

令和 2 年 6 月 8 日現在

機関番号：32612

研究種目：基盤研究(C) (一般)

研究期間：2016～2019

課題番号：16K03601

研究課題名(和文) 高頻度注文板データを用いた高速での株価形成に関する統計解析

研究課題名(英文) Statistical analysis on high-speed stock price formation with the use of high-frequency limit-order book data

研究代表者

林 高樹 (HAYASHI, Takaki)

慶應義塾大学・経営管理研究科(日吉)・教授

研究者番号：80420826

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：本研究は、高頻度注文板データに対する統計的データ解析により、高速化の進む今日の株式市場の株価形成に関する実証的知見を獲得し理解を深めることを目指すものであり、以下の成果を得た。(1) 国内株式市場の高速での株価間の先行遅行関係を探索的に分析し、実証的特徴を示した。(2) ウェーブレットを応用することで、証券価格の変動を異なる周波数成分毎に先行遅行時間を推定する統計理論・方法論を開発した。(3) 高速取引を行う市場参加者の取引行動に関する理論モデルを提案した。(4) 個別銘柄の高頻度領域での流動性を評価する方法論として、協調フィルタリングを応用するアプローチを検討した。

研究成果の学術的意義や社会的意義

(学術) (1) 高速での株価の先行遅行関係の実証分析は十分なされておらず、特に国内は希少であり新規性がある。(2) 提案手法は、複数の先行遅行時間を同時推定できる統計的方法論として新規性・貢献性が特に高い。理論のベースとなる確率過程モデルは、不均質市場仮説や、摩擦のある市場で無裁定性が成立するなどファイナンスの標準理論とも整合的である。(3) 高頻度取引業者(HFT)に関する実証的知見と整合的な性質を持つ理論モデルを考案した点で意義がある。(4) 推薦モデルの流動性推定問題への適用の点で新規性がある。  
(社会) 市場への理解の深化を通じて市場の機能を向上させる可能性がある。

研究成果の概要(英文)：The purpose of this study is to obtain empirical knowledge and deepen understanding of stock price formation in the stock markets with increasing speed of execution by statistical analysis of high-frequency limit order book data, and the following results are obtained. (1) An exploratory data analysis of high-frequency lead-lag relationships between stock prices in the domestic stock market is performed and empirical characteristics are demonstrated. (2) Based on wavelet theory, we develop a statistical theory and methodology for estimating the lead-lag times for each of the frequency components of securities price fluctuations. (3) A theoretical model is proposed to explain the trading behaviors of high-speed market participants. (4) A collaborative filtering-based methodology to evaluate the liquidity of individual securities in the high-frequency domain is investigated.

研究分野：計量ファイナンス

キーワード：高頻度データ 高頻度トレード (HFT) マーケット・マイクロストラクチャ 注文板市場 先行遅行分析 推薦システム Kyleモデル データサイエンス

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。

様式 C - 19、F - 19 - 1、Z - 19 (共通)

## 1. 研究開始当初の背景

近年の ICT 技術の飛躍的發展を背景に、株式市場は、売買の高速化、注文件数の増大、刻み値(ティックサイズ)の縮小など、急速な変化を遂げている。それに伴い、入手可能な取引データも、その項目の範囲や解像度が年々増している。例えば、2015 年半ば時点で、日本取引所グループ(JPX) 東京証券取引所の FLEX Full サービスでは 0.1 ミリ秒刻み、ニューヨーク取引所の Daily TAQ サービスは米国株式に関するマイクロ秒刻みの時間解像度を持つ高頻度データを提供している(2019 年度末時点では、前者はマイクロ秒、後者はナノ秒刻み)。

