

株主提案の要領におけるトピック別の文書特徴量と賛成率

久多里桐子

北九州市立大学経済学部

kirikokudari@kitakyu-u.ac.jp

概要

本研究は、株主提案の要領に関する基礎研究を行うことを目的とする。株主提案の要領である3つのテキスト(提案内容、提案理由、取締役会の意見)を独立した文書とみなし、各文書をトピックモデル(LDA)を用いて分類した上で、各トピックにおける文書の特徴量(文字数および文字数に占める数字の割合)と賛成率の関係を分析した。2011年～2019年の定時株主総会における株主提案議案1,350件を対象に分析を行った結果、提案内容や提案理由のトピックによって、得られる賛成率が異なることを確認した。本研究の分析結果は、他の株主が株主提案の要領に基づいて議決権を行使していることを示唆する。

1 研究の目的と意義

本研究は株主提案の要領に関する基礎研究を行うことを目的とする。2021年3月に施行された改正会社法では、株主が提出しようとする議案の数は10件までに制限された。一方で、株主提案の内容に制限を設けることは見送られた。しかしながら、2019年に公表された要綱案の段階では、内容に関する制限も含まれていた。

具体的には、提案が「専ら人の名誉を侵害し、人を侮辱し、若しくは困惑させ、又は自己若しくは第三者の不正な利益を図る」ことを目的としていた場合や、「株主総会の適切な運営が著しく妨げられ、株主の共同の利益が害されるおそれがあると認められる場合」には、株主提案権の行使を制限することが予定されていた([1])¹⁾。

その後、国会における法案審議において上記内容が削除された理由は、「権利の濫用に該当する株主提案権の類型についてさらに精緻に分析を深めなが

1) 株主提案権には議題提案権(会社法第303条)、議案提出権(同法第304条)、議案要領通知請求権(同法第305条第1項)が含まれる。目的等による議案の提案に対する制限は、議案提出権および議案要領通知請求権を対象としている。

ら、引き続き検討していくべき」([2], p. 5)であるからだ。つまり、2021年の改正会社法において、株主提案の内容に関する制限が見送られたことは、関連する知見が少ないことに起因する。そこで本研究は、株主提案権に関する今後の検討材料に資するという意義の下、株主提案権の内容に焦点を当てた分析を行う。

つづく第2節では、米国および日本の株主提案権について、提案内容および記述に関する制度上の規定について参照し、関連する先行研究を紹介する。第3節では、分析手法とサンプルについて述べる。第4節では分類したトピックの概要を確認し、第5節で主たる発見事項を提示する。最後に、第6節で本研究の要約を述べる。

2 制度の概要と先行研究

2.1 制度の概要

米国の株主提案権は制限が強い。例えば、提案内容が取締役の選任に関するものであったり、具体的な配当金額に関するものであったりした場合、企業側は提案を排除できる²⁾。また、記述量も補足説明を含めて500語以内に限られる。

一方で、日本では株主が提案できる内容は多様であり、記述量に関する法的な制限もない。一般的には、株主が記述した提案内容や提案理由の原文に加えて、議案に対する取締役(取締役会設置会社の場合は取締役会)の意見(以下、取締役会の意見)が、株主総会招集通知内の株主総会参考書類に記載される³⁾。

2) 17 C.F.R. § 240.14a-8 (2003)

3) ただし、提案の理由が明らかに虚偽である場合または専ら人の名誉を侵害し、若しくは侮辱する目的によるものと認められる場合は提案理由を削除されることもある。また、企業側が株主提案を拒否できる事由として明文化されているものは、法令もしくは定款に違反する場合や、実質的に同一の議案につき総株主の議決権の10の1以上の賛成を得られなかった日から3年を経過していない場合が該当する(会社法第304条ただし書き、第305条第4項)。

2.2 先行研究

本研究に関連した分析として、以下3つの研究を挙げる。[3]は、EDGARから取得した2004年から2013年の期間に提出されたForm 8-Kのうち、所定の要件を満たす73,986件を20のトピックに分類し、分類したトピックによって投資家の反応が異なることを確認している⁴⁾。

