

# 日本市場の注目業種クラスターとモーメンタム効果 —ホットスポットマイニングによる網羅的探索—

岡田 克彦  
羽室 行信

## 目 次

- |           |         |
|-----------|---------|
| 1. はじめに   | 3. 実証結果 |
| 2. 仮説と方法論 | 4. 結語   |

本稿では、統計的疫学調査で用いられる「空間スキャン統計量」という尤度比統計検定量を援用し、東証33業種において統計的有意に上昇している注目業種クラスターを特定した。この方法はホットスポットマイニングと呼ばれ、近年急速に発展したアルゴリズム技術によって超高速計算が可能となった。特定された業種クラスターネットワーク内においては、情報伝播速度の違いによるモーメンタムが存在し、年率換算10%程度のインパクトを持つことが分かった。

## 1. はじめに

米国においてはモーメンタム効果が顕著に観察されることはよく知られている。Jagadeesh and Titman [1993] は、過去3カ月から12カ月間のクロスセクションの収益率で上位10%の銘柄群を購入し、下位10%の銘柄群を空売りすることで、

平均年率12%の超過収益率が得られることを報告した。こうした効果は世界の多くの株式市場で観察されている。Rouwenhorst [1998] はモーメンタム効果を調査した欧州の12カ国全てにおいて、1980年から1995年の間に年率平均12%と、米国と同様の超過収益率が得られたことを報告している。Moskowitz and Grinblatt [1999] は、



**岡田 克彦 (おかだ かつひこ)**

関西学院大学大学院経営戦略研究科教授。神戸大学 博士 (経営学)。モルガン・スタンレー証券、UBS証券、ヘッジファンド運用会社Halberdier Capital Singapore Pte.を共同創業後、2005年からファイナンス研究者。12年にAI研究者らとともに機械学習を用いた解析で投資助言を行う(株)Magne-Max Capital Managementを創業。現在CEO / CIOを兼務。専門は行動ファイナンス。行動経済学会常任理事。



**羽室 行信 (はむろ ゆきのぶ)**

関西学院大学大学院経営戦略研究科准教授。データ解析ツール開発会社(株)NYSOL会長、(株)Magne-Max Capital Management顧問を兼務。米国保健衛生局 (NIH)、ERATO湊離散構造処理系プロジェクト、CRESTデータ粒子化プロジェクトなどの大型国家研究プロジェクトにおいて、データマイニング・機械学習応用に関する研究実績は多岐にわたる。企業との共同研究実績、国内外論文、国際学会発表、学会賞受賞多数。

モーメンタム効果の大部分は業種モーメンタムで説明できることを発見し、個別銘柄の過去収益率を基準にするより、業種分類に基づき、上位業種クラスターをロングし、下位業種クラスターをショートする戦略の利得が大きいと報告している。とりわけ短期の業種モーメンタム（1カ月以内）は強く、米国において1963年から1995年の期間で月次で105bpの効果を持つという。

このように、単純に過去収益率だけを基準に選択することで、世界の主要な株式市場においてモーメンタムする銘柄群を特定できるのだが、日本市場だけは例外であることが知られている（Iihara *et al.* [2004]）。先進国市場の中で、時価総額も大きく国際的に開かれた日本市場において、単純なモーメンタム効果が見られないことは多くの研究者を困惑させる。モーメンタム効果をリスク、あるいは投資家バイアスのいずれかで説明しようとする様々な議論も、エビデンスに一貫性がないことで止まってしまい、日本市場を常に例外として扱わなければならないからである。

そもそもモーメンタム効果とは、過去に良い（悪い）リターンを示した銘柄（業種）は、未来にも良い（悪い）リターンを示すという効果である。したがって、過去の情報は市場価格に反映されていると考えるウィークフォームの市場効率性の反証となってきた。この原因として、Hong *et al.* [2000] は、投資家に注目されていない銘柄は情報を十分に織り込めないからだと主張している。とりわけ、低注目銘柄への情報伝播速度は遅く、悪材料も織り込まれるのに時間がかかるという。Cohen and Frazzini [2008] も、投資家の注意力の限界により情報伝播に時間がかかり、これがモーメンタム効果を生み出していると主張している。彼らは企業間取引関係に注目し、取引関

係ネットワーク内で、カスタマー企業からサプライヤー企業へのモーメンタム効果を報告している。過去収益率の高い（低い）カスタマー企業の良い（悪い）情報は、取引関係ネットワーク内に存在する他のサプライヤー企業に遅く伝播し、モーメンタム効果を生むという。そのインパクトとして、前月の収益率の高い企業を顧客に持つサプライヤー企業群をロングし、低い企業を顧客に持つサプライヤー企業群をショートすることで、1984年から2004年の期間の米国市場で月次平均158bpの超過収益率が得られるという。この現象をカスタマーモーメンタムと呼ぶが、磯貝ほか [2019] や羽室・岡田 [2018] はカスタマーモーメンタムが日本市場においても存在することを確認している。

