

令和元年6月3日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K17105

研究課題名(和文) 確率過程アプローチによる価格発見の統計解析

研究課題名(英文) Statistical analysis of price discovery: a stochastic process approach

研究代表者

小池 祐太 (Koike, Yuta)

東京大学・大学院数理科学研究科・准教授

研究者番号：80745290

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：現代の金融市場では、単一の金融資産が複数の市場で取引されるのが普通である。このような状況下では、単一の金融資産に複数の価格がつくことになるが、これらの価格は超短期的には多少乖離しても、ある程度の時間スケールではほぼ同じ価格を取るべきである。この際に問題となるのは、どの市場における価格が基準となってその資産の価格が決定されているのか、という問題であり、これは価格発見の問題と呼ばれる。本研究では、金融高頻度データを連続時間確率過程の離散観測としてモデリングする方法(確率過程アプローチ)によって、モデルに特定の関数形を仮定せずにデータ解析することが可能となるような価格発見の統計モデルを構築した。

研究成果の学術的意義や社会的意義

本研究では、従来の価格発見モデルで利用されてきた離散時間モデルではなく、連続時間確率過程の離散観測モデルによって統計モデルおよび統計解析手法を構築しており、従来のモデルと比べて次のようなアドバンテージがある。

(1) パラメーターの時間変化を特定の関数形を仮定せずにモデルに組み込むことができる。時間変化の影響を考慮することでモデルの説明力が改善され、新たな知見が得られることが期待できる。

(2) 観測時刻の非同期性に対処することができる。従来モデルで広く使われる補間によるデータの同期化は重大なバイアスをもたらすことが指摘されており、この問題を解消して分析結果に生じるバイアスを排除できる。

研究成果の概要(英文)：In modern financial markets, it is common that a single asset is traded on multiple markets. In such a situation, a single asset can have multiple prices, but they should be almost identical at moderate time scales, although they might be different at a very short time horizon. In this case, one naturally asks which market determines the fundamental price of the asset. This is called the problem of price discovery. In this study, we construct a statistical model for price discovery, which enables us to analyze data without assuming a specific functional form of the model, by modeling high-frequency financial data as discretely observed continuous-time stochastic processes.

研究分野：計量ファイナンス

キーワード：リード・ラグ効果 高頻度データ マーケット・マイクロストラクチャー 漸近理論 非同期観測

## 1. 研究開始当初の背景

現代の金融市場では、単一の金融資産が複数の市場で取引されるのが普通である。このような状況下では、単一の金融資産に複数の価格がつくことになるが、「リスクなしで市場に対して超過収益をあげることはできない」という市場の無裁定性原理から、これらの価格は超短期的には多少乖離しても、ある程度の時間スケールではほぼ同じ価格を取るべきである。この際に問題となるのは、どの市場における価格が基準となってその資産の価格が決定されているのか、という問題であり、これは価格発見の問題と呼ばれる。従来の研究では、価格発見分析のための統計モデルとして、離散時間確率過程に基づく時系列モデルが広く利用されてきた。しかし、昨今の情報技術の進展に伴う取引の高速化に伴い、価格発見メカニズムを正確に捉えるためには、秒未満の単位で記録された超高頻度データの利用が必要となってきた。しかし、離散時間時系列に基づく統計モデルでは、金融超高頻度データに固有の特徴である観測の非同期性や複雑な日内周期性などの現象をうまくモデルに組み込めないという問題があった。このような問題点を背景として、本研究では、金融高頻度データを連続時間確率過程の離散観測としてモデリングする方法（確率過程アプローチ）によって、これらの特性を柔軟に組み込めるような価格発見の統計モデルの構築を目指した。

## 2. 研究の目的

上述の通り、本研究の目的は、従来の離散時間時系列モデルでは記述が困難であった金融超高頻度データ固有の特徴を柔軟にモデルに組み込めるような統計モデルの開発である。さらに、開発した統計モデルに基づき容易にデータ解析を実行することを可能にするような統計解析手法の開発もあわせて目指す。

## 3. 研究の方法

上記の目的を達成する手段として、本研究では、従来の研究で採用されてきた離散時間時系列によるデータのモデリングではなく、連続時間確率過程の離散観測としてデータをモデル化するアプローチ（確率過程アプローチ）を採用した。このアプローチを採用した背景には、2000年代以降の金融高頻度データ解析の分野において、ボラティリティや共分散の推定といった文脈で、確率過程アプローチが大きな成功をおさめたことがある。しかしながら、確率過程アプローチを価格発見の統計解析に適用した事例は先行研究において皆無であったため、本研究においてそのような方向性を新たに開拓した。