[5]では、2003年から2012年までの新聞記事77,814件を30のトピックに分類し、日ごとのトピックの割合と市場リターン(TOPIX)との間の関係を分析している。分析を通じて、トピックの割合と市場リターンの間に相関は見られなかった一方で、一部のトピックの割合と市場ボラティリティの間に相関があることを発見した。

[6]では2004年から2021年までの所定の要件を満たす有価証券報告書43,289件をもとに、MD&A(経営成績等の分析)の記述について分析を行い、近年のディスクロージャー制度改革がMD&Aの開示に影響を及ぼしていることを報告している。

3 分析手法とサンプル

3.1 分析の概要

分析の概要は以下の通りである。はじめに、株主提案の要領である3つのテキスト(提案内容、提案理由、取締役会の意見)を独立した文書とみなし、各文書の特徴量を取得する。つづいて、各文書をトピックに分類する。本研究の主たる分析として、分類したトピックごとに、文書の特徴量と賛成率の関係を確認する。

本研究では、文書の特徴量として以下2つを採用した。第1に、文字数である。定款で文字数に制限を設けている企業に対して、当該定款の削除を求める提案が提出される事例が存在していることから明らかに、記述量は賛成率に影響を与えると考えられる。第2に、文字数に占める数字の割合である。特に剰余金の分配に関する提案など、特定の議案においては具体的な数字を用いた提案の方が説得力が高まると考えられる。

以下は実際の提案理由に基づく特徴量であるが、例Aよりも例Bの方が賛成率は高いと予想する。

4) 米国上場企業は、投資家の意思決定に影響を及ぼす可能性のある重要な事象(決算発表、重要な契約の締結、倒産など)が生じた際、4営業日以内にForm 8-k(current report)を米国証券取引委員会(SEC)に提出する必要がある([4])。

例A:増配できる十分な利益がある。

– 文字数: 14 字

– 文字数に占める数字の割合: 0%

例B:貴社は、平成30年9月末の連結貸借対照表にて現預金及び投資有価証券の合計額350億円を有し、(省略)妥当なものであると考えます。

– 文字数: 278 字

– 文字数に占める数字の割合: 6.47%

議案からテキストを抽出する際の留意事項を下記に示す。提案内容が取締役等の役員選任議案の場合、当該候補者の経歴等は提案内容に含めない。また、会社注も含めない。さらに、提案内容と提案理由が明確に区別されていなかった場合、提案理由の記述はないものとみなしている⁵⁾。

取締役会の意見が複数の議案に対する集約的なものであった場合、各議案に同じ記述がなされたものとして扱った。ただし、個別の議案に対する意見に加えて、別立てで議案全体に対する取締役会の意見が付されていた場合は、個別の議案に対する意見のみを取締役会の意見として収集している。

3.2 前処理と形態素解析

トピックモデルの構築にあたり、下記の通りに前処理および形態素解析を行った。表記の統一を目的として、全角コロン「:」を読点「、」に、全角ドット「.」を句点「。」に変換した上で、全角の英数字と記号(括弧や%など)を半角に変換した。数値に関しては桁数区切りのコンマを除いた上で、0に置き換えた。そのため、例えば「100万円」は「0万円」、「10,000人」は「0人」に置き換わっている。

形態素解析にあたっては解析器にMeCab([7])を、辞書にmecab-ipadic-neologd([8])を利用した。品詞は名詞、形容詞、動詞を採用し、ストップワードとしてSlothLib([9])に頻度の高い単語を追加した⁶⁾。

3.3 LDAモデルの構築

トピックの分類においては、潜在的ディリクレ配分法(Latent Dirichlet Allocation; LDA([10]))によるトピックモデルを採用した。LDAモデルの構築において、辞書は出現文書数が13件(サンプルの1%)以下の単語および20%以上の文書に出現する単語は

5) 提案内容と混在している場合、可能であれば文章を分割した。

6) 具体的には、平成、昭和、令和、当社、貴社、会社、株主、提案、議案をストップワードとして採用している。また、ひらがな1、2文字の単語も除外した。

表 1 提案内容の各トピックにおける代表的な単語

トピック	単語	件数
Topic 1	報酬・規定	364
Topic 2	社外取締役	197
Topic 3	燃料・処理	279
Topic 4	変更	274
Topic 5	選任・候補者	247
Topic 6	配当	240

