

負け犬は誰だ？ 証券アナリストの格下げにより価値を失う企業の特徴について

Who is the Loser?: Stock Performance after Securities Analyst's Rating Downgrades – A Data Mining Approach –

岡田 克彦^{1*} 中元 政一² 東 高宏² 羽室 行信¹
K. Okada¹ S. Nakamoto² T. Azuma² Y. Hamuro¹

¹ 関西学院大学経営戦略研究科

¹ Institute of Business and Accounting

² 科学技術振興機構 湊離散構造処理系プロジェクト

² JST Minato Discrete Structure Manipulation Project

Abstract: We investigate characteristics of firms that lose market value in the post-downgrade period by securities analyst. We found higher pre-downgrade volatility is strongly associated with the negative return in the post-downgrade period. Among high volatility firms, small capitalization stocks and stocks with inferior sentiment are more likely to underperform.

1 はじめに

米国におけるアナリスト評価に関する研究では、株式推奨と株価の関係について網羅的に調査している研究が幾つか存在する。例えば、Stickel(1995)やWomack(1996)は、証券アナリストが個別株式に対して格上げレポートを書いた場合、発表後の正の株価効果は約1ヶ月持続するとしている[14][15]。一方、格下げレポートを書いた場合、発表後の負の株価効果は約6ヶ月間持続するとしている。また、太田・近藤(2010)は日本のサンプルで同様の検証を行い、短期的には格上げの場合も格下げの場合もほぼ対称的な株価変動している事、6ヶ月の期間で観察した場合、格下げのインパクトの方が大きい事を報告している[11]。Barber et. al(2001)は、株価反応を見るだけでなく、投資家の立場からアナリスト評価を運用に活かせるか否かを検討した[2]。その結果、証券アナリストの強気なコンセンサスに基づいてポートフォリオを構築した場合には、取引コストで超過リターンは消えてしまうが、弱気なコンセンサスに基づいて構築したポートフォリオは取引コストを勘案しても大きな超過リターンが得られると報告している。また、アナリストのコンセンサスが市場に伝わってからの投資家行動についても、弱気コンセンサスについては、行動が1ヶ月遅れたとしても、十分に負の超過リターンが得られるとしている。

では、なぜ弱気のコンセンサスに依拠した戦略の方が、超過リターンは大きくなるのだろうか。Miller(1977)はマーケットには多くの投資家が存在するので、理論的にはファンダメンタル価値から乖離するような価格は成立しないが、現実には空売り制約があるために、市場では銘柄が過大評価されるという形でミスプライスが発生すると言う[8]。そうであるならば、弱気コンセンサスに基づいた取引戦略が有効だとするBarber et al(2001)の報告[2]と整合的である。

本研究の目的は、過大評価される株式にはどのような特徴があるかについて、探索的アプローチを用いて、明らかにすることである。多くの先行研究によるイベントスタディの結果、アナリストの格下げについてはある程度の予測能力があることが明らかになっているが、ある特定の特徴を持つ銘柄群については、顕著に予測能力が高くなる場合や、別の特徴を持つ銘柄群については、予測能力が低くなる場合があるかもしれない。証券アナリストの格下げによって価値を大きく失う企業の特徴を、イベント発生前のリターン、変動率、流動性、出来高、業種、時価簿価比率、規模、モメンタム、アクルーアル、対象銘柄のセンチメント、属する業種のセンチメント、マーケット全体のセンチメント、アナリスト所属証券会社の評判、証券会社の規模、評価変動後の絶対評価水準の15の変数を手掛かりに探索する。

*連絡先：関西学院大学経営戦略研究科
〒662-8501 兵庫県西宮市上ヶ原一番町 1-155
E-mail: katsuoakada@kwansai.ac.jp

2 データ

本研究が対象とするのは、2000年1月から2011年5月までに発生した証券アナリストによるレーティングの格下げ（評価の下方修正）が行われた5365銘柄である。証券会社毎に様々なレーティングの表示方法が存在するが、証券アナリスト達の評価を筆者らは次の5段階に分類した。

- -2: 株価指数を大きく下回るという評価
- -1: 株価指数を下回るという評価
- 0: 株価指数と同等のパフォーマンスを示すだろうという評価
- +1: 株価指数を上回るという評価
- +2: 株価指数を大きく上回るだろうという評価

対象銘柄は、証券アナリストによって負の評価変動が起きた銘柄であり、絶対的な評価水準ではない。したがって、+2→+1、+1→0、0→-1、-1→-2、という変化の対象となった銘柄ということである。評価変動が発生した時点の翌日から44日後までのすべての期間についての累積株価リターンを計算し、それらを目的変数とする。

