

金融市場における価格変動メカニズムの分析： テキスト情報を用いた分析の試み

平成24年11月22日

高橋大志
慶應義塾大学

htaka@kbs.keio.ac.jp

内容

- はじめに
 - ◆ 分析の背景
 - ◆ テキスト情報の例
- 分析事例
 - ◆ 有価証券報告書
 - ◆ 金融市場に関する分析
- まとめ

2

利用可能な情報

数多くの情報伝達手段

- 数値情報
- 画像情報
- 音声情報
- テキスト情報



投資においては**いずれの情報も利用**

インターネットの普及等により情報量は飛躍的に増加

Web Facebook Twitter Blog

6

実証分析について

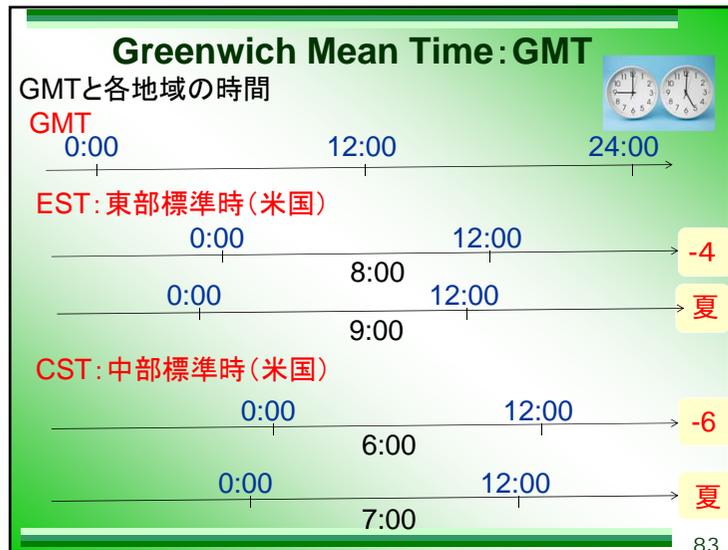
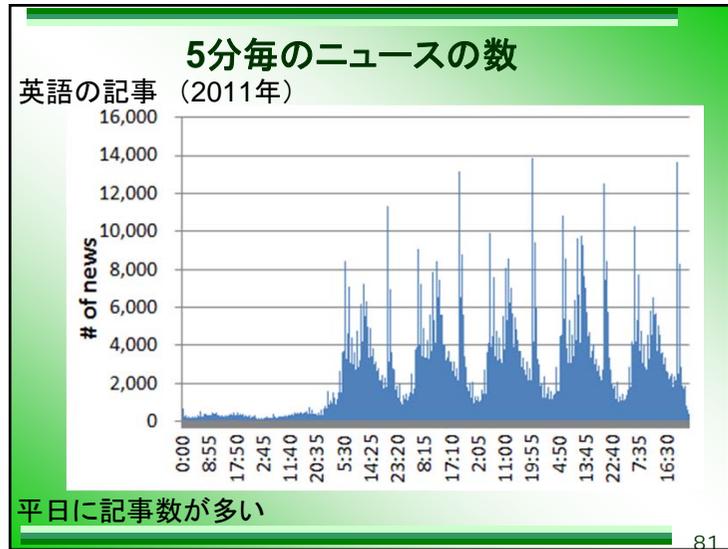
実証分析の多くは数値情報を用いた分析

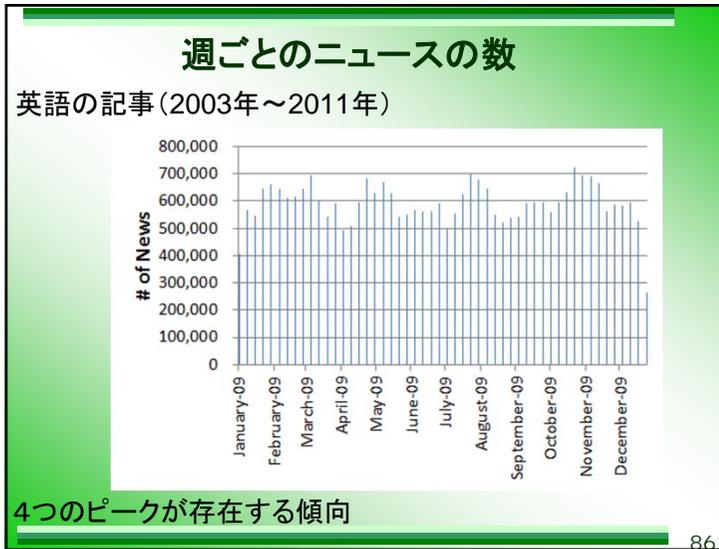
- × 株式価格, 債券価格, 為替レート
- × 取引量
- × 金利水準
- × 財務関連指標



