

選挙区レベルの質的な選挙情勢報道の統合と これを用いた2021年衆院選の選挙結果予測

Aggregating Qualitative District-Level Campaign Assessments to Forecast Election Results of the Japanese Lower House Election in 2021

梅田 道生¹

〔要旨〕

今日の先進民主主義国においては複数の選挙世論調査の結果を重みづけ平均等の手法を用いて統合し、これを選挙結果の予測のために用いているが、日本では従来こうした統合を行うことが困難であった。これは報道各社が法的な制約から選挙世論調査結果の数字そのものではなく、調査結果等に依拠した質的な選挙区の選挙情勢（優勢や接戦、不利など）のみを報道するためである。本稿ではUmeda（2023）が2017年衆院選の小選挙区での選挙結果予測のために用いた手法、すなわち報道各社の情勢報道を項目反応理論の段階反応モデルを用いて統合し、これを予測に利用する手法を2021年衆院選に対して用い、この手法が2017年衆院選以外でも有用かを検証した。この結果、この手法は2021年衆院選の選挙結果もかなり正確に予測できたことが明らかとなった。2021年衆院選では報道各社の一部が与党の獲得議席数を過小に評価したために、本稿の予測モデルはこれらの評価の影響を受け与党の獲得議席を一定程度低く予測した。他方小選挙区レベルでの勝敗予測に関しては、予測モデルは個別の報道各社の予測で最も正確なものを上回る成果を上げた。これらは本稿で用いる手法の長所と短所、すなわち複数の調査結果を統合することにより、それぞれの調査の標本誤差等を打ち消すことができる一方で、予測が各社の平均に近いものとなるため、調査法によるバイアスや選挙戦終盤のスイングの影響を受けやすいことを示したものと言える。

キーワード：選挙予測、世論調査統合、日本政治、項目反応理論

1 駒澤大学グローバル・メディア・スタディーズ学部

1. 導入

本稿の目的とは、1) 日本の報道各社が選挙期間中に行う衆議院小選挙区レベルでの選挙情勢報道の情報を統合すること、2)この統合された情報を用いて選挙結果を予測する手法を提示すること、また3) Umeda (2023) が第48回衆議院議員総選挙（以下2017年衆院選）の選挙結果の予測で用いた本手法の有効性を第49回衆議院議員総選挙（以下2021年衆院選）で再検証することにある。

近年複数の世論調査の結果を統合し、この統合した情報を用いて選挙結果を予測する世論調査統合（Poll Aggregation）の手法が幅広く利用されるようになってきている。この背景としては、さまざまな技術革新やその結果である調査の低廉化等により、投票日以前に有権者の投票意図を探る世論調査が数多く行われるようになったことがある（Traugott 2014）。もともとは米国の大統領選を中心にこの手法は発達したが、今日は他の先進民主主義国でも幅広く用いられている（例、ドイツを対象とするGraefe 2019）。この手法は不特定多数の参加者が結果予測を行う予測市場、また選挙前年の経済成長率などのマクロ的な要因から選挙結果の予測を試みる手法と並び、選挙予測における主要なアプローチの一つである（Lewis-Beck & Stegmaier 2014）。

しかし従来日本においてはこの世論調査統合の手法は用いられて来なかった。これは日本のマスメディアが選挙期間中に世論調査を行わないからではない。日本の大手報道各社は毎回の国政選挙で選挙区レベルでの選挙結果予測を目的とした大規模な調査を実施し、この結果に取材の結果等を加味して情勢報道を実施している。ただし法的な問題から、諸外国のように「候補Xへの投票意向が44%、候補Yへは40%」等の世論調査の数字を直接報じることはなく、これに代わり「X候補が優勢」「Y候補とZ候補は接戦」といった質的な評価のみを報道している。この結果として、日本における選挙予測を目的とした研究は、マスメディアの選挙戦報道や選挙の世論調査に依存しないもの（例えば、Lewis-Beck & Tien 2012）を含め、ごく少数にとどまってきた。

これに対し、Umeda (2023) は項目応答理論を用い、日本における報道各社の選挙区レベルでの質的な情勢評価を統合する新しい手法を提案した。そしてこの手法を2017年衆院選に対して用い、選挙結果をかなりの精度で予測可能なことを示した。例えばこの手法を利用した予測モデルは与党候補の小選挙区での期待獲得議席数を226と予測したが、実際に与党が獲得したのは223議席であった。

この手法は、日本における選挙予測に貢献するだけでなく、より一般的に質的な選挙区的情勢評価を統合する方法を提示することで、世論調査の数字を統合する既存の手法と合わせ、選挙予測の精度を向上させることが期待される。例えば米国における選挙区レベルの選挙戦に対する専門家の評価であるCook Political Report (2020) は日本の報道各社の情勢報道と類似した質的な表現で提供されており、他の専門家評価との統合に本手法が応用可能であろう。

しかしこのUmeda (2023) の手法が、2017年衆院選という特定の選挙の文脈を超えてどの程度有効かは明らかではない。2017年衆院選は例えばその前の2014年衆院選と比較して選挙戦の最終段階における情勢の変動が小さく、報道各社の情勢報道（の平均）が実際の選挙結果に近いところにあった幸運な事例とみなすことも可能である。そこで本稿では同様の手法を2021年衆院選に応用し、この手法の有効性を改めて検証することを目的とする。

2. 背景および先行研究

選挙情勢調査を含む社会調査において、調査対象である母集団の構成員全員に対する全数調査を行うことは（国勢調査などの少数の例外を除けば）ほぼ無い。全数調査に代わり、通常は母集団から数百ないし数千程度の調査対象者を抽出する標本調査を行うが、標本調査ではある標本での平均が母集団での「真の」平均から一定の範囲でずれる標本誤差が生じることは避けられない。これに対し、母集団から標本を複数回抽出して標本平均を平均した場合には、標本の誤差が互いに打ち消し合い、より母集団での平均に近い値が得られると期待できる。

