

第3章 2016年における予測の失敗と2020年大統領選挙

飯田 健

はじめに

2020年アメリカ大統領選挙の結果予測は全体としてバイデン勝利こそ言い当てたものの2016年と同様、失敗したと言われている。実際 FiveThirtyEight（以下、FTE）、New York Times（以下、NYT）、そしてコロンビア大学の統計学者および政治学者であるアンドリュー・ゲルマンを擁して新規参入した The Economist など主要な予測主体は世論調査結果の集計にもとづいてバイデン優勢の予測を出していたが、その勝ち方は当初予測されていたほどには圧倒的ではなかった。

これをふまえて本稿ではまず、2020年アメリカ大統領選挙の結果予測の失敗の実態について、どのような意味で失敗したと言えるのか、2016年との比較を通じて確認する。そのうえで予測の方法について説明し、予測の過程において誤差が生じた原因について先行研究等を参照しつつ検討を行う。

1. 予測の失敗

2020年大統領選挙はどのような意味で予測が外れたと言えるのだろうか。ここでは2016年においてトランプの勝利確率を29%と見積もり、主要な予測主体の中では最もまじだったFTEの予測¹を使って2つの点から検討したい。第1に、Brier Score（以下、BS）である。BSは確率予測の正確性の評価に用いられるもので、次の式によって与えられる。

$$BS = \frac{1}{N} \sum_{i=1}^N (f_i - o_i)^2$$

これをアメリカ大統領選挙のコンテキストに置き換えれば、 f_i は州*i*においてトランプが勝利すると予測された確率、 o_i はその州において実際に起こった結果（トランプ勝利の場合は1、トランプ敗北の場合は0をとる）、 N は結果予測の対象となる州の数となる。つまり、BSは州ごとの勝敗予測の誤差の2乗値の平均であり、この値が小さいほど予測誤差が小さいということを意味する。例えば、ある州でトランプが99%の確率で勝利すると予測され実際にトランプが勝利した場合、BSは $(0.99 - 1)^2 = 0.0001$ と予測誤差は非常に低くなるが、トランプ勝利の予測確率が80%でトランプが勝利した場合、トランプ勝利を的中させたとはいえ確信の度合いが弱かったためBSは $(0.80 - 1)^2 = 0.04$ と予測誤差は大きくなる。アメリカ大統領選挙の予測精度を評価するうえでは大統領選挙における州による重要性の違いを考慮するため、式にあるように単純平均をとるのではなく、各州に割り当てられた選挙人の票数で加重平均したBSを計算する。

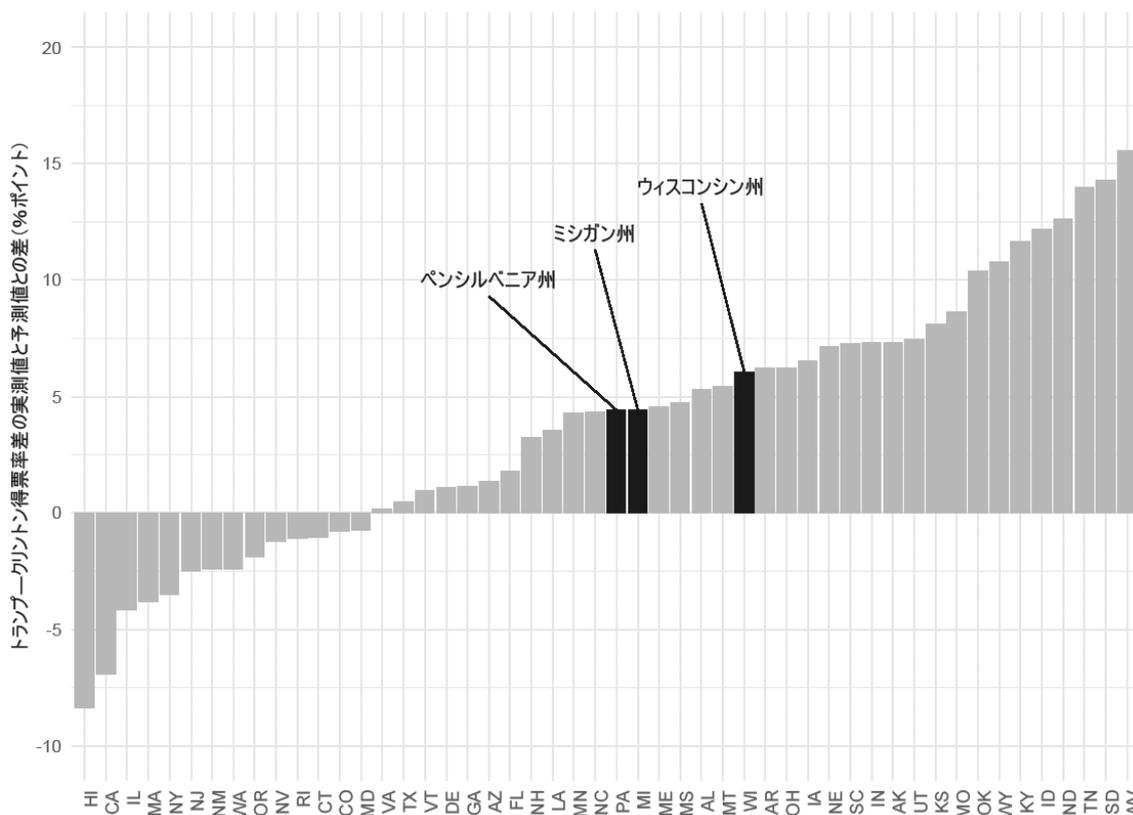
このようにして計算したFTEの2016年、2020年の予測のBSはそれぞれ0.094と0.066になる²。つまり、2016年に比べて2020年は全体として予測の誤差が小さかったと言える。実際、この両年のBSの差0.028は、2016年においてトランプ勝利確率を15%と見積もっ