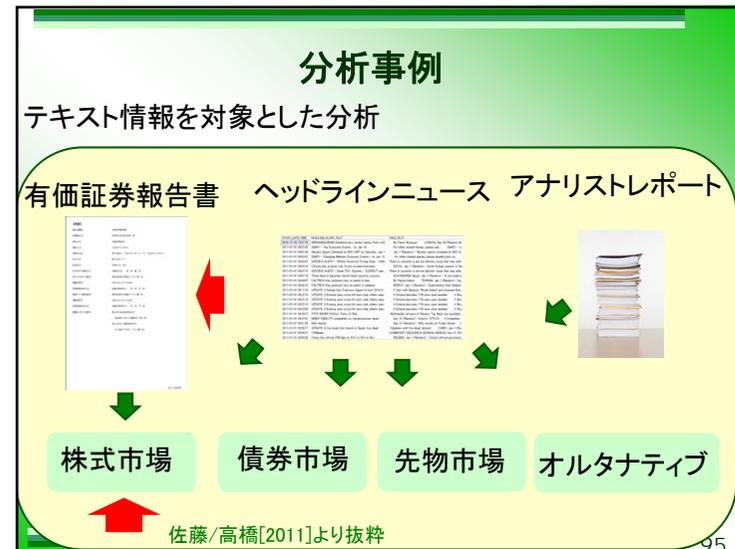
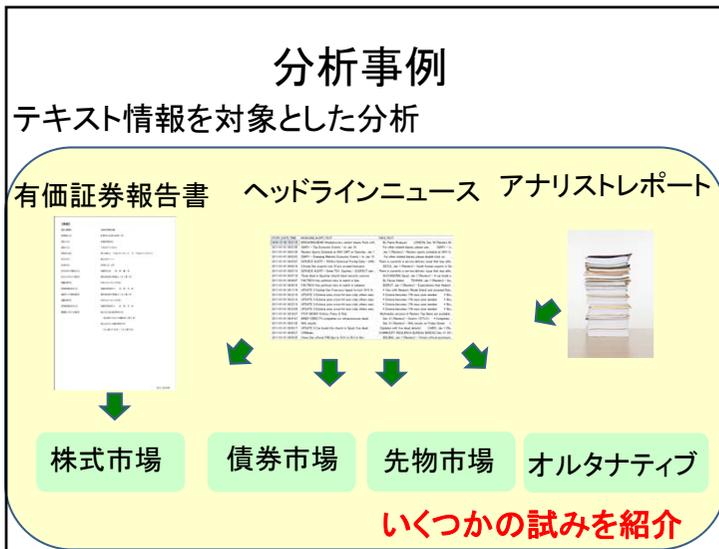
数値情報を用いた多様な分析

7





- ### 内容
- はじめに
 - ◆ 分析の背景
 - ◆ テキスト情報の例
 - 分析事例
 - ◆ 有価証券報告書
 - ◆ 金融市場に関する分析
 - まとめ
- 89



企業の自発的開示について 佐藤/高橋[2011]

企業が自発的開示を行う理由は？

自発的開示は、企業価値にマイナスの影響？？

- ①種々のコスト負担が発生する。
- ②競争優位の源泉が喪失する可能性がある。
- ③情報の非対称性の緩和なら制度開示の充実によっても可能である。

➡ ディスクロージャー優良企業は正の効果を獲得の可能性
制度開示と自発的開示

分類	情報開示		
	法定開示	適時開示	自発的開示
対象	株主・投資家・金融機関	同左	株主・投資家・アナリストなど マスコミ・ステークホルダー全般 一般消費者など
責任	法律などに基づく制度なので違反がなければ良い。会計監査が入ることもある	同左	企業が株式市場に影響を与えるので、経営トップの責任が大きい
内容	会社法(2005年5月施行)に基づく開示 計算書類(貸借対照表・損益計算書・株主資本等変動計算書)・参考資料・企業報告など 金融商品取引法(2007年3月施行)に基づく開示 有価証券報告書・四半期報告書・内部統制報告書など	有価証券上場規程に基づく開示 決算短信・業績予想の修正(売上高の10%以上、利益の30%以上など)	企業が自発的に行うIR活動を基にした開示 アニュアルレポート/ブック/決算説明会・スモールグループミーティング・工場見学・株主通信・株主便りなど 企業内容の理解促進に伴う開示 ニュースリリース・プレス・新聞・雑誌情報・CSR報告書など

佐藤/高橋[2011],佐藤[2004],木田[2007]

分析の目的 佐藤/高橋[2011]

自発的開示が企業に与える影響

1. 『自発的開示は株主資本コストを低下させる？』
2. 『自発的開示は株式収益率に正の影響？』
3. 『自発的開示の優良企業は、非優良企業と異なる特性を有する？』

データ

分析対象期間：6年間(2004年～2009年)

6つの表彰企業

表彰制度について 佐藤/高橋[2011]

表彰制度

- ・日本証券アナリスト協会「証券アナリストによるディスクロージャー優良企業選定」
- ・東京証券取引所「ディスクロージャー表彰制度」
- ・日本IR協議会「IR優良企業賞」
- ・大和インベスター・リレーションズ株式会社「インターネットIR・ベスト企業賞」
- ・日興アイ・アール株式会社「全上場企業ホームページ充実度ランキング」
- ・ゴメス・コンサルティング株式会社「Gomez IRサイトランキング」

2009年度 表彰企業の例 ※金融関連業を除く

表彰機関	表彰企業	表彰企業	表彰企業	表彰企業	表彰企業	表彰企業	表彰企業
日本証券アナリスト協会「証券アナリストによるディスクロージャー優良企業選定」	旭化成	アールシステム	住友化学工業	日本電産	日立製作所	大塚製作所	日本製鉄
東京証券取引所「ディスクロージャー表彰制度」	旭化成	アールシステム	住友化学工業	日本電産	日立製作所	大塚製作所	日本製鉄
日本IR協議会「IR優良企業賞」	旭化成	アールシステム	住友化学工業	日本電産	日立製作所	大塚製作所	日本製鉄
大和インベスター・リレーションズ株式会社「インターネットIR・ベスト企業賞」	旭化成	アールシステム	住友化学工業	日本電産	日立製作所	大塚製作所	日本製鉄
日興アイ・アール株式会社「全上場企業ホームページ充実度ランキング」	旭化成	アールシステム	住友化学工業	日本電産	日立製作所	大塚製作所	日本製鉄
ゴメス・コンサルティング株式会社「Gomez IRサイトランキング」	旭化成	アールシステム	住友化学工業	日本電産	日立製作所	大塚製作所	日本製鉄