また調査実施の際には、調査実施機関により異なる対象者を抽出する枠や回答拒否、あるいは質問の文言や順番等の影響から非標本誤差が生じうる。すべての実施機関が同じ方向に非標本誤差を持つ可能性もあるものの、複数の実施機関の調査結果を利用することで、通常は各機関の調査が持つバイアスを緩和することも期待できる。このため、標本誤差のみならず非標本誤差を減少させ、より「真の」値に近い調査結果を得るためには、複数調査の結果を統合して用いることは有益であると考えられる。

情勢報道が示す情報を統合し、これをもとに選挙結果の予測を図る先行する試みとしては、米国の大統領選やその他の選挙における試みを挙げることができる。米国では1930年代から選挙世論調査を実施し、これに基づき大統領選挙等の選挙結果を予測する試みがなされてきた（例. Gallup 1951）。しかし1980年代以降、コンピュータと電話の利用による世論調査の低廉化や効率化により、選挙世論調査の数そのものが大きく増加した。この結果として2000年代からReal Clear PoliticsやFiveThirtyEight等の世論調査データを収集して選挙戦の「現状」についてまとめ、選挙結果を予測するウェブサイトがいくつも生まれた。

米国大統領選は全国での一般得票の多寡により当選候補が決定されるのではなく、各州での得票に基づき（ほとんどの場合には勝者総取り形式で）州に配分された選挙人を候補へと配分し、より多くの選挙人を獲得した候補が当選する仕組みを持つ。このために、特に接戦が予想されるいわゆる「戦場州」の帰趨を予測するために数多くの選挙世論調査が実施される。また政党の公認候補決定のために同様に州ごとに予備選が実施されるため、実質的に1年半近くも選挙戦が繰り返されることになる。こうした中で、各州の有権者の動向は候補者や政党関係者に限らず、多くの人々の関心を集め、上記のウェブサイトの興隆につながった。

このようなウェブサイトにおける世論調査データの統合手法は多様である。例えば、Real Clear Politicsは2000年設立の政治ニュースサイトおよび世論調査データの収集サイトであるが、基本的には最新の数回の世論調査における有権者の投票意図の単純平均を用いて選挙戦の「現状」についてまとめ、選挙結果の予測を行っている（Jackson 2018）。

しかしこの単純平均という手法は問題含みである。例えば投票日直前と2週間前に実施された調査では、他の条件が等しければ前者のほうが最終的な結果の予測に有用と期待できるだろう。また標本規模の大小や調査対象枠、回収率等により調査結果の信頼性も同等ではないと考えられる。こうした問題に対処するために、例えば2006年設立の世論調査データの収集サイトであるPollster.com（現在ではHuffPost Pollster）では、調査手法について十分な情報を公開しない、あるいは適切な調査手法を用いていない世論調査を排除したうえで、局所的に推定された散布図平滑化（LOESS）を用い、直近の調査により大きな重みづけを行った（移動）平均を示している（HuffPost Pollster

n.d., Jackson 2018)。

また2008年に設立されたFiveThirtyEightは、各州で行われた世論調査の結果を独立に利用するのではなく、各州の有権者構成の特徴等の基礎的な情報と世論調査を組み合わせたより複雑な予測モデルを用いる。より具体的には、各調査会社の世論調査の数字に対し、これらが実施された時期だけでなく、過去に報告された結果の正確性や調査手法に対する透明性に依拠した調査の「信頼性」に基づく重みを用いた平均化を行う。さらにすべての州において信頼性の高い調査が頻繁に実施されるわけではないため、全国レベルや有権者の構成が類似した州における支持率の時間的推移を用いた補正も追加的に実施しているとされる(例. Jackson 2018; Silver 2020)

なお上記のHuffPost PollsterやFiveThirtyEight等のウェブサイトでは大統領選以外の選挙、特に連邦上院選や州知事選については世論調査結果を利用した予測が試みられている。ただし世論調査の実施数の少なさや全国的な関心の低さから、一般的に州知事選を扱うウェブサイトは少なく、さらに連邦下院選に関しては、個々の選挙区に対する世論調査数の少なさから予測は困難とされている(Jackson 2018)。

これに対して、日本は米国とは異なる政治および選挙制度、また独自の法規制からかなり異なった環境下にある。日本は二院制および議院内閣制を採用し、また二院を構成する衆議院と参議院ではそれぞれ多数決型と比例代表型の選挙制度を組み合わせた混合制選挙制度を用いている。本稿が分析の対象とする2017年と2021年の衆院選では、小選挙区制を用いて全国289の小選挙区(以下選挙区)から各1名合計289名を、また比例代表制を用いて全国11の選挙区(以下地域ブロック)から合計176名の議員を選出している。そのため選挙結果を予測するためには各選挙区における候補の勝敗、および11の地域ブロックにおける各政党の得票の予測が必要となる。

日本では衆議院の場合投票日の12日前の公示日から公式の選挙期間が始まるが、候補者も(公式には)この選挙期間中にのみ自らを当選させるための選挙活動を行うことができる。ゆえに報道各社の情勢報道もこの短い公式の選挙期間に集中する。またすべての選挙区に対する世論調査の実施には多額の費用を要することもあり、各社の実施する調査および情勢報道は選挙期間の序盤ないし中盤におけるほぼ1回のみ、少数の社が2度目の調査を(しばしば一部の選挙区に対してのみ)実施する程度である。

日本の選挙報道の特徴として、報道機関が世論調査結果の直接的な数字、例えば候補者Aへの投票予定者は44%、候補者Bは40%、未定が16%、のような形式で情勢報道を行うことはない点が挙げられるが、これは米国の一般的な報道とは大きく異なる。詳細については後述するが、日本の報道機関はそれぞれ世論調査を実施した上で、これを取材結果と合わせて「候補者Aが優勢」や「候補者BとCが接戦」、「候補者Dは独自の戦い」など質的な評価を用いて選挙区ごとの情勢報道を行う。ゆえに米国と同様の手法を用いて各社の評価を統合し、予測に用いることができない。

日本国内の報道機関による選挙報道自体については飯田(2007)、三春(2019)、福島(2015; 2018; 2022)などの先行研究が存在する。例えば飯田(2007)では2005年衆院選を事例とし、各報道機関の選挙情勢調査法や主要政党の獲得予測議席数、当落予測表現の差異についてまとめ、またそれぞれの予測の正確性についても論じている。また福島(2015; 2018; 2022)はそれぞれ2014年と2017年、2021年衆院選を扱うほかは、ほぼ同様の内容について論じている。しかしこれらの先行研究では、情勢報道情報の統合や、これを用いた選挙結果の予測手法等については論じられてい

