50

# Occasional Paper

 Internet Appendix for: 'Quantifying the High-Frequency Trading "Arms Race": A Simple New Methodology and Estimates'

Matteo Aquilina, Eric Budish and Peter O'Neill
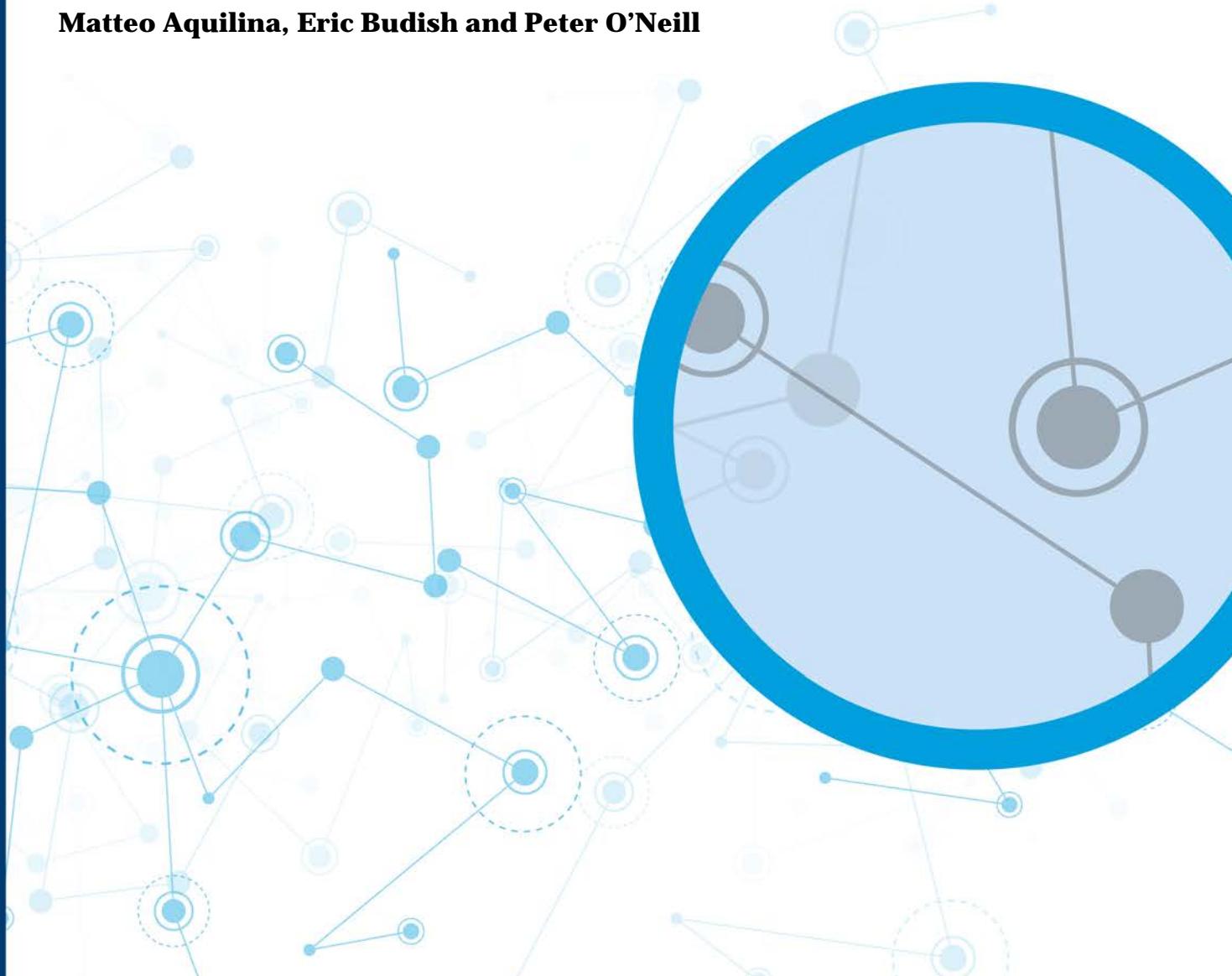
Online Appendix for "Quantifying the High-Frequency Trading 'Arms Race': A Simple New Methodology and Estimates"

Matteo Aquilina, Eric Budish, and Peter O'Neill

A   Additional Results

This online appendix contains additional results that are primarily alternate specifications of tables or figures reported in the main text. The results are presented in sequential order based on the location of the corresponding table or figure in the main text.

Additional Results Related to Computing the Information Horizon (Section 4.3)

Figure 4.1 in the main text reported the distribution of time between observed M1-M2 message pairs where M1 is an outbound message reporting a new limit order that has been added to the book, and M2 is an inbound message that is aggressive at the price level associated with M1. We used the spike in this distribution, at 29 microseconds, as an input into our computation of the information horizon.

The following figure reports an analogous analysis but with M1-M2 message pairs where M1 is an outbound message reporting that an existing limit order has been partially filled, and M2 is an inbound message that cancels the remainder of the limit order. The difference versus Figure 4.1 in the text is that in Figure 4.1 the response message M2 is sent by a different participant from M1, whereas in this figure, the participant who received M1 outbound then is the same participant who send M2 inbound. Thus, the difference in response times between this figure and Figure 4.1 reflects the difference in speed between reactions to a publicly disseminated book update, versus reactions to a privately received trade update. The former is more appropraite for computing the information horizon, but the latter may also be of interest and is reported here:

Figure A.1: **Distribution of Time between M1 Outbound partial fill → M2 Inbound Successful Cancel**
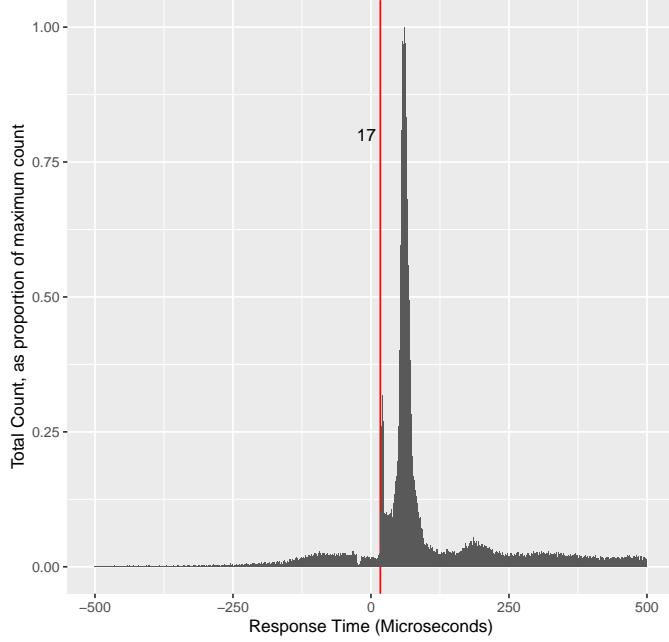
Notes: Over all regular-hour messages from four high-volume symbols, BP, GLEN, HSBA, VOD, we obtain all cases where some outbound message is a partial fill and a subsequent outbound message is a successful cancel. We then obtain the inbound cancel request message associated with the outbound cancel success message, and compute the difference in the message timestamp between the partial fill outbound message (M1) and the cancel request inbound message (M2). Note that this difference can be negative if M2's inbound is sent by the participant before M1's outbound is sent by the outbound gateway. The distribution depicted is a microsecond-binned histogram truncated at -500 microseconds and +500 microseconds. As described in the text of Section 4.3, we compute the start of the spike by computing the mean and standard deviation of the distribution in the period -100 microseconds to 0 microseconds, and then finding the first microsecond after 0 that is at least 5 standard deviations above this pre-0 mean.

Symbol-Date Version of Table 5.1

Table 5.1 reported the number of races per day at the symbol level averaged across all dates (Panel A), and at the date level averaged across all symbols (Panel B). The following table presents the number of races at the symbol-date level, i.e., without averaging across either symbols or dates.

Table A.1: **Number of Races Per Day Across Symbol-Dates**

Description	Mean	sd	Min	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99	Max
FTSE 100	537.24	542.96	29	73	152	215	346	629	1,194	2,635	7,014
FTSE 250	70.05	103.26	0	0	0	2	35	97	182	477	1,392
Full Sample	206.03	372.02	0	0	0	11	81	231	513	1,919	7,014

Notes: Please see Section 4.2 for a detailed description of the baseline race-detection criteria and Section 3 for details of the message data including how we classify inbound messages and how we maintain the order book. This table reports the distribution of the number of races detected at the symbol-date level. Table 5.1 in the main text reports the distribution at the symbol level and date level.

Total Time in Races

In the text of Section 5.1 we reported the distribution of the number of races per day (Table 5.1) and the distribution of the duration of races (Table 5.2). In this table we report the distribution of the total time in races per day. This is reported in seconds per day at the symbol-date level.

Table A.2: **Total Time in Races Across Symbol-Dates**

Description	Mean	sd	Min	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99	Max
FTSE 100	0.044	0.047	0.002	0.006	0.012	0.017	0.028	0.052	0.096	0.235	0.739
FTSE 250	0.005	0.008	0.000	0.000	0.000	0.000	0.002	0.007	0.013	0.036	0.093
Full Sample	0.016	0.032	0.000	0.000	0.000	0.001	0.006	0.018	0.042	0.153	0.739

Notes: For each race in our race records dataset (see notes for Table 5.1) we compute the difference in message timestamps between the first inbound message in the race that is a success and the first inbound message in the race that is a fail (success and fail are defined in Section 4.2.3). Denote these messages S1 and F1, respectively. The duration of a race is defined as the difference between F1's timestamp minus S1's timestamp, that is, by how long did the first successful message in the race beat the first failed message. For each symbol-date in our dataset, we sum all race durations and report the distribution. For example, the table indicates that in the mean FTSE 100 symbol-date, the sum of the duration of all races is 0.044 seconds.

Symbol-level Version of Table 5.3

Table 5.3 reported the percentage of volume and trades in races at the date level, i.e., averaged across all symbols in the FTSE 100, FTSE 250, and full sample respectively. In this table we report the percentage of volume and trades in races at the symbol level averaged across all dates.

Table A.3: **Volume and Trades in Races**

Panel A: Percentage of volume (value-weighted) in races across symbols

Description	Mean	sd	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99
FTSE 100	23.48	4.90	13.08	17.84	20.07	23.30	26.34	30.62	33.75
FTSE 250	11.33	8.48	0.00	0.61	1.99	12.69	18.48	22.07	27.30
Full Sample	14.86	9.40	0.00	1.11	5.79	17.20	22.02	25.78	33.06

Panel B: Percentage of number of trades in races across symbols

Description	Mean	sd	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99
FTSE 100	22.19	4.56	12.54	16.69	19.58	21.79	24.78	28.44	32.09
FTSE 250	11.31	8.37	0.00	0.55	2.00	13.21	18.32	21.63	27.31
Full Sample	14.48	8.95	0.00	0.87	6.05	16.70	21.36	24.67	31.16

Notes: Please see the notes for Table 5.3 in the main text. Table 5.3 reports the distribution of percentage of volume and trades in races at the date level. This table reports the same distribution but at the symbol level.

Additional Data on Messages Per Race

Table 5.4 reports the number of messages, takes, and cancels in the T microseconds after the start of a race for values of T between 50us and 1ms. In this table we break out the take messages into two types: immediate-or-cancels (IOCs) and limit orders. In many of the sensitivity analyses discussed in Section 6 we only allow for IOC take messages to count towards the 1+ fails requirement for race detection. We also report the number of participants in races. The number of participants can be strictly lower than the number of messages if some participants in races send multiple messages.