筆者らは以下の4つのファクターを単独、或いは組み合わせで考慮しリターンを予測する。一つは、証券アナリストによる評価変動前の1週間の株価リターン、変動率、出来高、売買代金などの株価ファクター。第2に、評価変動前のニュース記事の評価表現から構築したセンチメントファクター。第3に、証券アナリストの能力の代理変数となる所属証券会社の評判、規模、外資系、国内系別で見る評価者ファクター、これに加えて、ファイナンス研究で一般にコントロールされる格下げ対象企業の簿価時価比率、規模、モメンタムなどのリスクファクターである。

株式ファクターの中の変動率は、出来高、売買代金は、当該銘柄に対する市場評価の多様性の代理変数と考えられる。多くの投資家が異なる評価をしていれば、当該銘柄に対する出来高は増え変動率は高くなることが予測される。評価変動前のリターンについては、株価のランダムウォークを仮定するならば、説明変数にはなりえない。しかし、アナリストの評価変動についての情報漏洩やインサイダー取引が行われていれば、説明力をもつ可能性があると考え検証対象とした。次にセンチメントファクターであるが、岡田・羽室(2011)で、センチメントが株価変動に対して先行性を持つことを報告しているが[12]、その効果がアナリスト評価変動銘柄にも及んでいれば、説明力を持つと考えられる。評価者ファクターが説明力を持つ場合は、予測能力はそもそもアナリストが属する証券会社の規模やプ

レステージに依存するということである。最後に、リスクファクターであるが、時価簿価比率と規模については、Fama and Frenchにより、システマティックにリターンへ影響を与えるリスクファクターとして認識されている[10]。更に、De Bondt and Thalerによって明らかにされたモメンタムについても[3]、最近ではリスクファクターとしての認識が一般的になっているため、この3つの変数を採用する。

3 手法

証券アナリストがある企業の格下げを発表したとしても、必ずしもその企業の株価が下落するわけではない。それでは、どのような条件の時に、アナリストの格下げ発表に市場は反応するのであろうか？本研究ではデータマイニング手法を利用することで、証券アナリストによる格下げ発表とその後の株価変動に関する有力な仮説の取得を試みる。本節は、その方法論について論じる。そして次節で、得られた仮説の信頼性について、コントロールファーム法による検証を行い、またその仮説に基づいたトレーディングシミュレーションを行うことにより実運用の可能性についても検討する。

仮説の取得については、2つのマイニング手法を組み合わせることで実施する。一つは、格下げ発表前の時系列データから、発表後の株価変動と関係の強い系列パターンを抽出する手法で、BONSAIと呼ばれる分子生物学の分野で開発された手法を用いる。そして二つ目は、説明変数の全組み合わせの中から株価変動と関係の強いパターンを列挙する顕在パターン列挙手法である。いずれの手法においても、格下げ発表後の株価の大幅な下落変動とそうでない変動の二値を目的変数として用いる。

以下では、まず株価変動に関する目的変数の定義を与え、また説明変数として利用するセンチメント指数の導出法について論じる。そして、2つのマイニング手法、BONSAIと顕在パターンについて説明する。

3.1 目的変数

目的変数は株価の収益率によって定義する。個別銘柄 e についての格下げ発表が日 t に行われたとき、 τ 日後の株価収益率 $R(e, t, \tau)$ を以下のように定義する。ここで $close_{e,t}$ は、銘柄 e の日 t における終値である。

$$R(e, t, \tau) = \sum_{i=0}^{\tau-1} \left(\frac{close_{e,t+i+1}}{close_{e,t+i}} - 1 \right) \quad (1)$$

そして株価収益率を大幅に下落するかどうかの二値に離散化することで目的変数を定義する。本研究では、

$R(e, t, \tau)$ の二値離散化は、式 (2) に示すとおり、 -0.03 を下回るかどうかによって定める。この値は、著者らの経験に基づき恣意的に決めた値である。

$$R'(e, t, n) = \begin{cases} 0, & \text{if } R(e, t, \tau) < -0.03 \\ 1, & \text{else} \end{cases} \quad (2)$$

また本研究では日数 $\tau = 1, 5, 11, 22, 33, 44$ として実験を行った。日数は営業日に基づくため、実日数に換算すると、 $\tau = 5$ で約 1 週間、 $\tau = 44$ で約 2 ヶ月に対応する。

3.2 センチメント指数

センチメント指数は、ニュース記事テキストから作成される極性付き評価表現辞書に基づいて計算される。ここで、株価の上昇に影響を与えるであろう極性を肯定極性、下落に影響を与えるであろう極性を否定極性とし、それぞれの極性を伴った表現を肯定表現、否定表現と呼ぶことにする。「需要が伸びる」「株価が反発する」といった表現は肯定表現の代表例であり、「業績が悪化する」「売りが先行する」は否定表現の代表例である。これらの表現の出現頻度によって市場センチメントを定義しようということである。