ていたNYTの予測のBS(0.121)とFTEの予測のBSの差と同等であり³、実質的な予測誤差の改善と言えるだろう。

しかしながら、このようにBSで見た場合には改善があるとはいえ、2020年は2016年と比べてそれほど予測の精度が上がったという実感はないように思える。その理由の1つにはBSはトランプ勝利確率の過小評価も過大評価も予測誤差として同等に扱う指標であるということが考えられる。そこでより実質的に州ごとの得票率の予測誤差を用いて両年の予測の精度を比較する。

図1は、2016年について両候補者の得票率とFTEによる得票率予測値を用いて、トランプの得票率からクリントンの得票率を引いた得票率差の実測値と予測値の差を州ごとに計算したものである⁴。例えばトランプが勝ったウィスコンシン州の場合、トランプの実測の得票率は47.22%、クリントンの得票率は46.45%で、トランプ得票率からクリントン得票率を引いた得票率差は+0.77%ポイントであった。一方、このときのトランプの予測得票率は44.3%、クリントンの予測得票率は49.6%で、その得票率差は-5.3%ポイントであった。つまりクリントンが得票率差5.3%ポイントで勝つと予測されたが、実際には得票率差0.77%ポイントでトランプが勝ったということであり、得票率差の実測値と予測値の差は、 $0.77 - (-5.3) = +6.07\%$ ポイントとなる。この値が0よりも大きいほど予測はトランプを過小評価していたということになり、0よりも小さいほど予測はトランプを過大評価して

図1. 得票率の実測値と予測値との差(2016年)



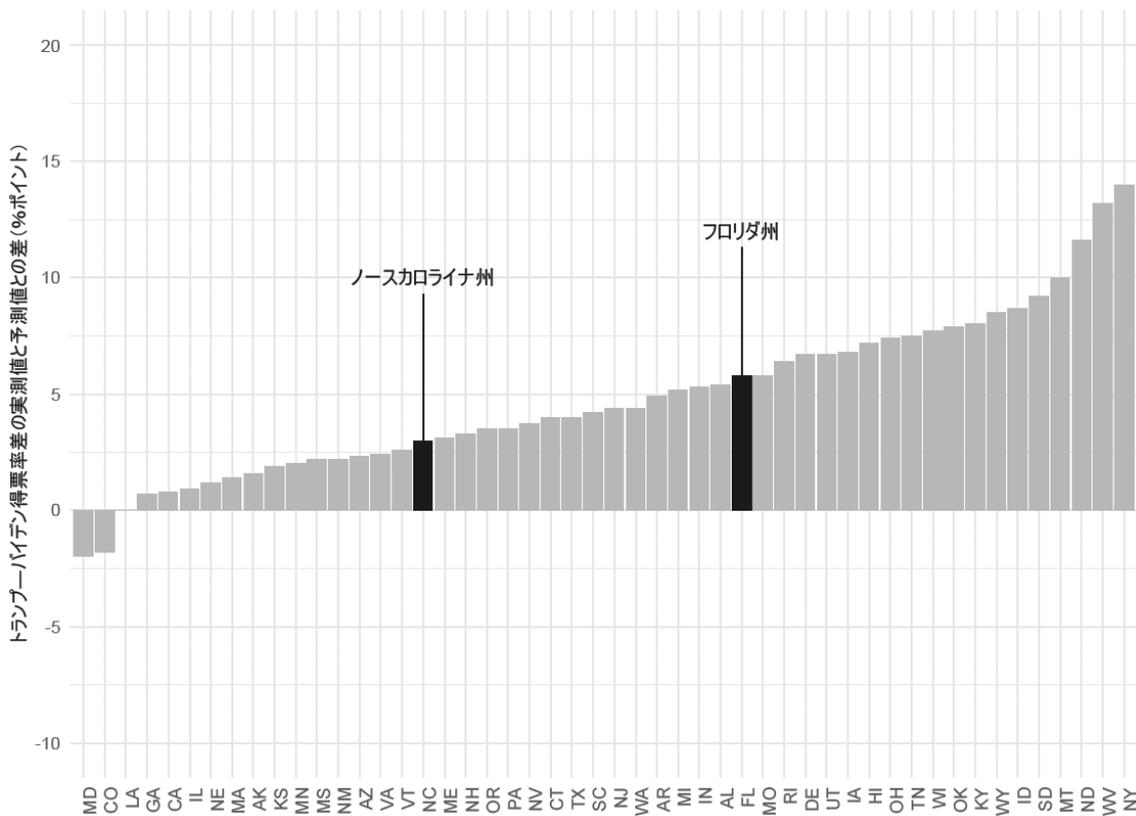
図出典：FiveThirtyEightのデータ等を用いて筆者作成。黒塗りはトランプが得票率差1%ポイント以上で負けると予測されたが、実際には勝った州を示す。