分析方法

表彰企業とその企業に対応する同業種同規模の企業を比較することにより分析



資本コスト
株式収益率
有価証券報告書

分析結果

自発的開示が企業に与える影響

1. 『自発的開示は株主資本コストを低下させる？』

表彰企業は株主資本コストの低下傾向があり、その傾向は企業規模が下位になるほど強くなることが示唆された。 **数値情報を基にした実証分析**

2. 『自発的開示は株式収益率に正の影響？』

表彰企業は正の超過収益を獲得している傾向があり、特に、企業規模3分位の中位と下位において傾向が顕著

3. 『自発的開示の優良企業は、非優良企業と異なる特性を有する？』

テキスト情報を対象とした分析

103

配当政策について

有価証券報告書は株式上場企業などに提出が義務付けられており、また、配当政策は各企業によって表現が異なる。よって、有価証券報告書の配当政策は、企業の自発的開示を分析するための最適なテキストとなり得る。また、本来、有価証券報告書は法定開示であるため自発的開示ではなく制度開示であるが、配当政策の表現方法は各企業に委ねられていることから、性格的に自発的開示に近いと判断したため分析対象とした。

元データの一例

表彰分類	企業規模	年度	業種	企業名	配当政策
Sample	上位	2004	輸送用機器	本田技研工業	当社は、株主の皆様に対する利益還元を経営の最重要課題のひとつとして位置づけ、グローバルな視野に立って世界各国で事業を展開し、企業価値の向上に努めております。 成果の配分にあたりましては、今後の資金需要などを総合的に考慮し、配当につきましては、長期的な視点に立ち連結業績を考慮しながら実施することを方針としております。 当期の配当につきましては、期末配当金を1株当たり23円といたしました。年間配当金では、中間配当金10円を合わせ42円といたしました。なお、次期の配当につきましては、中間配当金を4円増配の23円、期末配当金を23円、年間配当金では4円増配の46円とする予定であります。 内部留保資金につきましては、将来の成長に不可欠な研究開発や事業拡大のための投資および出資に充てることにより、業績の向上に努め、財務体質の強化をはかってまいります。また、資本効率の向上をはかるために、自己株式の取得も適宜実施し、今後とも株主の皆様のご期待にそぐべく努力してまいります。
Control	下位	2004	建設業	高松コンストラクショングループ	当社は、受注競争力と企業体質の強化を目指し、内部留保の充実による経営基盤の強化をはかるとともに、安定的配当を継続しつつ、業績に応じた利益還元をおこなうことを基本方針としております。 当期の配当につきましては、上記方針にもとづき、1株につき10円増配し、1株当たり23円の配当を実施することになりました。この結果、当期の配当性向は24.07%となりました。 なお、内部留保資金につきましては当社グループ全体における経営体質の向上の強化・充実に加えて今後の事業展開に役立てるとともに、株主還元としての自己株式の取得も弾力的に行うなど、中長期的な視点で投資効率を考慮して有効に活用してまいります。

3-1-1. 基本的情報(全データ)

基本情報		品詞別出現回数		
総行数	464	品詞	出現回数	
平均行長(文字数)	231	1 名詞	28,569	
総文字数	2,980	2 動詞	3,754	
平均文長(文字数)	36	3 接頭詞	1,243	
流べ単語数	34,421	4 副詞	331	
単語種別数	2,958	5 連体詞	311	
		6 記号	121	
		7 形容詞	85	
		8 感動詞	5	
		9 格助詞	3	
		- 接頭詞	0	
		- 引用助詞	0	
		- 複合助詞	0	
		- 係助詞	0	
		- 終助詞	0	
		- 接続助詞	0	
		- 連用助詞	0	
		- 副助詞	0	
		- 玉虫助詞	0	
		- 並立助詞	0	
		- 連体助詞	0	
		- 助動詞	0	
		- 間投詞	0	
		- グループ	0	
		- カテゴリ	0	
		- 欠損データ	0	
		- その他	0	

3-1-2. 頻度解析(全データ)

佐藤/高橋[2011]

単語頻度解析			掛り受け頻度解析		
順位	単語	頻度	順位	係り元単語	係り先単語
1	取締役会	392	1	剰余金	配当
2	中間配当	372	2	取締役会	決議
3	基本方針	271	3	定款	定める
4	1株当たり	267	4	配当	決定機関
5	総務	239	5	1株	つく
6	期末配当	233	6	配当	行う
7	利益還元	233	7	配当	実施
8	当事業年度	215	8	総合的	増大
9	内部留保	213	9	安定的	配当
10	年2回	199	10	当事業年度	係る
11	安定的	188	11	中間配当	行う+できる
12	内部留保資金	186	12	当社	定める
13	決定機関	184	13	財務体質	強化
14	株主総会	169	14	配当	継続
15	事業展開	169	15	強化	図る
16	減価	157	16	内部留保	充実
17	期末配当金	154	17	経営基盤	強化
18	配当性向	154	18	決議	行う+できる
19	総合的	152	19	会計法第454条第5項	規定
20	利益配分	146	20	配当性向	なる
21	財務体質	137	21	企業体質	強化
22	配当金	131	22	当期	配当
23	設備投資	118	23	減価	償する
24	総務部長	118	24	当事業年度	償する
25	年間配当金	105	25	企業価値	向上

