Internet Appendix: High Frequency Trading and Extreme Price Movements

This appendix includes two parts. First, it reports the results from the sample of EPMs defined as the 99.9th percentile of raw returns. Second, it details the estimation of the Lee and Mykland (2012) jump identification algorithm and reports the results for EPMs defined using this algorithm.

1. EPM sample based on the 99.9th percentile of raw returns

As an alternative to the procedure that defines EPMs as the 99.9th percentile of residuals from Eq. (1), we define EPMs as the 99.9th percentile of raw returns. The results from this sample are reported in Tables A1-A7. The results obtained from both samples are qualitatively similar to those discussed in the main manuscript.

2. EPM sample based on the Lee-Mykland jump identification algorithm

The Lee and Mykland (2012) algorithm (LM) identifies intervals with discrete price changes using the non-parametric approach based on realized volatility. First, the optimal sampling frequency is identified as k for the $(k-1)$-dependent noise. In our sample, the second-by-second midpoint return observations tend to follow an MA process with many statistically significant dependent AR lags. Nonetheless, the magnitude of the lag coefficients tends to drop sharply after 6–8 lags. To make the estimated results comparable across stocks, we select $k = 10$. The algorithm suggests that a researcher should pre-average over M k-sampled observations before computing the jump statistics as follows: $M = C\sqrt{(n/k)}$, where C is a parameter defined by the noise variance, and n is the number of observations in the jump

estimation period. Recognizing that in a high-frequency sample the realized variation of log-prices $P(t)$ converges to noise variance, we estimate the volatility of noise as:

$$\hat{q} = \sqrt{\left(\sum_{m=1}^{n'} (P(t_m) - P(t_{m+k}))^2\right)/2n'} \tag{A1}$$

Following LM, we split each trading day into seven intervals: 9:30–10:00, 10:00–11:00, 11:00–12:00, 12:00–13:00, 13:00–14:00, 14:00–15:00, and 15:00–16:00. The estimated noise variance for these intervals is reported in Table A8. Based on the estimates, we choose $C = 1/19$, and therefore the estimated M is close to one for all estimation periods.[1]

Next, the standardized statistic for jump detection is defined as follows:

$$\mathcal{X}(t_j) = \frac{\sqrt{M}}{\sqrt{V_n}} \mathcal{L}(t_j), \tag{A2}$$

where $\mathcal{L}(t_j) = P(t_{j+k}) - P(t_j)$, and V_n is the estimate of the total variance. The total variance is a sum of the estimated noise variance $\hat{q}$ and price volatility. We use the noise- and outlier-robust bipower variation of Christensen, Oomen and Podolskij (2014) as the measure of price volatility.

From LM's Theorem 1, the null hypothesis of no jump in a given interval is rejected if:

$$\frac{|\mathcal{X}(t_j)| - A_n}{B_n} > \beta^*, \tag{A3}$$

where

$$A_n = \sqrt{2log\frac{n}{kM}} - \frac{log\,\pi + log\left(log\frac{n}{kM}\right)}{2\sqrt{2log\frac{n}{kM}}} \tag{A4}$$

[1] Since we use midpoint prices, much of the noise coming from the bid-ask bounce is mitigated, and the remaining noise does not produce the estimate high enough to make pre-sampling useful.

and

$$B_n = \frac{1}{\sqrt{2log\frac{n}{kM}}}. \quad \text{(A5)}$$

The rejection threshold is selected from the standard Gumbel distribution, which implies that $\beta^* = -\log(-\log(1-\alpha))$. We use the significance level of $\alpha = 5\%$, leading to $\beta^* = 2.97$. The above procedure results in rejecting the no-jump hypothesis for 0.54% of sample observations.

We restrict the LM sample to include the same number of EPMs as the 99.9 sample based on raw returns, or 45,406 instances, using the highest-magnitude LM jumps for each stock. In Tables A9 through A15, we use the LM sample to replicate the results reported in the main manuscript. Overall, the results in the LM sample are similar to those discussed in the main manuscript.