ない。

なお選挙結果の予測については、2019年に早稲田大学において同年の参院選を対象にデータサイエンスコンペティションが実施されており、多様なデータを利用しこれを先進的な統計学的手法で分析する様々な予測モデルが検証された。しかしこのコンテストでは報道機関の情勢報道利用にはペナルティが設定されており、優勝チームも選挙区での結果予測に報道機関による情勢報道は利用していないため、本稿とは大きく異なるアプローチとなる（早稲田大学2019）。

3. 各社調査の概要と「正確性」

本稿では2017年および2021年衆院選の選挙期間中に実施された大手報道機関5社（読売新聞・朝日新聞・毎日新聞・産経新聞・時事通信）による選挙区の与党候補に対する選挙情勢報道を分析のために用いる。分析対象を与党候補に限るのは、少なくとも有力2候補が競合する選挙区における与党候補と野党候補に対する情勢判断（および誤差）は表裏一体であり、別々に扱う意味が薄いためである。

なお日経新聞は2017年と2021年共に独自の情勢報道を行っているが、調査は読売新聞と共同で実施しており、独自の判断も加わるものの情勢判断がかなり類似したものとなることが予想されるため、本稿では分析対象から外した。また2021年には共同通信は選挙戦序盤と終盤で調査を実施している模様（うち序盤は毎日新聞と共同）であるが、個別選挙区に対する情勢報道を全国レベルでは行っていない（少なくとも加盟各社が報道していない）ため、同様に分析対象から除いている。

表1aと1bは両衆院選における各社の調査の概要について、日経新聞と2017年の共同通信を含めてまとめたものである。2017年の調査方法は基本的に各社ともに無作為に発生させた電話番号を用いて調査対象を決定するRDDを用いた調査員による電話調査であり、回答者数は産経の4万から朝日の9万弱程度であった。これに対し2021年には、朝日新聞は事前に調査会社に登録された会員を対象としたウェブ調査に移行、他社は調査員ではなくオートコールを利用した電話調査を用い、回答者数も産経の15万から朝日の35万強と数倍になっている。各社の調査方法については上述の福島（2018; 2022）がより詳細に説明している。ただし時事通信についてはいずれの選挙においても著者の調査した範囲では調査方法は公開されておらず、上述の福島でも記述がない。

日本の報道機関は既に述べたように、各選挙区の情勢を具体的な数字、例えば「A選挙区においてX候補に投票する予定の有権者は44%、Y候補は40%で前者が4%リード、16%が未定と回答」などではなく、自社の世論調査の結果を、過去の選挙における世論調査の結果などと最終的な各候補の得票の関係に基づく予測モデルに投入し、この結果に取材等により得られた質的な情報を加味したうえで最終的な記事を作成する。

また記事作成の際には「X候補が有利」や「Y候補とZ候補が横一線」などの質的な表現を用いる。大まかに言って、情勢が有利な候補には「優勢」「安定」や「リード」、不利な候補には「追い上げ」「苦戦」、接戦が予想される場合には「横一線」などの表現が用いられ、また「横一線」等の表現が用いられた場合でも、前に記述された候補が後に記述されたものより先行していること等が知られている（飯田2007、三春2019）。

そこで表1aと1bではまず各社の情勢判断を、それぞれ与党候補の有利を伝える「与党有利」、接戦を伝えるものの与党候補を前に記述する「接戦/与党前」、後に記述する「接戦/与党後」、与党候

補の不利（ないし野党候補の有利）を伝える「与党不利」に分類し、各社ごとの数字をまとめた。既に飯田（2007）等でも指摘されているが、各社の判断基準がかなり異なることがわかる。2021年では読売新聞が与党候補を「有利」と判断した選挙区は125、「不利」（野党候補が「有利」と判断した選挙区は62と合計187に過ぎないのに対し、産経新聞はそれぞれ155と88、合計243選挙区となる。他方、産経新聞が接戦と判断した選挙区は合計で46であり、読売と比べて少ない。

それではこれらの情勢報道はほどの程度実際の選挙結果を予測することができたのだろうか。表1aと1bの下2行では2017年と2021年衆院選時の各社の勝敗予測の「正答率」と与党が勝利する選挙区数の予測を示した。ここでの「正答率」とは「与党有利」、「接戦/与党前」と報道されたのちに与党が勝利した、あるいは「接戦/与党後」ないし「与党不利」と報道したのちに与党が敗北したものを予測的中と評価し、この割合を示したものである。2017年には時事通信が93%（266）の選挙区で予測的中させてトップであり、これに読売の91%（261）が続き、全体的に9割前後の選挙区での勝敗の予測に成功している。他方、2021年では接戦の選挙区が2017年よりも多かったためか全体的にこの正答率は低く、毎日新聞が88%（253）の選挙区で予測的中させてトップとなり、朝日と産経が87%（252）とこれに続いた。

一方、選挙結果の予想においてはそれぞれの選挙区における勝敗予想よりも、全国で与野党ないしそれぞれの政党が獲得する議席数を予測できるほうが重要との議論もできる。特に接戦の選挙区における勝敗予測は運による面が大きいのに対し、調査にバイアスがなければ標本誤差からほぼ同数の選挙区が「接戦だが与党有利と予測したが敗北」「接戦で与党不利と予測したが勝利」となり、全体での議席獲得数の予想が実際の値に近づくと考えられるからである。

前述のとおり2017年衆院選では与党候補は223選挙区で勝利したが、これに最も近かったのは読売新聞の222であり、正答率が読売よりも高かった時事は219とこれより低い値を予測した。また2021年では与党候補は198の選挙区で勝利したが、正答率の最も高かった毎日新聞は「与党候補勝利と予測したが敗北」の選挙区が20、「与党候補敗北と予測したが勝利」の選挙区が16とあまり大きく変わらないために、与党獲得議席数に対する予測でも202と実際の198に近い結果を出した。他方、正答率では毎日新聞に次いだ朝日新聞と産経新聞はこれらがそれぞれ22と15、6と31であったため、与党獲得議席を7過大あるいは24過小に予想し、特に産経新聞は実際の数値とかなり離れた結果となった。