評価表現辞書の構築にはいくつかの方法論が提案されているが [7]、本研究では、周辺文脈の一貫性を利用した手法 [9] を用いることにした。この手法では、文書中で評価表現が存在すると、その周辺文脈でも同じ極性で一貫しているという仮定に基づいており、極性をはっきりした少数の表現を種表現として用意し、それらの種表現が含まれる文章の周辺文脈を調べることで、次々と新たな評価表現を獲得していく。手法の詳細については原著を参照されたい [9]。

この手法で対象とする評価表現は単純エントリと複合エントリに大別でき、単純エントリとは、「上昇する」「回復する」といった用言句からのみ構成される表現で、複合エントリとは「株価が上昇する」「景気が回復する」といった格助詞句と用言句のペアである。

本研究で筆者らが用意した種語は、肯定表現として「増益となる」「黒字となる」「株価が急騰する」「株価が反発する」、否定表現として「減益となる」「赤字となる」「株価が急落する」「株価が反落する」の計 8 つで、最終的に、合計 2,937 の評価表現を獲得することができた。その内訳を表 1 に、得られた評価表現の例を表 2 に示す。

センチメント指数は、評価表現辞書に登録された評価表現の出現頻度を日単位でカウントし、その線形加重移動平均として定義する。銘柄 e の日 t における過去 n 日間のセンチメント指数 $SI(e, t, n)$ は、式 (3) で示される通りである。

表 1: 獲得した評価表現の数

	単純エントリ	複合エントリ	合計
肯定	207	1,283	1,490
否定	252	1,195	1,447
合計	459	2,478	2,937

表 2: 獲得した評価表現

		評価表現
単純	肯定	上回る, 好調だ, 伸びる, 回復, 良好だ など
	否定	下回る, 落ちる, 赤字, 警戒, 減額する など
複合	肯定	(過去最高を, 更新), (業績が, 回復する) (計画を, 上回る), (期待が, 高まる) など
	否定	(赤字と, なる), (株価が, 続落) (計画を, 下回る), (ことが, 嫌気する) など

$$SI(e, t, n) = \sum_{i=0}^{n-1} \frac{n-i}{N} S_{e,t-i}^+ - \sum_{i=0}^{n-1} \frac{n-i}{N} S_{e,t-i}^- \quad (3)$$

ここで、 $N = \sum_{i=1}^n i$ で、 $S_{e,t}^+$ 、 $S_{e,t}^-$ は、それぞれ銘柄 e の日 t における肯定極性表現、否定極性表現の出現頻度である。

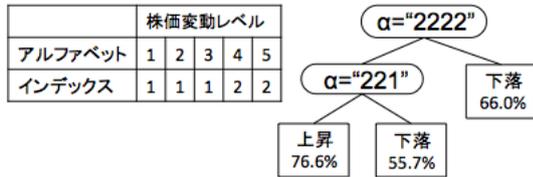
3.3 時系列解析

格下げ発表後の株価の下落に関連する要因として、発表前の株価の値動きや変動、さらにはマーケット全体のセンチメントの動向などの経時的傾向が関係していることは想像に難くない。格下げが発表されたとしても、過去数週間の収益率が安定的に増加していれば、市場が直ちに反応することはないかもしれない。

そこで、発表後の収益率の変化に対して、過去の時系列パターンがどのように影響しているかを探索することを考える。ここでは系列パターンを説明変数とした決定木モデルを構築する機械学習アルゴリズム BONSAI を利用する。BONSAI は分子生物学で得られた塩基配列やアミノ酸配列等のデータから規則性を発見することを主たる目的として開発されたアルゴリズムである [1, 13]。以下に、その概略を示す。

アルファベット Σ 上の文字列の集合が与えられているとする。各文字列には正、または負のラベルが付いている。正のラベルの付いている文字列集合を正事例集合、負のラベルのものを負事例集合ということにする。このとき、正事例集合と負事例集合を分類するための正規パターンを節点の条件とした決定木を構成するアルゴリズムが BONSAI である。

ここで正規パターン (regular pattern) とは、アルファベット Σ 上の n 個の文字列を $\alpha_1\alpha_2\dots\alpha_n$ とし、任意の $n+1$ 個の文字列を x_0, x_1, \dots, x_n とした時、 $x_0\alpha_1x_1\alpha_2x_2\dots\alpha_nx_n$ の形式で与えられる。本研究では、パターンの意味解釈のしやすさの観点から、正規パターンとして $x\alpha y$ (x, y は任意の文字列)、すなわち部分文字列 α に限定することにした。



