1. A Theory of Status Quo Competition Among Exchanges
2. Empirical Validation of the Theory
3. Incentives for Market Design Innovation
Empirical Validation of the Model

1. Evidence on Stage 3 Trading Game
   ▶ Stylized Fact #1: Many exchanges simultaneously at the best bid and best offer
   ▶ Stylized Fact #2: Linear depth-volume relationship
   ▶ Stylized Fact #3: Market shares interior and relatively stable

2. Evidence on Exchange Trading Fees ($f$)

3. Evidence on Exchange-Specific Speed Technology Fees ($F$)
Fact #1: Many Exchanges at the Same Best Price

Number of Exchanges at the Best Bid or Offer (% of Milliseconds, 2015)

Data: NYSE TAQ. Top 5 exchanges, Top 100 symbols, all milliseconds in 2015.
Fact #2: Linear Depth-Volume Relationship

Daily Volume Share vs. Depth Share – 2015

Data: NYSE TAQ. Top 5 exchanges, Top 100 symbols, all dates in 2015.
Fact #3: Exchange Shares Interior & Relatively Stable

Reg NMS Era Weekly Market Shares - Top 8 Exchanges

Data: NYSE TAQ. Top 8 exchanges, all symbols, from January 1, 2011 to December 31, 2015.
Empirical Validation of the Model

1. Evidence on Stage 3 Trading Game

2. Evidence on Exchange Trading Fees ($f$)
   - Stylized Fact #4: Average trading fees are economically small
   - Stylized Fact #5: Money-Pump constraint binds

3. Evidence on Exchange-Specific Speed Technology Fees ($F$)
Trading Fees: Complexity of Exchange Fees

Notes: From RBC Capital Markets, “Complexity of Exchange Pricing and Corresponding Challenges to Transparency and Routing,” (February 2016). RBC’s visual representation of exchange trading and routing fees by tier as of October 2015. Each path from the center node to an extreme node indicates a particular fee scenario, characterized by exchange, share price, transaction type (adding or removing liquidity), market participant type, market participant ADV, security listing exchange, other factors, and a fee or rebate. Red indicates rebate, blue indicates fee.
# Trading Fees: Complexity of Exchange Fees

## Bats BZX Exchange Fee Schedule
Effective April 12, 2017

### Transaction Fees:
- Rebates indicated by parentheses ( )
- The rates listed in the Standard Rates table apply unless a Member's transaction is assigned a fee code other than a standard fee code.
- Footnotes provide further explanatory text or where annotated to fee codes, indicate variable rate changes, provided the conditions in the footnote are met.
- Unless otherwise noted, all routing fees or rebates in the Fee Codes and Associated Fees table are for removing liquidity from the destination venue.

#### Standard Rates:
<table>
<thead>
<tr>
<th>Category</th>
<th>Adding Liquidity</th>
<th>Removing Liquidity</th>
<th>Routing and Removing Liquidity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Securities at or above $1.00</td>
<td>$(0.0325)</td>
<td>$0.0000</td>
<td>$(0.0325)</td>
</tr>
<tr>
<td>Securities below $1.00</td>
<td>Free</td>
<td>0.75% of total dollar value</td>
<td>0.50% of total dollar value</td>
</tr>
</tbody>
</table>

| Standard Fee Codes | R, V, Y | N, W, S/I | X |

### Add Volume Tiers
Applicable to the following fee codes: B, V and Y.

<table>
<thead>
<tr>
<th>Tier</th>
<th>Rebate Per Share to Add</th>
<th>Member has an AAE/IA as a percentage of TCV ≥</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier 1</td>
<td>$(0.0025)</td>
<td>0.15%</td>
</tr>
<tr>
<td>Tier 2</td>
<td>$(0.0028)</td>
<td>0.20%</td>
</tr>
<tr>
<td>Tier 3</td>
<td>$(0.0029)</td>
<td>0.25%</td>
</tr>
<tr>
<td>Tier 4</td>
<td>$(0.0030)</td>
<td>0.50%</td>
</tr>
<tr>
<td>Tier 5</td>
<td>$(0.0031)</td>
<td>1.50%</td>
</tr>
<tr>
<td>Tier 6</td>
<td>$(0.0032)</td>
<td>1.25%</td>
</tr>
</tbody>
</table>