4. 各社情勢報道情報の統合

本節ではこれらの日本の報道各社による情勢報道を統合し、選挙結果を予測するための手法について示す。上記のように、日本では情勢が「与党候補有利」や「接戦」というように質的に示されているため、米国のReal Clear PoliticsやFiveThirtyEightのように、各社が報道した候補別の支持率を（各社別の調査の信頼性や調査実施日の新しさ等により加重したうえで）平均する手法を採用することはできない。また同様の表現を用いていたとしても、各社の分類基準が同一ではないため、例えば「有利」は2、「接戦」だが前に記述された場合を1、といったように数値を割当て、この平均を算出するような手法も必ずしも効率的とは言えない。前節でも示したように読売新聞は情勢判断で慎重な基準を用いるため、読売の「有利」という表現は、他社の同様の表現に比べてより重く扱われるべきとなる。

そこで本稿ではUmeda（2023）同様に、以上のような日本における情勢報道の性質をふまえ、利用可能な情報を効率的に利用するために、項目反応理論の段階反応モデルを用いる。項目反応理論は潜在変数モデルの一種であり、少数の潜在変数が一群の観測された反応間の関係を説明すると仮定、従来の手法では直接的に測ることが難しい概念、例えば政治的態度を、この潜在変数を想定することで計測することを図る。

項目反応理論はもともと心理学や教育学等で試験を通じて「能力」を測定する手法として開発、発達したが（Paek & Cole 2019, p2）、近年政治学においても、例えば政治的知識（Delli Carpini & Keeter 1993）、議員の政策位置（Clinton 2006; Hirano et al. 2011; Shor & McCarty 2011）、最高裁判事の政治的位置（Martin & Quinn 2002; Tausanovitch & Warshaw 2013）の測定等で用いられている。

本稿では項目反応理論のうち、3 レベル以上の段階的な反応を用いることができる段階反応モデルを用いるが、最初に 2 項 (2パラメータ) モデルから説明する。2 項モデルは試験問題への解答（正答（1）と誤答（0））のように、反応が 2 値で与えられる場合に用いられ、調査対象の反応は、直接観測されない対象の特徴パラメータ θ_i （例えば受験者 i の能力）、項目 j の難易度パラメータ β_j および判別パラメータ α_j により与えられると想定する。

$$p(X_{i,j} = 1 | \theta_i, \beta_j, \alpha_j) = \frac{\exp[\alpha_j(\theta_i - \beta_j)]}{1 + \exp[\alpha_j(\theta_i - \beta_j)]}$$

$$\text{s. t. } i = 1, 2, 3, \dots, n; j = 1, 2, 3, \dots, m$$

項目ごとの難易度パラメータの違いにより、同じ「能力」を持つ対象でも項目により「正答率」は変化する。また判別パラメータの違いにより、同じ難易度を持つ問題の問題であっても、「能力」と「正答率」の関係が変化する。例えば、高い判別パラメータを持つ項目では、「能力」の高い対象はほぼ確実に正解、低い対象は不正解となるのに対し、このパラメータが低い場合には、「能力」と「正答率」の関係がそれほどはっきりとはしなくなる。

段階反応モデルはこの 2 項モデルを拡張し、反応が 3 レベル以上の順序尺度で与えられる場合に対応したものである。本稿の状況に即すれば、(研究者が) 直接観測できない選挙区 i の情勢 θ_i に基づき、報道機関 j が自社の調査と分類基準に従って（一定の誤差を含みつつ）「与党候補有利（4）」、「接戦：与党候補前（3）」、「接戦：与党候補後（2）」ないし「与党候補不利（1）」のいずれかに分類した情勢判断を研究者は観測する。この場合、 $p_{j1}(\theta)$ とは選挙区 i における情勢 θ_i をふまえた上で、 j 社の分類基準において、「与党不利」と判断される確率、 $p_{j2}(\theta)$ は同様に「接戦/与党後」、 $p_{j3}(\theta)$ は「接戦/与党前」、 $p_{j4}(\theta)$ は「与党有利」と判断される確率である。他方、研究者は各社の情勢報道およびその差異から、潜在的な各選挙区の情勢 θ および各社の判断基準の違い (α_j と β_{jk}) を推定する。本稿では報道機関 5 社による 4 段階での情勢判断を用いるため、数式は以下の通りとなる。

$$p_{ijk}^*(\theta) = \frac{\exp[\alpha_j(\theta_i - \beta_{jk})]}{1 + \exp[\alpha_j(\theta_i - \beta_{jk})]}$$

s. t. $i = 1, 2, 3, \dots, n; j = 1, 2, \dots, 5; k = 1, 2, 3$

$$p_{ij1}(\theta) = 1 - p_{ij1}^*(\theta)$$

$$p_{ij2}(\theta) = p_{ij1}^*(\theta) - p_{ij2}^*(\theta)$$

$$p_{ij3}(\theta) = p_{ij2}^*(\theta) - p_{ij3}^*(\theta)$$

$$p_{ij4}(\theta) = p_{ij3}^*(\theta)$$

本稿で段階反応モデルを用いるのには、以下の利点が存在する。まず報道各社が順序尺度で与える判断基準の差異をふまえた上で、それぞれの判断を統合して各選挙区の情勢 θ_i を間隔尺度で推定することができる。この段階反応モデルでは、各社の基準の違いは判別および難易度パラメータにより表現される。なお本稿の具体的な分析においては、Rの“mirt”パッケージを用いた(Chalmers 2020)。

本稿で分析を実施する際には、Umeda (2023) 同様に簡略化のためにいくつかの仮定を置いている。まず2017年と2021年衆院選では各社ともに同一の判断基準を用いると仮定する。すなわち各社の難易度および判別パラメータはこの2回の選挙において等しく、同一の情勢 θ_i に対し、2017年と2021年では各紙ともに同じ表現を用いて(より正確には同じ確率でいずれかの表現を用いて)情勢を報道すると仮定した。そして2017年衆院選のデータを用いて報道各社のパラメータを推定した上で、このパラメータを2021年のデータへと応用し、各選挙区の情勢 θ_i を推定している。