References:

Christensen, K., Oomen, R. C. A., Podolskij, M., 2014. Fact or friction: Jumps at ultra high-frequency. Journal of Financial Economics 114, 576–599.

Lee, S. S., Mykland, P. A., 2012. Jumps in equilibrium prices and market microstructure noise. Journal of Econometrics 168, 396–406.

Table A1
Summary statistics

The table reports summary statistics for the sample of extreme price movements (EPMs). $Absolute\ return$ is the absolute value of the 10-second midpoint return. $Total\ (HFT)\ trades$ is the number of (HFT) trades during the interval. $Dollar\ volume$ and $Share\ volume$ are the total dollar and share volume traded during the interval. $Quoted\ spread$ and $Relative\ spread$ are quoted and relative quoted NBBO spreads, respectively, in dollars and percentage points. All statistics are averaged over the 10-second sampling intervals.

	Mean	Median	Std. dev.
Absolute return, %	0.484	0.441	0.193
Total trades	73.0	43.0	88.7
Total HFT trades	57.6	33.0	73.2
Dollar volume	473,232	171,158	1,024,504
Share volume	15,595	5,431	31,734
Quoted spread, $	0.046	0.016	0.147
Relative spread, %	0.080	0.065	0.148
N	45,406		

Table A2
Liquidity supply and demand around EPMs

The table reports directional trading volume around extreme price movements. Time interval *t* is the 10-second EPM interval. In addition, we report the results for the two time intervals preceding the EPM and two subsequent time intervals. HFT^{D} ($nHFT^{D}$) is the difference in liquidity-demanding HFT (nHFT) volume in the direction of the EPM and liquidity-demanding volume against the direction of the EPM. HFT^{S} ($nHFT^{S}$) is the difference in liquidity-providing volume against the direction of the EPM and liquidity-providing volume in the direction of the EPM. HFT^{NET} ($nHFT^{NET}$) is the difference between HFT^{D} and HFT^{S} ($nHFT^{D}$ and $nHFT^{S}$). p-Values are in parentheses. *** and ** indicate statistical significance at the 1% and 5% levels.

	t-20	*t*-10	*t*	*t*+10	*t*+20
HFT^{NET}	1.5	45.7**	-299.3***	-122.5***	-42.7**
	(0.94)	(0.04)	(0.00)	(0.00)	(0.04)
HFT^{D}	30.6	163.4***	2215.2***	-279.0***	-99.1***
	(0.13)	(0.00)	(0.00)	(0.00)	(0.00)
HFT^{S}	-29.1	-117.6***	-2514.6***	156.5***	56.4***
	(0.14)	(0.00)	(0.00)	(0.00)	(0.00)
$nHFT^{NET}$	-1.5	-45.7**	299.3***	122.5***	42.7**
	(0.94)	(0.04)	(0.00)	(0.00)	(0.04)
$nHFT^{D}$	75.3**	326.7***	5576.3***	672.4***	317.0***
	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)
$nHFT^{S}$	-76.8**	-372.5***	-5277.0***	-549.9***	-274.3***
	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)

Table A3
Transitory and permanent EPMs

The table reports summary statistics for transitory and permanent EPMs. Transitory EPMs revert by more than 2/3 of the EPM return in the following 30 minutes. Permanent EPMs do not revert by more than 1/3 in the same interval. Because we exclude EPMs that revert by the amount between 1/3 and 2/3, the total number of EPMs in this table is 87.60% of that reported in Panel A of Table A2. Panel B reports HFT^{NET} around the two EPM types. Asterisks *** indicate statistical significance at the 1% level.

