

取引行動のクラスター分析：HFT 抽出の試み^a

宇野 淳^b 五島 圭一^c 戸辺 玲子^d

要約

1 銘柄 100 件の発注をする主体を high frequency trader (HFT)と見なして、クラスター分析により取引行動の近いトレーダーグループを各 10 組形成した。行動特性のプロキシは、リスク管理指標である在庫比率・注文取消率と高速取引特性を示す注文頻度・取引銘柄数という 2 組のプロキシを用いた。各々の結果でマーケットメイク (MM) 型と見なしたグループメンバーは 7 割が重複しており、リスク管理と高速取引インフラの間に共通するものがあることが確認された。これらのグループは、限定された銘柄を取引しており、取引銘柄の気配件数のシェアは最も高く、約定への貢献も高い。一方、重要な違いとしてはインフラ指標でグルーピングした HFT—MM は成行注文を使っていないメンバーだけが選ばれたという点である。より執行リスク管理を徹底しているメンバーだけでグループが構成されたことは興味深い結果であり、成行注文に対する考え方がグループ行動の基本と密接に関係している可能性を示唆している。

JEL 分類番号： G12, G13, G18, G28

キーワード：HFT, クラスター分析, 市場流動性, マーケットメイク

^a 本稿に示されている意見は、筆者たち個人に属し、日本銀行の公式見解を示すものではない。また、あり得べき誤りはすべて筆者たち個人に属する。

^b 早稲田大学 juno@waseda.jp

^c 日本銀行金融研究所・早稲田大学 keiichi.goshima@boj.or.jp

^d 早稲田大学 r.tobe@aoni.waseda.jp

1. イントロダクション

1.1. クラスター分析を用いる狙い

本研究の特徴は、データドリブンで株式市場での取引行動が似ている主体を抽出するという方法を採用している点にある。行動特性を示す指標である注文取消比率や在庫率でグループ分けする場合、100%と0%の中間のどこで線を引くのかを客観的に決定する方法は見当たらない。また、国や扱う金融商品によって、その水準は異なりうる。そこで、水準が近い主体を似たもの同士をクラスター分析によりグループ化し、グループとしての集団行動を観察する。クラスター分析では、行動特性を表す2組のプロキシを用い、利用する指標によりHFTとしてグループ化される主体がどのように変化するかを報告し、HFTの行動原理との関係を明らかにする。

1.2. 2組のプロキシ

具体的な2組のプロキシとは、在庫・注文取消と、注文頻度・銘柄数である。前者はHFTに関する多数の先行研究で、HFTは低在庫・高取消という行動を示すことが報告されていることにより選択した。実際、市場参加者全体について在庫と取消について調べると、HFTとは相容れない取引行動（例えば取引銘柄数が1000を超える）の主体でも、低在庫・高取消に該当するものがある。これに対して、後者の注文頻度・VS当たりの銘柄数は、相場動向やそのほかの偶然要因に左右されにくいという点に注目して選んだプロキシである。投資行動の原則や投資対象は予め決められており、相場動向等によって毎日変化するものではない。投資行動のビジネスモデルやリスク管理など基本的な投資方針を反映していると考えられる。こうした外形的な基準が似たグループが取消比率や在庫比率についてどのような特徴をもっているのかを確認することは、投資方針と投資インフラの関係を明らかにすることにも通じる。

2. 実証分析

2.1. クラスター分析の結果

本節では、分析に用いるデータとその前処理に関して記述する。本研究では、株式会社日本取引所グループから提供された、東証アローヘッドにおける株式の注文・約定データを用いる。データの期間は、2013年4月1日から5月31日までの42営業日を利用する。このデータの最大の特徴は、「仮想サーバー識別番号」と「個別注文識別番号」が付与されている点である。本研究では、上述の2組4指標に基づき、232の高頻度取引を実行しているVS-IDについて、クラスター分析の手法の一つであるk-meansを利用して分類を試