本稿では、注目業種クラスターというネットワークに着眼する。取引関係ネットワークのようにビジネス上の結びつきではなく、投資家の注目という意味における銘柄間の結びつきを、業種分類を用いてネットワーク表現しようという試みである。投資家の注目業種クラスターを、ホットスポットマイニングという手法で検知したものと定義し、その業種クラスターネットワーク内にモーメンタムが発生するかどうかを検証する。

検証の結果、注目業種クラスター内では、先の上昇した銘柄群に追随する形で出遅れ銘柄群が上昇するというモーメンタム効果が観察され、出遅れポートフォリオを取引することで月換算約80bpの $\alpha$ が得られることが明らかとなった。これは、注目業種クラスターという注目銘柄の所属業種群で結ばれた銘柄間に存在するネットワークを通じた情報伝播仮説と整合的な結果である。

以下、第2章では仮説と方法論、とりわけ空間スキャン統計量とそれを用いたホットスポットマイニングの検証方法について説明する。第3章で

は実証結果を示し、ホットスポットマイニングで特定できる業種クラスターと、単純な上昇業種の組み合わせの差異を比較する。第4章では結論を述べる。

## 2. 仮説と方法論

### (1) 業種クラスターネットワークを通じた情報伝播仮説

企業間には様々なネットワークが存在する。Cohen and Frazzini [2008] が扱ったサプライヤーとカスタマーという契約関係に基づくネットワークもあれば、Cohen and Lou [2012] らが着眼した、コングロマリット企業と専門企業のネットワークも存在する。取引関係のように企業間に契約があるわけではないが、コングロマリット企業と専門企業間には部分的には同じビジネス領域のプレーヤーであり、投資家視点に立てば両社間にはつながりがある。専門企業で起こったことは、そのネットワークを通じてコングロマリットに波及し、そこに情報伝播速度の相違があるためにモーメンタム効果が生まれる。Cohen and Lou [2012] では1977年から2009年の期間で平均月次で112bpのモーメンタム効果を確認している。

本稿では、正の超過収益率を示す銘柄が集中する業種クラスターを探索し、業種クラスター内にある銘柄間にネットワークが存在するとの仮説を検証する。正の超過収益率を示す銘柄が多い業種クラスターには、良い情報が入っていると考えられる。その良い情報は、業種クラスターネットワーク内に存在する企業群に遅れて伝播するとすれば、そのクラスター内に存在する企業群の中で、まだ十分に企業価値が上昇していない銘柄を保有することで、超過リターンが得られるはずである。こうした業種クラスター内におけるモーメンタム

効果の有無を検証する。

### (2) 空間スキャン統計量

注目業種クラスター（単業種だけでなく複数業種の組み合わせも含まれる）は、ホットスポットマイニングという手法で探索する。統計的疫学調査の研究分野では、病気や事故の原因を探るために発生頻度の高い領域（ホットスポット）を検出することが重要な問題となっている。こうした取組みをホットスポットマイニングと呼んでいる。空間スキャン統計量とは、ある領域内に起きた現象が偶然によるものか否かを検定し、発生頻度が有意に高いホットスポットを検出するための尤度比検定統計量である。本稿では疫学調査における「領域」を「業種」に、「疫病の発生」を「超過収益率の発生」に置き換えている。業種クラスター内に起きている超過収益率の発生が偶然によるものか否かを検定しながら、ホットスポットを特定する。

この問題に対して、Kulldorff [1997] は正の超過収益率の発生（疫病）の発生確率にベルヌーイ分布を仮定した以下に示すような方法を考えた。いま、対象とする全ての業種集合を $G$ 、その部分集合である業種クラスターを $Z \subset G$ とする。また $Z^c \cup Z = G, Z^c \cap Z = \emptyset$ とする。 $Z$ を構成する個別銘柄が正の超過収益率を示す確率を $p$ 、また $Z^c$ での確率を $q$ とすると、 $Z$ がホットスポットになるか否かを検定する帰無仮説は、 $H_0: p=q$ 、対立仮説は $H_1: p>q$ となる。ここで、集合 $Z$ のサイズ（すなわち業種クラスター $Z$ を構成する銘柄数）を $n(Z)$ 、そして、ある正の超過収益率を示す銘柄の数を $c(Z)$ で表す。 $Z$ と $Z^c$ における正の超過収益率の発生は全てお互いに独立だとすると、帰無仮説 $H_0$ の下での尤度関数は $L_0 = p^{c(Z)} (1-p)^{n(Z)-c(Z)}$ で表され、また対立仮説 $H_1$ の下で、ある $Z$ の尤度関数