Table A.4: **Number of IOC / Limit Takes and Number of Participants / Firms in Races**

Panel A: Number of take (IOC) messages

Description	Mean	sd	Min	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99	Max
IOC takes within 50us	1.56	0.99	0	0	1	1	1	2	3	5	14
IOC takes within 100us	1.80	1.10	0	0	1	1	2	2	3	5	15
IOC takes within 200us	2.20	1.32	0	0	1	1	2	3	4	6	17
IOC takes within 500us	2.81	1.73	0	0	1	2	2	4	5	8	29
IOC takes within 1000us	3.07	2.00	0	0	1	2	3	4	6	10	40

Panel B: Number of take (limit) messages

Description	Mean	sd	Min	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99	Max
Limit takes within 50us	0.10	0.32	0	0	0	0	0	0	0	1	5
Limit takes within 100us	0.13	0.39	0	0	0	0	0	0	1	2	6
Limit takes within 200us	0.17	0.45	0	0	0	0	0	0	1	2	7
Limit takes within 500us	0.25	0.60	0	0	0	0	0	0	1	3	11
Limit takes within 1000us	0.37	0.82	0	0	0	0	0	0	1	4	17

Panel C: Number of participants

Description	Mean	sd	Min	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99	Max
Participants within 50us	1.77	0.86	1	1	1	1	2	2	3	5	12
Participants within 100us	2.08	0.97	1	1	1	1	2	2	3	5	13
Participants within 200us	2.56	1.13	1	1	2	2	2	3	4	6	16
Participants within 500us	3.27	1.56	2	2	2	2	3	4	5	9	23
Participants within 1000us	3.64	1.94	2	2	2	2	3	4	6	11	26

Panel D: Number of firms

Description	Mean	sd	Min	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99	Max
Firms within 50us	1.55	0.69	1	1	1	1	1	2	2	4	7
Firms within 100us	1.77	0.76	1	1	1	1	2	2	3	4	8
Firms within 200us	2.12	0.82	1	1	1	2	2	3	3	4	8
Firms within 500us	2.60	1.01	1	1	2	2	2	3	4	6	10
Firms within 1000us	2.82	1.19	1	1	2	2	3	3	4	6	12

Notes: Please see the notes for Table 5.4 and the description in the text above this table.

Additional Versions of Percentage of 1st Successful and Failed Messages by Firm

Figure 5.2 reported the percentage of 1st successful and 1st failed messages in races, by firm, over all races in the FTSE 100. The following two figures report the same figure for the FTSE 250 and full sample.

Figure A.2: **Percentage of 1st Successful and Failed Messages by Firm**

Panel A: FTSE 250 Firm Dynamics

Panel B: Full Sample Firm Dynamics
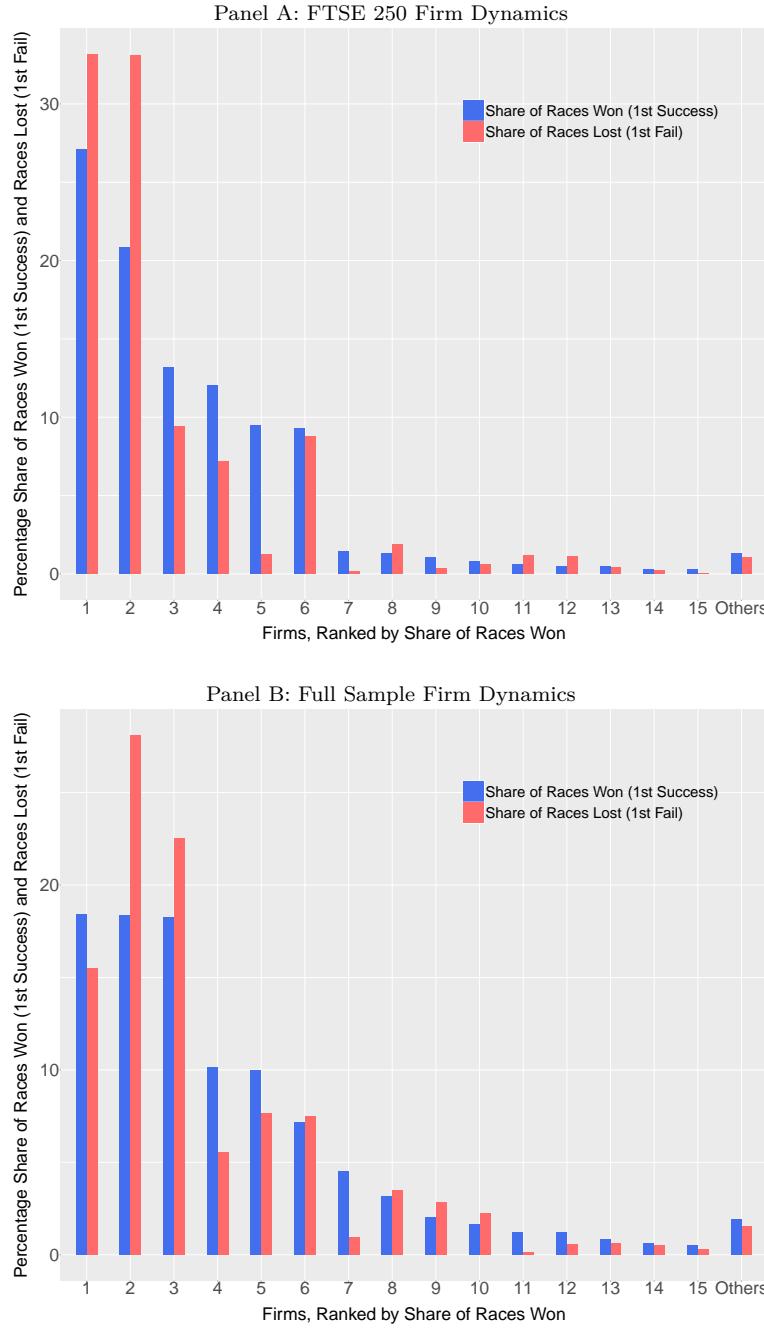
Notes: Please see the notes for Figure 5.2 and the description in the text above this figure.

Details for Expected Number of Races by Chance Analysis

In Section 5.2, "Expected Number of Races by Chance," we reported the number of times per day we would see N messages on the same side of the order book within T microseconds, by chance, if orders arrive randomly according to a Poisson process. Poisson processes are memoryless meaning that the arrival of a message at one point in time does not make it any more or less likely for other messages to arrive in the interval of time thereafter. We determined the Poisson arrival rate for each symbol-date based on the total number of potentially-race-relevant messages (i.e., marketable takes or cancels at the best bid or offer) for that symbol-date.