### Fee Codes and Associated Fees:

<table>
<thead>
<tr>
<th>Fee Code</th>
<th>Description</th>
<th>Fee (Rebate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Routined to NYSE Arca, adds liquidity (Tape B)</td>
<td>$(0.0025)</td>
</tr>
<tr>
<td>B</td>
<td>Routined to NYSE MKT, adds liquidity</td>
<td>$(0.0015)</td>
</tr>
<tr>
<td>C</td>
<td>Routined to NYSE Arca, adds liquidity (Tape A or C)</td>
<td>$(0.0015)</td>
</tr>
<tr>
<td>D</td>
<td>Routined to NASDAQ, adds liquidity</td>
<td>$(0.0015)</td>
</tr>
<tr>
<td>E</td>
<td>Routined to EDGA using ALIS routing strategy</td>
<td>$(0.0025)</td>
</tr>
<tr>
<td>F</td>
<td>Closing Auction, BZX listed security</td>
<td>FREE</td>
</tr>
<tr>
<td>G</td>
<td>Closing Auction, Bank of New York Mellon Open Order, BZX listed security</td>
<td>0.0015</td>
</tr>
<tr>
<td>H</td>
<td>Continuous Book Order that executed in the Opening or Closing Auction, BZX listed security</td>
<td>FREE</td>
</tr>
<tr>
<td>I</td>
<td>Opening, IPO or Hot Auction, BZX listed security</td>
<td>FREE</td>
</tr>
<tr>
<td>J</td>
<td>Opening, IPO or Hot Auction, Bank of New York Mellon Open Order, BZX listed security</td>
<td>0.0015</td>
</tr>
<tr>
<td>K</td>
<td>Routined to EDGX using ALIS routing strategy</td>
<td>0.0025</td>
</tr>
<tr>
<td>L</td>
<td>Routined to BZX using ALIS routing strategy</td>
<td>$(0.0015)</td>
</tr>
<tr>
<td>M</td>
<td>Routined to BZX using Destination Specific routing strategy (Tape B)</td>
<td>0.0025</td>
</tr>
<tr>
<td>N</td>
<td>Removes liquidity from BZX (Tape B)</td>
<td>0.0000</td>
</tr>
<tr>
<td>O</td>
<td>Routined to EDGA using TRIM or TRIM2 routing strategy</td>
<td>$(0.0015)</td>
</tr>
<tr>
<td>P</td>
<td>Routined using Destination Specific routing strategy unless otherwise specified</td>
<td>0.0020</td>
</tr>
<tr>
<td>Q</td>
<td>Routined to EDGX using Destination Specific, TRIM, TRIM2 or SLIM routing strategy</td>
<td>$(0.0015)</td>
</tr>
<tr>
<td>R</td>
<td>Routined to listing market closing process</td>
<td>0.0010</td>
</tr>
<tr>
<td>S</td>
<td>Routined to NASDAQ using Destination Specific or INET routing strategy</td>
<td>0.0020</td>
</tr>
<tr>
<td>T</td>
<td>Routined to BZX using Destination Specific, BZX MKT or Headset, adds non-displayed liquidity</td>
<td>FREE</td>
</tr>
<tr>
<td>U</td>
<td>Routined to any exchange not covered by Fee Code NA, adds non-displayed liquidity</td>
<td>0.0020</td>
</tr>
<tr>
<td>V</td>
<td>Routined to listing market opening or re-opening cross</td>
<td>0.0015</td>
</tr>
<tr>
<td>W</td>
<td>BZX Opening or Re-opening, non-BZX listed security</td>
<td>0.0000</td>
</tr>
<tr>
<td>X</td>
<td>Routined to EDGX, adds liquidity</td>
<td>$(0.0015)</td>
</tr>
<tr>
<td>Y</td>
<td>Re-routed by NYSE using ROOXT, ROOXR or Post to Away routing strategy</td>
<td>0.0020</td>
</tr>
<tr>
<td>Z</td>
<td>Routined to EDGA, adds liquidity</td>
<td>0.0020</td>
</tr>
<tr>
<td>A</td>
<td>Routined to NASDAQ BZX, adds liquidity</td>
<td>0.0020</td>
</tr>
<tr>
<td>B</td>
<td>Routined to NASDAQ MKT, adds liquidity</td>
<td>0.0020</td>
</tr>
<tr>
<td>C</td>
<td>Routined to EDGX using ROOXT routing strategy, adds liquidity</td>
<td>0.0020</td>
</tr>
<tr>
<td>D</td>
<td>Non-displayed order, adds liquidity using Supplemental Peg</td>
<td>$(0.0015)</td>
</tr>
<tr>
<td>E</td>
<td>Routined to BZX, adds liquidity</td>
<td>0.0015</td>
</tr>
<tr>
<td>F</td>
<td>Routined to OTC, adds liquidity</td>
<td>0.0020</td>
</tr>
<tr>
<td>G</td>
<td>Routined to BZX using Parallel T or SWAPAS routing strategy</td>
<td>0.0030</td>
</tr>
<tr>
<td>H</td>
<td>Routined to BZX using SLIM routing strategy except to BZX (NYSE)</td>
<td>0.0020</td>
</tr>
<tr>
<td>I</td>
<td>Routined to NASDAQ BZX using TRIM or TRIM2 routing strategy</td>
<td>$(0.0015)</td>
</tr>
<tr>
<td>J</td>
<td>Non-displayed order subject to price ailing that reserves price improvement, adds liquidity</td>
<td>FREE</td>
</tr>
<tr>
<td>K</td>
<td>Removes liquidity from BZX (Tape A)</td>
<td>0.0030</td>
</tr>
<tr>
<td>L</td>
<td>Routined to a displayed market to remove liquidity using Parallel O, Parallel 2D, ROUT, ROOXR or Post to Away routing strategy</td>
<td>0.0030</td>
</tr>
<tr>
<td>M</td>
<td>Routined to a displayed market to remove liquidity using Parallel 0, Parallel 2D, ROUT, ROOXR or Post to Away routing strategy</td>
<td>0.0030</td>
</tr>
<tr>
<td>N</td>
<td>Routined to a dark liquidity venue (except through SLIM)</td>
<td>0.0020</td>
</tr>
<tr>
<td>O</td>
<td>Routined to NYSE Arca, adds liquidity</td>
<td>$(0.0015)</td>
</tr>
<tr>
<td>P</td>
<td>Routined to NYSE MKT, removes liquidity</td>
<td>0.0020</td>
</tr>
</tbody>
</table>
Fact #4: Average Trading Fees are Small

