

ティックデータを用いたニュースと株価との関連性分析

Analyzing the Relationship between News and the Tokyo Stock Exchange Using Intraday Price Data

五島 圭一*¹ 高橋 大志*² 寺野 隆雄*³
Keiichi Goshima Hiroshi Takahashi Takao Terano

*¹*³東京工業大学 大学院総合理工学研究科
Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology

*²慶應義塾大学 大学院経営管理研究科
Graduate School of Business Administration, Keio University

This paper studies the relationship between news and the Tokyo Stock Exchange using intraday price data. Our two main findings are: news information almost reflects on stock prices within one second; positive news and negative news have an association with stock price volatility.

1. はじめに

ニュースと株式市場との関連性については、従来より数多くの報告が行われている。近年では、これらニュース分析に、テキスト分析が取り入れられている。ニュースの内容を評価し、テキストのポジネガ度合いを定量化することで、これまで観測が難しかったイベントを分析したり、投資家のセンチメントを計量したり、などの試みがなされている [Kearney 14][五島 16]。このようにテキスト情報を用いた様々な研究報告がある。しかしながら、これら先行研究は日次単位での分析が多く、日中データを用いた研究報告例はあるものの [Groß-Klußmann 11]、米国証券市場を対象とした分析であり、日本証券市場を対象とした研究報告例はない。そこで、本研究ではニュースデータをテキスト分析することで、ニュースのポジネガを定量化し、個別銘柄の株価データとティックデータを用いることで、秒単位でのニュースと日本株式市場との関連性の分析を試みる。次章は、本分析で用いるデータに触れ、3章では分析方法、4章では分析結果を記す。5章は、まとめである。

2. データ

本研究では、個別銘柄の株価リターンとマーケットファクター・リターンのデータを用いた。個別銘柄の株価データについては、Thomson Reuters Datastream から、日次データとティックデータを用いた。また、マーケットファクター・リターンのデータについては株式会社金融データソリューションズが提供する日本版 Fama-French ベンチマークを使用した。

ニュースデータについては、ロイターニュースを用いた。ロイターニュースは、世界で最も広く知られたニュース提供会社の一つであるトムソンロイター社が配信しているニュースである。本研究では、日本証券市場に関するニュースを分析対象とし、ニュースの本文を利用した。また、タグ情報については、ニュースの配信日時とニュースと関連する企業の証券コードを利用した。

3. 分析方法

3.1 ニュースの分類について

ニュースデータの選別について、本研究では、東証一部上場企業に関するニュースを分析対象とした。また、第一報のニュースのみを分析対象とするため、再送記事と訂正記事について分析対象外とし、さらに、ニュースのテキスト情報に注目したため、決算情報のみのニュースについても、分析対象外としている。加えて、ティックデータによる分析を行うため、営業日において9時から15時までに配信されたニュースのみを分析対象としている。ただし、学習データには2013年に配信された全てのニュースを用いた。

複数の企業の内容について報じているニュースは、付随している銘柄数の分だけニュースを増やし、一つのニュースに一つの企業を対応付け、分析を進めた。選別後のそれぞれのニュース数については、表1に示す*¹。

表 1: ニュース数

ニュース数	2,150
のベニュース数	2,991
銘柄数	787

ニュースを分類するために、過去の株価データとニュースを基にサポートベクター回帰によって、ニュースへのポジネガスコア付与とポジティブあるいはネガティブの度合いを表す値)を行った。SVRのパラメータについては、教師データを用いて10分割の交差検定を繰り返し、平均2乗誤差が最小になるように決定している。スコアが付与されたニュースを対象として、スコアの値が $z_{0.975}$ を超えたニュースをポジティブなニュースとして、 $-z_{0.975}$ を下回ったニュースをネガティブなニュースとして分類した。どちらにも分類されないニュースは、ニュートラルに分類した。分類後のニュース数は、表2に示す。

3.2 株価リターン・ボラティリティの算出について

前節にて分類した各ニュース群の関連する銘柄に関して、ニュース配信時刻の前後60分間における株価リターンとボラ

*¹ のベニュース数とは、一つのニュースに複数の証券コードが付随している場合、重複して計数した値である。

表 2: 各群のニュース数

Positive	579
Neutral	2,172
Negative	239

ティリティを算出する。ニュース配信時刻の前後 60 分間における株価リターンとして、ニュース配信から 1 秒後から 60 分後まで 10 分刻みで、また、60 分前から 1 秒前まで 10 分刻みで、さらに、1 秒前から 1 秒後までの各リターンを算出する。ニュース配信前の株価リターンの算出について、ちょうど値がなかった場合には直前の約定価格を用いて算出する。逆に、ニュース配信後の株価リターンについては、直後の約定価格を用いて算出する。また、ニュース配信時刻の前後 60 分間におけるボラティリティとして、ニュース配信から 1 秒後から 60 分後まで 10 分刻みで、また、60 分前から 1 秒前まで 10 分刻みで、ボラティリティを算出する。