Panel A: Summary statistics

	Transitory		Permanent	
	Mean	Std. dev.	Mean	Std. dev.
Absolute return, %	0.487	0.194	0.487	0.192
Total trades	71.39	88.17	70.43	85.85
Total HFT trades	56.02	71.73	55.65	71.86
Dollar volume	465,100	1,058,443	445,067	967,642
Share volume	14,716	29,412	14,652	29,098
Quoted spread, $	0.049	0.149	0.048	0.154
Relative spr., %	0.084	0.146	0.084	0.158
N	18,185		21,116	

Panel B: HFT^{NET}

	t-20	*t*-10	*t*	*t*+10	*t*+20
Transitory	-37.5	-24.5	-457.6***	-120.3***	-101.1***
Permanent	44.4	91.2***	-323.2***	-136.5***	2.0

Table A4
EPM magnitude quartiles

Panel A divides EPMs into quartiles by return magnitude, from smallest to largest. Panel B contains HFT^{NET} statistics. Asterisks ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A: Summary statistics

	Q1 (small)		Q2	
	Mean	Std. dev.	Mean	Std. dev.
Absolute return, %	0.387	0.094	0.419	0.103
Total trades	61.31	68.58	64.15	71.25
Total HFT trades	48.96	57.83	50.96	59.05
Dollar volume	378,141	798,985	408,766	897,376
Share volume	12,487	24,759	12,973	24,216
Quoted spread, $	0.042	0.134	0.043	0.111
Relative spr., %	0.074	0.086	0.075	0.083
N	11,358		11,327	
	Q3		Q4 (large)	
Absolute return, %	0.471	0.118	0.659	0.268
Total trades	70.81	80.88	95.75	120.11
Total HFT trades	55.90	66.96	74.48	98.58
Dollar volume	452,857	932,231	652,912	1,356,125
Share volume	15,070	30,330	21,842	43,031
Quoted spread, $	0.046	0.136	0.055	0.190
Relative spr., %	0.080	0.131	0.090	0.238
N	11,358		11,363	

Panel B: HFT^{NET}

	t-20	*t*-10	*t*	*t*+10	*t*+20
Q1	-29.8	-66.5	-110.8*	-125.0***	4.9
Q2	16.4	99.6***	-145.5***	-82.2**	-61.8*
Q3	24.9	66.2	-293.7***	-82.8*	-56.6
Q4	-11.1	82.5	-655.5***	-203.6***	-60.8

Table A5
Standalone and co-EPMs

Panel A divides EPMs into standalone and co-EPMs, with the latter group capturing EPMs that occur simultaneously in several stocks. Panel B contains HFT^{NET} statistics. Asterisks *** and ** indicate statistical significance at the 1% and 5% levels.

Panel A: Summary statistics

	Standalone		Co-EPMs	
	Mean	Std. dev.	Mean	Std. dev.
Absolute return, %	0.491	0.198	0.479	0.190
Total trades	89.30	107.05	60.83	69.54
Total HFT trades	68.60	87.76	49.34	58.72
Dollar volume	625,553	1,272,083	359,359	770,887
Share volume	21,368	40,535	11,280	22,092
Quoted spread, $	0.049	0.125	0.044	0.160
Relative spr., %	0.085	0.118	0.076	0.168
# Stocks			3.5	2.66
N	19,424		25,982	

Panel B: HFT^{NET}

	t-20	*t*-10	*t*	*t*+10	*t*+20
Standalone	-2.2	-32.1	-1296.9***	-128.4***	-40.9
Co-EPMs	4.4	103.9***	446.4***	-118.1***	-44.1**