株価変動の過去の時系列がアルファベット 1~5 の文字列として表現されている。インデックス化関数は、アルファベット 1,2,3 はインデックス 1 に、4,5 は 2 への写像をおこなう。決定木のトップノードは、過去の変動にインデックス 2 が 4 回連続出現すれば、すなわちアルファベット 4 もしくは 5 が 4 回連続出現すれば、格下げ発表後に株価が下落する確率が 66.0%であることを示している。

図 1: BONSAI による決定木の例

BONSAI ではアルファベットのインデックス化 (alphabet-indexing) というメカニズムを導入してモデルの分類精度を高めている。これは BONSAI の一つの大きな特徴である。元のアルファベット集合 Σ からサイズの小さな別のアルファベット集合 Γ への写像 ϕ (インデックス化関数) を考えて、元の文字列集合を ϕ で変換後、上記の決定木を構成するものである。通常 Γ のサイズは極めて小さいものを考える。決定木生成においては、様々なインデックス化関数を適用し、判別能力が最大となるような ϕ から得られる決定木を出力する。図 1 に BONSAI により構築された決定木の例を示す。決定木の構築に関する詳細や節点の条件となる候補パターンの導出方法は原著を参照されたい [1, 13]。

以上の方法によって得られたインデックス化関数 ϕ による Γ 上のパターンは、二値クラス判別の精度を高めるパターンであり、格下げ発表後の株価の推移について何らかの現実的な意味を含んでいることが期待できる。

本研究において、BONSAI を利用するのは、最終的な株価推移モデルの構築を目的としているのではなく、格下げ発表後の株価変動に関係のありそうな各種時系列パターンを仮説として抽出するために用いていることに注意されたい。

3.4 顕在パターン列挙

顕在パターン (emerging pattern) とはあるクラスに多頻度で、その他のクラスでは多頻度ではないような

パターンのことである。以下にその定義を与える。アイテム集合 $I = \{i_1, i_2, \dots, i_m\}$ について、その部分集合 $T \subseteq I$ をトランザクションと呼び、また、トランザクションの集合 \mathcal{D} をデータベースと呼ぶ。データベース \mathcal{D} において、あるアイテム集合 $\alpha \subseteq I$ を含むトランザクションの割合をサポート呼び、式 (4) で定義する。

$$\text{support}_{\mathcal{D}}(\alpha) = |\{t \in \mathcal{D} | \alpha \subseteq t\}| \quad (4)$$

いま異なる二つのクラスに属するデータベース $\mathcal{D}_0, \mathcal{D}_1$ について考える。本研究のケースでは、 \mathcal{D}_0 を格下げ発表後株価が下落するケース群、 \mathcal{D}_1 をそうでないケース群と考えるとよい。パターン α の \mathcal{D}_1 に対する \mathcal{D}_0 の増加率 (growth rate) $GR_{\mathcal{D}_0}(\alpha)$ を式 (5) の通り定義する。

$$GR_{\mathcal{D}_0}(e) = \begin{cases} \frac{\text{support}_{\mathcal{D}_0}(e)}{\text{support}_{\mathcal{D}_1}(e)} & , \text{support}_{\mathcal{D}_1}(e) \neq 0 \\ \infty & , \text{support}_{\mathcal{D}_1}(e) = 0 \end{cases} \quad (5)$$

そして、ユーザにより指定された最小サポート σ 、および最小増加率 γ について、 $\text{support}_{\mathcal{D}_0}(\alpha) \geq \sigma$ かつ $GR_{\mathcal{D}_0}(\alpha) \geq \gamma$ を満たすパターン α をクラス 0 の顕在パターンと呼ぶ。

本研究では、アナリストの格下げ発表をトランザクションの単位とし、またアイテムとして、表 6 に示す格下げに関連すると考えられる各種変数およびその値のペア (表記は「変数名=値」) を用いる。さらに、前節の BONSAI の結果から意味的に妥当と判断した系列パターンもアイテムとして扱う。表 3 にトランザクションデータベースのイメージを示す。

表 3: 顕在パターンを列挙する対象トランザクションの例

下落クラス \mathcal{D}_0		非下落クラス \mathcal{D}_1	
ID	アイテム集合	ID	アイテム集合
01	{afd=D, pre=2, big=1, acr=5, mom1=1, mom2=5, varPat=2222, ...}	11	{afd=F, pre=2, big=2, acr=5, mom1=2, mom2=1, liq1=1, ...}
02	{afd=F, pre=2, ...}	12	{afd=F, pre=1, ...}
⋮	⋮	⋮	⋮

アイテムを構成する変数名と値は表 6 に対応している。

4 仮説発見と考察

4.1 BONSAI による系列パターンの取得

BONSAI による有効な系列パターンを取得するために、表 4 に示される 5 種類の変数について、それぞれ 4 種類の時系列データを用意した。時系列データは、あ