投資家からの注文が証券会社を経由して取引所に到着すると、取引所の売買システム内で、銘柄ごとに注文の控え(板)に集められる。各銘柄の「注文板」は売り買い別、希望の売買価格(指値)ごとに配置され、各価格帯には、到着時間順に指値注文の待ち行列が形成される。市場が開き取引が開始されると、板上に提示された指値注文に対して、即時売買成立(約定)可能な反対側(売り指値注文に対しては買い、買い指値注文に対しては売り)の注文が到着するたびに、「価格優先・時間優先」などの取引所のルールに従って逐次付け合せが行われ(取引が成立し)、あるいは、注文変更や取消がなされ、注文板が更新されていく。

一般に、市場は取引の成立を通じて適正価格を市場内外に発信する働きを持つ(価格発見機能)。このような市場が“適正に”機能を果たす程度は市場クオリティと呼ばれ、取引コストの他、執行速度、変動性(ボラティリティ)、流動性などがこれに影響を与える要素とされる。先行研究では、高頻度取引業者(HFT)は平常時においては流動性供給を通じて、ボラティリティを下げるなど市場クオリティ向上に寄与していると報告されている。

上述のように、今日の株式市場の基盤である、大量の注文を高速に処理する取引所の売買システム内で形成される注文板は、多変量かつ時間的に状態が刻々と変化していく、解析の容易でないシステムである。注文板上でなされる価格形成は、需給のみならず、ティックサイズや取引コストなどの市場の仕組みや取引ルールなど(マーケット・マイクロストラクチャ)の影響も受ける。注文板に関する研究は、ファイナンス分野はもちろんのこと、統計学、計量経済学、計量ファイナンス、数理ファイナンス、金融工学、オペレーションズ・リサーチ、経済物理学など、多くの分野が関わる学際領域である。高頻度注文板データの利活用は、実務はもとより学術研究にも広がっている。

株式市場は、注文処理を行うホストコンピュータの処理スピードの向上が随時図られ、ティックサイズの変更など取引の仕組みも随時更新されている。高速化・微細化が加速的に進む株式市場において、価格形成がどのようになされ、市場クオリティがどのような制度下で維持されるかなどについて、研究が十分に行われ、理解が進んでいるとは言い難い。

## 2. 研究の目的

本研究は、国内市場を中心に高頻度注文板データに対する統計的データ解析を行うことで、売買の高速化、注文件数増大、ティックサイズ縮小など変化の著しい今日の株式市場に関する知識発見を行い、実証的理解を進めることを目的とする。

具体的には、(1)国内高速三市場の注文板データを用いて、高速化の進展する我が国株式市場の価格形成、変動性や流動性についての探索的データ解析を行い、次に、(2)(1)の結果を、高頻度・高速売買(HFT)に関する先行研究サーベイと重ねることで、HFT 業者などを中心に高頻度・高速領域における市場参加者の取引行動に関する仮説構築を試み、実証的に検証を行う。さらに、最終ステップとして、(3)、(1)(2)で得られた知見を全米市場データに適用することで、株式市場に対するより普遍的、一般的な知見を獲得することを目指す。

### 3. 研究の方法

上記の研究目的(1)-(3)を達成するための方法は、取引所より入手した高頻度注文板データの探索的分析、統計的方法論の検討・開発、資料収集と先行研究サーチによるHFTの取引行動に関する理論構築と検証(のサイクル)である。具体的には、研究経過と共に、随時見直しを図りながら、以下のように進めた。

研究を進める上で当初から中核に据えたテーマは、先行運行関係分析(方法論の開発と実証分析)である。まず、研究代表者の過去年度の成果を出発点に、先行運行関係分析をさらに進めることとした。推定方法として、ランダムに並ぶタイムスタンプのみの情報から先行運行推定量を構成するDobrev and Schaumburg(2016)法(DS法)を主要な方法論に据え、価格変化の情報も使用するHofmann-Rosenbaum-Yoshida(2013)法(HRY法)をベンチマークに据えた。高頻度注文板データを使い、DS法の修正版によって計測された先行運行時間を目的変数として、ボラティリティ、スプレッド株価比、気配更新件数、デプス(注文数量)などの共変量や、一日内時間帯や銘柄、日付などの因子(固定あるいは変量)を説明変数としたパネル回帰分析を実施し、2系列間の先行運行関係を説明する要因を実証的に調べた。

