High-Frequency Market Making to Large Institutional Trades

Robert A. Korajczyk, Northwestern University
Dermot Murphy, University of Illinois at Chicago

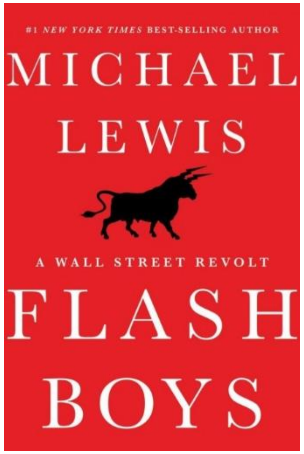
The HFT Debate

- In contemporary equity markets, HFTs have largely assumed the market maker role
- Substantial debate about their net effect on market quality, however
  - **Pros**: compete with each other to provide liquidity, leading to narrower bid-ask spreads and greater price efficiency
  - **Cons**: under no strict obligation to make markets, which can exacerbate volatility during times of stress or increased price pressure from large traders

The HFT Debate

- Warren Buffett on HFTs:
  - Small investors have "never had it so good"
  - Although the "big orders" are more costly

- The increased costs for these big orders through the "phantom liquidity" channel is also touched upon in Lewis' divisive book, *Flash Boys*

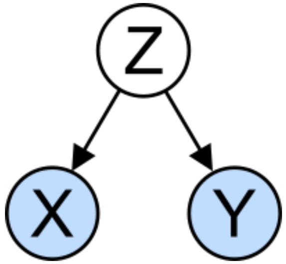


Our Central Questions

1) Are large institutional orders more costly in the presence of HFTs?
2) Do HFTs trade off of information inferred from institutional child orders?
3) How do HFT inventories change during the execution of a large institutional order?

Causality

- Establishing causality is challenging
  - Higher institutional execution costs and HFT activity may both be driven by outside forces
  - To address this, we require an event that directly affected HFT but not institutional trades



- Our event: on April 1, 2012, a Canadian regulation went into effect that taxed message activity, increasing the cost of HFT strategies
  - Regulation called the **Integrated Fee Model (IFM)**
  - HFT order submissions decreased by about 20% as a result

Our Central Results (1)

- Following the introduction of the IFM:
  - Price impact for institutional trades fell by 15%
  - The bid-ask spread increased by 3 bps
- Implication:
  - HFT is associated with higher costs for larger trades, lower costs for smaller trades
  - Trades above the $2.1 million break-even threshold benefited from the IFM

Our Central Results (2)

- Following the introduction of the IFM:
  - Price impact fell by about 28% for informationally-motivated trades compared to only 11% for liquidity-based trades
  - Implication: HFTs trade off information that they infer from the child orders of informed traders

- HFTs revert their inventories about 50% faster during a large trade execution
  - That is, about one-third of the inventory reversion can be attributed to information inferred from the large trade
  - Unconditional on their inventories, HFTs are also more likely to trade in the same direction as a large trade

Related Literature (Theory)

- Yang and Zhu (2019)
  - HFTs “back-run” institutional orders that are executed over two periods in a Kyle model setting
- Ait-Sahalia and Saglam (2017)
  - HFTs reprice limit orders if they anticipate an impatient low-frequency trader

Related Literature (Empirical)

- Van Kervel and Menkveld (2019)
  - HFTs in Sweden eventually trade in the same direction as an institutional order
- Malinova, Park, and Riordan (2018) (MPR)
  - The IFM increased the average retail effective spread
- Anand and Venkataraman (2016)
  - The liquidity provision of market-making HFTs depends on volatility, inventory risk, and their trading profits

Data

- Order-level data for all Canadian equities from January 2012 to June 2013
  - Access provided by IIROC, a non-governmental self-regulatory organization (like FINRA)
  - Each record provides masked identification of the trader submitting the order, allowing us to track individual traders over time
  - Approximately 60 billion observations

IIROC
Investment Industry
Regulatory Organization
of Canada

Classifying Market-Making HFTs

- For each user ID, we calculate their median time between order submission and cancellation
  - A trader is classified as an HFT if their median time is below 250 milliseconds
  - A trader is also classified as an HFT if they frequently trade in the first 500 milliseconds after 3:40pm, when information about the closing call auction is disseminated
  - Using this procedure, we classify 103 IDs as HFTs