いたということになる。すなわち +6.07% ポイントという数字は、2016年のFTEの予測がウィスコンシン州でのトランプの得票率を大幅に過小評価していたことを意味する。

図1によると、0を下回る値を示した州は14であるのに対して、0を上回る値を示した州は36と、3分の2弱の州で予測はトランプを過小評価している。また得票率差1%ポイント以上でトランプが敗北する（クリントンが勝利する）と予測したものの、結果的にトランプが勝利した州は、図中黒塗りで示したウィスコンシン州、ミシガン州、ペンシルベニア州の3州であったが、これらに割り振られた選挙人票はそれぞれ10、16、20と合計46にもなる。これは両候補者が獲得した選挙人の数の差に換算すると92に相当するが、この分を外したことが2016年において全体としてトランプ勝利を当てられなかった要因であると言えるだろう。

こうした2016年大統領選挙の予測の誤差は、同じく「失敗」とされた2020年と比べてどのように異なるのであろうか。図2は2020年について、トランプの得票率からバイデンの得票率を引いた得票率差の実測値と予測値の差を州ごとに計算したものである。これによると、ルイジアナ州、メリーランド州、コロラド州の3州を除くすべての州で数値は0を上回っており、FTEの予測は2016年よりも圧倒的に多くの州でトランプを過小評価していたことがわかる。実際、誤差の単純平均で見ると2016年は平均+3.80%ポイント、トランプの得票率を過小評価しているのに対し、2020年は+4.79%ポイントの過小評価となっ

図2. 得票率の実測値と予測値との差（2020年）



図出典：Federal Election Commission および FiveThirtyEight のデータを用いて筆者作成。黒塗りはトランプが得票率差1%ポイント以上で負けると予測されたが、実際には勝った州を示す。

ている。その度合いが最も大きかった、つまりトランプが予想よりも強かったのがニューヨーク州である。得票率29.1%ポイント差でトランプが負けると予想されたものの実際は15.1%差で敗北と、勝敗予測自体は外さなかったとはいえ14%ポイント分、予測されたよりも得票率差は開かなかった。

とはいえ2020年はニューヨーク州同様、ほとんどの州についてトランプ得票率が過小評価されたものの、トランプが予測よりも僅差で負けた、あるいは予測よりも大差で勝ったというだけで、勝敗予測自体は外れていない。得票率差1%ポイント以上でトランプ敗北と予測したものの、トランプ得票率を過小評価した結果予測が外れトランプが勝利した州は図中に黒塗りで示したフロリダ州およびノースカロライナ州の2つだけである。しかしそれぞれの選挙人票は29と15で、予測を外した選挙人票は合計44となっており、この点では2016年に予測を外した3州の選挙人票46とそれほど変わらない。

以上のことから、2020年の予測は2016年とは異なり大統領選挙全体の勝者を当て、誤差の方向性を考慮しないBSでも2016年よりも精度が高かったとはいえ、2016年と比べて明らかにトランプ得票率を過小評価した度合いが大きく、予測を外した選挙人票の数でも2016年と同程度であったと言える。これが実感として2016年と比べて2020年の予測が改善したとは思えない要因であろう。

2. 予測の方法

2016年、2020年とも予測においてトランプの過小評価が起きた原因について検討を行う前に、ここではそもそも予測がどのように行われるのか簡単に説明したい。アメリカの選挙予測は日本とは異なり世論調査を実施する調査主体と予測を行う主体とが必ずしも同一ではなく、別であることが普通である。FTEなどの予測主体は独自の調査は行わず、マスコミ、大学、調査会社などの各調査主体が州レベルで実施した世論調査における各候補者の予測得票率の平均値を算出することにより各州の勝者予測を行う⁵。

なぜ複数の調査における得票率の平均値をとるのかというと、理論的に誤差が小さくなるからである。例えば母集団のパラメータである、ある州の有権者全体でのトランプに投票する意図をもつ有権者の割合が知りたいとして、1,000人からなる1つの標本を母集団から無作為に抽出したとき、この標本におけるトランプに投票する意図をもつ有権者の割合がどの程度母集団のパラメータと乖離しているのかはわからない。仮にこの州の有権者から無作為に抽出した1,000人に対して調査を行ったところトランプに投票する意図を示す割合は52%だったとしよう。この値は有権者全体の値とほぼ一致しているのかもしれないし、極端に離れているのかもしれない。もしここで実は有権者全体の値が48%だった場合、この標本における52%という値1つを元に予測を行うと大きく結果を見誤ることになる。これは抽出誤差とよばれ無作為抽出した標本に対して調査を行う以上、避けられない問題である。しかしながら、複数の標本（例えば10個の世論調査）におけるトランプに投票する意図をもつ有権者の割合の値の平均をとることで、抽出誤差が原因で極端に上振れした値や、下振れした値が相殺され、推定の誤差を小さくすることができるのである。