除いた⁷⁾。コーパスの作成においては、TF-IDF による重み付けは実施しない。

トピック数に関して、提案内容は 6、提案理由は 6 (ただし、テキストが取得できなかった議案のみの Topic 7 を設定)、取締役会の意見は 5 に設定した⁸⁾。

3.4 サンプルとデータの取得源

分析対象は 2011 年～2019 年の定時株主総会において提出された株主提案議案 1,388 件である。ただし、賛成率を含めた分析においては、付議の要件を満たさなかったなどの理由で決議されなかった議案を除くため、1,350 件に減少する。

データの取得源は、以下のとおりである。株主提案の要領に関しては、eol (株式会社アイ・エヌ情報センター) から取得した株主総会招集通知の PDF から手作業でテキストを収集した。賛成率については日経 NEEDS-Cges 株主総会関連データ (日経メディアマーケティング株式会社) から取得している。

4 トピックの概要

表 1 および表 2 は、提案内容および提案理由の各トピックについて、頻出上位 5 単語から任意に選定した代表的な単語を示している⁹⁾。また、図 1 は提案内容と提案理由のトピックごとの該当件数を示している。図 1 において最も件数が多いのは、提案内容 Topic 6 と提案理由 Topic 5 のセルである。表 1・2 の単語をもとに解釈するならば、(現金などの)「保有」を理由とする「配当」提案が、比較的多く提出されていると考えられる。

7) 例えば「取締役会の意見」のテキストをもとに作成した辞書からは、「反対」「取締役会」などの定型的な単語が除かれる。

8) トピック数を 2～30 の範囲でシミュレーションを行い、Coherence^[11] および Perplexity の値をもとに決定した。また、コーパスを TF-IDF による重み付けを実施した場合と実施しなかった場合の両方でシミュレーションを行った結果、重み付けを実施した場合の方が Perplexity が著しく高い値であった。そのため、コーパスの作成においては、TF-IDF による重み付けは実施しないこととした。

9) 取締役会の意見の分類は、実質的に提案内容に基づく分類となっている。そのため、取締役会の意見については記載を割愛する。

表 2 提案理由の各トピックにおける代表的な単語

トピック	単語	件数
Topic 1	利益	337
Topic 2	株式	374
Topic 3	役員	399
Topic 4	株主総会	311
Topic 5	保有	318
Topic 6	原発	327
Topic 7	—	15

(注) Topic 7 は提案理由として明確な記述がなかったものを分類している。

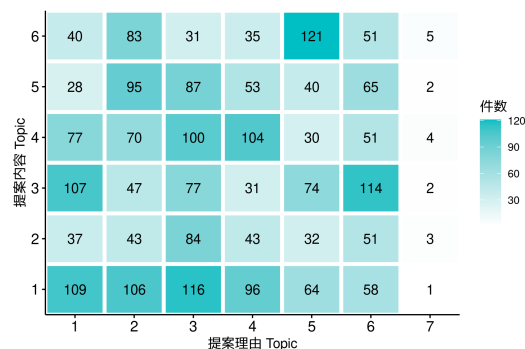


図 1 提案内容トピックと提案理由トピックの対応関係

5 分析結果

はじめに、株主提案議案をトピックに分類する前の、文字数ないし文字数に占める数字の割合と賛成率の関係について確認する。図 2 はサンプルの 1,350 議案について、文字数と賛成率の散布図と、被説明変数を賛成率、説明変数を文字数の自然対数とした回帰直線を、文書別に示している¹⁰⁾。

図 2 が示すように、取締役会の意見については、記述量と賛成率の間にはプラスの関係がある (回帰係数は 2.098)。つまり、賛成率が高いと予想される場合に、取締役会の意見の記述量が増える可能性があると考えられる。