具体的な研究方法を、理論的研究と実証的研究の二つに分けて述べる。理論的研究では、確率過程アプローチによって、先述した金融超高頻度データ固有の現象を柔軟に反映させられるようなモデルを構築したのち、モデルを解析するための統計手法を開発する。統計手法の理論的妥当性を担保するための統計理論は、通常、確率過程アプローチにならって、観測頻度が無限に大きくなるような漸近理論を考えた下で、推定量の一致性や漸近分布の導出を行う。これら理論的結果の導出には、確率過程アプローチにおける理論研究において広く利用される「Jacodの安定収束理論」を応用した。

実証的研究は、「人工データに基づく数値シミュレーションによるモデルと理論的結果の有用性の検証」および「実際の金融高頻度データへの適用」の二つに分けられる。前者においては、上述の通り統計理論の妥当性は観測頻度が無限大という仮想的な状況でしか担保されないため、十分大きい有限の観測頻度においても漸近理論が働くことを確認するというのが第一の目的である。もう一つの目的としては、開発した統計モデルが実際の金融高頻度データで観察される現象を正しく記述できるかという点を確認するというのが挙げられる。後者においては、実際に開発した統計モデルや統計解析手法が本当に実データ解析に適用可能かどうかを確認するというのが第一の目的にある。さらに、得られた実証結果を、経済理論や先行研究から示唆される結論と擦り合わせることで、開発した統計モデルから得られた新たな知見を報告して今後の実証研究の礎としていくということも目的として挙げられる。もう一つの重要な目的としては、上述した擦り合わせ作業の結果得られた理論と実データとの矛盾点を洗い出すことで、元のモデルにフィードバックしてモデルの改善をはかるということがある。

## 4. 研究成果

### 確率過程アプローチに基づく価格発見の統計モデルの構築

「金融高頻度データを連続時間確率過程の離散観測データとしてモデリングする」という確率過程アプローチによって、モデルに特定の関数形を仮定せずに、価格発見に対する統計推測を実行することが可能となるような統計モデルを構築した。構築したモデルでは、二つの市場で取引される資産間の価格発見スピードの違いを、資産の本源的価格が実際に観測される価格に反映されるスピードの違いとして計測する。計量ファイナンス分野では、このスピードの違いはリード・ラグ効果と呼ばれる。従って、提案モデルに基づき価格発見に対する統計解析を行うには、このリード・ラグ効果を計測すればよいことになる。ま

た、構築したモデルでは、ボラティリティや価格発見スピードなどの時間的変化を、関数形を仮定せずにモデルに組み込むことができる。これによって、金融高頻度データ固有の特徴である複雑な日内季節性のモデル化を回避することができる。さらに、金融超高頻度データ解析に基づく価格発見解析の際に重要となる非同期観測の問題についても自然に対処することができる。

#### 提案モデルに基づく価格発見に対する統計解析手法の構築

上述の通り、提案モデルにおける価格発見の解析は、モデルにおけるリード・ラグ効果の大きさを表すパラメーターの推定に帰着される。そこで、観測データに基づいてこのパラメーター(リード・ラグパラメーターと呼ぶことにする)に対する統計推測を実行するための統計解析手法を構築した。本モデルではリード・ラグパラメーターは時間的変動を持ちうるため、ある一定期間内のリード・ラグパラメーターの値を平均化して要約するための統計量(スペクトル・リード・ラグ指数)を提案し、その統計量に対する推定量を構成した。また、時間変化するリード・ラグパラメーター自身を時間の関数として推定するための推定量も構成した。あわせて、構成した推定量たちに対する安定型中心極限定理を確立することで、それら推定量の漸近的性質(一致性、漸近混合正規性および収束レートの最適性)を示した。さらに、確立した安定型中心極限定理を信頼区間の構成や統計的仮説検定へと応用するために必要となる漸近分散の推定量として、高頻度データに対するサブサンプリング法に基づく推定量を提案した。

#### 数値シミュレーションによる漸近理論の実用的妥当性の確認

前項で整備した統計理論は、あくまで観測頻度が無限大となるような仮想的な状況のもとでその理論的正当性が担保されるため、実データ解析を行う前に、現実に想定されるような(非常に大きい有限の)観測頻度において、実際に統計理論が仮想的状況とほぼ同様に働くかどうかを検証しておく必要がある。このことの確認のために、計算機を用いたシミュレーションによって発生させた人工データに基づくモンテカルロ実験を行なった。また、提案した推定量たちは(ある漸近的性質を満たす範囲で任意に選択できる)チューニングパラメーターをいくつか含むが、これらを実際のデータに適用する際にどのように選べばよいかという点についても、この数値シミュレーションにおいて検証した。