「取締役会」「中間配当」「剰余金」「安定的」「配当性向」「総合的」「企業価値」に着目して分析を進める。

3-1-3. 関連性分析①特徴語抽出(全データ)

佐藤/高橋[2011]

特徴語抽出とは
データに付随する属性ごとに、特徴的に出現する単語を抽出する分析のこと。

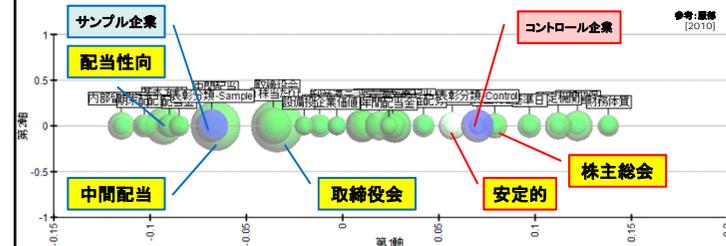
順位	サンプル企業	コントロール企業
1	連結業績	内部留保
2	連結ベース	法定繰上
3	基本方針	経営状況
4	長期的	財務状況
5	内部留保資金	財務体質
6	株主還元	安定配当
7	健全性	経営環境
8	中間配当	業績
9	研究開発	株主総会
10	収益向上	中長期的
11	取締役会	株主
12	積極的	安定的
13	直線的	剰余金
14	戦略的投資	期末配当
15	期末配当金	定時株主総会
16	中期的	企業体質
17	会計法第459条第1項	資金需要
18	連結当期純利益	株主利益
19	財務体質	投資的配当
20	フロー	利益配分
21	利益配当	株主各位
22	配当性向	株式分割
23	政経	ひとつ
24	当社グループ	年間
25	中間配当計画	利益

サンプル企業は「中間配当」「取締役会」「配当性向」と、
コントロール企業は「株主総会」「安定的」「剰余金」と関連性の高さが確認された。

3-1-4. 関連性分析②対応バブル分析(全データ)

佐藤/高橋[2011]

対応バブル分析とは
テキスト中の言葉や表現と属性の関係を、2次元または3次元空間上に分布させることにより、言葉や表現の使われ方が似ている場合は近くに配置され、それらの近さ遠さから、他の分析とは違った関係性が表される。

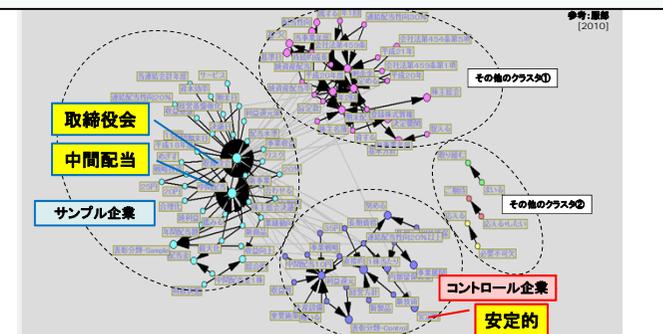


サンプル企業は「中間配当」「取締役会」「配当性向」と、
コントロール企業は「株主総会」「安定的」と関連性の高さが確認された。

3-1-5. 関連性分析③ことばネットワーク(全データ)

佐藤/高橋[2011]

ことばネットワークとは
テキスト全体から関連性の高い言葉同士をまとめて、いくつかのクラスタを作り、このクラスタをひとつの話題として捉えることにより、テキスト全体を大まかな話題ごとに分ける分析のこと。



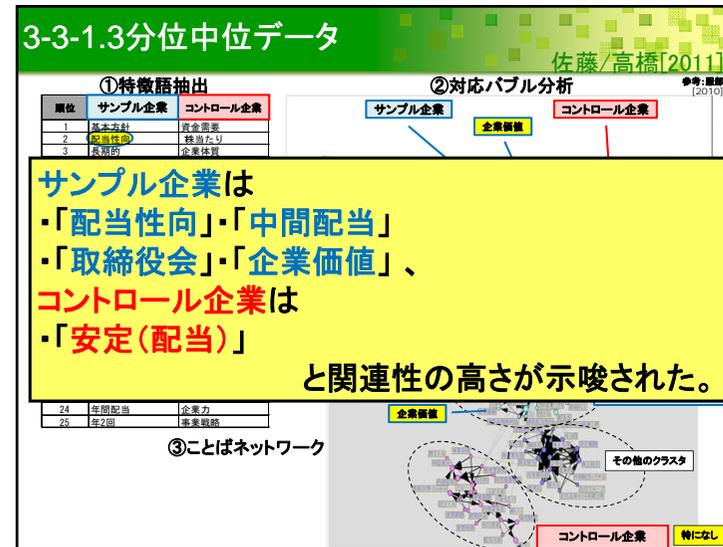
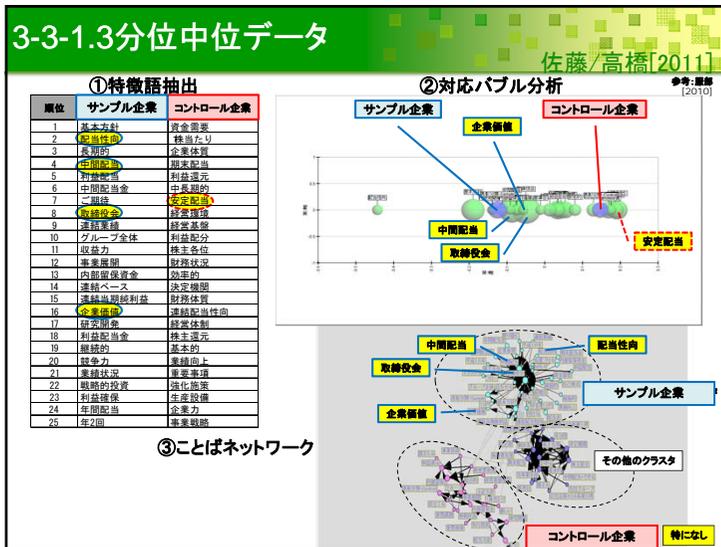
サンプル企業は「中間配当」「取締役会」と、
コントロール企業は「安定的」と関連性の高さが確認された。