Estimate of average regular-hours trading fees ("$f$")
(3 major exchange families)

<table>
<thead>
<tr>
<th>Exchange Family</th>
<th>$f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NYSE</td>
<td>$0.000128$</td>
</tr>
<tr>
<td>NASDAQ</td>
<td>$0.000105$</td>
</tr>
<tr>
<td>BATS</td>
<td>$0.000089$</td>
</tr>
</tbody>
</table>

Data: Exchange 10-K and S-1 filings, exchange fee schedules, NYSE TAQ, and other sources. Please see Appendix B.
### Fact #5: Money Pump Constraint Binds

**U.S. Equity Exchange Trading Fees Per Share ("f")**

<table>
<thead>
<tr>
<th>Exchange</th>
<th>Fee Type</th>
<th>Taker Fee</th>
<th>Maker Fee</th>
<th>Total fee per share per side</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Max</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>NASDAQ</td>
<td>0.00300</td>
<td>0.00300</td>
<td>-0.00325</td>
<td>-0.00150</td>
</tr>
<tr>
<td>BATS BZX</td>
<td>0.00300</td>
<td>0.00300</td>
<td>-0.00320</td>
<td>-0.00200</td>
</tr>
<tr>
<td>EDGX</td>
<td>0.00300</td>
<td>0.00300</td>
<td>-0.00320</td>
<td>-0.00200</td>
</tr>
<tr>
<td>NYSE</td>
<td>0.00270</td>
<td>0.00270</td>
<td>-0.00220</td>
<td>-0.00140</td>
</tr>
<tr>
<td>NYSE Arca</td>
<td>0.00280</td>
<td>0.00300</td>
<td>-0.00270</td>
<td>-0.00200</td>
</tr>
<tr>
<td>BATS BYX</td>
<td>-0.00160</td>
<td>-0.00160</td>
<td>0.00140</td>
<td>0.00180</td>
</tr>
<tr>
<td>EDGA</td>
<td>-0.00020</td>
<td>-0.00020</td>
<td>0.00030</td>
<td>0.00050</td>
</tr>
<tr>
<td>NASDAQ BX</td>
<td>-0.00150</td>
<td>-0.00040</td>
<td>0.00165</td>
<td>0.00200</td>
</tr>
</tbody>
</table>