Classifying Market-Making HFTs

- We use the "market-maker index" (MMI) from Comerton-Forde, Malinova, and Park (2018) to identify the subset of market-making HFTs

$$\text{MMI}_{i,j,d} = \left| \frac{\text{Passive Buy Order Volume}_{i,j,d} - \text{Passive Sell Order Volume}_{i,j,d}}{\text{Passive Buy Order Volume}_{i,j,d} + \text{Passive Sell Order Volume}_{i,j,d}} \right|$$

- An HFT is classified as a market-maker (HFTMM) is their median MMI is below 0.20
  - Using this procedure, we classify 68 IDs as HFTMMs

HFTMM Summary Statistics (Table 1)

	HFT summary statistics (N = 67,787)						
	Mean	Median	P5	P25	P75	P95	SD
Percentage of trade volume (%)	31.6	30.8	11.5	22.0	40.6	53.4	13.1
Percentage of orders (%)	55.4	56.0	21.0	41.3	69.2	85.9	22.9
Order-to-trade ratio	33.1	16.9	5.4	10.5	32.7	119.8	49.5
Aggressiveness (%)	27.8	26.9	7.8	18.2	36.2	50.9	13.3
Trade size (shares)	328	147	111	125	260	1,261	531
Trade value (dollars)	4,354	2,685	459	1,092	5,531	12,133	6,095
Inventory ($K)	3.7	1.3	-105.5	-16.9	23.6	119.4	72.9
Inventory (%)	2.5	0.2	-49.8	-3.3	5.8	63.6	52.1
ΔInventory ($K)	0.0	0.0	-55.6	-7.5	7.5	55.8	48.7
ΔInventory (%)	0.0	0.0	-100.0	-19.4	18.9	100.0	46.9

Classifying Institutional Trades

- An institutional trade is classified as follows:
  - At least $100 thousand of same-direction trades originating from the same user ID over one or more days
- For each institutional trade, we also calculate its implementation shortfall (IS), the main dependent variable in our analysis:

$$IS_{i,t} = \frac{\sum_{n=1}^{N} p_n x_{i,n} - p_0 x_{i,N}}{p_0 x_{i,N}} \times (\mathbf{1}_B - \mathbf{1}_S)$$

  - This measures the percentage difference between what the institution paid versus what they would have paid if all their shares were executed at the initial bid-ask midpoint
    - (For sells: what they would have received versus what they received)

Institutional Trade Summary Stats (Table 2)

Institutional trade statistics (N = 1,173,482)

	Mean	Median	P5	P25	P75	P95	SD
Trade size ($M)	0.72	0.28	0.11	0.16	0.64	2.53	1.91
Number of orders	234	48	1	11	178	855	2,207
Number of trades	118	50	3	20	124	438	261
Order-to-trade ratio	4.9	1.0	0.1	0.4	1.8	6.4	36.2
Aggressiveness (%)	57.0	61.1	0.0	22.1	96.5	100.0	36.6
Time to completion (hours)	3.0	1.7	0.0	0.1	5.3	6.5	4.0
Implementation shortfall (bps)	7.1	2.5	-97.9	-8.8	23.0	119.3	81.9

The Integrated Fee Model

- Our baseline empirical strategy involves examining implementation shortfall around the regulatory change on April 1, 2012 which especially affected HFTs

- **The Integrated Fee Model (IFM)**: traders would now be charged on a pro-rata basis for the messages they submit to exchanges
  - Why this regulation? Message traffic was steadily increasing over time, making it costlier for IIROC to monitor this traffic
  - The IFM was a way for IIROC to recoup some of these costs
  - MPR estimate this fee to be about $0.00026 per message

The Integrated Fee Model

- Pro-HFT commenters expressed concern about this regulation:
  - *“The regulation would extend an apparent bias against HFTs.”*
  - *“Taxing message traffic will disproportionately hurt HFTs.”*
- In response, IIROC stated that they developed the regulation “to be as neutral as possible between liquidity providers and liquidity takers.”