3-1-6. 全データにおける分析のまとめ

佐藤/高橋[2011]

サンプル企業は、
・「中間配当」「取締役会」「配当性向」、
コントロール企業は、
・「安定的」「剰余金」「株主総会」
と関連性の高さが示唆された。

これらの結果から、サンプル企業は「配当性向」をより意識して「中間配当」を高く位置付けており、その背後には「取締役会」の存在が挙げられる。また、「取締役会」は「中間配当」を決議し、「株主総会」における期末配当の議案を練ることから、配当政策に関して、**サンプル企業はコントロール企業より強いコーポレートガバナンスの機能を**有していると推察される。
一方、**コントロール企業は安定的な企業姿勢**であり、「剰余金」を「安定的」に配当へ充てていることが推察される。



まとめ

自発的開示が企業に与える影響

- (1) 表彰企業は株主資本コストの低下傾向があり、その傾向は企業規模が下位になるほど強くなることが示唆された。
- (2) 表彰企業は正の超過収益を獲得している傾向があり、特に、企業規模3分位の中位と下位において傾向が顕著
- (3) 企業規模3分位の中位と下位の表彰企業は、相対的に強力なガバナンス機能を有している可能性

123

考察

企業規模、中位と下位の企業において、自発的開示が企業価値にプラスの影響を与えている可能性

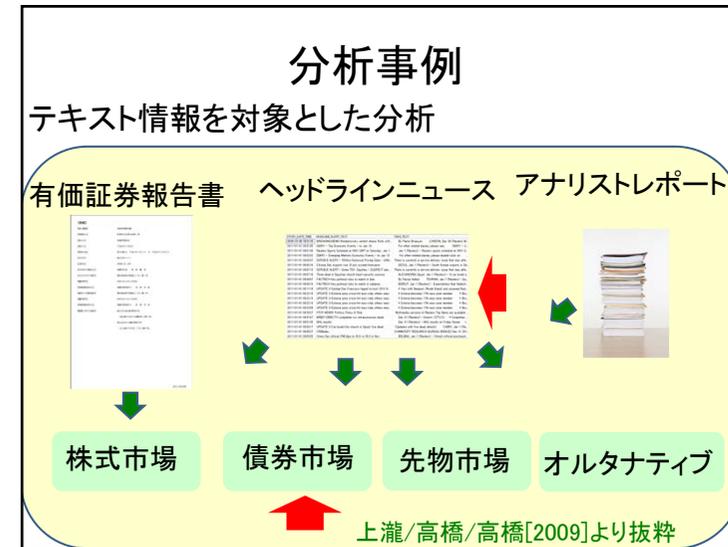
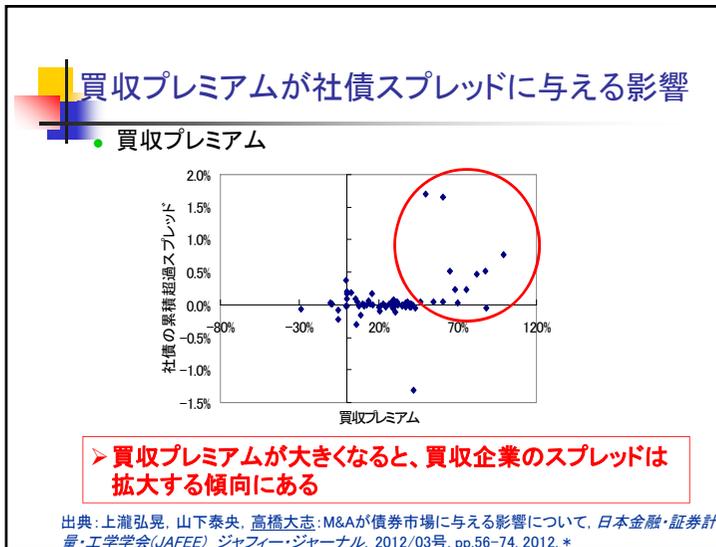
中堅中小企業は大企業よりも情報の非対称性が高いため、自発的開示により情報の非対称性を緩和する効果が高いことが一つの要因である可能性

企業規模、中位と下位の企業は、企業価値向上を念頭に強力なガバナンス機能を有し、その企業姿勢や機能が配当政策だけでなく自発的開示に大きな影響をもたらしている可能性

↓

中堅中小企業における自発的開示の意義と動機に関して興味深い示唆

124



- ### 研究目的
- 上瀧/高橋/高橋[2009]より
- ＞ クレジット市場においてヘッドラインニュースがもつ情報価値について分析
 - テキストマイニングの適用可能性を検討
 - ヘッドニュースがクレジット市場に与える影響を分析
 - 株価市場との関連性について分析
- 139