は $L(Z) = p_z^{c(Z)}(1-p_z)^{n(Z)-c(Z)} p_{z^c}^{c(Z^c)}(1-p_{z^c})^{n(Z^c)-c(Z^c)}$ と表される。そして $p, q$ に最尤推定量 $c(Z)/n(Z)$ 、 $c(Z^c)/n(Z^c)$ を当てはめ、比をとった(1)式がベルヌーイモデルにおける尤度比検定統計量 $\lambda(Z)$ で、最適な $Z$ の探索という文脈の中で特に空間スキャン統計量 (spatial scan statistic) と呼ばれる。

$$\lambda(Z) = \frac{L(Z)}{L_0} = \frac{\left(\frac{c(Z)}{n(Z)}\right)^{c(Z)} \left(1 - \frac{c(Z)}{n(Z)}\right)^{n(Z)-c(Z)} \left(\frac{c(Z^c)}{n(Z^c)}\right)^{c(Z^c)} \left(1 - \frac{c(Z^c)}{n(Z^c)}\right)^{n(Z^c)-c(Z^c)}}{\left(\frac{c(G)}{n(G)}\right)^{c(G)} \left(1 - \frac{c(G)}{n(G)}\right)^{n(G)-c(G)}} \quad (1)$$

スキャン統計量最大の業種クラスター (ホットスポット) を検出するためには、全業種 $G$ に含まれる全ての部分集合である業種クラスター $Z$ を求める必要がある。これは膨大な数の組み合わせである。東証33業種分類を使うとすれば約43億通りのスキャンを行う必要があり、全業種で全探索するには、現実的な計算時間では対応できない(注1)。これまで、最適な $Z$ 、すなわちスキャン統計量を最も高くするような部分集合 $Z$ の探索法は知られておらず、近似的な方法が用いられてきた。例えば空間的同心円状に連結された領域のみを探索するなど、探索対象を限定する中での極大解を探索する方法である(注2)。しかし、近年、Minato *et al.* [2019] により全探索による最適解を高速に求める方法が提案された。この方法は、スキャン統計量が、利用者が指定した閾値を上回るような部分集合 $Z$ を全列挙するもので、スキャン統計量が $c(Z)$ 及び $n(Z)$ に対して部分的に単調性を持つという性質を巧みに利用したアルゴリズム

であり、うまくいけば解の個数にほぼ比例する計算時間で探索が終了する(注3)。

以上の方法で $\lambda(Z)$ が最大のホットスポットを探索可能になるが、そのスキャン統計量が統計的に有意であるかを検定する必要がある。同じデータに対して多数の仮説(多数の $Z$ )を試して最も良いものを選ぶ場合には検定の多重性の問題が生じるので、ここではモンテカルロ法を用いている。モンテカルロ法とは、帰無仮説の下でランダムにデータを繰り返し生成し、観測データの検定統計量が、ランダム標本の上位何%に入っているかで $p$ 値を計算する方法である。本稿では有意水準を5%に設定し、ホットスポットの中でも有意なもののみを採用した。

### (3) 空間スキャン統計量の応用

統計的疫学調査において、空間スキャン統計量は伝染病の発生頻度の高い地域を都道府県の区画を用いた組み合わせで特定する際に活用されている。本稿では、ある特定の業種クラスターの株価が有意に上昇しているか否かを判定する手法として用いる。伝染病への罹患人数に該当する基準として、業種内において正の超過収益率を示す銘柄数を用いる。都道府県の人口比で罹患人数が多い地域には伝染病の蔓延が疑われるのと同様に、業種構成銘柄数に対して正の超過収益率を示す銘柄数が多い業種クラスターに投資家の関心が高まっていると考えられる。図表1に統計的疫学調査との比較をまとめておく。

対象銘柄は東京証券取引所第一部に上場する全

(注1) 33業種を用いた場合の、 $Z$ と $Z^c$ の組み合わせは $\sum_{i=1}^{32} \frac{33!}{i(33-i)!} \times \frac{1}{2} = 4,294,967,295$ 通りとなる。

(注2) 業種分類の場合は都道府県や州による区画などと異なり、隣接関係に意味がないため、この手法は援用できない。

(注3) 計算量は区画数により組み合わせ爆発が発生するため急速に拡大する。本稿の33業種の場合、全計算と比較すると約11,000倍の速度で計算可能となる (Minato *et al.* [2019] Table 3を参照)。