Table A6
Net HFT activity and EPMs

The table reports estimated coefficients from the following regression:

$$HFT^{NET}{}_{it} = \alpha_i + \beta_1 1_{EMP\,it} + \beta_2 Ret_{it} + \beta_3 Vol_{it} + \beta_4 Spr + \boldsymbol{Lags_{kit-\sigma}\gamma_{k\sigma}} + \varepsilon_{it},$$

where HFT^{NET} is the difference between HFT^{D} and HFT^{S}; the dummy 1_{EPM} is equal to one if a 10-second interval t is identified to contain an EPM and is equal to zero otherwise; $1_{EPM\text{-}TRANSITORY}$ and $1_{EPM\text{-}PERMANENT}$ are dummies that capture the two EPM types; $1_{EPM\text{-}STANDALONE}$ captures the standalone EPMs; $1_{CO\text{-}EPM}$ captures EPMs that occur simultaneously in two or more sample stocks; $1_{EPM\text{-}Q1}$ through $1_{EPM\text{-}Q4}$ identify four EPM quartiles by magnitude, from the smallest to the largest; Ret is the absolute return; Vol is the total trading volume; Spr is the percentage quoted spread; and $\boldsymbol{Lags_{kit-\sigma}}$ is a vector of σ lags of the dependent variable and each of the independent variables, with $\sigma \in \{1, 2, \ldots, 10\}$ and the variables indexed with a subscript k. All non-dummy variables are standardized on the stock level. Regressions are estimated with stock fixed effects. p-Values associated with the double-clustered standard errors are in parentheses. *** denote statistical significance at the 1% level.

	(1)	(2)	(3)	(4)
1_{EPM}	-0.818***			
	(0.00)			
$1_{EPM\text{-}TRANSITORY}$		-0.818***		
		(0.00)		
$1_{EPM\text{-}PERMANENT}$		-0.819***		
		(0.00)		
$1_{EPM\text{-}STANDALONE}$			-1.441***	
			(0.00)	
$1_{CO\text{-}EPM}$			-0.328***	
			(0.00)	
$1_{EPM\text{-}Q1}$				-0.490***
				(0.00)
$1_{EPM\text{-}Q2}$				-0.631***
				(0.00)
$1_{EPM\text{-}Q3}$				-0.807***
				(0.00)
$1_{EPM\text{-}Q4}$				-1.406***
				(0.00)
Ret	0.072***	0.072***	0.072***	0.073***
	(0.00)	(0.00)	(0.00)	(0.00)
Vol	0.081***	0.081***	0.081***	0.081***
	(0.00)	(0.00)	(0.00)	(0.00)
Spr	-0.010***	-0.010***	-0.010***	-0.010***
	(0.00)	(0.00)	(0.00)	(0.00)
Adj. R^2	0.02	0.02	0.02	0.02

Table A7
EPM determinants

The table reports the coefficients and the marginal effects from a probit model of EPM occurrence:

$$Prob\ (EPM = 1)_{it} = \alpha + \beta_1 HFT^{NET}_{it-1} + \beta_2 Ret_{it-1} + \beta_3 Vol_{it-1} + \beta_4 Spr_{it-1} + \varepsilon_{it},$$

where the dependent variable is equal to one if an interval t contains an extreme price movement and zero otherwise. All independent variables are lagged by one interval. HFT^{NET} is the share volume traded in the direction of the price movement minus the share volume traded against the direction of the price movement for all HFT trades, Ret is the absolute return, Vol is total traded volume, Spr is the percentage quoted spread. All variables are standardized on the stock level. The marginal effects are scaled by a factor of 1,000. p-Values are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

	All	Standalone	Co-EPMs	Permanent	Transitory
	(1)	(2)	(3)	(4)	(5)
Intercept	-3.232***	-3.438***	-3.380***	-3.426***	-3.464***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
HFT^{NET}_{t-1}	-0.003***	-0.006***	0.001	-0.002*	-0.005***
Marginal Effect	-0.008	-0.009	0.001	0.003	-0.006
	(0.00)	(0.00)	(0.42)	(0.09)	(0.00)
Controls	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	0.14	0.11	0.13	0.13	0.12

Table A8
Noise volatility

The table reports estimated noise volatility $\hat{q}$ as a percentage of price used for computation of the LM statistic. The noise variance is estimated by intraday period, day and stock.