4. 分析結果

はじめに、ニュース配信時刻の前後 60 分間における株価リターンについて考察する。

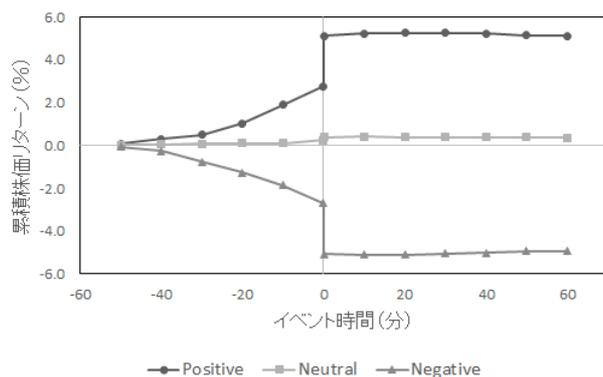


図 1: 累積株価リターンの推移

図 1 は、ニュース配信時刻の前後 60 分間におけるニュース群ごとの平均累積株価リターンの推移について、まとめたものである。縦軸は累積株価リターン、横軸はニュース配信時刻の 60 分前から 60 分後までのイベント時間を表している。横軸が 0 のときが、ニュースが配信された時刻となっている。ニュース配信時刻の 60 分前から 1 秒前までの累積株価リターンとして、ポジティブなニュースは 2.8%、ネガティブなニュースは -2.7% となっている。そして、どちらのニュース群もニュース配信の 1 秒前から 1 秒後において 2.4% ほど変化し、1 秒後から 60 分後では累積株価リターンの変化がほぼ見られなくなっている。この結果は、ニュース配信後 1 秒で、ニュースの内容が株価に織り込まれていることを示唆している。

次に、ニュース配信時刻の前後 60 分間におけるボラティリティについて考察する。

図 2 は、ニュース配信時刻の前後 60 分間におけるニュース群ごとの平均ボラティリティの推移について、まとめたものである。縦軸はボラティリティ、横軸はニュース配信時刻の 60 分前から 60 分後までのイベント時間を表している。横軸が 0

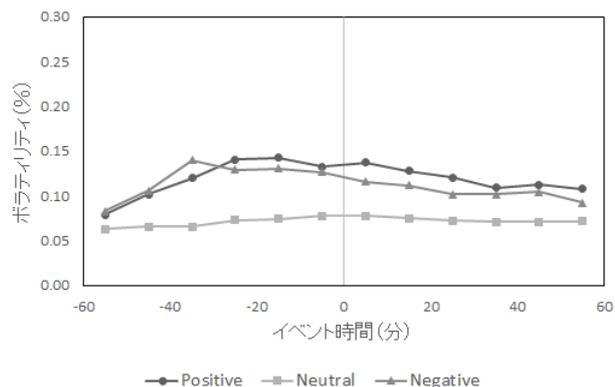


図 2: ボラティリティの推移

のときが、ニュースが配信された時刻となっている。ポジティブなニュースやネガティブなニュースはニュートラルなニュースより、ニュース配信時刻の前後 60 分間におけるボラティリティが高い傾向がある。とりわけ、ニュース配信時刻に近い方がボラティリティが高い傾向がある。具体的には、ニュートラルなニュースでは 0.07% で推移しているのに対して、ポジティブなニュースやネガティブなニュースでは、0.14% ほどまで上昇している。

5. まとめ

本研究では、ロイターニュースのポジネガ度合いを定量化し、ニュース配信時刻の前後 60 分間のニュースと株式市場との関連性の分析を試みた。分析の結果、ポジティブ及びネガティブなニュースについて、ニュースの配信に対し株価が短時間の間に反応していることを確認した。具体的には、1 秒前から 1 秒後において大きく株価が変動していることを見出した。また、ボラティリティについては、ニュース配信時刻に近いほど高い傾向があることが示された。マイクロ秒での分析や分析期間の拡張、他のニュース媒体の分析については今後の課題である。

参考文献

- [五島 15] 五島圭一・高橋大志：株価を用いたニュース評価に関する研究，日本ファイナンス学会第 23 回大会 予稿集，(2015)。
- [五島 16] 五島圭一・高橋大志：ニュースと株価に関する実証分析-ディープラーニングによるニュースの評判分析-，証券アナリストジャーナル，76-86 ページ (2016)。
- [Groß-Klußmann 11] Groß-Klußmann, A. and N. Hautsch：When Machines Read The News: Using Automated Text Analytics to Quantify High Frequency News-Implied Market Reactions, *Journal of Empirical Finance*, 18(2), pp. 321-340 (2011)。
- [Kearney 14] Kearney, C. and S. Liu：Textual Sentiment in Finance：A Survey of Methods and Models, *International Review of Financial Analysis* 33, pp. 171-185 (2014)。