The Integrated Fee Model

- After the regulation was implemented, there was a notable drop in HFT messages and trades

B. Daily number of HFT orders

	Pre-regulation	Post-regulation	Percentage change
Mean	116,783	91,778	-21.4%***
25th percentile	29,463	21,444	-27.2%
Median	60,590	50,989	-15.8%
75th percentile	161,291	137,355	-14.8%

  - There were also similar drops in the number of HFT cancellations and trades
  - Institutional trade activity was unaffected, however

Empirical Strategy

- We test the effect of the IFM on implementation shortfall (IS) using the following OLS regression model:

$$IS_{i,j,t} = \beta_1 \cdot \ln(TSize_{i,j,t}) + \beta_2 \cdot Fee_t + \beta_3 \cdot (Fee_t \times \ln(TSize_{i,j,t})) + \gamma \cdot X_{i,j,t} + \delta_j + \varepsilon_{i,j,t}.$$

- Key coefficients:
  - **$\boldsymbol{\beta_2}$: impact of the IFM (*Fee*) on the spread**
  - **$\boldsymbol{\beta_3}$: impact of the IFM (*Fee*) on price impact**

  - (X denotes control variables and δ_j denotes stock fixed effects)

Baseline Results (Table 4)

	[-3,+3] months		All months	Size > $500K	Size > $1M
	(1)	(2)	(3)	(4)	(5)
$\ln(TSize)$	6.742***	8.938***	9.114***	13.141***	15.569***
	(22.07)	(26.6)	(29.67)	(20.69)	(15.72)
Fee	3.002***	3.615***	2.859***	2.057**	2.261**
	(4.25)	(4.50)	(5.26)	(2.46)	(2.20)
$\ln(TSize) \times Fee$	-0.981**	-1.135**	-1.467***	-1.976***	-2.778**
	(-2.16)	(-2.36)	(-4.46)	(-2.73)	(-2.40)
SE clustering	Stock-date	Stock-date	Stock-date	Stock-date	Stock-date
Fixed effects	Stock	Stock	Stock	Stock	Stock
N	279,140	251,584	733,890	263,419	141,739
R-squared	0.061	0.071	0.063	0.077	0.085

Interpretation

- The post-IFM price impact for large institutional trades decreased by about **0.98/6.74 = 15%**
- The post-IFM spread increased by about **3 basis points**

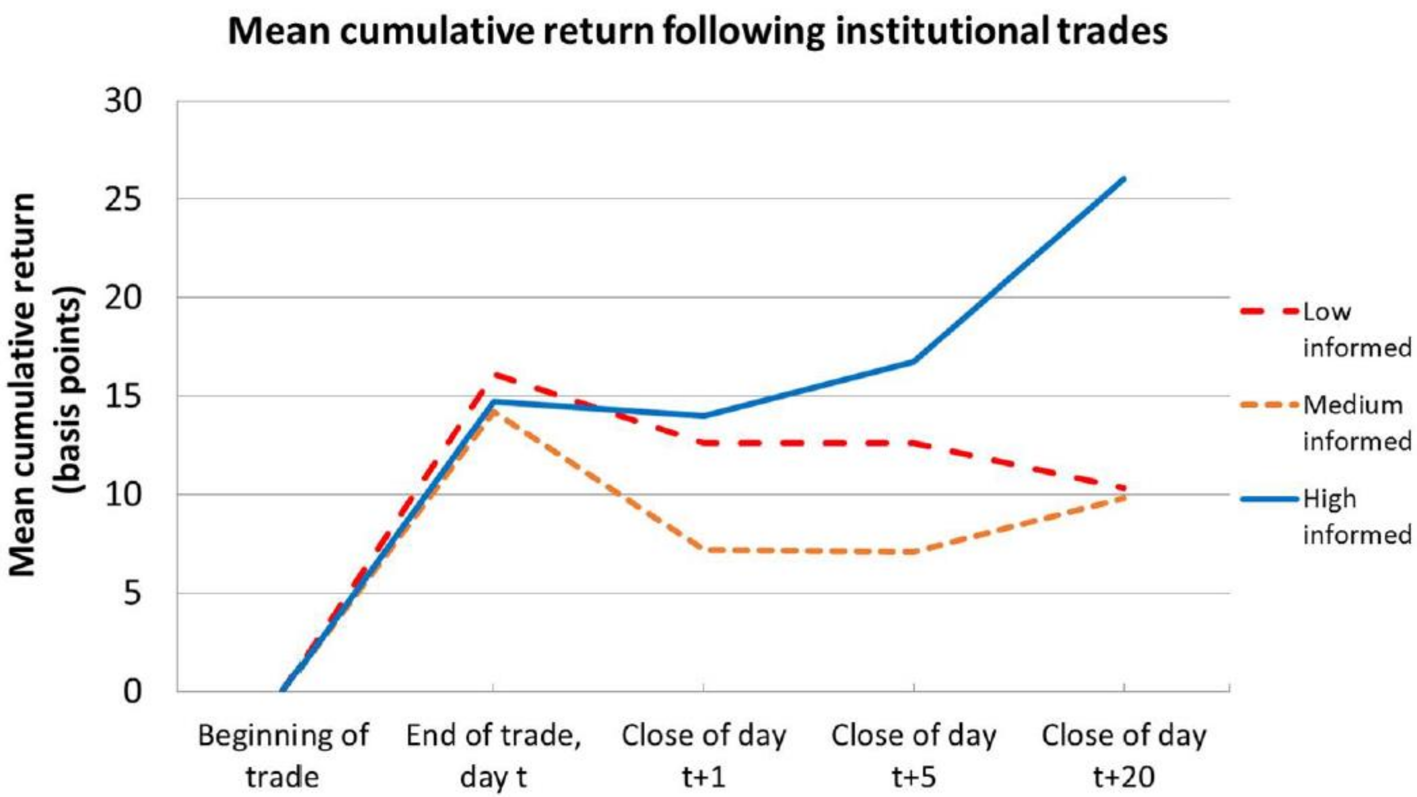
- The cost-reduction break-even point: **$2.1 million**
  - That is, post-IFM execution costs for trades above this size threshold decreased
  - Trades above this threshold account for **45%** of trading volume from our sample of institutional trades
  - Or about **$380 billion** of institutional trading volume