実際の予測では単純平均をとるのではなく加重平均をとる。なぜなら調査によってその信頼性や実施時期が異なっており、信頼性が高いほどあるいは実施時期が投票日に近いほど重みを付ける必要があるからである。この加重平均の取り方は予測主体によって異なり、

複数の同じ世論調査結果を用いていたとしても予測に違いが生じる。その意味で、各世論調査の結果をどのように評価し重み付けを行うかが各予測主体の腕の見せどころと言えるだろう。

例えば、FTEによる2020年のフロリダ州の最終予測⁶の場合、実施されたすべての世論調査のうち、過去の実績や標本規模から信頼性が高いと判断された複数の世論調査結果を元に実施時期も加味して加重平均を求めたところ、トランプの予測得票率46.6%に対し、バイデンの予測得票率49.1%となった。これにさらに投票先未決定の回答者をこの割合に従って配分し48.4%対50.9%となる⁷。場合によってはこれに世論調査に拠らない人口統計からの予測値(47.6%対51.7%)と経済状況や現職優位性などの「ファンダメンタルズ」からの予測値(52.7%対46.6%)も考慮されるが、最終的にこれらはFTEの最終予測には反映されず、前述の世論調査の加重平均の結果がそのまま予測結果となっている。

このようにして求めた両候補者の予測得票率の推定値とその誤差を元に選挙結果を数万回シミュレーションし、そのうちの何パーセントで各候補者が勝つかを計算する。FTEによるフロリダ州の予測の場合、4万回のシミュレーションを行い、そのうち69%でバイデンが勝利、31%でトランプが勝利すると出た。これがフロリダ州における各候補者の勝利確率の意味である。「今日選挙があった場合、69%の確率でバイデンが勝つ」と聞くとまずバイデンが勝つように感じるが、「3回に1回はトランプが勝つ」と聞くとトランプが勝つても決しておかしくないという印象になるのではないだろうか。

最後に、こうした各州における両候補者の勝利確率を元に大統領選挙での両候補者勝利確率を求める。これも各州の勝利確率と同様シミュレーションで求める。イメージとしては、州ごとに各候補者が勝利する確率が異なる「サイコロ」を同時に投げ、その結果に各州の選挙人票を掛け合わせてどちらの候補者が過半数を取るか確認するという作業を数万回繰り返し、そのうちの何パーセントで各候補者が当選するか算出するというものである⁸。FTEの場合、コンピュータ上での大統領選挙4万回の試行のうち89%でバイデン勝利、11%でトランプ勝利と出ている。

3. 失敗の原因

以上説明した予測の方法をふまえて、なぜ2016年、2020年において結果予測は失敗したのか、全米世論研究協会(the American Association for Public Opinion Research)の報告書であるKennedy et al. (2018)に拠りながら検討してみよう。2016年の失敗について、この報告書では主として4つの要因が検討されている。

第1に予測に用いられた世論調査における高学歴者の過大代表の問題である。一般的に大卒者ほど世論調査に協力しやすいという傾向が以前から知られているが、もし教育程度と投票する候補者との間に相関があれば得票率を予測するうえで大きな問題となる。しかし従来、大統領選挙において教育程度と投票選択との相関は必ずしも高くなかった。実際報告書によると、マスコミ各社が共同で実施している出口調査であるNational Election Pools (NEP)の2012年のデータを分析したところ、2016年において激戦州のウィスコンシン州、ペンシルベニア州、ミシガン州では学歴と民主党への投票傾向はU字型になっており、学歴が最も低いグループと最も高いグループで民主党に投票する傾向が高かった。そのため世論調査の標本において高学歴者が過大代表されていても、過小代表されている

低学歴者も民主党支持傾向が強いため、2012年には全体として予測に大きな誤差が生じなかったのである。

しかしトランプが共和党の候補者になった2016年には、これらの激戦州では、明らかに教育程度が高い有権者ほど民主党候補に投票するという直線的な傾向が見られた。その結果、高学歴者が過大代表された標本内でのクリントン支持が実際より高く出てしまい、予測がトランプを過小評価する方向で大きく歪んでしまったのである。

この問題に対処するためには通常、教育程度による回答の重み付け (weighting) を行う必要がある。例えばある州の世論調査の標本においてその50%を占める大卒以上の学歴をもつ回答者の中でのクリントン支持率が80%、残りの50%を占める大卒未満の学歴をもつ回答者の中でのクリントン支持率が30%だったとしよう。この標本内で普通にクリントン支持率を計算すると $0.5 \times 0.8 + 0.5 \times 0.3 = 0.4 + 0.15 = 0.55$ と、55%になる。しかしその州の有権者全体という母集団に占める大卒以上の学歴をもつ割合が40%、大卒未満の学歴をもつ割合が60%であり、標本においては大卒以上の学歴をもつ有権者が過大代表されていた場合、この標本における55%というクリントン支持率は母集団における値よりも高くなっている可能性がある。