みる。クラスター分析の対象とした 232 の VS-ID は、個別銘柄に毎日 100 件以上の注文を出している主体であり、高頻度取引を行っていると考えた。もちろん、これが一般に言われる「HFT」とは異なる参加者を含んでしまう可能性もあるが、むしろ、似たような取引スタイルなのに、HFT と非 HFT に分類してしまう可能性がないのかを確認することにも意義があると考えた。

表 1 G3、G8、G13、G18、G20 に選ばれた VS-ID の重複状況

在庫・取消基準 頻度・銘柄数基準	G13	G18	G20	計
G1	0	4	2	6
G2	1	0	0	1
G3	10	2	0	12
G5	0	1	2	3
G7	1	0	0	1
G8	2	3	1	6
G9	0	5	7	12
計	14	15	12	41

表 2 HFT グループの特徴

(A) 頻度・銘柄数基準で形成された 10 個の HFT グループ

頻度・銘柄数基準										
グループ 番号	銘柄数/VS	アクション数/銘柄	注文取消比率	在庫比率	VS数	指値価格乖離率	成行注文件数	在庫反転回数	注文修正 待機時間	注文取消待 機時間
1	10.0	564.4	69.4%	24.2%	6	0.20%	0.22	32.36	32.40	34.8
2	39.5	145.9	22.8%	69.3%	12	0.24%	75.43	11.19	20.48	114.1
3	1.8	2,761.9	81.3%	6.4%	8	0.15%	0.00	37.31	N/A	57.3
4	273.0	113.8	46.8%	77.9%	9	0.17%	296.35	20.87	130.52	278.6
5	14.5	146.0	61.0%	64.2%	13	0.20%	360.25	29.24	8.19	14.3
6	1.3	209.0	71.0%	72.3%	12	0.49%	13.25	3.32	1173.35	982.4
7	50.6	282.1	53.4%	85.5%	23	0.09%	2244.78	36.19	24.47	27.0
8	20.0	1,978.2	83.8%	5.5%	30	0.15%	1439.25	44.03	37.77	32.1
9	7.0	318.7	68.5%	21.1%	47	0.14%	57.10	20.51	32.92	29.2
10	35.9	186.2	16.1%	60.4%	70	0.24%	7.56	15.88	23.38	142.7

(B) 在庫・取消基準で形成された 10 個の HFT グループ

在庫・取消基準										
グループ 番号	銘柄数/VS	アクション数/銘柄	注文取消比率	在庫比率	VS数	指値価格乖離率	成行注文件数	在庫反転回数	注文修正 待機時間	注文取消待 機時間
11	40.4	175.3	8.6%	51.5%	15	0.25%	92.76	20.51	46.17	217.9
12	184.1	142.1	74.2%	100.0%	8	0.23%	108.42	1.44	26.88	11.6
13	13.3	2,583.1	77.9%	6.3%	11	0.16%	163.00	42.12	1.52	55.2
14	68.8	234.0	45.9%	98.4%	14	0.08%	384.18	11.26	20.80	15.8
15	41.0	175.1	8.4%	67.3%	31	0.25%	85.42	19.51	194.17	333.7
16	24.5	247.1	82.9%	45.3%	31	0.30%	399.45	21.51	49.60	109.6
17	165.4	152.4	1.6%	26.8%	12	0.03%	4325.06	51.21	75.00	100.8
18	8.6	792.7	95.5%	7.2%	50	0.14%	0.00	8.01	N/A	47.6
19	14.6	191.4	91.4%	79.6%	51	0.53%	0.50	2.93	71.81	18.8
20	8.4	421.2	41.4%	4.5%	9	0.05%	720.58	78.30	13.40	15.9

表 1 と表 2 は、クラスター分析によって抽出された各グループの特徴をまとめたものである。記載されている指標は各グループの平均値である。A 基準として 1 銘柄当たり注文頻度と 1VS 当り銘柄数、B 指標として在庫比率と注文取消比率を使った。前者は、取引インフラと行動特性の指標であり、後者は在庫リスク管理に関係する指標である。後者は多数の HFT 研究で指摘されている重要な特徴であり、これが取引インフラの設計とどう関係しているかという興味深い探求にもつながるものである。