**Data:** Exchange fee schedules obtained via Internet Archive dated from February 28, 2015 to September 1, 2015.
Empirical Validation of the Model

1. Evidence on Stage 3 Trading Game

2. Evidence on Exchange Trading Fees ($f$)

3. Evidence on Exchange-Specific Speed Technology Fees ($F$)

   ▶ Stylized Fact #6: Exchanges earn significant revenue from ESST

   ▶ Stylized Fact #7: ESST revenue has grown significantly in the Reg NMS era
Fact #6: Exchanges Earn Significant Revenues from Speed Technology

Estimated Market Data and Co-Location Revenues for U.S. Equities Market in 2015 (Millions of Dollars)

<table>
<thead>
<tr>
<th></th>
<th>BATS</th>
<th>NASDAQ</th>
<th>NYSE</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Data Revenue</td>
<td>114.1</td>
<td>222.4</td>
<td>218.9</td>
<td>555.4 – 623.0</td>
</tr>
<tr>
<td>Co-Location/Connectivity Revenue</td>
<td>64.3</td>
<td>121.0</td>
<td>251.6</td>
<td>436.8 – 484.8</td>
</tr>
<tr>
<td>Market Data + Co-Location Revenue</td>
<td>178.4</td>
<td>343.3</td>
<td>470.5</td>
<td>992.2 – 1107.8</td>
</tr>
<tr>
<td>CTA/UTP Tape Revenue</td>
<td></td>
<td></td>
<td></td>
<td>317.0</td>
</tr>
<tr>
<td>Market Data + Co-Lo Revenue net of Tape Revenue</td>
<td></td>
<td></td>
<td></td>
<td>675.2 – 790.8</td>
</tr>
</tbody>
</table>

**Data:** Exchange 10-K and S-1 filings, CTA fee-change filing. Please see Appendix D.
Fact #7: Exchange Speed Technology Revenue has Grown Significantly since Reg NMS

Exchange Market Data and Co-Location Revenue 2006 - 2017

Data: Exchange 10-K, S-1 and merger proxy filings. Please see Appendix D.
Empirical Validation: Summary of Main Results

- Overall fit is pretty good, especially for such a simple model.
- Also note that data is not consistent with other potential models of exchange competition (most not specifically tailored to US stock exchanges)
  - Exchanges will charge supra-competitive trading fees
  - Network effects often lead to tipping
- Models in which exchanges are meaningfully differentiated (Pagnotta and Philippon 2018, Baldauf and Mollner 2018)
  - Exchanges will charge supra-competitive trading fees
  - Differentiation leads to segmentation – who and what trade where
- Models in which tick-size frictions are central to understanding exchange fragmentation and competition (Chao, Yao and Ye 2019)
1. A Theory of Status Quo Competition Among Exchanges
2. Empirical Validation of the Theory
3. Incentives for Market Design Innovation
Incentives for Market Design Innovation

Now we study the incentives for market design innovation: “will the market fix the market?”