- ### 分析データについて
- ＞ 分析データ
 - 分析期間
 - 2006年8月10日～2006年11月24日
 - 分析対象企業
 - 東証1部上場企業のうち社債発行企業
 - 分析対象期間中に対象となる企業は185社
 - 分析対象ニュース
 - 期間中に時事通信から発信されたニュースの中から、分析対象企業の企業情報関連ニュースを選択
- 148

分析データについて

- 社債データ
 - 信用スプレッド
 - 社債利回りー国債利回り
 - 銘柄
 - 償還までの残存年数が3年以上5年未満
 - 同期間に複数銘柄が存在する場合は最も残存年数の長い銘柄
 - 時価
 - BPI時価

信用力悪化

信用力向上

信用リスクが高まるとスプレッドは拡大

150

分析データについて

- CDSデータ
 - スプレッド
 - Itraxx Japan
 - 銘柄
 - 5年

151

目次

- 研究目的
- 研究の背景
- 分析データ
- 分析方法
- 分析結果
- まとめ
- 今後の取り組み

152

分析方法①

- ニュースの分類
 - ヘッドラインニュースを内容により、「Good News」、「Bad News」、「Neutral News」の3つに自動分類
 - ニュースの自動分類には、テキストマイニングで用いられる、「ナイーブベイズ法」を使用
 - 「ナイーブベイズ法」は直感的な解釈が容易なテキスト分類法であり、かつ比較的高い分類精度を有する

153

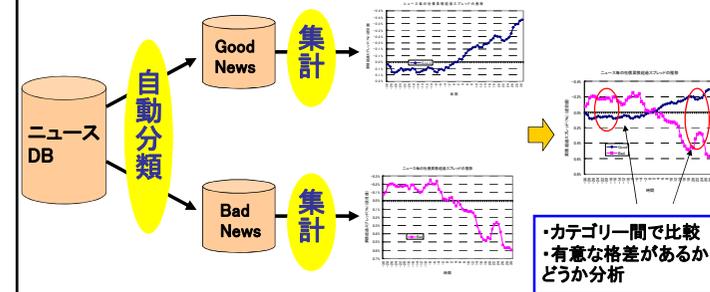
分析方法②

社債、CDSスプレッドの測定

- ニュースが発信されたタイミングを基準とし、対象企業のスプレッドを計測
 - スプレッドは、分析対象全銘柄のスプレッドの単純平均を計測し、その単純平均に対する超過スプレッドを分析対象とした
 - スプレッドが拡大 ⇔ 信用リスクが拡大
- ニュースのタイプにより、超過スプレッドに有意な差が発生するかどうか、統計的検定により分析
 - 検定には、Welch検定を使用

154

分析方法③



155

分析方法④

ナイーブベイズによる文章自動分類アルゴリズム

- テキストデータを特徴付けるキーワードがn個と仮定し、テキスト内にキーワードが存在するかどうかを用いて、テキストデータをベクトルで表現

$$\mathbf{x} = (x_1, x_2, \dots, x_n)$$

x_i はテキストデータの中にi番目のキーワードが存在するかどうかを表したもの

- 分類するクラスをcと仮定.
- テキストxが与えられた条件のもとで、xの属するクラスは、下式で求まる事後確率を最大化するクラスを求めることにより決定

$$\hat{c} = \arg \max_c (P(c | \mathbf{x})) = \arg \max_c \left(\frac{P(\mathbf{x} | c)P(c)}{P(\mathbf{x})} \right) = \arg \max_c (P(c)P(\mathbf{x} | c))$$

$P(\mathbf{x})$ はすべてのクラスで一定であること、およびベイズの定理を用いて算出

156

分析方法⑤

ナイーブベイズによる文章自動分類アルゴリズム

- 各キーワードの発生は、互いに独立であると仮定(ナイーブな仮定)すると、 $P(\mathbf{x} | c)$ は以下のように展開

$$P(\mathbf{x} | c) = P(x_1, x_2, \dots, x_n | c) \approx \prod_{i=1}^n P(x_i | c)$$

- xの属するクラスは、以下の式で表現

$$\hat{c} = \arg \max_c (P(c | \mathbf{x})) = \arg \max_c (P(c) \prod_{i=1}^n P(x_i | c))$$

- 学習データを選び、各クラスの相対頻度から $P(c)$ を決定し、また各クラスにおける x_i の相対頻度から $P(x_i | c)$ を決定
- テストデータを用いて、分類アルゴリズムの精度を検証
- 学習データとして、分析対象テキストデータの約3%にあたる、400のニュースを抽出し、またテストデータも同様に、分析対象テキストデータから400のニュースを抽出

157

分析結果

- テキスト分類アルゴリズムの作成
- 社債スプレッドを用いた分析
- CDSスプレッドを用いた分析
- 格付毎の社債, CDSスプレッド比較
- 市場間(社債, CDS, 株式)のリターン分析
- 社債と株式の関係

159

分析結果 ～テキスト分類アルゴリズムの精度～

- インサンプルテストによる分類
 - 学習データを用いて, 分類アルゴリズムを作成
 - 作成したアルゴリズムにより, 学習データを再度分類し, 精度を分析

		実際のタイプ					実際のタイプ		
		Good	Bad	Neutral			Good	Bad	Neutral
ベイズ学習結果	Good	179	24	24	ベイズ学習結果	Good	90.86%	24.74%	22.64%
	Bad	2	66	9		Bad	1.02%	68.04%	8.49%
	Neutral	16	7	73		Neutral	8.12%	7.22%	68.87%