Table A.5: **Number of Instances Per Day With N Participants Within T Microseconds, if Messages Arrive Randomly**

				FTSE 100				
N	T	Mean	sd	Pct01	Pct25	Median	Pct75	Pct99
2	50	0.35	0.80	0.01	0.04	0.09	0.32	3.28
2	100	0.71	1.60	0.02	0.08	0.18	0.64	6.56
2	200	1.42	3.20	0.03	0.15	0.37	1.29	13.13
2	500	3.55	7.99	0.08	0.38	0.91	3.22	32.81
2	1000	7.09	15.96	0.15	0.77	1.83	6.44	65.57
3	1000	0.00	0.02	0.00	0.00	0.00	0.00	0.05
Actual Number of Races								
All Races		537.24	542.96	73	215	346	629	2,635
Races with 3+ Participants within 1ms		368.47	352.03	51	155	249	438	1,709

				FTSE 250				
N	T	Mean	sd	Pct01	Pct25	Median	Pct75	Pct99
2	50	0.00	0.01	0.00	0.00	0.00	0.00	0.04
2	100	0.01	0.02	0.00	0.00	0.00	0.01	0.09
2	200	0.02	0.04	0.00	0.00	0.01	0.02	0.17
2	500	0.04	0.10	0.00	0.00	0.02	0.04	0.43
2	1000	0.08	0.20	0.00	0.01	0.03	0.08	0.86
3	1000	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Actual Number of Races								
All Races		70.05	103.26	0	2	35	97	477
Races with 3+ Participants within 1ms		59.25	77.76	0	8	35	82	355

Notes: This table details the distribution of the expected number of races that would occur by chance in a symbol-date given a Poisson arrival process for messages. For each symbol-date the arrival rate of potentially-race-relevant messages (see text of Section 5.2 for description) is calculated and the expected number of occurrences of N such messages within T microseconds, on the same side of the order book, is computed if messages arrive at this rate via a Poisson arrival process. We also report the actual number of races, both overall and conditioning on their being at least 3+ participants within the 1 millisecond following the start of the race. The analysis is reported separately for FTSE 100 and FTSE 250.

Distribution of the Bid-Ask Spread by Symbol and Date

Table 5.9 presents a decomposition of the bid-ask spread into price impact in races, price impact not in races, loss avoidance, and the realized spread. For context on this analysis, we present here the distribution of the bid-ask spread across symbol (averaged over all dates) and dates (averaged over all symbols). Spreads are presented based on both the time-weighted displayed spread (Panel A) and the quantity-weighted traded spread (Panel B); this latter quantity-weighted spread corresponds to the term effective spread utilized in the literature and in the text of Section 5.5. For each analysis, we present results in both ticks (sub-panel A) and basis points (sub-panel B); this latter measurement corresponds to the spread decomposition reported in the text. All spreads are reported as the "half-spread", i.e., half the distance between the bid and the offer, which corresponds to the difference between the tradable or traded price and the midpoint price. The half-spread is the standard measure in the literature.

Table A.6: **Spread by Date**

Panel A: Time-Weighted Average Half-Spread

Sub-Panel A: Ticks

Description	Mean	sd	Min	Pct10	Pct25	Median	Pct75	Pct90	Max
FTSE 100	0.97	0.06	0.86	0.92	0.93	0.96	1.00	1.04	1.20
FTSE 250	3.40	0.35	2.83	2.99	3.19	3.34	3.61	3.81	4.38
Full Sample	2.70	0.26	2.29	2.39	2.53	2.63	2.86	2.98	3.45

Sub-Panel B: Basis Points

Description	Mean	sd	Min	Pct10	Pct25	Median	Pct75	Pct90	Max
FTSE 100	3.77	0.20	3.42	3.54	3.66	3.76	3.82	3.97	4.39
FTSE 250	15.76	1.48	13.11	13.97	14.81	15.62	16.66	17.67	19.62
Full Sample	12.27	1.09	10.35	10.92	11.55	12.22	12.93	13.63	15.19

Panel B: Quantity-Weighted Average Half-Spread ("Effective Spread")

Sub-Panel A: Ticks

Description	Mean	sd	Min	Pct10	Pct25	Median	Pct75	Pct90	Max
FTSE 100	0.85	0.17	0.70	0.74	0.76	0.80	0.86	1.00	1.71
FTSE 250	1.44	0.13	1.15	1.31	1.37	1.44	1.47	1.53	1.82
Full Sample	0.93	0.15	0.77	0.83	0.85	0.88	0.95	1.06	1.66

Sub-Panel B: Basis Points

Description	Mean	sd	Min	Pct10	Pct25	Median	Pct75	Pct90	Max
FTSE 100	2.65	0.29	2.28	2.45	2.52	2.59	2.72	2.80	4.28
FTSE 250	6.76	0.58	5.72	6.24	6.44	6.66	6.95	7.19	8.97
Full Sample	3.17	0.27	2.74	2.92	3.06	3.12	3.22	3.38	4.52

Notes: Please see the description in the text above this table for a description of the spread variables. This table reports distributions of the spread at the date level, averaging over symbols.