図表1 ホットスポットマイニング手法の援用

	統計的疫学調査	本研究
目的	病気の伝染領域特定	資金流入領域の特定
区画	47 都道府県	東証33業種分類
組み合わせ数	70兆3,687億	43億
観察事象	罹患率	超過収益発生率

(出所) 筆者作成 (以下、同じ)

銘柄であり、超過収益率は東証株価指数 (TOPIX) との収益率の差分として定義する。閾値は任意に5%と設定した。したがって、前日の超過収益率が+5%以上ある銘柄数を日々カウントする。例えば、ある日において、5%以上の正の超過収益率が発生している銘柄群が業種Aと業種Bに集中しているとしよう。その日はAとBから構成される業種クラスターがホットスポット業種クラスターとして認識されることになる。

#### (4) モンテカルロ法による有意性判定とカレンダータイムポートフォリオの構築

Minato *et al.* [2019] の手法を援用し、43億通りの組み合わせの中で(1)式の $\lambda$ を最大化するような業種部分集合Zを求め、それが統計的に有意であるかどうかを検定するためにモンテカルロ法を用いるが、具体的には以下のような方法を実施した。各営業日における全体の超過収益率の発生率 ( $c(G)/n(G)$ ) に基づき、その発生率が期待値となるようにランダムに1,000回各業種の銘柄に超過収益率を発生させる。この1,000回の疑似的マーケットのそれぞれについて、 $\lambda$ が最大となる組み合わせを求める。ここで求められた1,000個の $\lambda$ の中で、観察されている $\lambda$ が上位50番目以内(5%有意に該当)であれば、ホットスポットだと認定する。ホットスポットか否かの判断は毎日

行う(注4)。

次に、ホットスポットだと認定された営業日においては、ホットスポットとなった業種クラスターに属する全銘柄を当日の収益率によって5分位に分割する。第1分位が最も収益率の小さい企業群であり、第5分位は収益率が最も高いものである。これら五つの分位ポートフォリオについて、ホットスポット発生直後からその収益率を観察する。本稿では、当日の収益率の最も低い第1分位ポートフォリオと、最も高い第5分位ポートフォリオの二つのポートフォリオの検証結果のみを示す。

#### (5) カレンダータイムポートフォリオの運用方法

上昇銘柄の多い業種クラスターにおいて、その後の株価収益率の予測可能性はあるのだろうか。ホットスポットとなった業種クラスターを、投資家が注目する業種ネットワークと考えた場合、そのネットワーク内でモーメンタム効果が見られるかもしれない。これについて、以下の検証方法で調査する。まず、時点t日にホットスポットとなった業種クラスターをユニバースとし、ユニバース内のt日の収益率で5分位に分割する。そして、第1分位と第5分位のそれぞれの分位ポートフォリオにおいてt+1日の引値から一定期間保有するカレンダータイムポートフォリオを作成する。ホットスポットは毎日探索されるため、連続してホットスポットとなる業種クラスターが見つかる場合がある。その場合、同じ業種クラスターから連続で銘柄が選択されることになるため、同じ銘柄を複数単位所有することになり、重なりのある銘柄のウェイトは高まることになる。どの程度の期間ウェイトが高まるかは、購入した銘柄群を何日

(注4) 計算量としては、43億通りの組み合わせ×1,000回のシミュレーション×観察日数となる。

間保有するかに依存する。本稿の実験では銘柄の保有期間 (Hp) を、2 営業日、4 営業日、5 営業日の 3 パターンに分けて行った。そのイメージを図表 2 に示しておく。

保有期間が 2 日 (Hp=2days) の場合は、ホットスポットが検知された t 日の翌日の引値で購入し、t+3 日の引値で売却する。今回の実験では、ポートフォリオの回転率が最も高くなる戦略である。ここで示している銘柄は分位ポートフォリオの構成銘柄である。Hp が 4 日の場合は、4 日間銘柄を保持し、t+5 日の引値で売却する。同様に、Hp が 1 週間 (5 日) の場合は、t+6 日の引値で売却する。図表 2 の例では、簡単化のため、t+4 日、t+5 日、t+6 日にはホットスポットは検知されていないものとして記載した。Hp が長く

なるほど、回転率は下がり、ポートフォリオに組み入れられる銘柄数は増加する。ただ、業種クラスター内のモーメンタムが短期間で終わってしまうのであれば、回転率が高いポートフォリオの方が成績は良くなるはずである。第 1 分位と第 5 分位をそれぞれ三つの異なる保有期間で作成した六つのカレンダータイムポートフォリオを、Fama-French の 3Factor Model に従って  $\alpha$  を計算する (注 5)。具体的には (2) 式で時系列に回帰し、 $\hat{\alpha} \hat{\beta} \hat{\gamma} \hat{\delta}$  を求めた。

$$E(r_p) = \alpha_p + \beta_p(r_m - r_f) + \gamma_p(SMB) + \delta_p(HML) + \varepsilon \quad (2)$$