なお分析に先立ってFiveThirtyEightが行うような各社の「信頼性」ないし調査日時による重みづけは行っていない。基本的には投票日(ないし情勢判断を行う)直前に実施した調査、また標本規模が大きい調査ほど投票日ないし「現状」に対する判断には信頼を置けると考えられるが、日本では調査が選挙告示後に集中的に実施され、調査の間隔が最大で1週間以内に留まるため、重みづけを用いるメリットは小さいものと考えられる。

また標本規模が大きい社の報道が相対的に信頼できると考えられるが、この2回の選挙における報道を見る限りでは、調査の標本規模と予測の正確性の関係は必ずしも明確ではない。これは僅差の選挙区については偶然によるものかもしれない。他方、各社の調査方法は必ずしも同一ではないため、非標本誤差の大きさによる可能性もある。

以上の前提を置くこともあり、本稿のモデルにより推定される各選挙区の情勢 θ_i とは、各社の判断(の背後にある各社の情勢評価)における共通因子を代用指標としたものである。各社がそれなりに正確な評価を下す能力を持つ一方で、各選挙区における標本誤差等から判断のばらつきが生じている場合には、項目反応理論を用いて推定される情勢評価は、誤差が相互に打ち消された結果として各社の予測を正確性で上回ることが期待できる。また仮にある社が(標本規模の小ささ等による誤差から)他社と比較的異なる判断を行う場合には、その社の判断は情勢評価の推定に対して割り引かれた影響を与える。その一方で、調査実施後に全国レベルで有権者の投票意図が変化した等の理由により、多数の報道各社が実際の選挙結果に対して系統的にバイアスを持った評価を下した場合には、推定された情勢評価はこのバイアスをそのまま取り込むことになる。

なお項目反応理論を用いた分析では一般的に「局所独立の仮定」、すなわち潜在特性値 θ を特定の値に固定した場合には、各項目に対する回答は相互に独立するという仮定が置かれる（Paek & Cole 2019, p4）。本稿の文脈では情勢 θ_i が特定の値を取る場合には、各社の判断は他社とは独立するという仮定となるが、本稿でもこれを踏襲する。なお表2は段階反応モデルを用いて推定された各社の難易度パラメータ β_{jk} および判別パラメータ α_j である。

それではこの情勢 θ_i はどの程度実際の選挙結果と相関していたであろうか。図1aと図1bはX軸に θ_i 、Y軸に与党候補と最も得票の大きい野党候補の得票差（両者の差を両者の合計で割ったもの）を2017年と2021年衆院選について図示したものである。情勢 θ_i が0.78と-2.01付近の位置に観測値が多いが、これはすべての社が与党候補の「有利」ないし「不利」を報じた選挙区である。これは一種の「天井効果」ないし「床効果」（Cramer & Howitt 2005）と解釈することができ、両者の関係に本来存在する線形性を損なう影響がある。

図1aと図1bは情勢 θ_i が得票差と強い相関関係を持つことを示しており、相関係数は2017年では0.855、2021年には0.877となっている。しかし（天井および床効果の影響もあり）この程度にとどまるとも言える。この水準では、接戦の選挙区においてどの候補者が先行しているかを誤って評価する可能性は決して小さくないと考えられる。

それでは衆院選の選挙期間中に各社の情勢報道が出そろった時点で、過去の衆院選における情勢報道と実際の選挙結果との関係を用いることで、選挙区の結果をどの程度正確に予測できるだろうか。本節で推定した情勢 θ_i は相対的な評価であり、値が大きければ与党候補が多く得票を得、また勝利する可能性が高いことを想定できるが、どの値を超えた場合に与党候補が勝利すると予測できるのか、また値に対応した与党候補の得票や勝利確率を直接示すことができるわけではない。次節ではこの予測の問題について論じる。

5. 選挙結果の予測

本節では前節で統合した情報を用いて選挙結果をどの程度正確に予測できるかの検証を行う。より具体的には2017年衆院選における情勢報道（およびこれから推定された情勢 θ_i ）と選挙結果との関係をモデルの「学習」に用い、2021年衆院選における各社の情勢報道が出そろった時点で、個別選挙区における与野党候補の得票差および与党候補の勝利確率の予測を行い、さらにこの個別選挙区における勝利確率を合計することで与党候補の勝利選挙区数の予測を行う。

2021年衆院選において、各社の情勢報道を用いて各選挙区における与野党候補得票差を予測する手段として以下の方法を用いる。

1.1 回帰分析を用いて、2017年衆院選における各社の情勢報道から推定された情勢 θ_i と与党候補と野党候補（のうち最大得票候補）の得票差 Y_{17} の関係を分析する。以下ではこの回帰分析を2017得票回帰モデルとする。

1.2 2017得票回帰モデルの回帰係数を用い、2021年衆院選の各社情勢報道から推定した情勢 θ_i から、2021年衆院選における与野党候補の得票差 Y_{21} を予測する。

なお2017年の衆院選では情勢報道（の平均）が与党候補の勝利選挙区数をほぼ適切に予測しているため、報道各社が与党の勝利を過大に予測した2014年衆院選に対してUmeda（2023）が行ったような予測と実際の結果の関係に対する調整を行う必要はないものと思われる。

具体的には2017年得票回帰モデルにおいて、与野党候補の得票差 Y_{i17} と情勢 θ_i の関係は $Y_{i17} \sim 0.0833 + 0.1040 \theta_i + e_i$ と推定されるため、2021年衆院選の与野党候補の得票差についても、

$$\hat{Y}_{i21} \sim 0.0833 + 0.1040 \theta_i$$

という予測モデルを作成することができる。

さらにこれと同様の手法を応用し、2021年衆院選における各社の情勢報道を用いて、各選挙区における与党候補の勝利確率および与党の合計勝利選挙区を予測する。このための手段としては以下の方法を用いる。