Table A.7: **Spread by Symbol**

Panel A: Time-Weighted Average Half-Spread

Sub-Panel A: Ticks

Description	Mean	sd	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99
FTSE 100	0.97	0.32	0.56	0.64	0.83	0.92	1.02	1.32	2.14
FTSE 250	3.40	3.00	0.83	1.09	1.53	2.57	3.94	6.52	16.73
Full Sample	2.70	2.76	0.58	0.85	1.01	1.79	3.25	5.67	12.86

Sub-Panel B: Basis Points

Description	Mean	sd	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99
FTSE 100	3.77	1.56	1.09	1.70	2.56	3.77	4.85	5.49	7.59
FTSE 250	15.76	13.67	3.38	6.36	7.74	11.32	17.92	29.90	59.41
Full Sample	12.27	12.76	1.21	3.09	4.95	8.10	15.04	27.07	56.01

Panel B: Quantity-Weighted Average Half-Spread ("Effective Spread")

Sub-Panel A: Ticks

Description	Mean	sd	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99
FTSE 100	0.80	0.27	0.52	0.55	0.64	0.73	0.89	1.17	1.71
FTSE 250	2.09	1.42	0.60	0.84	1.13	1.75	2.58	3.80	6.62
Full Sample	1.72	1.34	0.54	0.66	0.81	1.32	2.17	3.21	6.38

Sub-Panel B: Basis Points

Description	Mean	sd	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99
FTSE 100	3.27	1.22	1.22	1.75	2.28	3.18	4.13	4.91	5.79
FTSE 250	11.61	9.53	2.66	4.90	5.99	8.22	13.67	22.96	47.35
Full Sample	9.18	8.90	1.29	2.59	4.21	6.26	10.38	18.47	40.07

Notes: Please see the description in the text above this table for a description of the spread variables. This table reports distributions of the spread at the symbol level, averaging over dates.

Spread Decomposition with Different Time Horizons

Table 5.9 in the main text reports results of our spread decomposition (Section 5.5, Approach #1) using a 10 second mark-to-market time horizon for calculating price impact and loss avoidance. In this appendix we report the same decomposition but using 100 millisecond and 1 second time horizons instead. Notably, while the overall sample realized spread is slightly negative at 10 seconds, it is slightly positive at 100 millisecond and 1 second. This is because price impact is smaller at shorter time horizons as discussed in **?**.

Table A.8: **Spread Decomposition - 100ms**

Panel A: FTSE 100 by Symbol

Description	Mean	sd	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99
Effective spread paid (basis points)	3.27	1.22	1.22	1.75	2.28	3.18	4.13	4.91	5.79
Effective spread paid - in races (basis points)	3.18	1.22	0.99	1.70	2.21	3.17	4.05	4.89	5.98
Effective spread paid - not in races (basis points)	3.29	1.22	1.25	1.78	2.30	3.17	4.15	4.96	5.71
Price impact (basis points)	3.18	1.25	1.16	1.71	2.18	3.06	3.96	5.06	5.82
Price impact in races (basis points)	1.11	0.61	0.35	0.45	0.66	0.99	1.50	2.10	2.54
Price impact not in races (basis points)	2.07	0.70	0.81	1.22	1.52	2.09	2.47	3.02	3.60
Loss avoidance (basis points)	0.00	0.01	-0.01	0.00	0.00	0.00	0.00	0.01	0.02
Realized spread (basis points)	0.09	0.27	-0.44	-0.21	-0.02	0.05	0.18	0.37	1.07
PI in races / PI total (%)	33.26	6.28	21.27	25.97	29.36	31.77	37.35	43.12	46.06
PI in races / Effective spread (%)	32.49	7.56	18.81	23.89	28.30	30.96	36.37	43.84	49.45

Panel B: FTSE 250 by Symbol

Description	Mean	sd	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99
Effective spread paid (basis points)	8.06	3.81	2.65	4.63	5.59	7.14	9.84	13.10	19.10
Effective spread paid - in races (basis points)	6.74	3.03	2.42	4.32	4.97	6.08	7.63	9.96	15.62
Effective spread paid - not in races (basis points)	8.22	3.87	2.72	4.70	5.72	7.31	9.94	13.34	19.55
Price impact (basis points)	5.99	2.47	2.24	3.58	4.34	5.44	7.09	9.23	14.30
Price impact in races (basis points)	1.20	0.56	0.16	0.48	0.84	1.18	1.49	1.88	2.83
Price impact not in races (basis points)	4.79	2.40	1.65	2.63	3.17	4.03	5.68	7.89	12.48
Loss avoidance (basis points)	0.00	0.02	-0.05	-0.02	-0.01	0.00	0.00	0.01	0.06
Realized spread (basis points)	2.07	1.69	-0.04	0.45	1.18	1.83	2.57	3.50	6.97
PI in races / PI total (%)	21.82	9.31	2.14	7.49	15.08	23.34	28.22	32.29	39.41
PI in races / Effective spread (%)	17.14	8.59	1.30	4.59	10.97	17.30	22.54	27.63	37.15

Panel C: Full Sample by Date

Description	Mean	sd	Min	Pct10	Pct25	Median	Pct75	Pct90	Max
Effective spread paid (basis points)	3.17	0.27	2.74	2.92	3.06	3.12	3.22	3.38	4.52
Effective spread paid - in races (basis points)	2.99	0.13	2.64	2.84	2.90	2.99	3.06	3.16	3.28
Effective spread paid - not in races (basis points)	3.22	0.32	2.77	2.95	3.10	3.17	3.29	3.44	4.90
Price impact (basis points)	2.88	0.16	2.54	2.71	2.79	2.90	2.95	3.13	3.18
Price impact in races (basis points)	0.91	0.09	0.69	0.79	0.86	0.90	0.97	1.02	1.09
Price impact not in races (basis points)	1.98	0.11	1.72	1.85	1.91	1.97	2.03	2.12	2.28
Loss avoidance (basis points)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
Realized spread (basis points)	0.28	0.23	0.10	0.17	0.20	0.24	0.30	0.39	1.65
PI in races / PI total (%)	31.43	2.31	24.08	28.54	30.40	31.69	32.47	34.07	36.64
PI in races / Effective spread (%)	28.77	3.12	15.24	26.47	27.76	29.26	30.37	31.92	34.52

Notes: Please see the notes for Table 5.9 in the main text. This table is the same except that price impact and loss avoidance are calculated based on mark-to-market at 100 milliseconds instead of 10 seconds.