この場合、「ある属性の母集団での比率」／「ある属性の標本での比率」によって与えられるウェイトを計算し、それによって標本における得票率の推定値を重み付けすることになる。この例の場合のウェイトは大卒以上の学歴をもつ回答者について $0.4/0.5=0.8$ 、大卒未満の学歴をもつ回答者について $0.6/0.5=1.2$ となり、これで重み付けすると標本におけるクリントンの支持率は $0.4 \times 0.8 + 0.15 \times 1.2 = 0.50$ で50%に下方修正される。このように理論的には、ある州の有権者全体という母集団における実際の大卒以上／大卒未満の比率を用いて世論調査における標本の数値を重み付け補正することで母集団の数値に近づけることができるのである。

ところが報告書によると、2016年にはこうした教育程度による重み付け補正を行った激戦州の世論調査は半分にも満たなかった。その結果、世論調査での標本においてクリントンを支持する傾向の強い高学歴者が過大代表されているのが放置されてしまうこととなった。このような投票選択に影響を与える属性によって重み付け補正されていない世論調査結果を予測主体が加重平均したところでその結果も歪んでしまう。報告書によると、これが2016年の結果予測が外れた有力な原因の1つである。このように母集団に含まれる特定の属性をもつ集団が調査に協力しないことによって生じるバイアスは非回答バイアス (nonresponse bias) とよばれ、世論調査方法論においては対処すべき最も重要な問題の1つとなっている (Groves et al. 2001)。

さらに報告書においては、予測が失敗した要因として実際に投票する有権者を特定し損ねたという可能性についても検討されている。世論調査で回答した有権者全員が投票するわけではないため、選挙結果を予測する上では回答者のうち「投票しそうな有権者 (likely voters)」の属性を特定し、それによって重み付けを行うことで得票率を予測する必要がある。例えば2016年においてこれまであまり投票しなかった属性をもつ有権者がトランプに投票するなど投票傾向の大幅な変化が起きると、各調査主体がそれを考慮することができず予測が失敗する可能性がある⁹。

しかしながらこの点について報告書によると、ウィスコンシン州、ペンシルベニア州、ノー

スカロライナ州、フロリダ州といった激戦州の世論調査においては、調査主体が投票しそうな有権者の属性で重み付けをすることによって予測得票率差の誤差が概して小さくなっていた。つまり投票しそうな有権者の属性を特定し損ねたことによって予測結果に誤差が生じたとは考えにくい。

次に報告書で検討されているのは、直前でトランプに投票することを決めた有権者の存在である。2016年の選挙でのNEPの出口調査データによると、トランプが僅差で勝った激戦州においては11～15%の投票者が選挙戦最終週に誰に投票するかを決めたと回答していたが、その中ではトランプがクリントンを得票率で実質的に上回っていた(Blake 2016)。つまり、仮に世論調査はそれが行われた時点では正確であったとしても、選挙戦最終盤まで誰に投票するかを決めなかった有権者が最後にトランプに投票したことを世論調査は捕捉できなかつたため、予測に誤差が生じた可能性があると考えられる。

最後に報告書が検討するのは、いわゆる「隠れトランプ支持者(“hidden Trump supporters”)¹⁰」の仮説である。この仮説では、トランプの過激な言動からトランプ支持を口にすることは望ましくないとの意識が働き、世論調査において実際はトランプに投票しようと思っている回答者が何らかの形でウソをついた結果、予測においてトランプの過小評価が起きたとされる。このように、世論調査において回答者が自らの本音ではなく社会的に望ましいとされる回答をしたり、逆に社会的に望ましくないとされる回答をしなかったりすることによって生じるバイアスを社会的望ましきバイアス(social desirability bias)とよぶ。

報告書ではこうした社会的望ましきバイアスがトランプ支持について存在していたのかを確認するために、調査員を介する電話調査と、比較的本音を回答しやすいと考えられる調査員を介さないオンライン調査の結果について、2016年9月1日に以降に実施されたすべての世論調査を比較している。それによると、電話調査とオンライン調査とはトランプ支持にほとんど差が無く、電話調査が一貫してトランプ支持を過小評価しているという証拠は無かった。これは世論調査による調査期間の違い、1回きりの調査あるいは継続調査の違い、回答者全体と投票しそうな回答者の違いを考慮しても同様であった。つまり社会的望ましきバイアスによる、トランプ支持者による虚偽の回答によって予測が歪められた可能性は低いと言える。