#### 実際の金融高頻度データへの提案手法の適用

提案手法を実際の金融データに適用した際に妥当な結論を得られるかという点について検証するために、S&P500 指数とその先物・ETF 間の価格発見(リード・ラグ効果)の解析を行なった。この金融データを検証対象として採用した理由は、多くの先行研究において先物の価格発見が指数を先行しているという事実が指摘されているため、提案手法から得られる結論の妥当性を検証する上で都合がよいからである。それぞれのペアについて日次のスペクトル・リード・ラグ指数を計算してプロットしたグラフを図1に示す。xで示された値がスペクトル・リード・ラグ指数の計算値で、バーは95%信頼区間を表す。赤で表されている値は5%水準で統計的に有意であるような値である(95%信頼区間が0を含まないということ)。左2つの図は指数と先物・ETF間の解析結果であり、スペクトル・リード・ラグ指数がいずれの日も有意に負であることを示しており、これは指数が先物・ETFと比べて遅行していることを示している。最も右にある図は先物とETF間の解析結果であり、概して先物の方が先行しているという結果が出てはいるが、左2つの図と比べて差は明瞭ではない。これらの結果は先行研究の結論と比較して妥当なものとなっている。他方、図2はリード・ラグパラメーターを時間の関数として推定した際の、日内の時間的変化のプロットを示している(サンプル期間中での平均を取っている)。図から読み取れるように、リード・ラグパラメーターはU字型の形状をしている。これはボラティリティの時間変化をプロットした際に頻繁に観察される形状であり、金融市場における取引の活発さの度合いを反映していることが予想される。

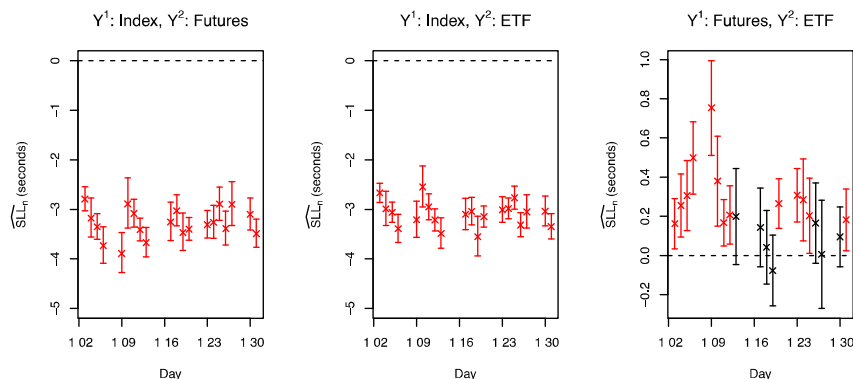


図1 S&P500 指数とその先物・ETF間のスペクトル・リード・ラグ指数

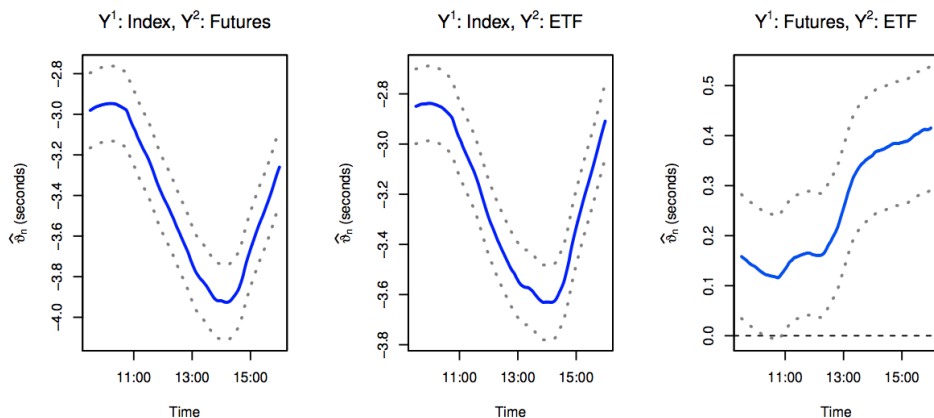


図2 S&P500 指数とその先物・ETF 間のリード・ラグパラメーターの日内変動

#### 提案推定量の予期せぬ振る舞いとその解決案の提示

本研究の成果をまとめた論文を査読付き雑誌に投稿したところ、査読者の一人から、推定量に含まれるあるパラメーターを変動させた際の振る舞いが、理論上予測される結果と実データで実際に観測される振る舞いとで齟齬があるという点を指摘された。実際、手元にある複数のデータで確認したところ、いずれのデータにおいても同様の現象が観察された。この現象を説明するために元の統計モデルの修正案を検討した結果、次の成果を得た。もしモデルにリード・ラグ効果をもつ成分とまたない成分とが共存し、リード・ラグ効果を持たない成分がより支配的である場合、上述の齟齬を理論的に説明できることがわかった。このような状況を考えることは、金融市場のファクター構造を考えるとごく自然であり、推定量の異常な振る舞いを説明する上で有力な候補である。しかし、この修正によって上の齟齬を完全に解消できるかどうかはまだ不明瞭であり、さらなる検証が必要である。

## 5 . 主な発表論文等

### 〔雑誌論文〕(計 3 件)

Yuta KOIKE, Zhi, LIU, “Asymptotic properties of the realized skewness and related statistics”, *Annals of the Institute of Statistical Mathematics* (査読有), in press.