まず 10 グループからマーケットメイク (MM) 型といえるグループを選ぶと、A 基準では G3 と G8 が選ばれた。両者は、高頻度・小銘柄数で、かつ高取消・低在庫の特徴をもっているが、成行注文の使用に関する判断で明確に異なっている。G3 は全く成行注文を使っておらず、G8 は 10 グループで 2 番目に成行注文件数が多い。一方、B 基準では G13、G18、G20 が高取消・低在庫の特徴をもっている。このなかでは G20 の取消比率は 41% と他の 2 つよりもかなり低い。どこまでの水準で取消比率の足切をするのか、明確な基準がないなかでは検討対象に加えた。この 3 つの HFT は MM 型の特徴を示しているだけでなく、高頻度・小銘柄という特徴を共有している。頻度は 2583 件から 421 件と幅があるが、上位 3 位以内である。ここでも成行注文の有無に注目すると、G18 は一切利用がなく、G13 と G20 は利用している。

2 つの異なる基準で選ばれた G3 と G13 のメンバーを見ると、両者は 10 個の VS-ID を共有しており、全メンバーの 7-8 割が同じである。つまり、全く異なる基準を使ったにもかかわらず、過半数のメンバーが同じグループに選ばれていたということは、取引インフラのデザインとリスク管理方針の間には重要な関係があることが確認されたと考えている。

2.2.2 段階回帰分析による流動性寄与分析

本節では、HFT の注文・約定シェアが市場の流動性にどう影響したのかを分析する。流動性指標としては、各銘柄の時間加重ビッド・アスク・スプレッドと時間加重デプスを用いる。用いたモデル・変数の詳細については紙面の都合上ここでは割愛するが、推計に当たっては銘柄固定効果モデルで推計した。

注文シェアとスプレッドの結果によれば、G2 と G3 の注文増加はスプレッドの縮小につながっている反面、G1、G7 および G8 の注文増加はスプレッドの拡大につながっている。また、注文シェアとデプスの結果によれば、G3 の注文シェアの増加はデプスの増大につながっているが、G7 と G8 は減少につながっている。流動性供給という視点で、G3 はスプレッドの縮小とブックの厚みへの貢献が確認された。

前節の結果を踏まえ、HFT-MM の存在を実際の注文・約定シェアでとらえたところ、G3

と G13 の注文シェアが最も高かった。また約定シェアでも 1 位 2 位を争う位置にある。シェアの大きさが、取引対象銘柄のスプレッドやデプスの変化と関係するかをみたところ、G3 はスプレッドの低下とデプスの増大に寄与していることが確認された。一方、G8 は G3 とは対照的な影響を与えており、G3 が最も流動性供給に貢献していると結論をくだせる。

以上で G3 について述べたことは、G13 についても同様であり、クラスター分析を 2 組の異なる指標で行ったものの、どちらからもコアとなるグループ (G3 と G13) が選ばれた。在庫・取消比率は、HFT を見分けるうえで有用な指標であり、取引インフラに関する情報がない場合でも、同程度の精度でグループを形成できる可能性を示唆するものである。しかし、G3 と G13 では成行注文に対する方針が異なっており、このことが持つ意味については、今後の研究課題としたい。

表 3 2 段階パネル分析 (銘柄固定効果モデル)

(A) 注文シェア

被説明変数	スプレッド			対数(デプス)		
	回帰係数	標準誤差	t値	回帰係数	標準誤差	t値
注文シェアG1(IV)	0.077	0.042	1.84 *	0.019	0.538	0.04
注文シェアG2(IV)	-0.148	0.112	-1.32	0.606	1.602	0.38
注文シェアG3(IV)	-0.039	0.016	-2.34 ***	0.375	0.220	1.70 *
注文シェアG7(IV)	0.008	0.006	1.38	-0.056	0.073	-0.77
注文シェアG8(IV)	0.158	0.062	2.54 ***	-1.731	0.883	-1.96 **
1/VWAP	79.734	4.467	17.85 ***	203.394	73.583	2.76 ***
切片	-0.230	0.166	-1.39	19.753	1.769	11.16
調整済み決定係数	0.933			0.922		