Table A.9: **Spread Decomposition - 1s**

Panel A: FTSE 100 by Symbol

Description	Mean	sd	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99
Effective spread paid (basis points)	3.27	1.22	1.22	1.75	2.28	3.18	4.13	4.91	5.79
Effective spread paid - in races (basis points)	3.18	1.22	0.99	1.70	2.21	3.17	4.05	4.89	5.98
Effective spread paid - not in races (basis points)	3.29	1.22	1.25	1.78	2.30	3.17	4.15	4.96	5.71
Price impact (basis points)	3.39	1.29	1.27	1.85	2.34	3.34	4.15	5.20	6.30
Price impact in races (basis points)	1.18	0.62	0.39	0.48	0.69	1.08	1.57	2.22	2.60
Price impact not in races (basis points)	2.21	0.74	0.90	1.33	1.64	2.21	2.71	3.10	3.87
Loss avoidance (basis points)	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.02
Realized spread (basis points)	-0.12	0.25	-0.58	-0.39	-0.25	-0.16	0.00	0.12	0.76
PI in races / PI total (%)	33.29	6.26	20.88	25.73	29.49	32.11	37.49	42.69	46.16
PI in races / Effective spread (%)	34.74	7.42	19.79	26.20	30.94	34.06	39.08	44.93	49.85

Panel B: FTSE 250 by Symbol

Description	Mean	sd	Pct01	Pct10	Pct25	Median	Pct75	Pct90	Pct99
Effective spread paid (basis points)	8.06	3.81	2.65	4.63	5.59	7.14	9.84	13.10	19.11
Effective spread paid - in races (basis points)	6.74	3.03	2.42	4.32	4.97	6.08	7.63	9.96	15.62
Effective spread paid - not in races (basis points)	8.22	3.87	2.72	4.70	5.72	7.31	9.94	13.34	19.55
Price impact (basis points)	6.71	2.83	2.43	4.14	4.95	5.98	7.79	10.34	17.10
Price impact in races (basis points)	1.33	0.61	0.18	0.53	0.95	1.30	1.70	2.08	3.17
Price impact not in races (basis points)	5.38	2.77	1.82	2.95	3.52	4.54	6.30	8.96	15.35
Loss avoidance (basis points)	0.00	0.01	-0.04	-0.01	0.00	0.00	0.00	0.01	0.07
Realized spread (basis points)	1.35	1.45	-0.50	0.05	0.59	1.11	1.73	2.65	5.66
PI in races / PI total (%)	21.79	9.41	2.10	6.72	15.03	23.58	28.40	32.31	39.77
PI in races / Effective spread (%)	19.03	9.41	1.61	5.19	12.08	19.61	25.39	30.01	41.32

Panel C: Full Sample by Date

Description	Mean	sd	Min	Pct10	Pct25	Median	Pct75	Pct90	Max
Effective spread paid (basis points)	3.17	0.27	2.74	2.92	3.06	3.12	3.22	3.38	4.52
Effective spread paid - in races (basis points)	2.99	0.13	2.64	2.84	2.90	2.99	3.06	3.16	3.28
Effective spread paid - not in races (basis points)	3.22	0.32	2.77	2.95	3.10	3.17	3.29	3.44	4.90
Price impact (basis points)	3.10	0.17	2.72	2.90	3.00	3.11	3.21	3.36	3.44
Price impact in races (basis points)	0.97	0.10	0.74	0.85	0.91	0.96	1.03	1.09	1.16
Price impact not in races (basis points)	2.13	0.13	1.85	1.98	2.06	2.11	2.20	2.30	2.47
Loss avoidance (basis points)	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.01	0.01
Realized spread (basis points)	0.07	0.21	-0.11	-0.06	-0.03	0.02	0.11	0.19	1.30
PI in races / PI total (%)	31.24	2.41	23.10	28.32	30.29	31.69	32.37	33.99	36.59
PI in races / Effective spread (%)	30.71	3.37	16.41	28.06	29.47	31.27	32.89	34.03	36.64

Notes: Please see the notes for Table 5.9 in the main text. This table is the same except that price impact and loss avoidance are calculated based on mark-to-market at 1 second instead of 10 seconds.

Sensitivity Analysis: 5+ Race Participants

In the text of Section 6.2 we reported a sensitivity analysis for requiring 3+ participants in a race, and discussed results from a sensitivity analysis that requires 5+ participants in a race. The following table reports the full results from the 5+ participants sensitivity analysis. It is analogous to Table 6.2 in the main text.

Table A.10: **Sensitivity Analysis: 5+ Race Participants**

Measure	Baseline	InfoHor	5+ Race Participants Within						
			50μs	100μs	200μs	500μs	1ms	2ms	3ms
Frequency and Duration of Races									
Races per day									
FTSE 100 - per symbol	537.24	37.83	5.96	13.58	35.27	121.76	202.00	268.66	297.78
FTSE 250 - per symbol	70.05	4.91	0.88	2.03	5.21	16.36	26.67	33.77	36.62
Mean race duration (microseconds)	78.65	73.23	11.14	23.94	61.66	170.05	304.84	469.80	582.24
% of races with wrong winner	4.30	5.62	14.93	9.48	4.98	2.32	1.84	1.45	1.30
% of volume in races									
FTSE 100	22.15	3.39	0.38	0.99	2.70	10.36	17.94	23.31	25.51
FTSE 250	16.90	2.23	0.33	0.78	2.12	7.65	12.81	16.43	17.87
Full Sample	21.46	3.24	0.37	0.97	2.62	10.01	17.27	22.41	24.52
Mean number of messages within 500 μs	3.46	7.04	7.37	7.37	7.06	6.23	4.79	4.11	3.90
Per-Race Profits									
Per-share profits									
ticks	0.55	1.01	1.02	0.98	0.92	0.84	0.83	0.86	0.87
GBX	0.17	0.34	0.29	0.31	0.30	0.28	0.27	0.27	0.27
basis points	1.66	3.39	3.25	3.30	3.13	2.79	2.69	2.66	2.64
Per-race profits GBP									
displayed depth	1.85	6.30	4.52	5.14	5.01	4.89	4.82	4.63	4.58
qty trade/cancel	1.76	6.29	4.28	4.91	4.96	5.06	5.03	4.84	4.80
Aggregate Profits and LA Tax									
Daily Profits									
FTSE 100 - per symbol	1,047	262	29	77	195	637	1,037	1,310	1,433
FTSE 250 - per symbol	108	21	3	7	19	63	102	129	139
Full Sample - aggregate	132,378	31,663	3,699	9,609	24,265	79,717	129,773	163,927	178,855
Latency Arbitrage Tax, All Volume (bps)									
FTSE 100	0.38	0.10	0.01	0.03	0.07	0.24	0.38	0.48	0.53
FTSE 250	0.66	0.13	0.02	0.04	0.12	0.38	0.62	0.79	0.85
Full Sample	0.42	0.10	0.01	0.03	0.08	0.25	0.41	0.52	0.57
Latency Arbitrage Tax, Non-Race Volume (bps)									
FTSE 100	0.49	0.12	0.01	0.03	0.09	0.38	0.68	0.91	1.03
FTSE 250	0.80	0.16	0.02	0.05	0.14	0.58	1.02	1.36	1.50
Full Sample	0.53	0.13	0.01	0.04	0.10	0.40	0.73	0.97	1.09
Spread Decomposition									
Price impact in races / All price impact %	30.58	6.01	0.63	1.72	4.61	16.26	27.53	35.79	39.63
Price impact in races / Effective spread %	32.82	6.46	0.68	1.85	4.96	17.46	29.56	38.43	42.55
Loss avoidance / Effective spread %	0.19	0.06	0.01	0.03	0.10	0.42	0.92	1.38	1.61
Implied Reduction in Cost of Liquidity									
% Reduction in liquidity cost									
FTSE 100 - by symbol	19.95	3.88	0.50	1.15	2.91	10.46	18.79	25.71	29.26
FTSE 250 - by symbol	11.93	2.76	0.77	1.18	2.40	7.29	12.09	15.75	17.14
Full Sample - by date	16.73	3.31	0.38	0.99	2.55	8.94	15.82	21.39	24.19