る格下げ発表 r の発表日より t 期前の値 q_t^r (式 (6)) についての系列 $Q^r = \langle q_t^r, q_{t-1}^r, \dots, q_1^r \rangle$ で表わされる。

$$q_t^r = \frac{1}{m} \sum_{i=m \cdot (t-1)+1}^{m \cdot t} a_i^r \quad (6)$$

ここで、 m はひとつの期を構成する日数で、 a_i^r は格下げ発表 r の発表日より i 日前の値を表す。

そして、系列 Q^r を BONSAI が扱えるアルファベクトルに離散化された系列 $Q^{r'} = \langle q_t^{r'}, q_{t-1}^{r'}, \dots, q_1^{r'} \rangle$ に変換する。いま、全体で n 件の格下げ発表があり、それぞれの発表日より t 期前における全ての値を小さい順に並べた $q_t^1, q_t^2, \dots, q_t^n$ を k 個のグループ G_1, G_2, \dots, G_k にできるだけサイズが等しくなるように分割する (詳細は省略)。そして q_t^i がグループ G_j に属するとき、 $q_t^{r'} = j$ と離散化する。

表 4 に示された 5 種類それぞれに対して $m = 1, 2, 3, 4$ とした系列変数 20 個について、それぞれ BONSAI によって決定木モデルを構築した。

表 4: パターン抽出に利用した変数一覧

変数名	内容: 式 (6) における a_i
収益率系列 (ret)	$\frac{close_t}{close_{t-1}} - 1$
変動系列 (var)	$\frac{(close_t - avg)^2}{avg^2}$, $avg = \frac{1}{22} \sum_{t=1}^{22} close_{t-1}$
出来高系列 (vol)	$\frac{volume_t}{発行済株式総数_t}$
市場センチメント 指数系列 (asi)	$\sum_e \sum_{i=1}^{22} \frac{22-i+1}{253} SI_{e,t-i}$ 22 日間の線形加重移動平均
業界センチメント 指数系列 (isi)	$\sum_{e \in ind} \sum_{i=1}^{22} \frac{22-i+1}{253} SI_{e,t-i}$ ind : 格下げ対象企業が属する業界集合

そして正答率の高かった上位のモデルについて、ファイナンスにおける意味的妥当性を検討し、表 5 に示す 3 つの系列、計 5 つのパターン (a)~(e) を選択した。その他の多くのパターンは、正答率が極端に低い、もしくは意味的妥当性が低いために除外した。

系列 var は、株価の過去の変動率から推定している。株式価値は株主に帰属するキャッシュフローの現在価値として決まるため、事業ポートフォリオからあがるキャッシュフローの変動率が高い企業、リスクの高い事業に取り組んでいる企業ほど株価の変動率は高くなる。逆に、株価変動率の低い企業は、公益株の様に事業ポートフォリオから生まれるキャッシュフローが安定している企業と言えるだろう。したがって、リスクな事業に取り組んでいる企業のファンダメンタルズが悪化し、それが証券アナリストの格下げというイベント

表 5: パターン抽出に利用した変数一覧

変数	τ	m	acc.	妥当なパターン (indexing)
var	11	3	70.0	(a) 2222 (1={1,2,3}, 2={4,5})
asi	44	4	55.7	(b) 2111 (1={1,2}, 2={3,4,5}) (c) 2121 (1={1,2}, 2={3,4,5})
isi	33	3	55.1	(d) 121 (1={1,2}, 2={3,4,5}) (e) 21111 (1={1,2}, 2={3,4,5})

var の決定木モデルは図 1 に示されている。

表の見方 (一行目): 変動系列 (var) について 3 日を一期とする時系列データを説明変数とし、発表後 11 日後を予測する決定木モデルを構築したところ、その正答率 (acc.) は 70% であった。そしてその決定木モデルから意味的に妥当なパターン $\alpha = "2222"$ を選択した。パターンを構成するインデックスは、元のアルファベクトル 1,2,3 が 1 に、4,5 が 2 に対応している。

で公開情報となった時に株価は大きく下落すると考えられる。

系列 asi は、評価辞書から構築した個別銘柄のセンチメントを総合することで市場全体のセンチメントを表す指標である。投資家は同じファンダメンタルズ情報であっても、その評価をムードによって変えていることが知られている [6]。市場全体のムードが悪化している時期においては、格下げイベントに対する市場評価は、実態以上の下落率を伴うと考えられる。