161

分析結果 ～テキスト分類アルゴリズムの精度～

- アウトオブサンプルテストによる分類
 - 学習データを用いて作成したアルゴリズムを, テストデータに適用し, 分類精度を分析

		実際のタイプ					実際のタイプ		
		Good	Bad	Neutral			Good	Bad	Neutral
ベイズ学習結果	Good	168	27	35	ベイズ学習結果	Good	90.81%	30.00%	28.00%
	Bad	3	58	5		Bad	1.62%	64.44%	4.00%
	Neutral	14	5	85		Neutral	7.57%	5.56%	68.00%

162

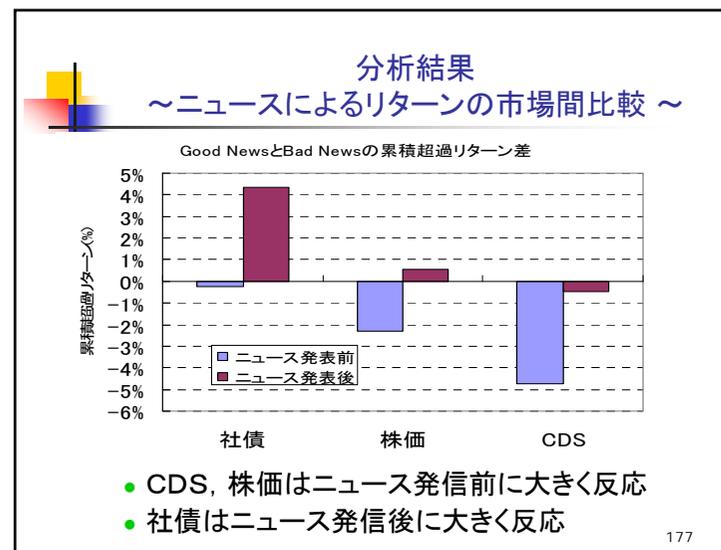
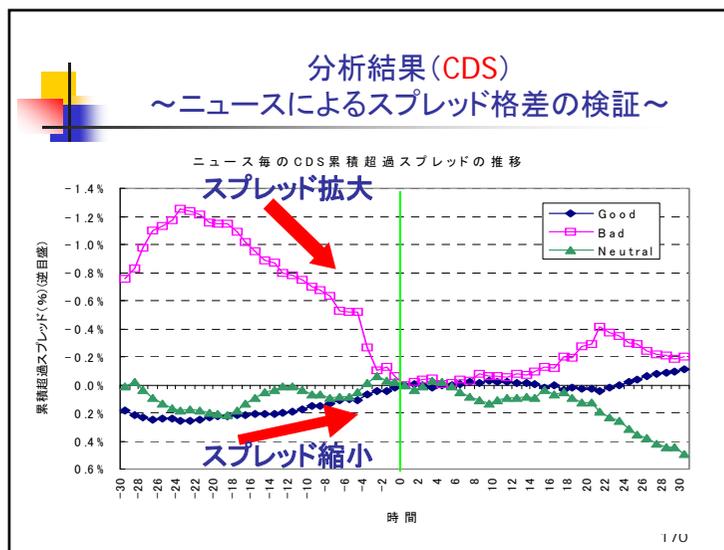
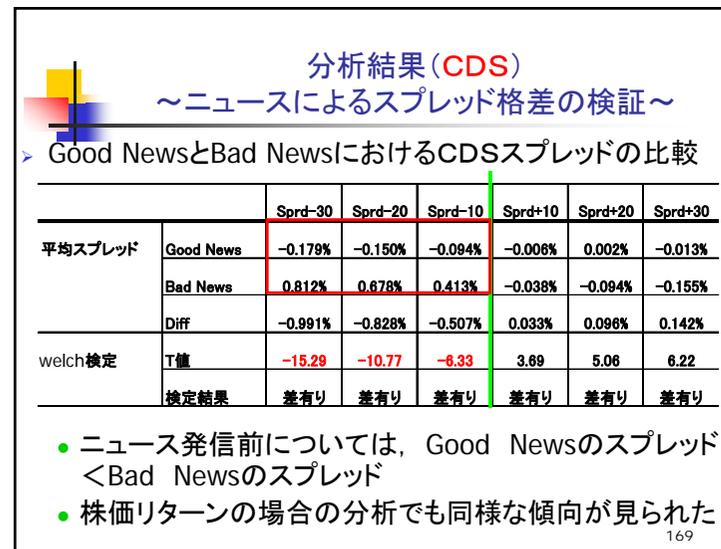
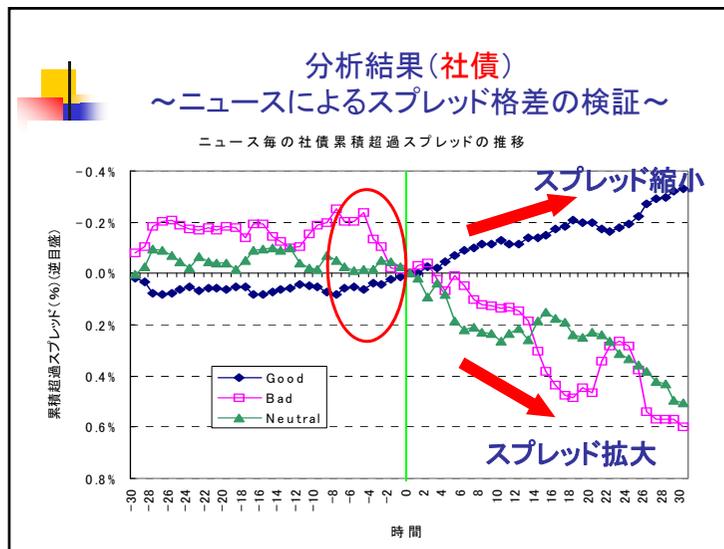
分析結果(社債) ～ニュースによるスプレッド格差の検証～