- What are the private incentives for an exchange to adopt a new market design that eliminates latency arbitrage and the speed race?
- Focus on frequent batch auctions (“Discrete”) with very short time interval

We analyze this question in 3 steps:

1. “Continuous” vs. “Discrete”
3. Use analyses of Continuous vs. Continuous (status quo), Continuous vs. Discrete, and Discrete vs. Discrete to study the incentives to adopt. Consider both de novo and incumbents.
Frequent Batch Auctions ("Discrete")

- **Time in discrete units**
  - In context of competition among exchanges, time interval very fast, e.g. 1 millisecond or potentially even finer
  - Long enough for a computer to batch process if multiple TFs act at "the same time" in response to public information
  - Short enough that an investor does not care about delay per se, only the price they pay

- **Orders are just as in continuous market**
  - Price, quantity, buy/sell
  - Remain outstanding until either executed or canceled

- At end of each discrete interval, FBA processes new orders alongside outstanding orders from before using uniform-price auction.
  - Priority is price then (discrete) time
  - Information policy: at end of each time interval, publicly announce any trades and updated state of the order book
Continuous vs. Discrete: Model

- **STAGE 1**: Continuous and discrete choose trading fees $f$. Continuous chooses ESST fees $F$; discrete does not sell ESST.

- **STAGE 2**: $N$ TFs choose whether to purchase ESST on continuous. On discrete, no decision to make.

- **STAGE 3 [Repeated]**: Trading game, essentially as before.
  - **Only modeling difference**: if multiple orders received at the same time (e.g., after public news), discrete processes these orders in batch, using a uniform-price auction, rather than serially. This eliminates latency arbitrage.
  - **Same as before**: Investors / Informed traders can trade on all venues before others act in response. Justified by fast discrete time interval (e.g., 1 millisecond).
Continuous vs. Discrete: Bid-Ask Spreads

- Recall: equilibrium spread for Continuous with fee of zero (or set of multiple competing Continuous) is characterized by

\[ \lambda_{\text{invest}} \frac{s^*_{\text{continuous}}}{2} = (\lambda_{\text{public}} + \lambda_{\text{private}})L(s^*_{\text{continuous}}) \]

where \( L(s) \equiv Pr(J > \frac{s}{2})E(J - \frac{s}{2}|J > \frac{s}{2}) \)

- If a single Discrete exchange operated in isolation w/ trading fee of zero, equilibrium spread:

\[ \lambda_{\text{invest}} \frac{s^*_{\text{discrete}}}{2} = (\lambda_{\text{private}})L(s^*_{\text{discrete}}) \]

where \( s^*_{\text{discrete}} < s^*_{\text{continuous}} \). Eliminating sniping reduces the spread. Like eliminating a tax on liquidity.
Continuous vs. Discrete: Equilibrium

Proposition: In any equilibrium of the multi-exchange game among a single Continuous and a single Discrete exchange:

- Discrete gets 100% share: in every iteration of the trading game, exactly one unit of liquidity is provided on Discrete, and no liquidity is provided on Continuous
- Continuous earns zero profits
- Discrete charges strictly positive trading fees and earns expected per-trading-game profits that exceed $N - 1/N \Pi_{\text{continuous}}$

Intuition:
- Eliminating latency arbitrage like eliminating a tax on liquidity
- Frictionless search and access -> if two markets operate in parallel, one with a tax and one without, the one without the tax wins.

Caveat: please do not take 100% literally (tick constraints, agency frictions. See Appendix E).
**Proposition**: In *any* equilibrium of the multi-exchange game among a single Continuous and a single Discrete exchange:

- Discrete gets 100% share: in every iteration of the trading game, exactly one unit of liquidity is provided on Discrete, and no liquidity is provided on Continuous
- Continuous earns zero profits
- Discrete charges strictly positive trading fees and earns expected per-trading-game profits that exceed $\frac{N-1}{N} \Pi^*_{\text{continuous}}$

**Intuition**:
- Eliminating latency arbitrage like eliminating a tax on liquidity
- Frictionless search and access $\rightarrow$ if two markets operate in parallel, one with a tax and one without, the one without the tax wins.

