Identifying HFT Activity without Proprietary Data proprietary Data

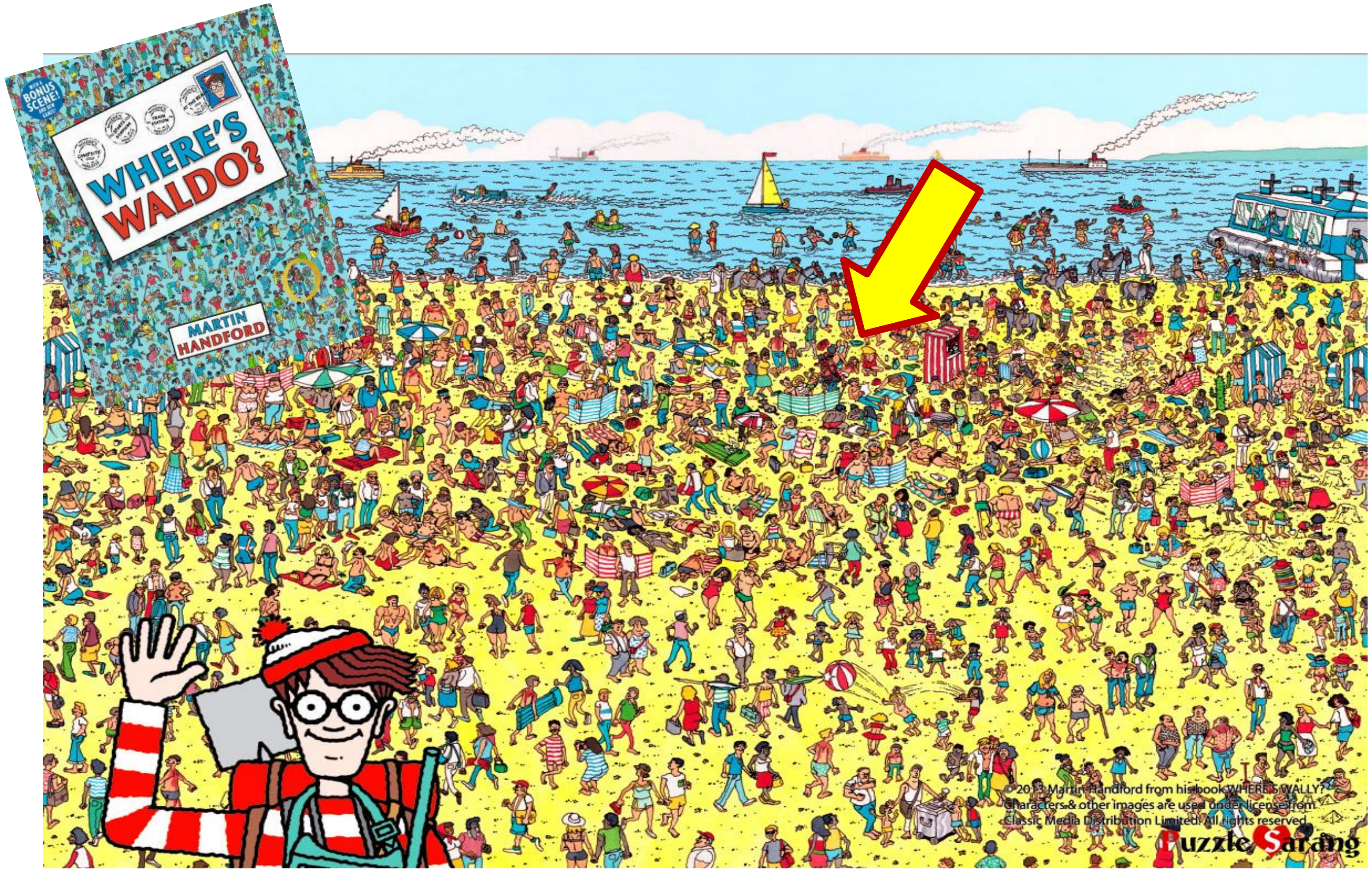
Bidisha Chakrabarty
(Saint Louis University, USA)

Carole Comerton-Forde
(Univ of Melbourse, Aust.)