一方で、提案内容および提案理由については、記述量が増えると賛成率は低くなる傾向が見られる (回帰係数はそれぞれ -1.064 と -1.749)。他の株主にとっては、記述量が増えるほど提案内容等を吟味するためのコストが高くなるため、賛成票を投じる可能性が低くなるのかもしれない。

類似した傾向は、図 3 でも確認できる。図 3 は 1,350 議案について、文字数に占める数字の割合と賛成率の散布図と、被説明変数を賛成率、説明変数を文字数に占める数字の割合とした回帰直線を、文

10) 外れ値を除くため、文字数が 0 字以上 4,000 字未満である文書に限定した。

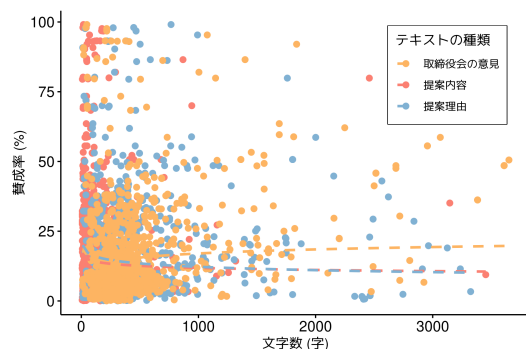


図2 文書別の文字数と賛成率

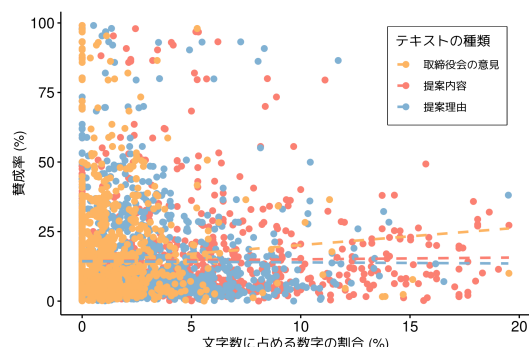


図3 文書別の文字数に占める数字の割合と賛成率

書別に示している¹¹⁾。回帰係数はそれぞれ、提案内容が 6.566、提案理由は -4.106、取締役会の意見が 62.700 である。

上記の発見事項は、議案を分類していない場合の株主提案に関する全体的な傾向である。表3と表4では、提案内容を6つのトピックに分類した場合の文書の特徴量と賛成率の関係を確認できる。具体的には、表3は被説明変数を賛成率、説明変数を文字数の自然対数とした単回帰の係数を示している。また、表4は被説明変数を賛成率、説明変数を文字数に占める数字の割合とした単回帰の係数を提示している。

表3・4から、トピックによって文書の特徴量と賛成率の関係は異なることがわかる。注目すべきは、表4において、特に Topic 6 の提案理由で記述に占める数字の割合が高いほど賛成率が高いという点である。Topic 6 の代表的な単語が「配当」であることに鑑みると、提案理由として会計数値等を用いた方が、高い賛成率を得られると考えられる。

最後に、図4において、提案内容と提案理由のトピックごとの、賛成率の中央値を示す。まず、提案理由の Topic 7 で高い賛成率が確認できる。これは、

表3 提案内容トピック別の各文書と文字数の関係

	提案内容	提案理由	取締役会の意見
Topic 1	0.319	-0.242	2.170
Topic 2	-1.658	-10.370	0.695
Topic 3	0.888	2.883	1.647
Topic 4	-1.782	-3.442	4.367
Topic 5	-2.732	4.358	4.076
Topic 6	3.441	0.925	2.642

表4 提案内容トピック別の各文書と文字数に占める数字の割合の関係

	提案内容	提案理由	取締役会の意見
Topic 1	0.1944	0.3532	0.373
Topic 2	0.3858	-0.9396	2.433
Topic 3	0.0058	-0.0951	0.943
Topic 4	-0.2588	0.4061	2.403
Topic 5	0.6036	0.0051	1.924
Topic 6	0.1308	0.5230	-0.184