Notes: Please see the notes and surrounding text for Table 6.2. This table is identical except it conditions on 5+ participants in a race instead of 3+ participants.

Sensitivity Analysis: 2+ Unique Firms

Our baseline method requires that a race contains at least 2 unique participants as determined by their UserID in our data. As discussed in the text, some firms use different UserIDs for different trading desks. Typically, this will be the case when the trading desks are operated sufficiently separately that if they happen to trade with each other the firm would not be in violation of wash-trade requirements. This economic separation is the reason why our baseline uses UserIDs as the measurement of the number of participants. The following table provides results if the requirement is changed from 2+ unique participants to 2+ unique firms. The format is analogous to Table 6.2 in the main text, and the results can also be compared to the first 8 columns of Table 6.1 in the main text.

Table A.11: **Sensitivity Analysis: 2+ Participating Firms**

Measure	Baseline	InfoHor	2+ Participating Firms Within						
			50μs	100μs	200μs	500μs	1ms	2ms	3ms
Frequency and Duration of Races									
Races per day									
FTSE 100 - per symbol	537.24	479.32	247.25	332.99	462.39	736.14	818.92	871.31	891.49
FTSE 250 - per symbol	70.05	60.44	32.74	43.39	59.97	102.41	116.32	122.08	124.89
Mean race duration (microseconds)	78.65	81.59	16.08	31.24	74.03	196.91	306.40	447.04	552.85
% of races with wrong winner	4.30	4.67	9.46	7.22	4.57	2.05	1.73	1.48	1.38
% of volume in races									
FTSE 100	22.15	20.08	8.15	11.51	17.62	35.79	42.20	45.86	47.26
FTSE 250	16.90	15.19	6.89	9.54	14.37	31.57	36.83	39.52	40.63
Full Sample	21.46	19.44	7.98	11.25	17.19	35.23	41.49	45.02	46.39
Mean number of messages within 500 μs	3.46	3.52	3.52	3.54	3.58	3.47	3.08	2.90	2.83
Per-Race Profits									
Per-share profits									
ticks	0.55	0.54	0.51	0.50	0.50	0.53	0.56	0.57	0.58
GBX	0.17	0.17	0.16	0.16	0.16	0.16	0.17	0.17	0.17
basis points	1.66	1.65	1.62	1.58	1.56	1.62	1.67	1.69	1.70
Per-race profits GBP									
displayed depth	1.85	1.93	1.58	1.60	1.65	1.94	2.04	2.08	2.09
qty trade/cancel	1.76	1.83	1.40	1.45	1.56	1.94	2.05	2.10	2.12
Aggregate Profits and LA Tax									
Daily Profits									
FTSE 100 - per symbol	1,047	971	404	553	793	1,482	1,744	1,889	1,945
FTSE 250 - per symbol	108	98	46	62	87	176	205	221	226
Full Sample - aggregate	132,378	122,218	52,221	70,992	101,416	192,912	226,603	245,049	252,001
Latency Arbitrage Tax, All Volume (bps)									
FTSE 100	0.38	0.36	0.15	0.20	0.29	0.54	0.64	0.69	0.71
FTSE 250	0.66	0.60	0.29	0.38	0.53	1.08	1.26	1.35	1.39
Full Sample	0.42	0.39	0.17	0.23	0.32	0.61	0.72	0.77	0.80
Latency Arbitrage Tax, Non-Race Volume (bps)									
FTSE 100	0.49	0.46	0.17	0.24	0.36	0.87	1.13	1.31	1.39
FTSE 250	0.80	0.72	0.31	0.43	0.64	1.63	2.06	2.32	2.43
Full Sample	0.53	0.49	0.18	0.26	0.40	0.97	1.26	1.45	1.53
Spread Decomposition									
Price impact in races / All price impact %	30.58	28.12	10.66	15.27	23.26	48.00	57.26	62.91	65.21
Price impact in races / Effective spread %	32.82	30.18	11.43	16.38	24.97	51.50	61.44	67.50	69.97
Loss avoidance / Effective spread %	0.19	0.19	0.07	0.13	0.26	0.53	0.94	1.32	1.48
Implied Reduction in Cost of Liquidity									
% Reduction in liquidity cost									
FTSE 100 - by symbol	19.95	18.15	6.55	9.30	14.28	34.09	45.35	53.50	57.24
FTSE 250 - by symbol	11.93	10.49	4.97	6.68	9.52	22.75	28.20	30.26	31.44
Full Sample - by date	16.73	15.15	5.66	7.99	12.18	28.90	37.77	44.04	46.86

Notes: Please see the description in the text above this table for a description. The table is identical to Table 6.1 in the main text except it conditions on 2+ unique firms in a race whereas the baseline conditions on 2+ unique participants.