DOI: 10.1007/s10463-018-0659-8

Yuta KOIKE, “Gaussian approximation of maxima of Wiener functionals and its application to high-frequency data”, *Annals of Statistics* (査読有), Vol. 47, No. 3, 2019, pp. 1663-1687.

DOI: 10.1214/18-AOS1731

Yuta KOIKE, “Mixed-normal limit theorems for multiple Skorohod integrals in high-dimensions, with application to realized covariance”, *Electronic Journal of Statistics* (査読有), Vol. 13, No. 1, 2019, pp. 1443-1522.

DOI: 10.1214/19-EJS1553

### 〔学会発表〕(計 24 件)

小池 祐太, “超高頻度データに対するリード・ラグ効果の推定について”, 慶應義塾大学経済研究所計量経済学ワークショップ, 2016 年.

Yuta KOIKE, “Statistical analysis of price discovery: a stochastic process approach”, IMS-APRM 2016, 2016 年.

小池 祐太, “高頻度時系列データに対するリード・ラグ効果の統計解析”, 2016 年度統計関連学会連合大会, 2016 年.

Yuta KOIKE, Takaki HAYASHI, “Wavelet-based methods for high-frequency lead-lag analysis”, *Stochastic Analysis and Statistics* 4, 2016 年.

小池 祐太, 吉田 朋広, “統計的リード・ラグ効果”, 高精度情報抽出のための統計理論・方法論とその応用, 2016 年.

Yuta KOIKE, Nakahiro YOSHIDA, “Statistical lead-lag effects”, TMU Workshop on Financial Mathematics and Statistics 2016, 2016 年.

Yuta KOIKE, “Time varying lead-lag effect”, CMStatistics 2016, 2016 年.

Yuta KOIKE, Nakahiro YOSHIDA, “Statistical lead-lag effects”, *Portfolio dynamics and limit order books*, 2016 年.

小池 祐太, 林 高樹 “高頻度データにおけるリード・ラグ効果のウェーブレット解析”, 応用統計ワークショップ, 2017 年.

Yuta KOIKE, “Capturing heterogeneous lead-lag effects from ultra high frequency data”, ASC2017, 2017 年.

Yuta KOIKE, “Capturing heterogeneous lead-lag effects from ultra high frequency data”, EcoSta 2017, 2017 年.

Yuta KOIKE, “Capturing heterogeneous lead-lag effects from ultra high frequency data”, SoFiE 2017, 2017 年.

Yuta KOIKE, Takaki HAYASHI, “Wavelet-based methods for high-frequency lead-lag analysis”, IAAE 2017, 2017 年.

Yuta KOIKE, “On the asymptotic structure of Brownian motions with a small lead-lag effect”, Asymptotic Statistics of Stochastic Processes and Applications XI, 2017 年.

小池 祐太, “Lead-lag analysis of non-synchronously observed time series with R”, 2017 年度統計関連学会連合大会, 2017 年.

Yuta KOIKE, “Lead-lag analysis of non-synchronously observed time series with R”, CMStatistics 2017, 2017 年.

Yuta KOIKE, “Testing the absence of lead-lag effects in high-frequency data”, EcoSta 2018, 2018 年.

Yuta KOIKE, “Gaussian approximation of maxima of Wiener functionals and its application to high-frequency data”, 10<sup>th</sup> World Congress of the Bachelier Finance Society, 2018 年.

小池 祐太, “On the asymptotic structure of Brownian motions with a small lead-lag effect”, 2018 年度統計関連学会連合大会, 2018 年.

小池 祐太, 林 高樹, “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”, 2018 年度中之島ワークショップ, 2018 年.

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕

- 統計ソフトウェア R のパッケージ `yuima` に、統計的に有意なリード・ラグ関係の存在を見するための関数 `llag.test` を実装した。
- 平成 30 年 9 月, 「On the asymptotic structure of Brownian motions with a small lead-lag effect」と題する研究論文に対して, 「第 3 2 回小川研究奨励賞」を受賞した。

## 6 . 研究組織

(1)研究分担者  
なし

(2)研究協力者  
なし

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。