Information-Based Trading

- Our results indicate that HFTs are associated with higher execution costs for large institutional trades
- Are HFTs trading off information inferred from institutional trades?
  - To answer this question, we analyze the differential effect of the IFM on the execution costs for informed versus uninformed traders
  - For each month, we place each institutional trader into an "informed" tercile (HIGH, MEDIUM, or LOW) based on the average five-day return performance of their trades

Information-Based Trading (Fig. 1)

- First, we show that informed traders profitably trade out of sample, suggesting some degree of skill (and not luck)



IFM Effect by Informed Type (Table 5)

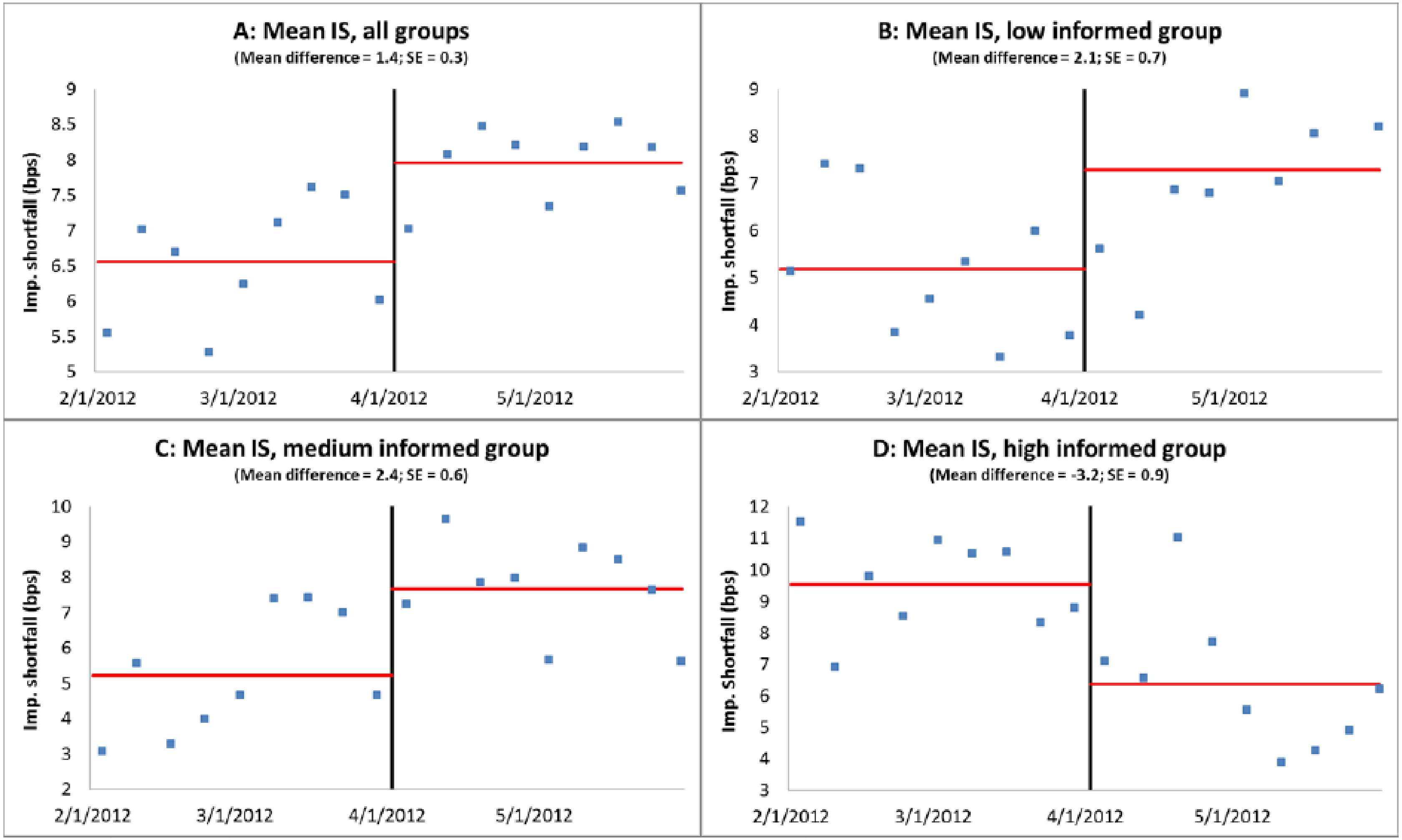
	Trader informativeness			
	High	Medium	Low	Pooled
	(1)	(2)	(3)	(4)
$\ln(TSize)$	9.790***	8.797***	8.581***	9.067***
	(14.98)	(19.17)	(10.50)	(24.02)
Fee	0.424	2.129***	4.579***	0.156
	(0.35)	(3.07)	(3.43)	(0.14)
$Fee \times \ln(Tsize)$	-2.693***	-0.889*	-1.069	-2.342***
	(-3.74)	(-1.86)	(-1.26)	(-4.97)
$Fee \times \mathbf{1}_M$				2.138*
				(1.80)
$Fee \times \ln(TSize) \times \mathbf{1}_M$				1.203***
				(3.56)
$Fee \times \mathbf{1}_L$				4.202***
				(2.81)
$Fee \times \ln(TSize) \times \mathbf{1}_L$				1.333***