**Caveat**: please do not take 100% literally (tick constraints, agency frictions. See Appendix E).
Proposition: In any equilibrium of the multi-exchange game with at least two Discrete exchanges:

- At least one Discrete charges zero trading fees
- Volume split across Discrete exchanges that charge zero
- All exchanges and trading firms earn zero profits
- Interior equilibria just like C vs. C, also with Bertrand competition on fees ... but now no ESST fees.
Proposition: In any equilibrium of the multi-exchange game with at least two Discrete exchanges:

- At least one Discrete charges zero trading fees
- Volume split across Discrete exchanges that charge zero
- All exchanges and trading firms earn zero profits
- Interior equilibria just like C vs. C, also with Bertrand competition on fees ... but now no ESST fees.
Proposition: Prisoner’s Dilemma

- Exchanges’ economic profits as a function of their market designs constitute a prisoner’s dilemma:
  - All Continuous: each exchange $j$ earns $NF^*_j$ per trading game (Proposition 3.2)
  - Single Discrete, rest Continuous: Discrete exchange earns economic profits $\Pi_D^* > N - 1/N \Pi^*_\text{continuous}$, Continuous earn zero (Proposition 5.2)
  - Multiple Discrete: all exchanges earn zero economic profits (Proposition 5.3)
Proposition: Prisoner’s Dilemma

▶ Exchanges’ economic profits as a function of their market designs constitute a prisoner’s dilemma:

▶ All Continuous: each exchange $j$ earns $NF_j^*$ per trading game (Proposition 3.2)
Proposition: Prisoner’s Dilemma

- Exchanges’ economic profits as a function of their market designs constitute a prisoner’s dilemma:
  - All Continuous: each exchange \( j \) earns \( NF_j^* \) per trading game (Proposition 3.2)
  - Single Discrete, rest Continuous: Discrete exchange earns economic profits \( \Pi_D > \frac{N-1}{N} \Pi_{continuous}^* \), Continuous earn zero (Proposition 5.2)
Proposition: Prisoner’s Dilemma

- Exchanges’ economic profits as a function of their market designs constitute a prisoner’s dilemma:
  - All Continuous: each exchange \( j \) earns \( NF_j^* \) per trading game (Proposition 3.2)
  - Single Discrete, rest Continuous: Discrete exchange earns economic profits \( \Pi^D > \frac{N-1}{N} \Pi^*_\text{continuous} \), Continuous earn zero (Proposition 5.2)
  - Multiple Discrete: all exchanges earn zero economic profits (Proposition 5.3)
Proposition: Prisoner’s Dilemma

- Exchanges’ economic profits as a function of their market designs constitute a prisoner’s dilemma:
  - All Continuous: each exchange $j$ earns $NF_j^*$ per trading game (Proposition 3.2)
  - Single Discrete, rest Continuous: Discrete exchange earns economic profits $\Pi^D > \frac{N-1}{N} \Pi^*_{continuous}$, Continuous earn zero (Proposition 5.2)
  - Multiple Discrete: all exchanges earn zero economic profits (Proposition 5.3)
Prisoner’s Dilemma Among Incumbents

- Frank Hatheway, at this same seminar in Nov 2013 when asked whether NASDAQ would consider trying frequent batch auctions:

  Technologically, we could do it. The big issue, one of the big issues for us, when I talked about cost, the cost we would bear, would be getting [the SEC] to approve it, which would take a lot of time and effort, and if we got it approved, it would immediately be copied by everybody else. ... So we would have essentially no first-mover advantage if we put it in there, we would have no incentive to go through the lift of creating [the new market design].
Frank Hatheway, at this same seminar in Nov 2013 when asked whether NASDAQ would consider trying frequent batch auctions:

*Technologically, we could do it. The big issue, one of the big issues for us, when I talked about cost, the cost we would bear, would be getting [the SEC] to approve it, which would take a lot of time and effort, and if we got it approved, it would immediately be copied by everybody else. ... So we would have essentially no first-mover advantage if we put it in there, we would have no incentive to go through the lift of creating [the new market design].*
Discussion of Policy Implications

Analysis suggests two key wedges between the private and social incentives to adopt market designs that address latency arbitrage:

1. Innovator only earns profits commensurate with social value for the period of time before the market design innovation is imitated. Market gets a benefit of $\Pi$ in perpetuity, innovator only gets a benefit until imitated. Applies whether innovator is de novo or incumbent.