(B) 約定シェア

被説明変数	スプレッド			対数(デプス)		
	回帰係数	標準誤差	t値	回帰係数	標準誤差	t値
約定シェアG1(IV)	0.315	0.242	1.30	-4.360	2.221	-1.96 **
約定シェアG2(IV)	0.209	0.176	1.18	-2.902	1.759	-1.65 *
約定シェアG3(IV)	0.049	0.021	2.32 **	-0.682	0.206	-3.32 ***
約定シェアG7(IV)	0.103	0.048	2.12 **	-1.405	0.498	-2.82 **
約定シェアG8(IV)	-0.020	0.011	-1.78 *	0.379	0.117	3.25 ***
1/VWAP	85.209	4.140	20.58 ***	143.235	67.629	2.12 **
切片	-1.157	0.517	-2.24 **	32.752	5.222	6.27 ***
調整済み決定係数	0.937			0.926		

注 : White cross-section standard errors & covariance (d.f. corrected)

引用文献

保坂豪, 2014. 東京証券取引所における High-Frequency Trading の分析. JPX ワーキングペーパー. <http://www.jpx.co.jp/corporate/news-releases/20140520-01.html>

- Aït-Sahalia, Y. and M. Saglam, 2013. High Frequency Traders: Taking Advantage of Speed. NBER Working Paper, 19531 (<http://www.nber.org/papers/w19531>).
- Arthur, D. and S. Vassilvitskii, 2007. K-means++: The Advantages of Careful Seeding. In Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms, 1027–1035.
- Autorité des Marchés Financiers (AMF), 2017. Study of the Behaviour of High-Frequency Traders on Euronext Paris: Risks and Trends. (http://www.amf-france.org/en_US/Publications/Lettres-et-cahiers/Risques-et-tendances/Archives?docId=workspace%3A%2F%2FSpacesStore%2F619a82db-11ea-4544-a03e-991b7af2864d)
- Bellia, M., L. Pelizzon, M. G. Subrahmanyam, J. Uno, and D. Yuferova, 2017. Low-Latency Trading and Price Discovery: Evidence from the Tokyo Stock Exchange in the Pre-Opening and Opening Periods. (<https://ssrn.com/abstract=2841242>).
- Brogaard, J., T. Hendershott, and R. Riordan, 2014. High-Frequency Trading and Price Discovery. *Review of Financial Studies*, 27(8), 2267–2306.
- Brogaard, J., B. Hagströmer, L. Nordén, and R. Riordan, 2015. Trading Fast and Slow: Colocation and Liquidity. *Review of Financial Studies*, 28(12), 3407–3443.
- Easley, D., M. M. L. de Prado, and M. O'Hara, 2012. Flow Toxicity and Liquidity in a High Frequency World. *Review of Financial Studies*, 25(5), 1457–1493.
- Hastie, T., R. Tibshirani, and J. Friedman, 2001. *The Elements of Statistical Learning*. Springer.
- Kirilenko, A., A. S. Kyle, M. Samadi, and T. Tuzun, 2017. The Flash Crash: High-Frequency Trading in an Electronic Market. *Journal of Finance*, 72(3), 967–998.
- Malinova, K., A. Park, and R. Riordan, 2018. Do Retail Investors Suffer from High Frequency Traders. (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2183806).
- Menkveld, A. J., 2013. High Frequency Trading and The New Market Makers. *Journal of Financial Markets*, 16(4), 712–740.
- O'Hara, M., 2015. High Frequency Market Microstructure. *Journal of Financial Economics*, 116 (2), 257–270.