- 2.1 プロビット分析を用いて、2017年衆院選における各社の情勢報道から推定された情勢 θ_i と与党候補の勝敗 W_{i17} の関係を分析する。以下ではこのプロビット分析を2017勝敗プロビットモデルとする。
- 2.2 2017勝敗プロビットモデルの回帰係数を用いて、2021年衆院選の各社情勢報道から推定した情勢 θ_i から、2021年衆院選における与党候補の勝利確率 $\hat{P}(W_{i21} = 1)$ を予測する。

2017年勝敗プロビットモデルにおいては、与党候補の勝敗 W_{i17} と情勢 θ_i の関係は、 $P(W_{i17} = 1) \sim F(2.414 + 2.965 \theta_i)$ と推定され²、2021年衆院選の与党候補の予測勝利確率についてもこのプロビットモデルで推定された回帰係数を用いて、

$$\hat{P}(W_{i21} = 1) \sim F(2.414 + 2.965 \theta_i)$$

とすることができる。

それではこれらの予測モデルはどの程度実際の選挙結果を正確に予測することができたであろうか。図2はまず上記の得票差の予測モデルを用いた2021年衆院選時の予測と選挙結果の関係を示したものである³。

この図は2017年の情勢報道と実際の選挙結果の関係をを用いた予測モデルが、2021年の選挙結果（与野党候補得票差の実現値）を比較的正確に予測できたことを示している。2021年衆院選の予測値を説明変数、実際の選挙結果を結果変数として回帰分析を実施したところ、切片は0.0137、回帰係数は0.9855と理論上の値である0と1に近い数値となった。ただし2021年衆院選における報道各社の予測が平均的に与党候補の得票を過小評価していたため、切片がやや0よりも大きく（1.37%）なっている。参考までにこの回帰分析の結果に基づく予測直線を図2へと追加した。

2 なお $F(\alpha + \beta x_i) = \int_{-\infty}^{\alpha + \beta x_i} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$

3 結果の見やすさの観点から、与党候補が予想得票差0.22を超えて勝利した選挙区を図の枠には含めていないが、2021年衆院選では33選挙区存在した。

なおこの図2の第一象限に位置する選挙区は与党候補の勝利が予測され、また実際に与党候補が勝利した選挙区、第四象限の選挙区は同様に与党候補の敗北が予測され、また実際に与党候補が敗北した選挙区であるために、本稿のモデルが予測的中させた選挙区となる。こうした選挙区は全体で255存在し、全体の88%となった。これは2021年衆院選で最も「正答率」が高かった毎日新聞の88%（253議席）を上回る結果である。ここから、本稿での手法は報道各社の情勢報道を統合することにより、各社の調査が持つ標本誤差等の影響を打ち消し、より正確な予測が可能であることが2021年衆院選の事例からも示されたものと評価できるだろう。

一方、第二象限に位置する選挙区は与党候補の敗北が予想されたのに実際には勝利した選挙区、また第三象限に位置する選挙区は与党候補の勝利が予想されたのに実際には敗北した選挙区であるが、両者の数にはかなりの差があり、前者が26であったのに対し後者は8であった。ここから勝敗のみの予測に関しては、本稿のモデルは実際の選挙結果（与党が198選挙区で勝利）と比較すると、与党が勝利する選挙区を18過小評価した180と予測することとなる。これはこの選挙における読売新聞や産経新聞等の与党獲得議席に対する低い予測の影響を受けたものと思われる。

ただしすでに述べた通り、接戦が予想された選挙区での勝敗は偶然によるものもあり、勝敗予測の的中のみで評価することは必ずしも適切ではないだろう。例えば与党候補が52%で勝利すると予測された選挙区で与党候補が敗北したとしても、特に不思議はない。

これに対し、プロビットモデルを用いて推定された各選挙区における与党候補の勝利確率は、これを合計することで与党の期待勝利選挙区数を予測することもできる。この予測勝利確率を合計する手法では、上記の与党候補が52%の確率で勝利すると予測した選挙区での与党の獲得期待議席数は0.52となる。この手法を用いた場合には、与党は184議席を獲得すると予測された。これでも実際の選挙結果と比較すると過小評価ではあるが、二値的な勝敗予測よりは実際に近い結果となった。

また機械学習等で用いられる二乗平均平方根誤差（RMSE）を用いて与野党候補得票差に対する予測値の正確性を評価すると誤差は5.8%であったが、全ての社が与党候補有利ないし不利と判断したために床効果および天井効果の影響を受ける選挙区を除いた場合には、3.8%となった。

6. 結論と議論

本稿ではUmeda（2023）が2017年衆院選の選挙結果を予測したのと同様の手法を用い、2021年衆院選の選挙結果予測を試みた。より具体的には、日本の報道各社が選挙期間中に行う衆議院小選挙区レベルでの選挙情勢報道の情報を項目反応理論の段階反応モデルを用いて統合し、この統合された情報を用いて選挙結果予測を行った。

この結果、この手法は2017年衆院選ほどではないにせよ、2021年衆院選の選挙結果もかなり正確に予測できたことが明らかとなった。2021年衆院選では情勢報道を実施した報道各社の一部が与党獲得議席をかなり過小に予想したために、本稿の予測モデルも与党の獲得議席をこれらの影響を受け低く予測することとなった。例えば、各選挙区の与党候補の勝利確率予測値を合計した与党の期待獲得議席数は184となり、これは実際に与党が獲得した198をかなり下回った。これは報道各社の評価を統合し平均的な値を予測値として用いる本稿のモデルの限界を示すものであろう。

他方、勝敗予測に関しては、本稿のモデルは88%の選挙区（255）での勝敗を正しく予想したが、これは各社の「正答」率を上回るものとなった（最高が毎日新聞の88%、253選挙区）。これは本

稿のモデルが勝敗ラインの推定では各社の平均からある程度バイアスを持っていたとしても、複数の調査結果を統合することにより、比較的小規模な標本に依拠した小選挙区レベルでの各社調査の持つ標本誤差等を打ち消し、より正確な予測を行うことができるためと思われる。これは本稿のモデルの持つ強みが2021年選挙の予測においても発揮された結果と評価することができるだろう。