系列 isi は、評価辞書から構築した個別銘柄を業界単位で総合した業界全体のセンチメントを表す指標である。証券アナリストは証券分析のプロであるが、投資家がその格下げ発表に素直に従い売却するわけではない。投資家心理には保有効果が働きやすいことは Kahneman et. al (1990) の実験で確認されており、保有している銘柄を手放すには追加的誘引が必要である。証券アナリストが格下げする企業の業界全体に芳しくないニュースが多い場合、(例えば、円高が急激に進行した時期の自動車製造業界等) 当該企業を保有する投資家にとっては手放そうという追加的誘引となるものと考えられる。

4.2 顕在パターンによる仮説の導出

BONSAI によって得られた系列パターンを含め、表 6 に示される変数について顕在パターンを列挙した。

最小サポート $\sigma = 0.01$, 最小 $GR\gamma = 2.0$ の条件で、1,2 アイテムから構成される顕在パターンを全列挙し、そこから意味的に妥当なパターンを選択した。列挙されたパターン数は、1 アイテムで構成されるパターンが 1 件、2 アイテムで構成されるパターンが 128 件であった。表 7 に、それらのパターンのうち、意味的に妥当と判断した 3 つのパターンを示す。

表 7 は得られた顕在パターンと、予測期間、サポート、予測精度 (確率比) を示したものである。一つの明

表 6: パターン抽出に利用した変数一覧

変数名	値
アナリスト国内外 (afd)	D(国内),F(国外)
アナリスト信頼度 (pre)	(低) 1,2 (高)
アナリスト名 (brk)	56 企業
アナリスト規模 (big)	(小) 1,2 (大)
accrual rank(acr)	(低) 1,2,3,4,5 (高)
市場規模 rank(mv)	(低) 1,2,3,4,5 (高)
時価簿価比率 rank(bpr)	(低) 1,2,3,4,5 (高)
格下げ後評価 rank(to)	(低) -2,-1,0,1 (高)
業界 (ind)	30 業界
過去 1,5,10,22 日 個別センチメント rank (si1,si5,si10,si22)	(低) -2,-1,0,1,2 (高)
過去 44,66,125,250 日 モメンタム rank (mom1,mom5,mom10,mom22)	(低) 1,2,3,4,5 (高)
過去 1,5,10,22,44,66 日 流動性 rank (liq1,liq5,liq10,liq22,liq44,liq66)	(低) 1,2,3,4,5 (高)
変動系列 varPat=2222	0:なし 1:あり
市場センチメント系列 asiPat=2111 asiPat=2121	0:なし 1:あり
業界センチメント系列 isiPat=121 isiPat=21101	0:なし 1:あり

表 7: 得られた妥当な顕在パターン (単位は%)

size	pattern	τ	support	GR	prior	Post.
1	varPat=2222	11	53.8	2.56	43.7	66.4
2	varPat=2222 & mv=1	2	2.67	4.18	30.6	64.8
	varPat=2222 & si10=-1	11	2.92	3.31	43.7	72.0

らかなパターンは、イベント前のボラティリティの高い企業群の方が格下げ後の下落確率がかなり高く (高ボラティリティでない企業群の下落確率の 2.56 倍) ということである。ボラティリティが高い企業は、事業ポートフォリオの不確実性が高い企業であるため、格下げ情報が市場に伝わった場合に、下落する確率が高くなると考えられる。格下げイベント発生前 15 日間の中で、12 日以上連続でボラティリティの高い企業群については、12 日間の累積リターンで見た場合 66.4% の確率で下落している。これは prior の 43.7% と比較すると大幅な改善である。またサポート数も 1327 と、トレーディング戦略にのせるに十分な銘柄数を確保できている。

表 7 の下段に示すのは、2 つのファクターの組み合わせ列挙である。まず、時価総額で見た場合に最も企業規模が小さいグループ (mv-Rank-1) ほど下落確率が高い (4.18 倍) ということになる。11 日間のイベント後の累積リターンで言えば、小型株に限定することで、事前確率 30.6% に対して 64.8% の確率で下落する。小型株は大型株に比較してリスクが高いことが報告されているが [10]、格下げイベントは小型株に対してより大きな影響を与えていると言える。

同様に、センチメントと変動系列の組み合わせで列挙した場合も、大幅な確率の向上が観察される。

5 検証

5.1 コントロールファーム法による検証

これまでの探索結果から、格下げイベントを起こした企業群の中でも、変動系列の高い銘柄群については、その下落確率が高いことが明らかになった。本節では、そのインパクトを調査するために、コントロールファーム法を用いて超過リターンを算出し、顕在パターンでスクリーニングした結果 (match) としなかった結果 (unmatch) を比較する。

表 8 に示すように、変動系列で match した場合と unmatch の場合の差は有意であり、ボラティリティの高い銘柄群がイベント発生後約 2ヶ月にわたってより大きな下落率を示し、ドリフトを起こしていることがわかる。