図表 2 ホットスポットが特定された日に組成される保有期間別時系列ポートフォリオの構成イメージ

保有期間 2 日 (Hp=2days)					
t+1	t+2	t+3	t+4	t+5	t+6
銘柄 1 購入		銘柄 1 売却			
銘柄 2 購入		銘柄 2 売却			
銘柄 3 購入		銘柄 3 売却			
	銘柄 4 購入		銘柄 4 売却		
	銘柄 5 購入		銘柄 5 売却		
		銘柄 6 購入		銘柄 6 売却	
保有期間 4 日 (Hp=4days)					
t+1	t+2	t+3	t+4	t+5	t+6
銘柄 1 購入				銘柄 1 売却	
銘柄 2 購入				銘柄 2 売却	
銘柄 3 購入				銘柄 3 売却	
	銘柄 4 購入				銘柄 4 売却
	銘柄 5 購入				銘柄 5 売却
		銘柄 6 購入			(t+7 で売却)
保有期間 5 日 (Hp=1wk)					
t+1	t+2	t+3	t+4	t+5	t+6
銘柄 1 購入					銘柄 1 売却
銘柄 2 購入					銘柄 2 売却
銘柄 3 購入					銘柄 3 売却
	銘柄 4 購入				(t+7 で売却)
	銘柄 5 購入				(t+7 で売却)
		銘柄 6 購入			(t+8 で売却)

(図表注) 簡単化のため、t+4、t+5、t+6 ではホットスポットは存在しなかったとする。

(注 5) 3Factor Model の日本市場の値については、Kenneth French のデータライブラリーから取得した。  
[https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)

### 3. 実証結果

#### (1) ホットスポット認定された業種クラスタの分布

本稿の検証期間は2008年から18年9月末までであるが、紙幅の関係で図表3には18年7月20日から8月30日までのホットスポットを記載している。濃い色で示されているのがホットスポットとなった業種クラスタ、つまり(1)式のλが最大となり、かつモンテカルロ法で5%有意だと判断された業種クラスタである。

小さくしか描けていないが、図表3では日付、東証33業種が水平に、二つのブロックに分かれて並んでいると解釈いただきたい。左側の区画に33業種、右側の区画にも33業種、1業種が小さ

い枠で表されている。左側にはホットスポットと認知された業種クラスタを示している。色が濃い枠は、その日にホットスポット業種クラスタとして認定された業種群である。一方、右側の区画は33業種を単独で観察し、正の超過収益率を示す銘柄数が10%以上になった業種を色が濃い枠として示したものである。各業種を小さなセルで表現しているため、どの枠がどの業種に当たるのかはこの図表からは読み取れないが、ホットスポットマイニングを行って認識する業種クラスタと、業種を単独に観察した場合では全く異なることが理解できるだろう(注6)。

例えば、18年7月31日であれば、単純に業種のみを観察して正の超過収益率の発生比率を10%で判断すると「海運」業種のみがホットだ

図表3 2018年7月20日～8月30日の期間におけるホットスポットの分布



(図表注) 東証33業種を「水産農林業」からサービス業までを左から順番に並べたものを二つのブロックに分けて表示している。 $n_i$ は1業種に含まれる銘柄数であり、 $c_i$ は其中で超過リターンを5%以上示した銘柄数である。

(注6) 閾値を10%と20%で実験したが、ホットスポットマイニングで選択される業種クラスタとは大きく異なった。

ったということになる。しかし、業種の組み合わせを考慮した業種クラスターで考えた場合、「海運」を含む8業種に正の超過収益率を示す銘柄が偏っていることが分かる。

本手法のもう一つの特徴は、単独の業種で考える場合のように10%という任意の閾値を設ける必要がないことである。(1)式の $\lambda$ を最大にする業種クラスターを探索し、そのクラスターがモンテカルロ法で統計的に有意な偏りがあるかを判断するだけで列挙可能である。

## (2) ホットスポット業種クラスターにおけるモーメンタム効果の検証結果

業種クラスターネットワーク内をユニバースとした五つの分位ポートフォリオのうち、第1分位と第5分位についてのみ報告する。それぞれ保有期間(Hp)を2営業日、4営業日、5営業日(1wk)として観察した計六つのカレンダータイムポートフォリオを、(2)式で回帰し得られた $\hat{\alpha}$   $\hat{\beta}$   $\hat{\gamma}$   $\hat{\delta}$ を図表4にまとめている。