研究目的(1)において、当初、国内で現物株式の高速売買が行われている3市場(東証、Japannext PTS, Chi-X Japan)において同時に取引されている同一銘柄の株価の先行運行関係を調査した(異市場間・同一銘柄分析)。しかし、東証が国内の取引シェアの9割を占めていること、また、これらの3市場の注文データには注文者を特定できるIDがついておらず、高頻度取引業者(HFT)の市場をまたがった行動を把握できないことなどから、2年目以降はPTS市場は対象から外し、株価指数(TOPIXや日経平均株価)に連動するETF(上場投資信託)間の先行運行関係分析(同一市場内異銘柄分析)や、株価指数先物と株価指数連動ETF間の先行運行関係分析(異市場間・異銘柄分析)を行った。

さて、先行運行関係分析を進める中、2つの推定法、HRY法とDS法による推定の結果が大きく異なることが明らかになった。そこで、金融証券市場が資金の質や取引目的、リスク許容度など多様な市場参加者によって構成されるという「不均質市場仮説」(Heterogeneous Market Hypothesis)にも整合的なアプローチを、信号処理分野を中心に学術研究や実用化が進んでいる「ウェーブレット」を先行運行関係推定に応用する研究を開始することにした。

一方、目的(2)については、3年目よりマイクロストラクチャ分野における高頻度取引業者(HFT)の行動に関する先行研究サーチを行い、先行研究や(1)の国内市場の先行運行関係に関する探索的分析の観察結果とも整合的な、HFTの売買行動に関する理論モデルの構築に着手した。ベースとなるモデルは、市場参加者間の情報の非対称性をモデル化したものとして代表的なKyle(1985)モデルである。

同(3)については、期間中の進捗を踏まえ、当初目的(3)に掲げたような株式市場に関する(実証的)知見の一般化を目指すのではなく、まずは国内市場の分析に注力するよう軌道修正した。そして、米国市場は調査中の方法論の有効性を示す際の対象市場の一つとして扱うに止めることとした。

当初より、本研究においても、今日研究発展の著しい機械学習・人工知能系のアプローチを積極的に活用することを視野においていた。過年度先行運行分析に焦点を当てていた(1)において、新たな切り口からこれを進めるため、3年目より、高頻度注文板データを使う、「推薦システム」、特に「協調フィルタリング」を応用した市場クオリティ評価のための枠組みを検討した。

#### 4. 研究成果

期間中に得られた研究成果は、(着手の順に)大きく4つに分けられる。まず、国内株式市場の高速での株価間の先行運行関係を探索的に分析し、実証的特徴を示したことである。二つ目は、「ウェーブレット」を応用することで、証券価格の変動を異なる周波数成分毎に先行運行時間を推定するための統計理論・方法論を開発したことである。三つ目は、高頻度・高速取引を行う市場参加者の取引行動に関する理論モデルを提案したことである。四つ目は、個別銘柄の高頻度領域での流動性を評価することを目的として、「協調フィルタリング」を応用した新たなアプローチを検討したことである。

##### 4.1 国内株式市場における高頻度先行運行関係の実証分析

3市場間の先行運行の実証分析を、Dobrev and Schaumburg(2015)による先行運行時間推定法(DS法)を改良した方法で行った。DS法は価格に含まれるマイクロストラクチャノイズの影響を受けにくい“ロバストな”推定が可能な一方、提案されて日が浅く応用研究は少ない。TOPIX100構成銘柄を2年分、ミリ秒刻みの高頻度注文板データを用いて3市場間の先行運行関係を実証分析した。同一銘柄・異市場の先行運行分析は、国内外でいまだに希少である。2014年に実施された東証における2回のティックサイズ変更前後の変化に関する観察も含めて、分析結果をまとめた単著論文を国内学術誌に掲載した(林(2017))。