Period	$\hat{q}$	Max	Std. Dev.
9:30–10:00	0.057%	0.095%	0.014%
10:00–11:00	0.040%	0.068%	0.010%
11:00–12:00	0.030%	0.051%	0.008%
12:00–13:00	0.025%	0.043%	0.006%
13:00–14:00	0.026%	0.042%	0.006%
14:00–15:00	0.030%	0.047%	0.007%
15:00–16:00	0.034%	0.054%	0.008%

Table A9
Summary statistics

The table reports summary statistics for the sample of extreme price movements (EPMs). $Absolute\ return$ is the absolute value of the 10-second midpoint return. $Total\ (HFT)\ trades$ is the number of (HFT) trades during the interval. $Dollar\ volume$ and $Share\ volume$ are the total dollar and share volume traded during the interval. $Quoted\ spread$ and $Relative\ spread$ are quoted and relative quoted NBBO spreads, respectively, in dollars and percentage points. All statistics are averaged over the 10-second sampling intervals.

	Mean	Median	Std. dev.
Absolute return, %	0.363	0.317	0.193
Total trades	72.5	47.0	80.8
Total HFT trades	55.0	34.0	64.3
Dollar volume	531,054	216,249	1,007,435
Share volume	16,688	6,189	31,989
Quoted spread, $	0.039	0.013	0.109
Relative spread, %	0.062	0.053	0.100
N	45,400		

Table A10

Liquidity supply and demand around EPMs

The table reports directional trading volume around extreme price movements. Time interval t is the 10-second EPM interval. In addition, we report the results for the two time intervals preceding the EPM and two subsequent time intervals. HFT^D ($nHFT^D$) is the difference in liquidity-demanding HFT (nHFT) volume in the direction of the EPM and liquidity-demanding volume against the direction of the EPM. HFT^S ($nHFT^S$) is the difference in liquidity-providing volume against the direction of the EPM and liquidity-providing volume in the direction of the EPM. HFT^{NET} ($nHFT^{NET}$) is the difference between HFT^D and HFT^S ($nHFT^D$ and $nHFT^S$). p-Values are in parentheses. *** and ** indicate statistical significance at the 1% and 5% levels.

	t-20	t-10	t	t+10	t+20
HFT^{NET}	30.5**	40.6**	-892.6***	-65.1***	7.1
	(0.04)	(0.02)	(0.00)	(0.00)	(0.66)
HFT^D	8.8	175.6***	2641.6***	-363.0***	-146.1***
	(0.53)	(0.00)	(0.00)	(0.00)	(0.00)
HFT^S	21.7	-135.1***	-3534.2***	297.9***	153.2***
	(0.14)	(0.00)	(0.00)	(0.00)	(0.00)
$nHFT^{NET}$	-30.5**	-40.6**	892.6***	65.1***	-7.1
	(0.04)	(0.02)	(0.00)	(0.00)	(0.66)
$nHFT^D$	5.2	282.5***	7805.2***	478.9***	145.8***
	(0.84)	(0.00)	(0.00)	(0.00)	(0.00)
$nHFT^S$	-35.7	-323.0***	-6912.6***	-413.8***	-152.9***
	(0.14)	(0.00)	(0.00)	(0.00)	(0.00)

Table A11
Transitory and permanent EPMs

The table reports summary statistics for transitory and permanent EPMs. Transitory EPMs revert by more than 2/3 of the EPM return in the following 30 minutes. Permanent EPMs do not revert by more than 1/3 in the same interval. Because we exclude EPMs that revert by the amount between 1/3 and 2/3, the total number of EPMs in this table is 87.60% of that reported in Panel A of Table A2. Panel B reports HFT^{NET} around the two EPM types. Asterisks *** and ** indicate statistical significance at the 1% and 5% levels.