2. Incumbents face an additional wedge: lose the net present value of economic rents from the speed race. Empirics suggest this is large. These gaps suggest it is possible that private-market forces alone may not solve the problem. Private < social. On the other hand, analysis suggests that if there is an entrant, will gain share. Implication: “push”.
Discussion of Policy Implications

Analysis suggests two key wedges between the private and social incentives to adopt market designs that address latency arbitrage:

1. Innovator only earns profits commensurate with social value for the period of time before the market design innovation is imitated.
   - Market gets a benefit of $\Pi_{continuous}^*$ in perpetuity, innovator only gets a benefit until imitated.
   - Applies whether innovator is de novo or incumbent
Discussion of Policy Implications

Analysis suggests two key wedges between the private and social incentives to adopt market designs that address latency arbitrage:

1. Innovator only earns profits commensurate with social value for the period of time before the market design innovation is imitated.
   - Market gets a benefit of $\Pi^{\text{continuous}}$ in perpetuity, innovator only gets a benefit until imitated.
   - Applies whether innovator is de novo or incumbent

2. Incumbents face an additional wedge: lose the net present value of economic rents from the speed race
   - Empirics suggest this is large.
Discussion of Policy Implications

Analysis suggests two key wedges between the private and social incentives to adopt market designs that address latency arbitrage:

1. Innovator only earns profits commensurate with social value for the period of time before the market design innovation is imitated.
   - Market gets a benefit of $\Pi^*_\text{continuous}$ in perpetuity, innovator only gets a benefit until imitated.
   - Applies whether innovator is de novo or incumbent

2. Incumbents face an additional wedge: lose the net present value of economic rents from the speed race
   - Empirics suggest this is large.

These gaps suggest it is possible that private-market forces alone may not solve the problem. Private $<<<<$ social.

On the other hand, analysis suggests that if there is an entrant, will gain share.

Implication: “push”.
Concluding Remarks

- SEC Chair White asked “whether trading venues have sufficient opportunity and flexibility to innovate successfully with initiatives that seek to deemphasize speed as a key to trading success”
Concluding Remarks

- SEC Chair White asked “whether trading venues have sufficient opportunity and flexibility to innovate successfully with initiatives that seek to deemphasize speed as a key to trading success”

- Our model clarifies why, even if allowed, exchanges may not want to innovate: exchanges profit from the speed race generated by the current market design
Concluding Remarks

- SEC Chair White asked “whether trading venues have sufficient opportunity and flexibility to innovate successfully with initiatives that seek to deemphasize speed as a key to trading success”

- Our model clarifies why, even if allowed, exchanges may not want to innovate: exchanges profit from the speed race generated by the current market design

- Modest proposals to better align private incentives with social interests. A push could help the market fix the market.
Concluding Remarks

▶ SEC Chair White asked “whether trading venues have sufficient opportunity and flexibility to innovate successfully with initiatives that seek to deemphasize speed as a key to trading success”

▶ Our model clarifies why, even if allowed, exchanges may not want to innovate: exchanges profit from the speed race generated by the current market design

▶ Modest proposals to better align private incentives with social interests. A push could help the market fix the market.

▶ Standalone contribution: IO model of the modern U.S. stock exchange industry
Concluding Remarks

- SEC Chair White asked “whether trading venues have sufficient opportunity and flexibility to innovate successfully with initiatives that seek to deemphasize speed as a key to trading success”

- Our model clarifies why, even if allowed, exchanges may not want to innovate: exchanges profit from the speed race generated by the current market design

- Modest proposals to better align private incentives with social interests. A push could help the market fix the market.

- Standalone contribution: IO model of the modern U.S. stock exchange industry

- Also potentially useful: “virtual single platform” idea for platforms literature, incentives for market design innovation
Concluding Remarks

- SEC Chair White asked “whether trading venues have sufficient opportunity and flexibility to innovate successfully with initiatives that seek to deemphasize speed as a key to trading success”
- Our model clarifies why, even if allowed, exchanges may not want to innovate: exchanges profit from the speed race generated by the current market design
- Modest proposals to better align private incentives with social interests. A push could help the market fix the market.
- Standalone contribution: IO model of the modern U.S. stock exchange industry
- Also potentially useful: “virtual single platform” idea for platforms literature, incentives for market design innovation
- Many directions for future research