また、同じDS法の改良型により、国内株式市場について日経平均連動型ETFと日経平均指数先物との先行運行関係分析を行い、HFTの1日内の行動変化や、ETFタイプの違いによる組成やヘッジ方法の相違を反映していると思われる実証的特徴を調査した(異市場・異銘柄分析)。指数先物とETFの連動性や先行運行関係を調べた価格発見機能に関する実証研究は古くからあるが、HFTやコロケーションなど今日の市場構造が反映された新しいデータセットを用いたものはいまだに少ない。

今回行ったこれらの実証分析は、現時点においては探索的分析に止まっており、得られた実証的知見に対する理論的考察や仮説の構築・検証などの課題が残された。

##### 4.2 ウェーブレットによる高頻度先行運行時間推定手法の開発

ステップ(1)のほか(2)(3)にも利用可能な新たな統計的方法論の開発に着手した(東京大学小池祐太准教授との共著)。これは信号処理分野を中心に応用が進む「ウェーブレット法」により、証券価格の変動を周期の異なる成分に分解し、各成分毎に先行運行時間を推定する方法論である。金融市場は多様な市場参加者によって構成されているという「不均質市場仮説」(Heterogeneous Market Hypothesis)にも整合的なアプローチである。

また、提案方法論のベースとして導入した確率過程モデル(ガウス過程の一種)に関する性質を調べた他、推定法の改良について検討するなど、理論面での整備を継続した。同手法の適用事例として、米国株式データを用いた同一銘柄・異市場分析をNASDAQ 100銘柄について実行した。進行中の実証分析によれば、先行研究では報告されていなかった高速領域における少なくとも2つの先行運行時間の存在可能性が示唆され、提案方法論の有効性を示すものとなった。国内の指数連動型ETFの先行運行関係の実証分析にも適用を試みた。

本手法は、複数の先行運行時間を同時推定できる新しい統計的方法論を提案するものでありオリジナリティが特に高い。導入した確率過程モデルは、不均質市場仮説の他、摩擦

のある市場で無裁定性が成立するなどファイナンスの標準理論とも整合的であるなど、市場価格の変動を記述するモデルとしての妥当性も有している。今後の課題としては、先行遅行推定量に関する統計的性質に関する調査や、計算処理速度を向上させるための改良方法の検討、推定法の実際のデータに応用した実証研究などが挙げられる。

研究代表者は、2016年7月、数理ファイナンス分野での2年に一度の世界大会Bachelier World Congress (ニューヨーク)におけるプレナリー講演者の一人に選ばれ、また、同年12月には、高頻度データ分析・マイクロストラクチャ分析の最先端の研究者が集まる2年に一度の国際会議 (パリ)「Market Microstructure Confronting Many Viewpoints #4」で招待講演者の一人として、初期段階の研究成果を世界に発信する機会を得た。その他にも、国内外の研究集会・会議において多数回研究報告を行った。

小池氏との共著論文は国際学術誌に掲載された(Hayashi and Koike(2018), Hayashi and Koike(2019))。別のワーキングペーパーは投稿準備中である (arXiv 登録済み)。

#### 4.3 高頻度・高速領域における市場参加者の取引行動に関する理論構築の研究

高頻度・高速領域における市場参加者の取引行動に関する理論構築に向けて、Kyle(1985)型の市場モデルを検討した (一橋大学西出勝正教授との共著)。この理論モデルは市場に高速と低速の2種類のトレーダーが存在すると仮定する。数値計算により、取引頻度に優位性がある高速トレーダーが、自身の期待利潤を最大化するための戦略的行動の結果として、二者間の取引頻度の差が大きい場合、低速トレーダに対する流動性供給者としての役割を果たすことを示した。単純な設定にも拘らず、実証分析におけるHFTの一日内行動とも整合的と思われる戦略的行動パターンの出現が確認された。HFTの市場機能への影響を理解する手掛かりを得る上で、トレーダー間の取引頻度の差異の効果を調べた先行研究は稀有であり、新規性があると考える。研究初期段階にあり今後研究の拡張を図る必要がある。国内外の会議で1件ずつ報告した。また、国内の報告が1件確定しているほか、ワーキングペーパーをSSRNに登録した。