Panel A: Summary statistics

	Transitory		Termanent	
	Mean	Std. dev.	Mean	Std. dev.
Absolute return, %	0.362	0.191	0.361	0.189
Total trades	69.96	78.69	68.32	75.25
Total HFT trades	52.74	61.26	51.95	60.71
Dollar volume	513,793	1,014,166	481,990	914,796
Share volume	15,400	28,963	14,929	27,001
Quoted spread, $	0.040	0.110	0.039	0.112
Relative spr., %	0.064	0.105	0.064	0.095
N	18,249		21,523	

Panel B: HFT^{NET}

	t-20	*t*-10	*t*	*t*+10	*t*+20
Transitory	27.7	5.7	-973.2***	-75.1***	-3.4
Permanent	18.7	-4.6	-1063.9***	-61.6**	1.8

Table A12
EPM magnitude quartiles

Panel A divides EPMs into quartiles by return magnitude, from smallest to largest. Panel B contains HFT^{NET} statistics. Asterisks ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels.

Panel A: Summary statistics

	Q1 (small)		Q2	
	Mean	Std. dev.	Mean	Std. Dev.
Absolute return, %	0.250	0.059	0.287	0.069
Total trades	58.83	56.15	63.78	63.44
Total HFT trades	44.90	45.22	48.73	50.48
Dollar volume	429,399	763,089	458,932	778,772
Share volume	12,861	22,658	14,437	27,024
Quoted spread, $	0.033	0.082	0.035	0.085
Relative spr., %	0.054	0.039	0.058	0.041
N	11,313		11,366	
	Q3		Q4 (large)	
Absolute return, %	0.348	0.087	0.566	0.269
Total trades	70.83	71.83	96.66	113.67
Total HFT trades	53.74	56.34	72.77	91.05
Dollar volume	512,954	919,937	722,441	1,403,664
Share volume	16,194	28,929	23,240	44,130
Quoted spread, $	0.039	0.104	0.048	0.151
Relative spr., %	0.063	0.057	0.074	0.182
N	11,355		11,366	

Panel B: HFT^{NET}

	t-20	*t*-10	*t*	*t*+10	*t*+20
Q1	13.4	35.1	-648.8***	-41.5	-7.3
Q2	81.7***	43.7	-748.6***	-1.5	-30.4
Q3	35.0	62.0*	-859.2***	-77.3**	25.7
Q4	-8.5	21.3	-1312.5***	-140.1***	40.5

Table A13
Standalone and co-EPMs

Panel A divides EPMs into standalone and co-EPMs, with the latter group capturing EPMs that occur simultaneously in several stocks. Panel B contains HFT^{NET} statistics. Asterisks *** and ** indicate statistical significance at the 1% and 5% levels.

Panel A: Summary statistics

	Standalone		Co-EPMs	
	Mean	Std. dev.	Mean	Std. dev.
Absolute return, %	0.366	0.194	0.358	0.190
Total trades	76.64	84.24	64.77	73.24
Total HFT trades	56.96	66.13	51.41	60.44
Dollar volume	590,320	1,091,870	418,678	812,154
Share volume	18,342	33,310	13,552	29,064
Quoted spread, $	0.043	0.111	0.031	0.106
Relative spr., %	0.066	0.088	0.054	0.118
# Stocks			3.0	2.18
N	29,724		15,676	

Panel B: HFT^{NET}

	t-20	*t*-10	*t*	*t*+10	*t*+20
Standalone	30.2	6.3	-1811.9***	-62.2**	27.9
Co-EPMs	31.0	105.4***	850.7***	-70.6***	-32.4