表 8: 変動系列=“2222” のコントロールファーム法による収益率比較 (単位は%)

	r-1	r-5	r-11	r-22	r-33	r-44
match	-1.66	-3.68	-4.56	-5.41	-5.96	-6.78
unmatch	-0.77	-0.86	-0.76	-1.07	-1.52	-1.83
差	-0.89	-2.83	-3.79	-4.33	-4.45	-4.95

match 件数: 1,988 unmatch 件数: 3,770

列項目タイトルの数字 r-n は、格下げ発表 n 日後の収益率を表している。match と unmatch の平均収益率の差の有意確率はいずれも $P < 10^{-12}$ 。

表 9 は変動系列に時価総額 (mv) を追加した場合の match と unmatch の差のみを抽出して一覧表にしたものである。時価総額が大きいもの (mv-5) から小さいもの (mv-1) になるに従って下落率が大きくなっていく様子示されている。小型株効果の影響が色濃くできていると言えよう。

表 10 は変動系列に個別銘柄センチメントランクを追加した場合の match と unmatch の差を抽出したものである。時価総額ほどセンチメントランク通りにはなっ

表 9: 変動系列=“2222” に市場規模 (mv) を追加した場合のコントロールファーム法による収益率比較 (単位は%)

mv	r-1	r-5	r-11	r-22	r-33	r-44
1	-1.74	-4.96	-7.76	-12.11	-14.95	-16.81
2	-1.62	-4.04	-5.55	-6.74	-6.91	-8.54
3	-1.14	-3.37	-5.63	-6.28	-5.96	-6.34
4	-0.35	-1.82	-2.41	-1.80	-2.16	-2.68
5	-0.33	-1.25	-2.33	-2.11	-1.99	-2.34

条件にマッチするケースの平均収益率-マッチしないケースの平均収益率のみ示している。太字は 5%有意。

ていないものの、概ねセンチメントが悪化するほど下落率が大きくなっている。

表 10: 変動系列=“2222” に過去 10 日間の個別センチメントランク (si10) を追加した場合のコントロールファーム法による収益率比較 (単位は%)

si10	r-1	r-5	r-11	r-22	r-33	r-44
-2	-1.03	-3.42	-6.47	-7.09	-6.40	-6.74
-1	-0.86	-2.97	-4.38	-3.43	-3.62	-3.51
0	-1.09	-3.20	-4.34	-5.27	-6.39	-7.35
1	+0.33	-0.62	-1.09	-2.93	-4.88	-4.96
2	-0.07	+0.35	+0.81	+2.00	+0.09	+0.23

条件にマッチするケースの平均収益率-マッチしないケースの平均収益率のみ示している。太字は 5%有意。

5.2 シミュレーション

以上の結果より、過去のボラティリティが一定期間連続する条件が、格下げ発表後に株価が大きく下落することと関係していることが明らかとなった。そこで、この条件を加味してシミュレーションを行うことで実務の有効性を確認する。表 9,10 に示された市場規模および個別センチメントランクについては、対象となる件数が少ないのでシミュレーションでは扱わない。

投資戦略としては、個別銘柄に関するイベントドリブンによるマーケットニュートラル手法を用いる。今回対象となるイベントは証券アナリストの格下げ発表である。ある銘柄において格下げ発表が発生した場合、その銘柄を売却しショートポジションを取る。今回のケースでは、ロングポジションをとるイベントが少ないため、ショートポジションと同額の日経 225 先物をロングすることでマーケットニュートラルを保つ。

期間は、2000 年 1 月~2011 年 5 月で、初期資産 10 億円で運用を開始する。損切りルールを適用し、収益率が-0.03 を超過すると強制的に精算する。マーケットインパクトを回避するために、前日までの 5 日平均取



図 2: 変動系列ルール適用前後の結果

引高の 10%を超える取引は行わない。ただし、最大 3 日間の分割購入を実施する。

以下では 2 つのシミュレーション実験を行った。一つは、全ての格下げ発表イベントに対してショートポジションをとる戦略で、他方は、格下げイベントのうち、本研究で明らかにした過去の変動系列が系列パターン“varPat=2222”を含むときのみショートポジションをとる戦略である。結果は図 2 および表 11 に示されている。

図 2 の左に示されているように、アナリストの格下げ情報に基づくカレンダータイムポートフォリオを組成した場合、その価値は右肩上がり上昇する。これは証券アナリストによる格下げにより、株価の下方ドリフトが起こっているという大田・近藤 (2010) の報告内容の別表現である [11]。図 2 の右に示すのは、変動系列でスクリーニングにした 1327 銘柄を対象にカレンダーポートフォリオを作成した結果である。まず、下落確率が高まっていることから、相対パフォーマンスで負ける銘柄がすくないため、ポートフォリオ価値のドロウダウンが少ないことがわかる。