- Good NewsとBad Newsにおける社債スプレッドの比較

		Sprd-30	Sprd-20	Sprd-10	Sprd+10	Sprd+20	Sprd+30
平均スプレッド	Good News	-0.057%	-0.056%	-0.051%	-0.071%	-0.117%	-0.159%
	Bad News	0.158%	0.154%	0.156%	0.056%	0.201%	0.280%
	Diff	-0.214%	-0.210%	-0.206%	-0.127%	-0.317%	-0.440%
welch 検定	T値	-19.91	-14.39	-7.55	-5.06	-7.32	-10.88
	検定結果	差有り	差有り	差有り	差有り	差有り	差有り

- Good Newsのスプレッドは縮小(Bad Newsは拡大)
- すべての計測時点において, Good NewsがBad Newsより有意にスプレッドが小さい

165



分析結果 ～社債スプレッドから株価の影響を控除～

社債のニュース発信後の超過スプレッドがニュース発表前の株価超過リターンでどの程度説明できるか分析

$$ExcessSpread_{社債,+30day} = \beta \cdot Excess Return_{株価,-30day} + \varepsilon$$

説明変数	β	t
累積株価リターン	-0.0032	-22.2
R-square	0.07	
Adjusted R-square	0.07	
N	6172	

- 株価の累積超過リターンの係数は有意

株価で説明できない部分をヘッドラインニュースで説明可能？

180

分析結果 ～株価の影響を控除した後の社債スプレッド～

Good NewsとBad Newsの累積超過スプレッド差

銘柄格	株価リターンの影響排除前	株価リターンの影響排除後
BBB以下	-5.0%	-3.5%
A	-1.0%	-0.5%
AA以上	-0.5%	-1.0%

- BBB格以下の銘柄については株価リターンによる調整後もヘッドラインニュースの効果が残る

181

まとめ

- ナイーブベイズ分類法により、ヘッドラインニュースを自動分類し、約80%の分類精度を持つアルゴリズムを構築
- Good News, Bad Newsに分類されたニュースの社債スプレッドに違いが見られた
- 社債スプレッドはニュースが発信された後に大きく反応する
- 社債スプレッドを、株価リターンで調整した後も、ヘッドラインニュースの効果が残っている可能性がある

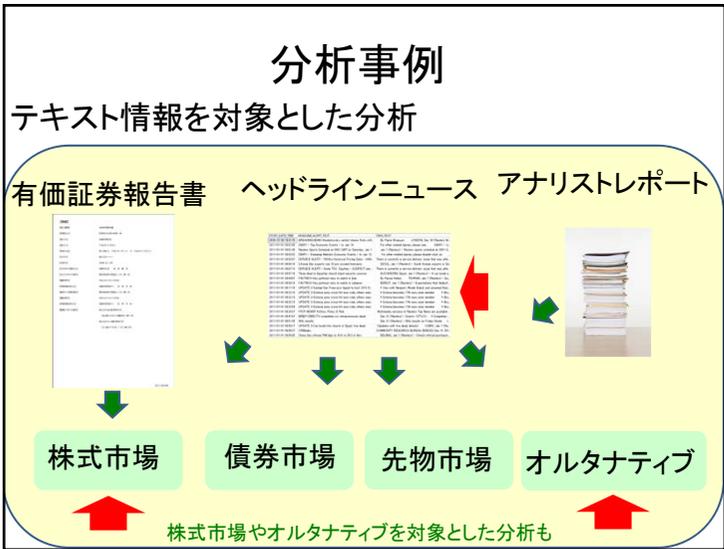
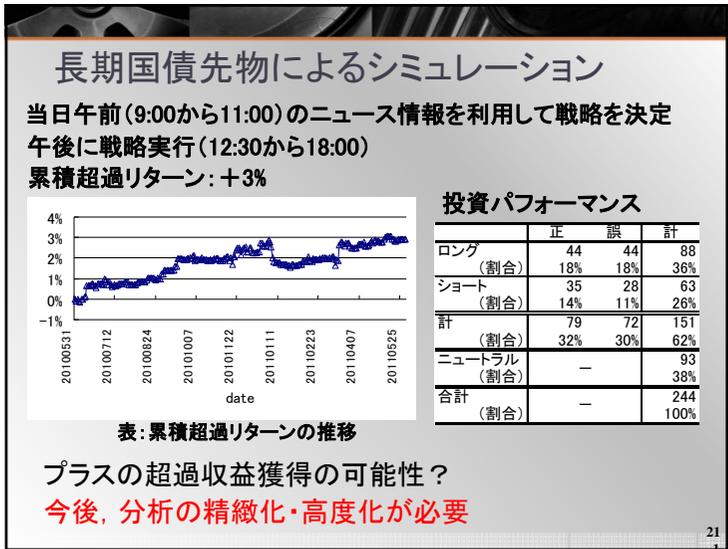
184

分析事例

テキスト情報を対象とした分析

有価証券報告書 ヘッドラインニュース アナリストレポート

山下/上瀧/高橋[2012]より抜粋



- ### まとめ
- テキスト情報を対象とした分析事例の紹介
 - ◆ 有価証券報告書
 - ◆ 金融市場に関する分析
- テキスト情報も実証分析の対象
- 分析の位置づけを認識しておく必要
- いくつかの注意点

御清聴ありがとうございました