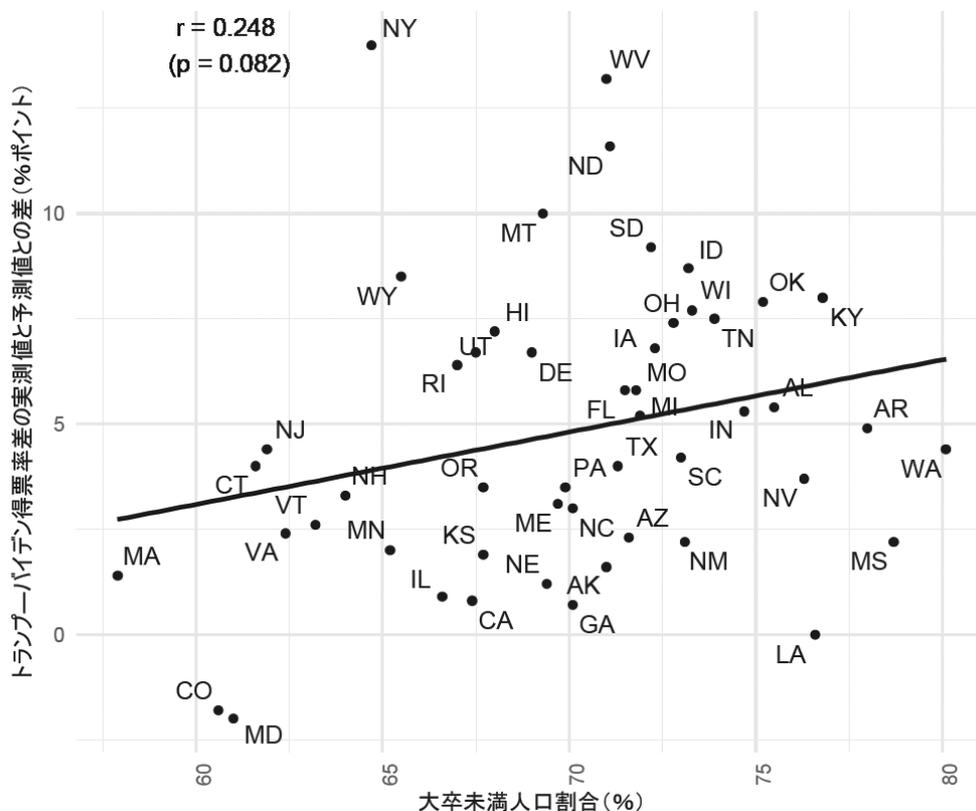
以上まとめると、報告書の結論としては、2016年に予測が失敗した原因として可能性が高いのは、世論調査の標本における高学歴者の過大代表を重み付けにより修正できなかったこと、そして最後まで誰に投票するか迷っていた有権者の存在を考慮できなかったこと、ということになる。これらの点をふまえて2020年の予測のあり方はどのように異なっていたのであろうか。まず高学歴者の過大代表の問題については、2016年より多くの世論調査で教育程度による重み付けが行われるようになっており(Cohn 2020)、FTEでも複数の世論調査結果の加重平均をとる際、教育程度による重み付けが行われている調査に対しより大きな重みを与えるなど改善が見られる(Rakich 2020)。直前に投票先を決めた有権者の存在についても、Morning Consultantの調査によると選挙1週間前の時点で誰に投票するか決めかねていた有権者の割合も2016年に11%だったのが2020年には3%に減っているうえ(Yokley 2020)、最終盤に投票先を決めた有権者の間でトランプに投票する傾向が強いということもなかったことから(Enns and Schuldt 2020)、大きな問題にはならなかったと

がまたしても失敗したことを説明するのは難しいように思える。では何が2016年と2020年の予測の失敗を説明するのだろうか。

1つ考えられる原因は、重み付けによっては解決されない非回答バイアスである¹²。先に述べたような重み付けによって母集団と標本における属性の比率を調整することで非回答バイアスを軽減し誤差を小さくするという方法が機能するためには、標本に含まれる特定の属性をもつ有権者の意見と標本に含まれなかったその属性をもつ有権者の意見が同じ傾向をもつという重要な前提が満たされている必要がある。例えばトランプ支持者ほどそもそも調査を拒否する傾向が強いことにより、標本に含まれる非大卒者のトランプ支持率が標本に含まれない非大卒者のそれよりも有意に低く、両者の意見がずれているのであれば、いくら標本に含まれる非大卒者の回答を重み付けによって「水増し」しても母集団の支持率には近づかない。2016年の反省をふまえて世論調査における大卒者の過大代表の問題に対処するために2020年には重み付けが改善されたものの、それでも予測の誤差が縮小しなかったのはこれが1つの原因と考えられる。

このように重み付けで解決できない非回答バイアスによって予測の誤差が大きくなったという可能性についてごく簡単に検討してみたい。図4は大卒未満人口割合と、図2で用いた2020年の得票率差の実測値と予測値の乖離（トランプ得票率過小評価の度合い）との関係を示したものである¹³。先にも述べたとおり学歴はトランプ支持と強い相関をもつ属性である。したがってトランプ支持者ほど調査を拒否するという傾向が存在する場合¹⁴、