表 11: 変動系列ルール適用前後のシミュレーション結果の比較

指標	導入前	導入後
Sharpe 比	1.73	2.63
平均年利	8.52%	7.24%
標準偏差	4.91%	2.76%
平均レバレッジ	51.4%	11.1%
最大下落率 (日)	-4.17%	-3.70%
最大下落率 (月)	-4.77%	-4.10%
最大下落率 (年)	-17.75%	-17.32%

表 11 に 2 つのアプローチの比較を載せている。ポートフォリオが単位当りのリスクに対してどの程度のリターンを達成しているかを見るのが Sharpe 比であるが、劇的な改善がみられるのがわかる。変動系列のスクリーニングによって、明らかに優れたポートフォリオの構築が実現できている。

6 むすび

本研究では、証券アナリストによる企業の格下げ変更の発表が、その企業の株価に与える影響について分析を進めた。方法論としては、時系列解析とパターン列挙のデータマイニング手法を用いることで、格下げ発表後の株価変動を説明する仮説を導出した。それらの中で、格下げ発表前のボラティリティが高い状態で推移すると、発表後の株価が大きく下落することが検証された。またボラティリティに市場規模もしくはセンチメントを組み合わせることで、より正確に株価の下落を説明できることが明らかとなった。さらに、格下げ変更をイベントとしたトレーディングシミュレーションの結果、ボラティリティを考慮するトレーディングは、考慮しないトレーディングに比べて、飛躍的な好成績を達成できることも示された。

謝辞

本研究の一部は、ERATO 湊離散構造処理系プロジェクト、および科学研究費補助金基盤研究 C(課題番号 22530330) の研究助成を受けている。

参考文献

- [1] Arikawa, S., Miyano, S., Shinohara, A., Kuhara, S., Mukouchi, Y. and Shinohara, T., A machine discovery from amino acid sequences by decision trees over regular patterns, *New Generation Computing*, Vol.11, pp. 361–375, 1993.
- [2] Barber, B., R. Lehavy, M. McNichols, and Trueman, B., Can Investors Profit from the Prophets? Securities Analyst Recommendations and Stock Returns, *Journal of Finance*, Vol. 56, pp. 793–805, 2001.
- [3] De Bondt, F.M.W. and R. Thaler, Does the Stock Market Overreact? *Journal of Finance*, Vol. 40, pp. 531–564, 1985.
- [4] Dong, G. and Li, J., Efficient mining of emerging patterns: discovering trends and differences, *Proc. of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 43–52, 1999.
- [5] Fama, E.F. and K. French, Common risk factors in returns on stocks and bonds, *Journal of Financial Economics*, Vol. 33, pp. 3–56, 1993.
- [6] Hirshleifer, D and T. Shumway, Good Day Sunshine: Stock Returns and the Weather, *Journal of Finance*, Vol. 58, pp. 1009–1032, 2003.
- [7] 乾孝司, 奥村学, 「テキストを対象とした評価情報の分析に関する研究動向」 自然言語処理, Vol.13, No.3, pp. 201–241, 2006.
- [8] Miller, Edward M., Risk, Uncertainty and divergence of opinion, *Journal of Finance*, Vol.32, pp. 1151–1168, 1977.
- [9] 那須川哲哉, 金山博 「文脈一貫性を利用した極性付評価表現の語彙獲得」 情報処理学会自然言語処理研究会 (NL-162-16), pp. 109–116, 2004.
- [10] Fama, E.F. and K. French, The cross-section of expected returns, *Journal of Finance*, Vol.47, pp. 427–466, 1992.
- [11] 大田浩司・近藤江美 「株式レーティングの公表に対する市場反応—株価と出来高の検証—」 経営財務研究, 第 29 巻, pp. 50–84, 2010.
- [12] 岡田克彦・羽室行信 「相場の感情とその変動—自然言語処理で測定するマーケットセンチメントとボラティリティ—」 証券アナリストジャーナル, Vol.49 No.8 pp .37–48, 2011.
- [13] Shimozono, S., Shinohara, A., Shinohara, T., Miyano, S., Kuhara S. and Arikawa, S., Knowledge Acquisition from Amino Acid Sequences by Machine Learning System BONSAI, *Trans. Information Processing Society of Japan*, Vol.35, pp. 2009–2018, 1994.
- [14] Stickel, Scott E., 1995, The anatomy of the performance of buy and sell recommendations, *Financial Analysts Journal*, Vol.51, pp. 25–39.
- [15] Womack, Kent L., 1996, Do brokerage analysts' recommendations have investment value? *Journal of Finance*, Vol. 51, pp. 137–167.