Roberto Pascual
(University of the Balearic Islands, Spain

- NSE NYU Conference on Indian Capital Markets
  December, 2021

Identifying HFT without identifers in the data



Identification strategies

- **Proprietary databases**
  (e.g., Brogaard, et al., 2014; 2017; Kirilenko et al., 2017; Comerton-Forde et al., 2018; Boehmer et al., 2019)
  - Availability; Replicability, Coverage; Based on proxies (e.g., EUROFIDAI, IIROC)
- **Latency-changing exogenous events**
  (e.g., Hendershott et al., 2011; Riordan and Storkenmaier, 2012; Boehmer et al., 2020; Shkilko and Sokolov, 2020)
  - Unable to directly identify HFT activity; combined with proxies

  - **HFT proxies**

Proxy Metrics

- Message Traffic (*Mess*)
- Cancellations (*Can*)
- Monitoring Intensity (*MonInt*)
- Fleeting Orders (*FleetOrd*)
- Quotation intensity (*QuoteInt*)
- Quote flickering (*Flick*)
- Speed of response (*SResp*)
- Strategic runs (*SRuns*)

Our study: research questions

- Comprehensive examination of the **8** most popular HFT proxies.
- RQ#1: Reliability
- RQ#2: Type of HFT activities
- RQ#3: Identify HFTs but not other traders

Precise HFT identifiers – Our Data

- National Stock Exchange (NSE) of India: 4th (10th) largest Exchange by #trades (dollar volume)

- Ranked just below the TSX in market cap;
- TSX 2.6 trillion USD NSE USD 2.55 trillion USD

In 2018, AT share of volume is almost 50%

Trader types

- Two internal flags:

	"Client account"	
"Order entry mode"	Proprietary	Agency
Algorithmic trader (AT)	High-frequency traders (**HFTs**)	Agency Algorithmic Traders (**AATs**)
Non-AT	Non-algorithmic traders (**NATs**)	

- Algorithmic order entry for prop trading = SEC definition of HFT

Metrics

- We consider 8 metrics used to proxy for HFT activity,
- Two versions:
  - **“True”** HFT metrics, using the HFT messages only
  - **“Proxy”** HFT metrics, using all messages (ignoring trader indification)

RQ1: How reliable are the proxies? Correlations between true & proxy metrics

	Metric							
Interval	*Mess*	*Can*	*MonInt*	*FleetOrd*	*QuoteInt*	*Flick*	*SResp*	*SRuns*
30s	96.77 ***	78.21 ***	97.13 ***	92.28 ***	88.93 ***	67.94 ***	93.81 ***	98.02 ***
60s	96.92 ***	79.56 ***	97.40 ***	92.97 ***	89.72 ***	71.53 ***	94.84 ***	97.83 ***
300s	96.93 ***	80.39 ***	97.67 ***	93.62 ***	90.33 ***	77.70 ***	95.72 ***	98.00 ***
900s	96.72 ***	80.02 ***	97.74 ***	93.49 ***	89.71 ***	78.69 ***	95.69 ***	98.30 ***
1500s	96.60 ***	79.56 ***	97.72 ***	93.43 ***	89.36 ***	78.75 ***	95.50 ***	98.45 ***

HFTs' contribution (*HFTCont*) to liquidity supply/demand

- Liquidity supply:
  - Trade-based metric: % trades in which HFTs are on the passive side
  - LOB-based metrics: best quotes, depth, top 5 levels of LOB
- Liquidity demand:
  - Trade-based contribution metric: % trades initiated by HFTs

$$HFTCont_{i,t} = \alpha + \beta HFTProxy_{i,t} + \gamma HFTCont_{i,t-1} + \lambda_o Open_{i,t} + \lambda_c Close_{i,t} + \delta_i + \varepsilon_{i,t}$$