Sensitivity Analysis: 1+ Cancels and 2+ Takes

Table 6.3 in Section 6.3 presents sensitivity analysis for requiring 1+ cancel in a race and, separately, for requiring 2+ takes in a race. The former rules out races with 0 cancels (and hence 2+ takes, at least one of which succeeds and one of which fails); the latter rules out races with 1+ cancels and exactly 1 take. The following table presents sensitivity analysis for requiring both criteria simultaneously. This rules out races with either 0 cancels, or with 1+ cancels and exactly 1 take.

Table A.12: **Sensitivity Analysis: 1+ Cancels and 2+ Takes**

Measure	Baseline	1+ Cancel and 2+ Takes Within				
		InfoHor	50μs	500μs	1ms	3ms
Frequency and Duration of Races						
Races per day						
FTSE 100 - per symbol	537.24	59.32	16.56	145.02	208.36	285.82
FTSE 250 - per symbol	70.05	5.27	1.69	14.29	21.39	29.83
Mean race duration (microseconds)	78.65	86.51	17.41	218.12	384.76	754.95
% of races with wrong winner	4.30	3.60	7.57	1.59	1.22	0.79
% of volume in races						
FTSE 100	22.15	3.74	0.72	9.18	14.32	20.28
FTSE 250	16.90	1.60	0.38	4.85	7.88	11.57
Full Sample	21.46	3.46	0.67	8.62	13.49	19.16
Mean number of messages within 500 μs	3.46	4.72	4.68	4.42	3.53	3.00
Per-Race Profits						
Per-share profits						
ticks	0.55	0.59	0.44	0.47	0.50	0.52
GBX	0.17	0.20	0.15	0.15	0.15	0.16
basis points	1.66	1.80	1.44	1.42	1.46	1.51
Per-race profits GBP						
displayed depth	1.85	3.42	2.36	2.59	2.80	3.02
qty trade/cancel	1.76	3.23	2.01	2.54	2.77	3.02
Aggregate Profits and LA Tax						
Daily Profits						
FTSE 100 - per symbol	1,047	220	43	402	626	920
FTSE 250 - per symbol	108	11	3	26	42	65
Full Sample - aggregate	132,378	24,881	4,925	46,952	73,558	109,059
Latency Arbitrage Tax, All Volume (bps)						
FTSE 100	0.38	0.08	0.02	0.15	0.23	0.34
FTSE 250	0.66	0.07	0.02	0.16	0.26	0.41
Full Sample	0.42	0.08	0.02	0.15	0.23	0.35
Latency Arbitrage Tax, Non-Race Volume (bps)						
FTSE 100	0.49	0.10	0.02	0.24	0.41	0.66
FTSE 250	0.80	0.08	0.02	0.24	0.43	0.71
Full Sample	0.53	0.10	0.02	0.24	0.41	0.67
Spread Decomposition						
Price impact in races / All price impact %	30.58	5.42	0.98	11.82	18.82	28.48
Price impact in races / Effective spread %	32.82	5.82	1.05	12.70	20.21	30.58
Loss avoidance / Effective spread %	0.19	0.16	0.06	0.59	1.09	1.76
Implied Reduction in Cost of Liquidity						
% Reduction in liquidity cost						
FTSE 100 - by symbol	19.95	3.10	0.59	6.18	10.35	16.88
FTSE 250 - by symbol	11.93	1.28	0.41	2.86	4.49	7.02
Full Sample - by date	16.73	2.61	0.50	5.22	8.68	14.05

Notes: Please see the description in the text above this table and in Section 6.3 for a description. The table is similar to Table 6.3 in the main text except that it conditions on both 1+ cancels and 2+ takes.

15

Additional Extrapolation Models

Table 7.1 presented regressions of daily latency arbitrage profits on volume and 1-minute realized volatility. These regressions were used for the purpose of out-of-sample extrapolation in Section 7. The following table presents analogous regressions using additional volatility variables, as discussed in the text. Columns (1)-(4) are analogous to Columns (3)-(6) in Table 7.1, but using 5-minute realized volatility instead of 1-minute realized volatility. Columns (5)-(8) are analogous to the same columns in Table 7.1, but using midpoint distance traveled (**?**) as the volatility measure. As discussed in the main text, the fit is worse with 5-minute realized volatility than with 1-minute realized volatility, and is slightly better with midpoint distance traveled. We nevertheless utilize 1-minute realized volatility in the main text since it is more easily interpreted, and its measurement does not depend on the number of significant digits of the trading index (or the tick size if using a futures contract price for the index) in the way that distance traveled does.

Table A.13: **Extrapolation Models (Appendix)**

			Dependent variable:					
				Latency Arbitrage Profits (GBP)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Volume (10,000 GBP)			0.4237*** (0.0583)	0.4123*** (0.0320)			0.2561*** (0.0790)	0.2833*** (0.0578)
Volatility (5 min) * Average Volume	0.0147*** (0.0020)	0.0276*** (0.0013)	0.0004 (0.0024)	0.0006 (0.0022)				
Volatility (Midpoint Distance Travelled) * Average Volume					0.0072*** (0.0006)	0.0090*** (0.0002)	0.0032** (0.0013)	0.0030** (0.0012)
Constant	68,085*** (9,796)		−2,768 (11,717)		28,891*** (8,771)		5,464 (10,709)	
Observations	43	43	43	43	43	43	43	43
R²	0.561	0.134	0.811	0.811	0.791	0.742	0.835	0.834

*p<0.1; **p<0.05; ***p<0.01

Notes: Please see the description in the text above this table and the notes for Table 7.1 in the main text. 5-minutes volatility is computed as realized 5-minute volatility for the FTSE 350 index in percentage points, using TRTH data. The distance travelled for each day is calculated as the sum of the absolute value of the change in midpoint on each update of the FTSE350. The FTSE350 is disseminated 4 times a second, or every 250 milliseconds.