本稿での予測モデルはすでに論じたように報道各社の実施した予測のバイアス、ないし各社調査が終了した後の選挙戦終盤における情勢の変化に加え、特に2021年に近畿地方の選挙区で見られたような三つ巴の競争等に対応するのが困難である。また今後2021年衆院選と同様に複数の報道社が共同で選挙調査を実施し、これを用いて情勢報道を実施することで利用可能な（他社から独立した）情勢判断の数が減少するかもしれない。それでも本稿の予測モデルは比較的簡便に報道各社の情勢報道を統合し、選挙予測を実施するための手法として有効であることが示されたと考えられる。

〈参考文献〉

- 飯田良明（2007）『新聞の選挙情勢報道の分析—第44回総選挙を事例として—』「実践女子大学人間社会学部紀要」第3集, 19-42頁
- 亀ヶ谷雅彦（2001）『選挙予測のアナウンスメント効果に関する先行研究の概観—アナウンスメント効果の低位効果の拡張に向けて』「山形県立米沢女子短期大学紀要」36号 71-86頁
- 福島靖男（2015）『第47回衆議院議員選挙 新聞・テレビはどう伝えたか』「よろん」116号 7-17頁
- 福島靖男（2018）『第48回衆議院議員選挙 新聞・テレビはどう伝えたか』「よろん」121号 68-81頁
- 福島靖男（2022）『第49回衆議院議員選挙 新聞・テレビはどう伝えたか』「よろん」129号 84-96頁
- 三春充希（2019）「武器としての世論調査—社会をとらえ、未来を変える」筑摩書房
- 早稲田大学（2019）「早稲田大学第一回データサイエンスコンペティション」<https://dswaseda.com>（2021年4月20日アクセス）
- Chalmers, Phil (2020) Package 'mirt'. <https://cran.r-project.org/web/packages/mirt/mirt.pdf> (2021年3月11日アクセス).
- Clinton, Joshua D (2006) "Representation in Congress: Constituents and Roll Calls in the 106th House". *The Journal of Politics*, 68 (2): 397-409.
- Cook Political Report (2020). *House Race Ratings*. <https://cookpolitical.com/ratings/house-race-ratings> (2021年7月3日アクセス).
- Cramer, Duncan, and Dennis Howitt (2005). *The SAGE Dictionary of Statistics: A Practical Resource for Students in the Social Sciences* (3rd ed.). Sage Publications.
- Delli Carpini, Michael, and Scott Keeter (1997). *What Americans Know About Politics and Why It Matters*. New Haven: Yale University Press.
- Gallup, George (1951). "The Gallup Poll and the 1950 Election." *Public Opinion Quarterly*, 15 (1): 16–22.
- Graefe, Andreas (2019). "Accuracy of German Federal Election Forecasts, 2013 and 2017." *International Journal of Forecasting*, 35 (3): 868–877.
- Hirano, Shigeo, Kosuke Imai, Yuki Shiraito and Masaki Taniguchi (2011) Policy Positions in Mixed Member Electoral Systems: Evidence from Japan. *Unpublished Manuscript*. <https://imai.fas.harvard.edu/research/japan.html> (2018年7月18日アクセス).
- Huffpost Pollster (n.d.) "The Huffington Post Pollster Policy for Including Polls in Charts". <https://elections.huffingtonpost.com/pollster/faq> (2021年4月20日アクセス).

- Jackson, Natalie (2018) . “The Rise of Poll Aggregation and Election Forecasting” . Atkeson, Lonna Rae and Alvarez, R. Michael. Ed. *The Oxford Handbook of Polling and Survey Methods* (pp. 609–632). Oxford University Press.
- Lewis-Beck, Michael S. and Stegmaier, Mary (2014) . “US Presidential Election Forecasting.” *PS, Political Science & Politics*, 47 (2) : 284–288.
- Lewis-Beck, Michael S. and Tien, Charles (2012) . “Japanese Election Forecasting: Classic Tests of a Hard Case.” *International Journal of Forecasting*, 28 (4) : 797–803.
- Martin, Andrew D., and Kevin M. Quinn (2002) “Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for The U. S. Supreme Court, 1953–1999.” *Political Analysis*, 10 (2) : 134–153.
- Paek, Insu and Ki Cole (2019) . *Using R for Item Response Theory Model Applications*. Taylor and Francis.
- Shor, Boris, and Nolan McCarty (2011) “The Ideological Mapping of American Legislatures.” *American Political Science Review*, 105 (3) : 530–551.
- Silver, Nate (2020) . “Our New Polling Averages Show Biden Leads Trump By 9 Points Nationally” . <https://fivethirtyeight.com/features/our-new-polling-averages-show-biden-leads-trump-by-9-points-nationally> (2021年4月20日アクセス) .
- Tausanovitch, Chris, and Christopher Warshaw (2013) . “Measuring Constituent Policy Preferences in Congress, State Legislatures, and Cities.” *The Journal of Politics*, 75 (2) : 330–342.
- Traugott, Michael W. (2014) . “Public Opinion Polls and Election Forecasting.” *PS: Political Science and Politics*, 47 (2) : 342–344.
- Umeda, Michio (2023) . “Aggregating Qualitative District-Level Campaign Assessments to Forecast Election Results: Evidence from Japan.” *International Journal of Forecasting*, 39 (2) : 956–966.