Results: At all levels of aggregation **β** is positive and highly significant

High vs low HFT liquidity demand/supply

- Unusually high (low): >75 (<25) percentile of the corresponding indicator:
  - (Dhigh, Shigh): High demand & High supply
  - (Dhigh, Slow): High demand & Low supply
  - (Dlow, Shigh): Low demand & High supply
  - (Dlow, Slow): Low demand & Low supply

$$HFTMetric_{i,t} = \alpha + \beta_{LL} DlowSlow_{i,t} + \beta_{LH} DlowShigh_{i,t} + \beta_{HL} DhighSlow_{i,t} + \beta_{HH} DhighShigh_{i,t} + \\ + \lambda_o Open_{i,t} + \lambda_c Close_{i,t} + \delta_i + \varepsilon_{i,t}$$

High vs low HFT liquidity demand/supply

(proxy metrics; supply = % time at the best quotes)

	Mess	Can	MonInt	FleetOrd	QuoteInt	Flick	SResp	SRuns
(Dlow, Slow)	-245.98***	-10.55***	-98.83***	-29.46***	-83.43***	-0.18***	-22.10***	-2.67***
(Dhigh, Slow)	-113.00***	-5.74***	-43.60***	-11.56***	-44.15***	-0.07***	-7.26***	-1.32***
(Dlow, Shigh)	-114.26***	-5.94***	-41.27***	-13.36***	-30.18***	-0.07***	-0.45	-0.96***
(Dhigh, Shigh)	231.16***	9.69***	92.78***	29.98***	94.96***	0.20***	33.35***	2.58***

Can the proxies isolate HFTs from others?

- 1st stage: for each stock *i* and "true" metric

$$HFTtrue_{i,t} = \alpha + \beta_{AAT}AATtrue_{i,t} + \beta_{NAT}NATtrue_{i,t} + \widetilde{HFTtrue}_{i,t}, \forall \, \mathrm{i} = \{1, \ldots, 50\}$$

- 2nd stage: Pooled regression

$$\widetilde{HFTtrue}_{i,t} = \alpha + \beta HFTproxy_{i,t} + \lambda_o Open_{i,t} + \lambda_c Close_{i,t} + \delta_i + \varepsilon_{i,t}$$

- We repeat the process above for AATs and NATs true metrics

Can the proxies isolate HFTs from others?

2nd stage estimates - 30s intervals

Trader type	Statistic	*Mess*	*Can*	*MonInt*	*FleetOrd*	*QuoteInt*	*Flick*	*SResp*	*SRuns*
HFT	Coef.x100	35.02 ***	31.05 ***	51.28 ***	39.66 ***	52.29 ***	28.05 ***	34.94 ***	77.59 ***
	t-stat	(22.52)	(13.08)	(31.60)	(14.97)	(19.24)	(4.55)	(9.90)	(40.97)
	R^2 (2nd stage)	0.35	0.31	0.51	0.40	0.52	0.05	0.35	0.78
AAT	Coef.x100	2.42 ***	33.93 ***	2.25 ***	6.20 ***	7.54 ***	22.61 ***	4.69 ***	0.89 ***
	t-stat	(11.66)	(10.24)	(10.23)	(12.20)	(7.93)	(14.83)	(5.21)	(7.15)
	R^2 (2nd stage)	0.04	0.34	0.03	0.08	0.08	0.12	0.05	0.01
NAT	Coef.x100	0.32 ***	1.17 ***	0.17 ***	1.71 ***	0.27 ***	13.97 ***	0.37 ***	0.03 ***
	t-stat	(6.79)	(5.26)	(4.64)	(6.26)	(5.66)	(8.13)	(2.70)	(7.62)
	R^2 (2nd stage)	0.02	0.01	0.00	0.03	0.02	0.12	0.00	0.00
	Obs.	2419500	2419500	2419500	2413936	2400679	2419500	2231961	2419500

Conclusions

- Proxies of HFT activity:
  a) Perform well in identifying HFT activity
  b) Are highly correlated with each other
  c) Are good at identifying HFT liquidity demand as well as supply, but <u>cannot differentiate</u> them
  d) Their performance is not dependent on the level of time aggregation
  e) Hasbrouck and Saar's (2013) strategic runs outperform other proxies in capturing HFT-specific activity

Thank you!

Comments & suggestions welcome:

- bidisha.chakrabarty@slu.edu
- carole.comerton-forde@unimelb.edu.au
- rpascual@uib.es