#### 4.4 推薦システムの応用による個別株式銘柄の高頻度領域での流動性評価手法の研究

市場クオリティの測定に関する方法論の確立と実証分析を目的とした研究に着手した(法政大学高橋慎准教授との共著)。近年普及の著しい「推薦システム」を応用することで、個別株式銘柄の高頻度領域での流動性を評価する方法論を検討した。スパースで相関性の高い高頻度データを使った個別銘柄の流動性の推定問題は、「協調フィルタリング」が有効に働く可能性がある。本研究では、欠損が著しい「コールドスタート」に対応するため、Agarwal and Chen (2009)による回帰ベース潜在因子モデルの枠組みを採用し、銘柄属性や市場全体に関する情報を活用したモデルの構築を試みた。東証一部の高頻度注文板データを用いて、検討手法の有効性を調査した。研究初期段階にあり、今後も研究を継続する必要があるが、協調フィルタリングを流動性評価に応用する研究は先行研究がサーベイした限りでは見当らなく、アプローチに新規性があると考えている。なお、今回の研究では流動性評価を具体的に上げたが、方法論自体には汎用性があることから、ボラティリティなど他の市場クオリティ指標の評価や目的に対しても、検討手法の応用が可能である。国内会議2件(内1件は会議中止・発表成立)報告した。また、国内外1件ずつの報告が確定しているほか、ワーキングペーパーは投稿準備中である。

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 1件）

1. 著者名 Takaki Hayashi and Yuta Koike	4. 巻 154
2. 論文標題 No arbitrage and lead-lag relationships	5. 発行年 2019年
3. 雑誌名 Statistics & Probability Letters	6. 最初と最後の頁 108530
掲載論文のDOI（デジタルオブジェクト識別子） <a href="https://doi.org/10.1016/j.spl.2019.06.006">https://doi.org/10.1016/j.spl.2019.06.006</a>	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Takaki Hayashi and Yuta Koike	4. 巻 9
2. 論文標題 Wavelet-based methods for high-frequency lead-lag analysis	5. 発行年 2018年
3. 雑誌名 SIAM Journal on Financial Mathematics	6. 最初と最後の頁 1208
掲載論文のDOI（デジタルオブジェクト識別子） 10.1137/18M1166079	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 林 高樹	4. 巻 65
2. 論文標題 高頻度注文板データの統計解析：異市場・同一株式価格間の先行遅行関係	5. 発行年 2017年
3. 雑誌名 統計数理	6. 最初と最後の頁 113-139
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計18件（うち招待講演 3件／うち国際学会 8件）

1. 発表者名 林高樹, 高橋慎
2. 発表標題 協調フィルタリングを応用した注文板市場の流動性評価
3. 学会等名 第52回2019年度冬季JAFEE大会（中央大学）（大会中止・発表成立）
4. 発表年 2020年

1. 発表者名 林高樹, 高橋慎
2. 発表標題 協調フィルタリングを応用した注文板市場の流動性計測
3. 学会等名 第7回 統計数理研究所 リスク解析戦略研究センター 金融シンポジウム (東京) (招待講演)
4. 発表年 2019年

1. 発表者名 Takaki Hayashi and Katsumasa Nishide
2. 発表標題 Strategic Liquidity Provision in High Frequency Trading
3. 学会等名 Quantitative Methods in Finance Conference 2019, Sydney (国際学会)
4. 発表年 2019年