検証期間内に選択される銘柄数は第1分位で73,869銘柄、第5分位で77,981銘柄であった。Hp=2daysとすると平均で1日当たり約57.5銘柄

選択されていることになる。注目されるのは、第1分位における取引については保有期間にかかわらずおおむね有意な $\alpha$ (Hp=1wk以外では全て信頼区間5%で有意)が観察される点である。一方、第5分位では $\alpha=0$ の帰無仮説を棄却できない。ホットスポットとなっている業種クラスター内でも、当日に注目され上昇した銘柄( $c(Z)$ にカウントされる銘柄)を含む第5分位のポートフォリオよりも、上昇銘柄数は少ないものの、業種クラスターネットワーク内にある第1分位のポートフォリオ、すなわち、当日のパフォーマンスが相対的に悪い分位ポートフォリオにおいて、有意に正の $\alpha$ が観測される。

ここから明らかのように、 $\beta$ 、規模、簿価時価比率をコントロールしても、保有期間(Hp)2日と4日の場合において、年率で10%の $\alpha$ が獲得できている。この結果は、Cohen and Frazzini [2008] や磯貝ほか [2019]、羽室・岡田 [2018] のサプライヤー・カスタマーネットワーク上に存在したモーメンタムと同様、業種クラスターネットワーク上に存在するモーメンタムの可能性を示唆している。情報が業種クラスターネットワークを通じて伝播するというメカニズムが働いている

図表4 ホットスポット業種クラスター内の分位ポートフォリオ別パフォーマンス

		Hp=2days		Hp=4days		Hp=1wk		
		n	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
第1分位	$\alpha$		0.0004	0.0130	0.0004	0.0390	0.0003	0.0650
	Rm-rf	73,869	0.5000	0.0000	0.6630	0.0000	0.7180	0.0000
	SMB		0.1900	0.0000	0.3140	0.0000	0.3630	0.0000
	HML		0.0400	0.2146	0.0675	0.0672	0.1107	0.0022
第5分位	$\alpha$		0.0002	0.2780	0.0001	0.7020	0.0000	0.9090
	Rm-rf	77,981	0.4810	0.0000	0.5530	0.0000	0.7310	0.0000
	SMB		0.0960	0.0052	0.2390	0.0000	0.3250	0.0000
	HML		0.0650	0.1100	0.1112	0.0081	0.1898	0.0000

(図表注) 時点tにおいてホットスポットと判定された業種群の組み合わせを対象とし、当該業種群の構成銘柄の中で、収益率の上位20% (第5分位) 及び下位20% (第1分位) の銘柄をt+1日に購入し、Hp日間保有した場合のカレンダータイムポートフォリオ収益率をFama-Frenchの3 Factor Modelで評価した。太字は信頼区間5%で有意な水準を示す。nは検証期間中選択された銘柄総数を示す。