図4. 大卒未満人口割合と予測の誤差（2020年）

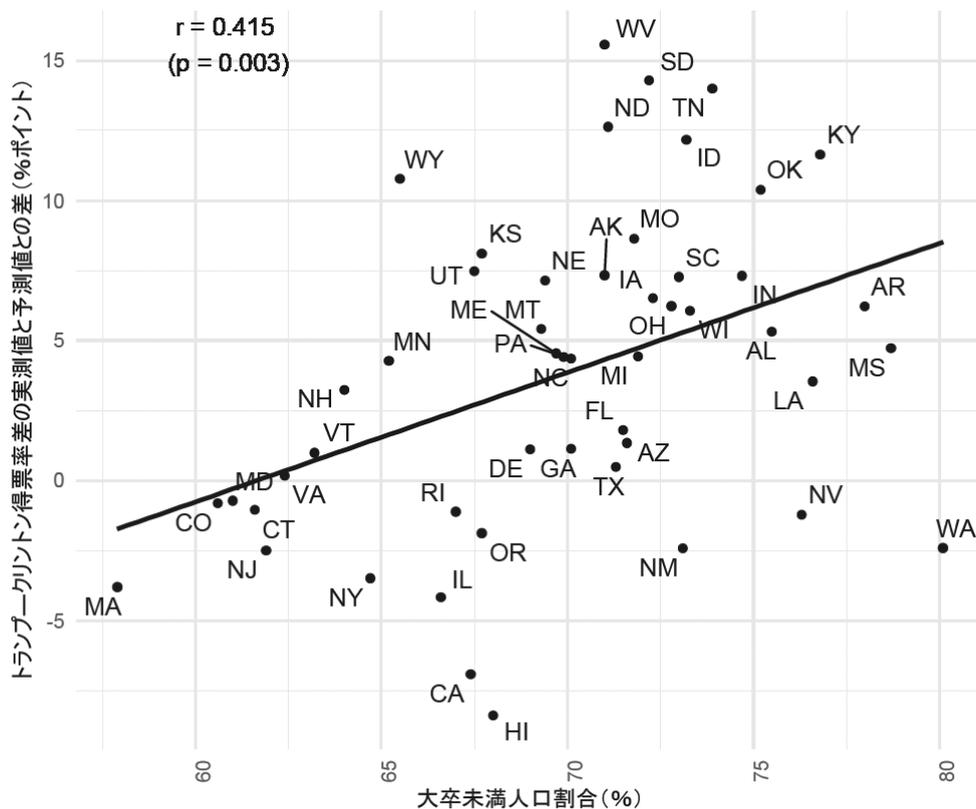


図出典：American Community Survey および FiveThirtyEight のデータを用いて筆者作成。

そもそも標本に含まれる非大卒者におけるトランプ支持率の方が母集団に含まれる非大卒者におけるそれより低いため、大卒未満人口割合が高い州ほどトランプ得票率の過小評価が起きる度合いは大きくなると考えられる。

図4によると実際、大卒未満人口割合が高い州ほど予測はトランプ得票率を過小評価する傾向があることがわかる。両変数の正の相関は10%水準で統計的に有意であるが、高卒未満人口割合が低い割に誤差が大きく外れ値となっているニューヨーク州を除くと、5%水準で統計的に有意となる。また、こうした大卒未満人口割合が高い州ほどトランプ得票率を過小評価する度合いが強いという関係は図5に示したとおり2016年にも見られる。さらなる厳密な検証が必要ではあるが、2020年に教育程度による重み付けが改善されてもなお2016年と同様の傾向が見られるということは、やはり重み付けによっては解消されない非回答バイアスが存在する可能性がある。つまりトランプ支持者には調査でウソをつく傾向があるのではなく、そもそも調査自体を拒否する傾向があるため、母集団における非大卒者の間でのトランプ支持が世論調査に適切に反映されていないことが2016年と2020年に共通する失敗の原因であると言えるのではないだろうか。

図5. 大卒未満人口割合と予測の誤差 (2016年)



図出典：American Community Survey および FiveThirtyEight のデータを用いて筆者作成。

また2020年に予測が大きく外れたフロリダ州ではその原因として、ラティーノの有権者の中でもキューバ系が今回大きくトランプ投票に傾いたことを考慮できなかったということが指摘される。これもラティーノの中でもトランプ支持の強いキューバ系が世論調査に

協力しなかったという非回答バイアスにより、標本に含まれるラティーンにおけるトランプ支持率と母集団に含まれるラティーンにおけるトランプ支持率との間に乖離が生じ、全体としてトランプ得票率の過小評価が起きた可能性が考えられる。

いずれにせよ、こうした予測の対象である候補者への支持が世論調査への協力の有無と相関をもつために重み付けによって非回答バイアスが解消できないという問題は、分極化するアメリカの選挙結果予測において今後も重要な課題となり続けるであろう。

参考文献

日本語

小椋郁馬. 2020年. 「世論調査は外れたのか? : 2020年アメリカ大統領選挙」ikeの日記, 11月16日. <https://ike-og.hatenablog.com/entry/2020/11/16/102935> (2021年3月11日最終閲覧)

英語

Blake, Aaron. 2016. "How America Decided, at the Last Moment, to Elect Donald Trump." *The Washington Post*, November 17. <https://www.washingtonpost.com/news/the-fix/wp/2016/11/17/how-america-decided-at-the-very-last-moment-to-elect-donald-trump/> (2021年3月10日最終閲覧)

Cohn, Nate. 2020. "Are State Polls Any Better Than They Were in 2016?" *The New York Times*, May 12. <https://www.nytimes.com/2020/05/12/upshot/polls-2020-trump-biden.html> (2021年3月10日最終閲覧)

Enns, Peter and Jonathon Schuldt. 2020. "Did late deciders for Donald Trump impact the election polls this year?" *The Hill*, Nov. 20. <https://thehill.com/opinion/white-house/527259-did-late-deciders-for-donald-trump-impact-the-election-polls-this-year> (最終閲覧2021年3月10日)