支配株主が可決を見越して十分な説明なしに提案を行っていることを示している。

つづいて、提案内容 Topic 5 と提案理由 Topic 4 のセルや、提案内容 Topic 1 と提案理由 Topic 1 のセルで高い賛成率 (17.1% と 15%) を確認することができる。つまり、「利益」を理由とした「報酬」提案などが高い賛成率を得ている。分析結果は、提案内容や提案理由のトピックによって、得られる賛成率が異なることを示している。

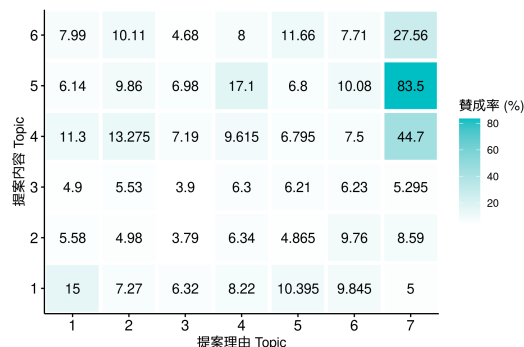


図4 提案内容と提案理由のトピック別賛成率(中央値)

6 本研究の結論

本研究では、株主提案の要領に関する基礎研究を行った。分析の結果、提案内容や提案理由のトピックによって、得られる賛成率が異なることを確認した。本研究で得られた結果は、株主提案の要領が他の株主の議決権行使に影響を及ぼしている可能性を示唆する。

11) 外れ値を除くため、文字数に占める数字の割合が 20% 未満である文書に限定した。

謝辞

本報告は JSPS 科研費 JP20K13642 の助成を受けた研究成果の一部である。また、本報告は日本経済会計学会第 3 回秋季大会 (2022 年 12 月 17 日開催) の自由論題報告「自然言語処理を用いた株主提案の要領に関する分析」に一部修正を施したものである。

参考文献

- [1] 法務省. 会社法制 (企業統治等関係) の見直しに関する要綱案, 2019. 法制審議会会社法制 (企業統治等関係) 部会 第 19 回会議 (2019 年 1 月 16 日開催) <http://www.moj.go.jp/content/001279742.pdf>.
- [2] 竹林俊憲, 遑英基, 坂本佳隆, 蘭牟田泰隆, 青野雅朗, 若林功晃. 令和元年改正会社法の解説-ii-. 旬刊商事法務, Vol. 2223, pp. 4–10.
- [3] Stefan Feuerriegel and Nicolas Pröllochs. Investor reaction to financial disclosures across topics: An application of latent dirichlet allocation. **Decision Sciences**, Vol. 52, No. 3, pp. 608–628, 2021.
- [4] U.S. Securities and Exchange Commission. How to read an 8-k, 2021. <https://www.sec.gov/oiea/investor-alerts-and-bulletins/how-read-an8-k>.
- [5] 竹澤譲, 伊藤真利子, 大西立顕. Lda で求めた新聞記事のトピック割合と株価指数の関係性. 人工知能学会第二種研究会資料, Vol. 2021, No. FIN-026, pp. 1–8, 2021.
- [6] 中野貴之, 五十嵐未来, 湯浅大地. Md&a 情報における『トピック』の分析. 証券アナリストジャーナル, Vol. 60, No. 10, pp. 27–35, 2022.
- [7] 工藤拓. Mecab: Yet another part-of-speech and morphological analyzer. <https://taku910.github.io/mecab/>.
- [8] 佐藤敏紀, 橋本泰一, 奥村学. 単語分かち書き辞書 mecab-ipadic-neologd の実装と情報検索における効果的な使用方法の検討. 言語処理学会第 23 回年次大会 (NLP2017), pp. NLP2017–B6–1. 言語処理学会, 2017.
- [9] Slothlib. <http://svn.sourceforge.jp/svnroot/slothlib/CSharp/Version1/SlothLib/NLP/Filter/StopWord/word/Japanese.txt>.
- [10] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. **Journal of machine Learning research**, Vol. 3, No. Jan, pp. 993–1022, 2003.
- [11] David Mimno, Hanna Wallach, Edmund Talley, Miriam Leenders, and Andrew McCallum. Optimizing semantic coherence in topic models. In **Proceedings of the 2011 conference on empirical methods in natural language processing**, pp. 262–272, 2011.