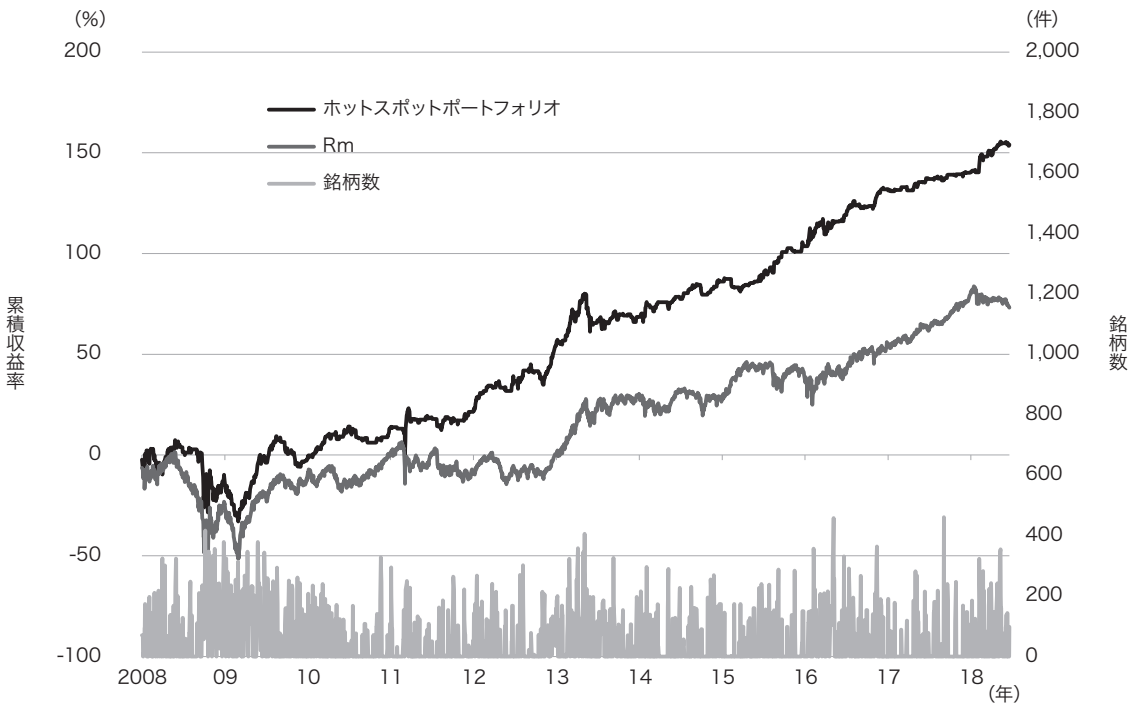


とすると、保有期間 (Hp) が4営業日のものでも有意になっていることから、出遅れ銘柄群への注目は約1週間弱のタイムラグをもって継続していると解釈できる。

次に、業種クラスターによるモーメンタム効果が、一時期に集中することにより統計的に有意な $\alpha$ をもたらしている可能性について確認しておく。**図表5**はカレンダータイムポートフォリオの累積収益率とマーケットの累積収益率と比較したものである。**図表5**に描いたポートフォリオ組成は、保有期間 (Hp) を2日としたものを示しているが、4日でも5日でも大きな差異はない。比較のためのマーケットポートフォリオ (Rm) で運用した場合の累積収益率は、Kenneth Frenchのウェブ

サイトから取得したものである。また、時系列に変化する保有銘柄数の推移は、**図表5**下部の棒グラフで示しており、本検証期間では平均57銘柄程度保有していた。保有銘柄数がゼロとなっている日は、保有期間の2日を超えてホットスポットが発生しなかった時期である。**図表5**から明らかなように、特定の時期の異常値により大きな $\alpha$ が検出されているのではなく、業種クラスターネットワーク上においては継続的にモーメンタム効果が存在し、その効果を利用することで $\alpha$ を達成する可能性があることが示唆された。

図表5 ホットスポット検知による時系列ポートフォリオと銘柄数推移 (Hp=2days)



(図表注) RmはKenneth Frenchのウェブサイトに掲載されているマーケットポートフォリオの収益率を2008年1月から累積した。ホットスポットポートフォリオは保有期間を2日として計算した時系列ポートフォリオの収益率を累積した。銘柄数(右目盛)は各時点における構成ポートフォリオの保有銘柄数を示す。

### (3) 注目業種の単純な組み合わせによるモーメントの有無

最後に、ホットスポットマイニングを用いて特定される業種クラスターというネットワークは、業種だけを観察することでは特定できない点について確認しておく。業種の組み合わせを考慮しながら尤度比を計算せずとも、単独業種を観察することで特定できるのであれば、ホットスポットマイニングを行う必要はない。そこで、単独業種を観察することでネットワークの特定を試み、 $\alpha$ が得られるかどうかを検証する。

そのため、任意の閾値を10%として、業種の構成銘柄のうち1割以上の銘柄数が5%以上の正の超過収益率を示すのであれば、注目業種であると仮定する。そして、各営業日に注目業種内でネットワークが形成されていると考える。図表3における右側の区画において濃い色の柁として示されている業種の単純な組み合わせである。そして、注目業種群をユニバースとして、当日の収益率に基づいて五つの分位ポートフォリオを考え、前節で行ったと同様に(2)式で回帰する。その結果得られた $\hat{\alpha} \hat{\beta} \hat{\gamma} \hat{\delta}$ を図表6にまとめている。

ここから明らかのように、保有期間にかかわら

ず全てのカレンダータイムポートフォリオについて有意な $\alpha$ は観察されていない。また、紙幅の関係上掲載していないが、超過収益率の定義を3%に変更した場合や、任意の閾値を20%と注目業種の基準をより厳しくした場合の実験も行ったが、有意な $\alpha$ は観察されず、本結果がパラメータに依存的でないことを確認している。このことから、単純に正の超過収益率が集中する業種という区分に投資家が注目しているのではなく、投資家は別の視点で銘柄選別を行っており、その一端をホットスポットマイニングによる業種クラスターが捉えている可能性がある。

## 4. 結語

日本の株式市場には、過去に高いリターンを示した株式がその後も高いリターンを示すという単純なモーメント効果は存在しない。しかし、カスタマーモーメントのようにネットワーク上でつながった株式間においては、カスタマー企業の情報がサプライヤー企業に遅れて伝播し、モーメント効果が見られることは確認されている。本稿では、取引関係ネットワークのように明確に定