Table A14
Net HFT activity and EPMs

The table reports estimated coefficients from the following regression:

$$HFT^{NET}{}_{it} = \alpha_i + \beta_1 1_{EMP\,it} + \beta_2 Ret_{it} + \beta_3 Vol_{it} + \beta_4 Spr + \boldsymbol{Lags_{kit-\sigma}\gamma_{k\sigma}} + \varepsilon_{it},$$

where HFT^{NET} is the difference between HFT^{D} and HFT^{S}; the dummy 1_{EPM} is equal to one if a 10-second interval t is identified to contain an EPM and is equal to zero otherwise; $1_{EPM\text{-}TRANSITORY}$ and $1_{EPM\text{-}PERMANENT}$ are dummies that capture the two EPM types; $1_{EPM\text{-}STANDALONE}$ captures the standalone EPMs; $1_{CO\text{-}EPM}$ captures EPMs that occur simultaneously in two or more sample stocks; $1_{EPM\text{-}Q1}$ through $1_{EPM\text{-}Q4}$ identify four EPM quartiles by magnitude, from the smallest to the largest; Ret is the absolute return; Vol is the total trading volume; Spr is the percentage quoted spread; and $\boldsymbol{Lags_{kit-\sigma}}$ is a vector of σ lags of the dependent variable and each of the independent variables, with $\sigma \in \{1, 2, \ldots, 10\}$ and the variables indexed with a subscript k. All non-dummy variables are standardized on the stock level. Regressions are estimated with stock fixed effects. p-Values associated with the double-clustered standard errors are in parentheses. *** denote statistical significance at the 1% level.

	(1)	(2)	(3)	(4)
1_{EPM}	-1.014***			
	(0.00)			
$1_{EPM\text{-}TRANSITORY}$		-1.036***		
		(0.00)		
$1_{EPM\text{-}PERMANENT}$		-0.989***		
		(0.00)		
$1_{EPM\text{-}STANDALONE}$			-1.595***	
			(0.00)	
$1_{CO\text{-}EPM}$			0.174	
			(0.07)	
$1_{EPM\text{-}Q1}$				-0.582***
				(0.00)
$1_{EPM\text{-}Q2}$				-0.798***
				(0.00)
$1_{EPM\text{-}Q3}$				-1.000***
				(0.00)
$1_{EPM\text{-}Q4}$				-1.737***
				(0.00)
Ret	0.072***	0.072***	0.071***	0.073***
	(0.00)	(0.00)	(0.00)	(0.00)
Vol	0.083***	0.083***	0.083***	0.083***
	(0.00)	(0.00)	(0.00)	(0.00)
Spr	-0.010***	-0.099***	-0.094***	-0.010***
	(0.00)	(0.00)	(0.00)	(0.00)
Adj. R^2	0.02	0.02	0.02	0.02

Table A15
EPM determinants

The table reports the coefficients and the marginal effects from a probit model of EPM occurrence:

$$Prob\ (EPM = 1)_{it} = \alpha + \beta_1 HFT_{it-1}^{NET} + \beta_2 Ret_{it-1} + \beta_3 Vol_{it-1} + \beta_4 Spr_{it-1} + \varepsilon_{it},$$

where the dependent variable is equal to one if an interval t contains an extreme price movement and zero otherwise. All independent variables are lagged by one interval. HFT^{NET} is the share volume traded in the direction of the price movement minus the share volume traded against the direction of the price movement for all HFT trades, Ret is the absolute return, Vol is total traded volume, Spr is the percentage quoted spread. All variables are standardized on the stock level. The marginal effects are scaled by a factor of 1,000. p-Values are in parentheses. *** and * indicate statistical significance at the 1% and 10% levels.

	All	Standalone	Co-EPMs	Permanent	Transitory
	(1)	(2)	(3)	(4)	(5)
Intercept	-3.135***	-3.256***	-3.430***	-3.344***	-3.388***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
HFT^{NET}_{t-1}	0.000	-0.002***	0.001	0.002*	-0.002
Marginal effect	0.001	-0.004	0.002	0.003	-0.002
	(0.58)	(0.00)	(0.21)	(0.06)	(0.18)
Controls	Yes	Yes	Yes	Yes	Yes
Pseudo-R^2	0.04	0.04	0.03	0.03	0.04