File, Thom. 2017. "Voting in America: A Look at the 2016 Presidential Election." U.S. Census Bureau, May 10. https://www.census.gov/newsroom/blogs/random-samplings/2017/05/voting_in_america.html (2021年3月10日最終閲覧)

Isakov, Michael and Shiro Kuriwaki. 2020. "Towards Principled Unskewing: Viewing 2020 Election Polls Through a Corrective Lens from 2016." *Harvard Data Science Review*, 2(4). <https://doi.org/10.1162/99608f92.86a46f38> (2021年3月11日最終閲覧)

Kennedy, Courtney, Mark Blumenthal, Scott Clement, Joshua D Clinton, Claire Durand, Charles Franklin, Kyle McGeeney, Lee Miringoff, Kristen Olson, Douglas Rivers, Lydia Saad, G Evans Witt, Christopher Wlezien. 2018. "An Evaluation of the 2016 Election Polls in the United States." *Public Opinion Quarterly* 82 (1): 1-33.

Matthews, Dylan. 2020. "One pollster's explanation for why the polls got it wrong." *Vox*, Nov. 10. <https://www.vox.com/policy-and-politics/2020/11/10/21551766/election-polls-results-wrong-david-shor> (2021年3月11日最終閲覧)

Groves, Robert M., Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, eds. 2001. *Survey Nonresponse*. Wiley-Interscience.

Rakich, Nathaniel. 2020. "How To Read Polls In 2020." *FiveThirtyEight*, Jun. 15. <https://fivethirtyeight.com/features/how-to-read-2020-polls-like-a-pro/> (2021年3月10日最終閲覧)

Yokley, Eli. 2020. "Trump and Biden Look Set to Split the Few Remaining Undecided Voters of 2020." Oct. 27. <https://morningconsult.com/2020/10/27/trump-biden-undecided-voters-polling/> (2021年3月10日)

— 注 —

¹ FTEの2016年と2020年の予測値はそれぞれ以下のURLから参照することができる。
<https://projects.fivethirtyeight.com/2016-election-forecast/>
<https://projects.fivethirtyeight.com/2020-election-forecast/> (2021年3月5日最終閲覧)

- 2 BSの計算にあたっては、全州と州内各下院選挙区に議席が割り振られるメイン州とネブラスカ州については全州の議席予測のみを使用した。なお、FTEは2016年については勝利確率を小数点第1位まで出しているが、2020年は小数点まで出しておらず、この点で若干の違いがある。
- 3 NYTのBSについては、BuzzfeedNewsのデータエディターであるJeremy Singer-Vineによる計算を参照した。Jeremy Singer-Vine. “Grading The 2016 Election Forecasts.” BuzzfeedNews, November 28, 2016, <https://www.buzzfeednews.com/article/jsvine/2016-election-forecast-grades>. (2021年3月5日最終閲覧)
- 4 各州における両候補者の得票率は、Federal Election Commissionが集計したデータを参照した。<https://www.fec.gov/> (2021年3月11日最終閲覧)
- 5 NYTは例外的にシエナ大学と共同で州レベルの世論調査を実施しているが、これにしてもあくまで平均値算出の元となる調査の1つに過ぎない。
- 6 <https://projects.fivethirtyeight.com/2020-election-forecast/florida/> (2021年3月5日最終閲覧)
- 7 合計が100%にならないのは、リバタリアン党、緑の党などの泡沫候補がいるため。
- 8 ただし各「サイコロ」の目が独立していると仮定するのではなく、The Economist誌の予測のように隣接州どうし（例えば激戦州である中西部の州どうし）で相関をもつと仮定することもできる。“How The Economist presidential forecast works.” The Economist, <https://projects.economist.com/us-2020-forecast/president/how-this-works> (2021年3月5日最終閲覧)
- 9 実際2016年には、2012年と比べてとりわけ黒人の間で投票率が下がるなど投票率の変化はトランプに有利に働いていた (File 2017)。
- 10 英語では、「恥ずかしがり屋のトランプ支持者 (shy Trump supporters)」とされることも多い。
- 11 投票率はUnited States Elections Projectが集計した“VEP Turnout Rate (Total Ballots Counted)”のデータを参照した。<http://www.electproject.org/home> (2021年3月11日最終閲覧)
- 12 なおこのような属性が同じでも標本に含まれる有権者と標本に含まれない有権者の間でトランプ支持傾向が体系的に異なることにより、重み付けによる修正が機能しない可能性についてはすでに多く指摘されている。例えばIsakov and Kuriwaki (2020)、Matthews (2020)、小椋 (2020) など。
- 13 大卒未満人口割合は2013年から2017年の間に国勢調査局 (U.S. Census Bureau) によって実施されたAmerican Community Surveyの回答にもとづく集計値を参照した。<https://www.census.gov/programs-surveys/acs> (2021年3月11日最終閲覧)
- 14 紙幅の関係でここでは紹介しないが、筆者によるオンライン調査によるとトランプ支持者はとりわけリベラル系メディアによる世論調査を拒否する傾向がある。