				(3.54)

IFM Effect by Informed Type

- Price impact reductions from the pooled regression in column (4) indicate that the reduction was highest for the high-informed institutional traders:

Informed Type	Price Impact Reduction
High	25.8%
Medium	12.6%
Low	11.1%

IFM Effect by Informed Type (Fig. 2)



HFT Inventory Dynamics

- We have established that HFT is associated with higher execution costs for large trades, especially those submitted by informed traders
- To better understand the mechanism through which this effect occurs, we examine HFT trading and inventory dynamics during large trade executions

HFT Inventory Dynamics

- We use the methodology in Hansch, Naik, and Viswanathan (1998) (HNV) to analyze HFT inventory dynamics

$$\Delta I_{m,j,t} = \alpha + \sum_{k=1}^{4} \beta_k D^k I_{m,j,t-1} + \varepsilon_{m,j,t-1}$$

  - $\Delta I_{m,j,t}$: normalized inventory change for HFT m in stock j and 15-minute period t
  - $I_{m,j,t-1}$: HFT inventory level in previous 15-minute period
  - D^k : indicator variable for extremity of inventory position, $k \in \{1,2,3,4\}$ (D^4: most extreme)
  - $L_{j,t-1}$: signed indicator variable for inst. trade (next slide)

HFT Inventory Dynamics (Table 6)