〈図表〉

表1a. 第48回来議院議員衆院選（2017年10月22日投票）各社情勢報道⁴

	読売	日経	朝日	毎日	共同	時事	産経
記事	10/12	10/12-13	10/14-15	10/16	10/13	10/16	10/17
調査日	10/10-11		10/10-13	10/13-15	10/10-11	NA	10/12-15
回答者数	78,285		88,152	73,087	90,261	NA	39,944
与党有利	156	188	200	204	213	189	209
接戦／与党前	64	25	32	27	17	32	17
接戦／与党後	38	31	26	20	25	36	12
与党不利	28	42	28	35	31	29	48
勝敗予想正答率	91%	91%	90%	91%	87%	93%	90%
与党勝利数予測	222	214	232	231	230	219	227

表1b. 第49回来議院議員衆院選（2021年10月31日投票）各社情勢報道⁵

	読売	日経	朝日	毎日	時事	産経
記事	10/21	10/22	10/26	10/22	10/26	10/27-28
調査日	10/19-20		10/23-24	10/19-20	NA	10/23-24
回答者数	186,683		353,868	174,024	NA	155,045
与党有利	125	132	170	168	148	155
接戦／与党前	39	36	35	34	46	18
接戦／与党後	63	59	43	27	47	28
与党不利	62	62	41	60	48	88
勝敗予想正答率	85%	85%	87%	88%	85%	87%
与党勝利数予測	164	168	205	202	194	173

4 読売（と日経）は10月20日に二度目の調査結果を報道しているが、およそ半分程度の選挙区のみを調査の対象としていること、また非標本誤差が1回目の調査と強い相関を持つことが予想されるために、分析対象から外した。

5 2017年衆院選と同様に、読売（と日経）は10月29日に二度目の調査結果を報道しているが、こちらも半分程度の選挙区のみを調査対象としていること、また非標本誤差が1回目の調査と強い相関を持つことが予想されるために、分析対象から外している。

表 2. 各社判別および難易度パラメータ

	読売	朝日	毎日	時事	産経
α_j	7.32	4.16	3.88	7.13	4.73
β_{j1}	9.82	5.93	4.97	9.15	5.01
β_{j2}	5.78	4.06	3.68	5.55	4.24
β_{j3}	0.842	2.37	2.34	2.86	3.23

図1a. 2017年衆院選の情勢報道に基づく選挙情勢推定値と与野党候補得票差

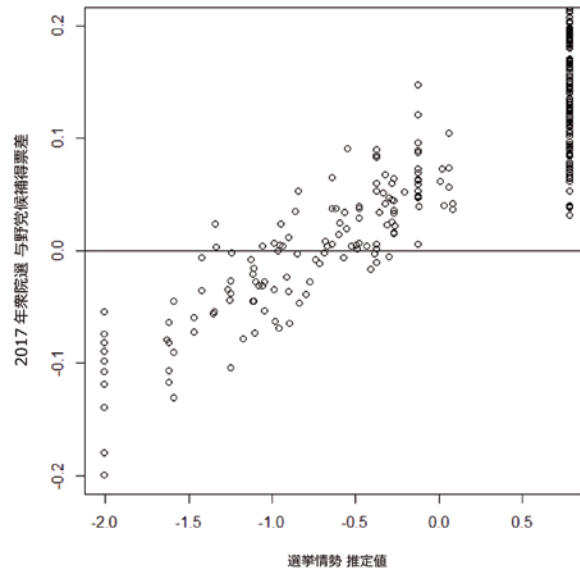


図1b. 2021年衆院選の情勢報道に基づく選挙情勢推定値と与野党候補得票差

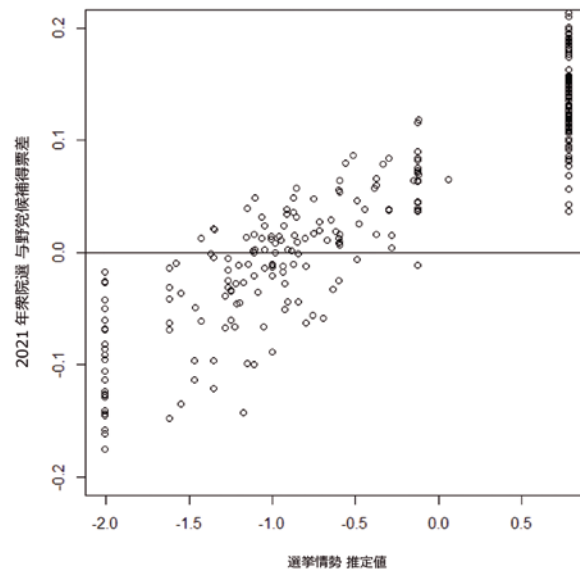
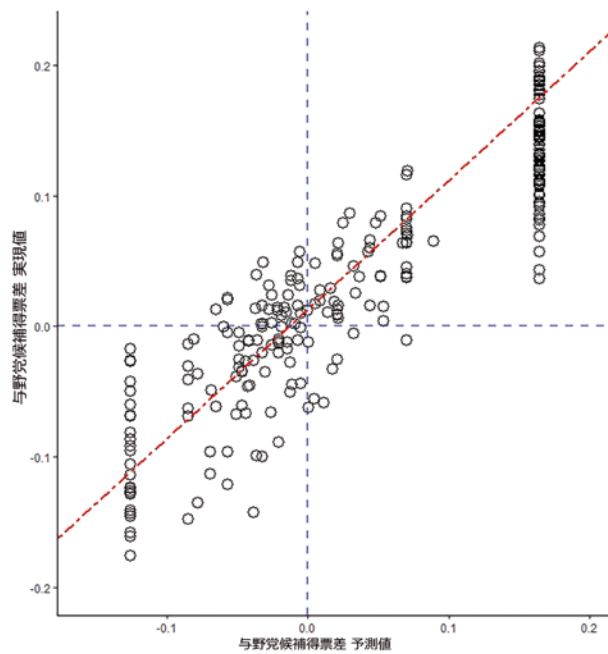


図2. 2021年衆院選の情勢報道に基づく与野党候補得票差予測値と実現値



別表. 記事における表現とその分類例

判断	表現
有利	「安定」, 「優位」, 「優勢」, 「手堅い」, 「盤石」, 「圧倒」, 「突き放し」, 「独走」, 「引き離す」, 「幅広く支持」, 「有力」, 「先行」, 「リード」, 「抜け出した」 (相手に) 〈不利〉な表現
接戦	「接戦」, 「伯仲」, 「混戦」, 「激しく争う」, 「競り合う」, 「激戦」, 「横並び」, 「五角」, 「死闘」, 「デッドヒート」, 「横一線」, 「三つ巴」, 「しのぎを削る」, 「つばぜり合い」, 「一歩も譲らぬ」
不利	「懸命」, 「追う」, 「猛追」, 「肉薄」, 「追い上げを図る」, 「支持拡大を図る」, 「浸透が課題」, 「出遅れが響く」, 「伸び悩む」, 「巻き返し」 (相手に) 〈有利〉な表現