1. 発表者名 林高樹, 西出勝正
2. 発表標題 Strategic Liquidity Provision in High Frequency Trading
3. 学会等名 日本ファイナンス学会第1回秋季研究大会 (大阪大学)
4. 発表年 2019年

1. 発表者名 林 高樹, 小池祐太
2. 発表標題 高頻度金融市場におけるリード・ラグ関係の 多時間スケール解析
3. 学会等名 2018年度統計関連学会連合大会 (中央大学)
4. 発表年 2018年

1. 発表者名 Takaki Hayashi and Yuta Koike
2. 発表標題 No arbitrage and lead-lag relationships
3. 学会等名 10th World Congress of The Bachelier Finance Society, Dublin (国際学会)
4. 発表年 2018年

1. 発表者名 Takaki Hayashi and Yuta Koike
2. 発表標題 No arbitrage and lead-lag relationships
3. 学会等名 第48回 2017年度冬季JAFEE大会 (武蔵大学)
4. 発表年 2018年

1. 発表者名 林 高樹, 小池祐太
2. 発表標題 Wavelet-based methods for high-frequency lead-lag analysis
3. 学会等名 日本ファイナンス学会第25回大会 (千葉工業大学)
4. 発表年 2017年

1. 発表者名 Takaki Hayashi and Yuta Koike
2. 発表標題 Wavelet-based methods for high-frequency lead-lag analysis
3. 学会等名 10th Annual SoFiE Conference, New York (国際学会)
4. 発表年 2017年



1. 発表者名 Takaki Hayashi and Yuta Koike
2. 発表標題 Multi-scale analysis of lead-lag relationships in high-frequency financial markets
3. 学会等名 釧路・経済統計キャンプ2017 (新しい時系列計量分析の理論と応用) (釧路公立大学)
4. 発表年 2017年

1. 発表者名 林 高樹, 小池祐太
2. 発表標題 ウェーブレット法による高頻度金融時系列間の先行遅延関係分析
3. 学会等名 2017年度統計関連学会連合大会 (南山大学)
4. 発表年 2017年

1. 発表者名 林 高樹
2. 発表標題 高頻度データによる証券価格間の先行遅延関係分析: 代替的アプローチ
3. 学会等名 ワークショップ「証券市場の諸問題」(大阪大学)
4. 発表年 2017年

1. 発表者名 林 高樹
2. 発表標題 高頻度データを用いた金融証券価格間の先行遅延時間推定方法の比較
3. 学会等名 第46回 2016年度冬季JAFEE大会 (武蔵大学)
4. 発表年 2017年

1. 発表者名 Takaki Hayashi and Yuta Koike
2. 発表標題 Wavelet-based methods for high-frequency lead-lag analysis
3. 学会等名 Asymptotic Statistics and Computations 2017, University of Tokyo (国際学会)
4. 発表年 2017年

1. 発表者名 Takaki Hayashi and Yuta Koike
2. 発表標題 A Multiresolution Approach to High-Frequency Lead-Lag Analysis
3. 学会等名 Market Microstructure Confronting Many Viewpoints 2016, Paris (招待講演) (国際学会)
4. 発表年 2016年

1. 発表者名 Takaki Hayashi and Yuta Koike
2. 発表標題 Wavelet-based methods for high-frequency lead-lag analysis
3. 学会等名 TMU Workshop on Financial Mathematics and Statistics 2016, Tokyo Metropolitan University (国際学会)
4. 発表年 2016年

1. 発表者名 Takaki Hayashi
2. 発表標題 On lead-lag analysis with high-frequency data
3. 学会等名 Keio Symposium on Risk Assessment, Keio University (国際学会)
4. 発表年 2016年

1. 発表者名 Takaki Hayashi
2. 発表標題 From covariance estimation to lead-lag analysis with high-frequency data: methods and issues
3. 学会等名 9th World Congress of The Bachelier Finance Society, New York (招待講演) (国際学会)
4. 発表年 2016年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----