	HFT net inventory change ($\Delta I_{m,j,t}$)			
	Full sample	Inst. buy	Inst. sell	Full sample
	(1)	(2)	(3)	(4)
$I_{m,j,t-1} \times D^1$	-0.276***	-0.323***	-0.329***	-0.216***
	(-126.8)	(-101.4)	(-102.9)	(86.9)
$I_{m,j,t-1} \times D^2$	-0.281***	-0.330***	-0.332***	-0.218***
	(-111.2)	(-105.9)	(-110.5)	(-77.4)
$I_{m,j,t-1} \times D^3$	-0.312***	-0.360***	-0.362***	-0.254***
	(-75.6)	(-67.8)	(-72.0)	(-53.7)
$I_{m,j,t-1} \times D^4$	-0.377***	-0.398***	-0.418***	-0.347***
	(-25.4)	(-22.6)	(-26.4)	(-16.6)
$(I_{m,j,t-1} \times D^1) \times \lvert L_{j,t-1} \rvert$				-0.109***
				(-38.0)
$(I_{m,j,t-1} \times D^2) \times \lvert L_{j,t-1} \rvert$				-0.112***
				(-42.0)
$(I_{m,j,t-1} \times D^3) \times \lvert L_{j,t-1} \rvert$				-0.106***
				(-22.4)
$(I_{m,j,t-1} \times D^4) \times \lvert L_{j,t-1} \rvert$				-0.069***
				(-2.8)
$L_{j,t-1}$				0.020***
				(19.1)

Column (1) takeaways:

- For nonextreme positions (D^1), HFTs revert 27.6% of their position in the following 15 minutes
- For extreme positions (D^4), HFTs revert 37.7% of their position in the following 15 minutes

HFT Inventory Dynamics (Table 6)

HFT net inventory change ($\Delta I_{m,j,t}$)

	Full sample	Inst. buy	Inst. sell	Full sample
	(1)	(2)	(3)	(4)
$I_{m,j,t-1} \times D^1$	-0.276***	-0.323***	-0.329***	-0.216***
	(-126.8)	(-101.4)	(-102.9)	(86.9)
$I_{m,j,t-1} \times D^2$	-0.281***	-0.330***	-0.332***	-0.218***
	(-111.2)	(-105.9)	(-110.5)	(-77.4)
$I_{m,j,t-1} \times D^3$	-0.312***	-0.360***	-0.362***	-0.254***
	(-75.6)	(-67.8)	(-72.0)	(-53.7)
$I_{m,j,t-1} \times D^4$	-0.377***	-0.398***	-0.418***	-0.347***
	(-25.4)	(-22.6)	(-26.4)	(-16.6)
$(I_{m,j,t-1} \times D^1) \times \lvert L_{j,t-1}\rvert$				-0.109***
				(-38.0)
$(I_{m,j,t-1} \times D^2) \times \lvert L_{j,t-1}\rvert$				-0.112***
				(-42.0)
$(I_{m,j,t-1} \times D^3) \times \lvert L_{j,t-1}\rvert$				-0.106***
				(-22.4)
$(I_{m,j,t-1} \times D^4) \times \lvert L_{j,t-1}\rvert$				-0.069***
				(-2.8)
$L_{j,t-1}$				0.020***
				(19.1)

Column (4) takeaways:

- HFTs revert 21.6% of their position in the following 15 minutes
- Increases to 32.5% when an inst. trade (L) is underway
- HFTs trade in the same direction as the inst. trade, even unconditional on inventory

HFT Quote Dynamics (Table 7)

	HFT net order submission ($Q_{m,j,t}$)			
	(1)	(2)	(3)	(4)
$I_{m,j,t-1} \times D^1$	-0.096***	-0.094***	-0.094***	-0.078***
	(-41.9)	(-41.3)	(-41.0)	(-24.6)
$I_{m,j,t-1} \times D^2$	-0.102***	-0.100***	-0.100***	-0.082***
	(-54.6)	(-54.3)	(-54.2)	(-36.5)
$I_{m,j,t-1} \times D^3$	-0.112***	-0.110***	-0.109***	-0.085***
	(-40.3)	(-39.9)	(-39.9)	(-24.2)
$I_{m,j,t-1} \times D^4$	-0.124***	-0.123***	-0.122***	-0.093***
	(-14.5)	(-14.5)	(-14.5)	(-9.6)
$L_{m,j,t-1}$		0.040***	0.020***	0.019***
		(25.7)	(14.0)	(13.4)
$(I_{m,j,t-1} \times D^1) \times \lvert L_{m,j,t-1}\rvert$				-0.028***
				(-6.4)
$(I_{m,j,t-1} \times D^2) \times \lvert L_{m,j,t-1}\rvert$				-0.031***
				(-10.2)
$(I_{m,j,t-1} \times D^3) \times \lvert L_{m,j,t-1}\rvert$				-0.044***
				(-9.1)
$(I_{m,j,t-1} \times D^4) \times \lvert L_{m,j,t-1}\rvert$				-0.066***
				(-4.5)

Column (1) takeaway:

- HFTs quote more aggressively in the opposite direction to their inventory position

Column (4) takeaway:

- Opposite direction quoting even more aggressive when an institutional trade is underway

Buying and Selling by HFTs (Table 8)

	Buy volume	Sell volume	Buy orders	Sell orders
	(1)	(2)	(3)	(4)
$\mathbf{1}_B$	0.061***	0.026**	0.048***	0.027**
	(5.2)	(2.3)	(3.7)	(2.2)
$\mathbf{1}_S$	0.022*	0.056***	0.017	0.036***
	(1.9)	(4.6)	(1.3)	(2.7)
SE clustering	Stock-Date	Stock-Date	Stock-Date	Stock-Date
N	1,576,111	1,576,111	1,576,111	1,576,111
R-squared	0.184	0.184	0.117	0.118

- The previous two tables showed that HFTs' net trading and quoting activities tend to be in the same direction as an institutional order
- This table focuses on buy and sell activity separately and shows that:
  - HFT buy volume/buy orders are more likely during an institutional buy ($\mathbf{1}_B$)
  - HFT sell volume/sell orders are more likely during an institutional sell ($\mathbf{1}_S$)

Signal Processing by HFTs

- What signals do HFTs use to detect and compete with institutional trades?
- To answer this question, we first test a probit regression model to identify predictors of institutional trades:

$$\Pr(\mathbf{1}_{j,t,z}|\cdot) = \sum_{k=1}^{4} \beta_{k,z} r_{j,t-k} + \lambda_{k,z} y_{j,t-k} + \phi_{k,z} LOIB_{j,t-k} + \delta_j + \varepsilon_{j,t-1}$$

- RHS variables: past returns (r), trade imbalances (y), limit order imbalances ($LOIB$)

Inst. Trade Predictors (Table 9)

Probit regression

	Inst. buy	Inst. sell
	(1)	(2)
r_{t-1}	0.057***	-0.052***
r_{t-2}	0.047***	-0.041***
r_{t-3}	0.039***	-0.034***
r_{t-4}	0.036***	-0.030***
y_{t-1}	0.004**	-0.001
y_{t-2}	0.006***	-0.004**
y_{t-3}	0.008***	-0.006***
y_{t-4}	0.008***	-0.007***
$LOIB_{t-1}$	0.023***	-0.025***
$LOIB_{t-2}$	0.011***	-0.013***
$LOIB_{t-3}$	0.007***	-0.008***
$LOIB_{t-4}$	0.008***	-0.009***

Main takeaway:

- Positive past returns, trade imbalances, and limit order imbalances predict institutional buy orders
- Negative past returns, trade imbalances, and limit order imbalances predict institutional sell orders

Signal Processing by HFTs

- We use these variables to predict L, a signed indicator variable for an institutional trade execution, where $L \in \{-1,0,1\}$
- Then, we test the effect of the predicted value of L on the contemporaneous change in the HFT inventory level
- We also run the same test using only the subset of *aggressive* institutional trades
  - Definition: institutional trades in which at least 57% of the child orders are executed using marketable limit orders (this number represents the median aggressiveness)
  - Aggressive trades should be easier to detect by other traders

Inst. Trade Predictors (Table 9)

	$\Delta I_{m,j,t}$	Agg. trades
	(4)	(5)
$I_{m,j,t-1} \times D_1$	-0.277***	-0.279***
$I_{m,j,t-1} \times D_2$	-0.284***	-0.286***
$I_{m,j,t-1} \times D_3$	-0.314***	-0.315***
$I_{m,j,t-1} \times D_4$	-0.378***	-0.379***
Pred. $L_{j,t}$	0.161***	0.338***
Clustering	Stock-date	Stock-date
N	1,490,416	1,490,416
R-squared	0.113	0.118

- The anticipated direction of an institutional trade is predictive of the change in HFT inventory
- This is especially true for aggressive institutional trades, which are easier to detect by other traders

Conclusion

- In this study, we provide evidence that HFT is associated with:
  - Higher execution costs for large, informed trades via reduced depth
  - Lower execution costs for smaller trades via reduced spread
- The Integrated Fee Model provides an exogenous shock to HFT that helps to establish the causality of these findings
- Additional evidence indicates that HFTs are more likely trade in the same direction as large institutional trades

- Overall, our findings indicate that HFT has a multifaceted effect on market quality
- That is, there are market quality tradeoffs associated with large speed differentials across market participants