図表6 単業種のみを用いた場合の分位ポートフォリオ別パフォーマンス

	n	Hp=2days		Hp=4days		Hp=1wk	
		Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
第1分位	$\alpha$	0.0002	0.5542	0.0003	0.3907	0.0002	0.5110
	Rm-rf	0.6664	0.0000	0.8221	0.0000	0.8753	0.0000
	SMB	0.3318	0.0000	0.4012	0.0000	0.4746	0.0000
	HML	0.2509	0.0003	0.3011	0.0000	0.2850	0.0000
第5分位	$\alpha$	-0.0003	0.3750	-0.0003	0.4498	-0.0003	0.4105
	Rm-rf	0.6833	0.0000	0.8527	0.0000	0.9343	0.0000
	SMB	0.2526	0.0002	0.4126	0.0000	0.5591	0.0000
	HML	0.3112	0.0001	0.3266	0.0000	0.4661	0.0000

(図表注) 時点tで単業種内の構成銘柄の10%以上が超過収益率を生んでいる場合、当該業種構成銘柄のt時点の収益率の上位20% (第5分位) 及び下位20% (第1分位) をt+1日に購入し、Hp日間保有した場合のカレンダータイムポートフォリオ収益率をFama-Frenchの3 Factor Modelで評価した。太字は信頼区間5%で有意な水準を示す。nは検証期間中に選択された銘柄総数を示す。

義される関係性ではないが、投資家が注目し超過収益率が発生している銘柄群と、それらの所属業種ネットワークでつながっている出遅れ銘柄群を特定することで、時間差をおいて情報が伝播し、モーメンタム効果が見られるのではないかと着想した。

このネットワークの探索のために、ホットスポットマイニングを高速で解く最新のアルゴリズムを実装し、投資家の注目度が高まっている業種クラスターを日次で特定した。そして、特定した注目業種クラスターをユニバースとしてポートフォリオを構築した。その結果、当該ユニバースの出遅れ銘柄群に投資することで、 $\beta$ や、規模、簿価時価比率をコントロールした後でも年率換算10%の $\alpha$ が得られることが明らかになった。これは、業種を区画として実施したホットスポットマイニングが探索した業種クラスターネットワークが、何らかの銘柄間のつながりをとらえていることと整合的である。

本研究は、CRESTデータ粒子化プロジェクト（宇野毅明代表）の資金援助を受けています。使用したアルゴリズムの実装については京都大学湊真一教授にご指導いただきました。また、匿名のレフェリー2名の先生から貴重なコメントをいただきました。ここに記して深謝申し上げます。

〔参考文献〕

- 磯貝明文・川口宗紀・小林寛司 [2019] 「サプライヤー・カスタマーのつながりに基づくクロスモメンタムの株価予測可能性」、『現代ファイナンス』40、25-48ページ。
- 羽室行信・岡田克彦 [2018] 「情報伝播速度の相違が生み出す投資機会：サプライチェーン・ネットワークを利用した個別銘柄投資」、『人工知能基本問題研究会SIG-FPAI』105、40-45ページ、2018-01-28。
- Cohen, L. and Frazzini [2008] “Economic links and predictable returns,” *Journal of Finance* 104, pp.383-400.
- Cohen, L. and D. Lou [2012] “Complicated firms,” *Journal of Financial Economics* 104, pp.383-400.
- Hong, H., T. Lim and J. Stein [2000] “Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies,” *Journal of Finance* 55, pp.265-295.
- Iihara, Y., H. Kato and T. Tokunaga [2004] “The winner-loser effect in Japanese stock returns,” *Japan and the World Economy* 16, pp.471-485.
- Jagadeesh, N. and S. Titman [1993] “Returns to buying winners and selling losers: Implications for stock market efficiency,” *Journal of Finance* 48, pp.65-91.
- Kulldorff, M. [1997] “A spatial scan statistic,” *Communications in Statistics, Theory and Methods* 26, pp.1481-1496.
- Minato, S., J. Kawahara, F. Ishioka, M. Mizuta and K. Kurihara [2019] “A fast algorithm for combinational hotspot mining based on spatial scan statistic,” *SIAM International conference on Data Mining, SDM 2019*, pp.91-99.
- Moskowitz, T. and M. Grinblatt [1999] “Do industries explain momentum?” *Journal of Finance* 54, pp.1249-1290.
- Rouwenhorst, G. [1998] “International momentum strategies,” *Journal of Finance* 53, pp.267-284.
- (この論文は投稿論稿を採用したものです。)