

第24回参議院選挙の議席数予測を振り返る

—Yahoo! JAPAN ビッグデータレポート—

Revisiting the Big Data-Driven Forecast of Japan's 24th Upper House Election by
Yahoo! JAPAN

安宅 和人

Kazuto Ataka

1. 2016 参議院議員選挙の結果と予測との比較
2. 前提条件と実際の結果の比較検証
3. 振り返りのまとめ

<要旨>

ヤフーにおける検索クエリデータ（注目度）をもとに第24回参院選の予測を行った。前回の参院選は96%の一致だったが、今回は91%の一致に終わった。ズレの8割は野党連合による勝利を当てられなかったことによるものだった。党別の検索ボリュームと獲得票数の相関は低下したが、データに基づきコンバージョン補正を行ったところほぼ成功した。選挙前の注目度の盛り上がりはどの党もほぼ予測どおりであった。注目度の多寡によるベースライン補正は今回の選挙でも有効だった。公明党の得票率には周期性があるが今回は乖離が大きく、選挙区においては候補者数と得票率が直結していることを発見した。

We conducted a forecast of Japan's 24th Upper House election based on Yahoo! JAPAN's search query data. The result was 91% match, lower than 96% match achieved in the 2013 election. Opposition alliance was a major cause of the mismatch. Conversion correction by party, as well as base-line adjustment by total related query volume worked fine. We also found that ballot earning rate of New Komeito in non-proportional district highly correlated with the number of candidates.

7月10日（日）に第24回参議院選挙が行われました。それに先立ち、ビッグデータレポートでは「ビッグデータが導き出した第24回参議院選挙の議席数予測」の中で今回の参議院選挙の獲得議席数予測を実施し、その数値を発表しました。今回のレポートでは、予測結果の数値と実際の選挙結果の数値を比較したいと思います。

1. 2016 参議院議員選挙の結果と予測との比較

まずは実際の選挙結果を確認してみましょう。

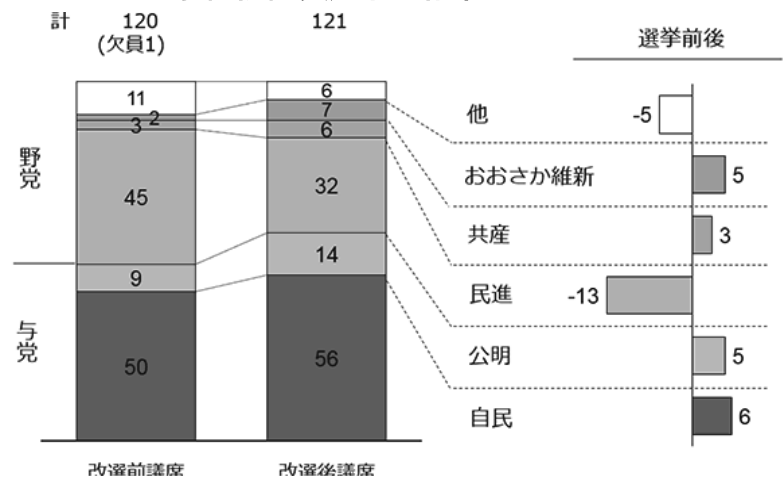
与党はそれぞれ改選前に比べて議席数を増加させ、自民党が6議席増加の56議席、公明党は5議席増加の14議席となりました。一方、野党は民進党が改選前の45議席から大幅減の32議席となりましたが、共産党は3議席増加の6議席、おおさか維新の会は5議席増加の7議席といった結果となりました。

次に、ビッグデータレポートの予測と選挙結果の比較です。

一致率は、比例区で94%、選挙区で82%（全体で91%）となり、選挙区予測での乖離が大きい結果となりました

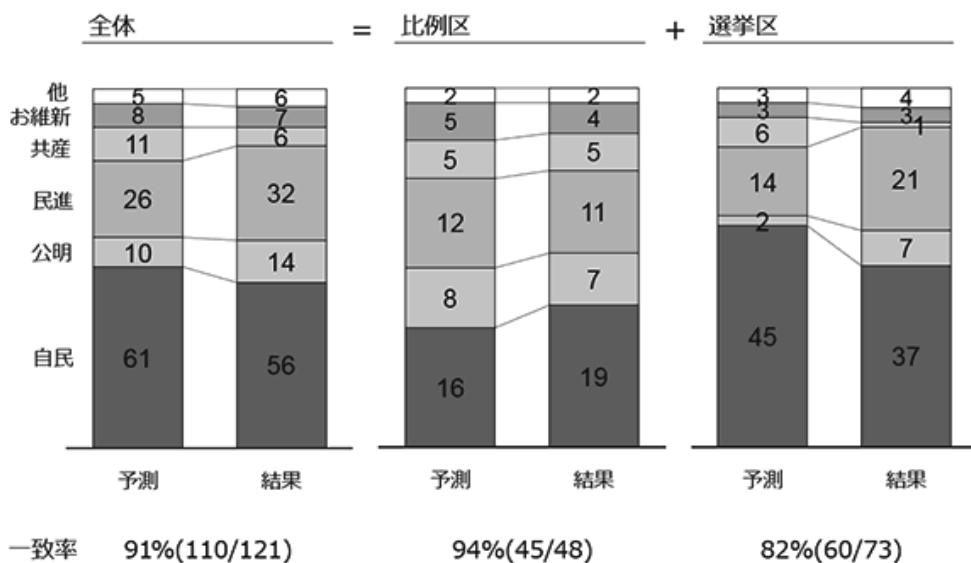
（2013年の参院選では96%の一致）。今回のビッグデータレポートでは、過去の分析結果から導き出された次の前提に基づき予測数値を算出しております。

図1. 2016 参議院議員選挙の結果



（資料：参議院HP，選挙結果データ）

図2. 2016 参院選最終予測と結果比較



（資料：「Yahoo!検索」データ，選挙結果データ）

- 前提 1:** 公示日から投票日前日までのネット上の注目度は、政党の得票数に直接的に相関する
- 前提 2:** ネット上の注目度が政党の得票にどの程度つながるかは、一部の例外を除き党ごとに一定である
- 前提 3:** 前提 2 で表したつながりやすさは、自民党が上昇傾向に、共産党が下降傾向にある
- 前提 4:** 公明党の得票率はネット上の注目度に影響されず、一定幅の中で周期的に変動する
- 前提 5:** 投票日までの注目度の盛り上がり度合いは党によって一定である
- 前提 6:** 予測の基準としている選挙と比べて公示後の注目度にズレがあった場合は、基準としている選挙と同水準の注目度になるようベースラインを補正する必要がある

この前提条件に当てはめて予測を行いました。結果との乖離があったということはいずれかの前提条件に変化が生じた、もしくは適用に問題があった可能性を示唆しています。

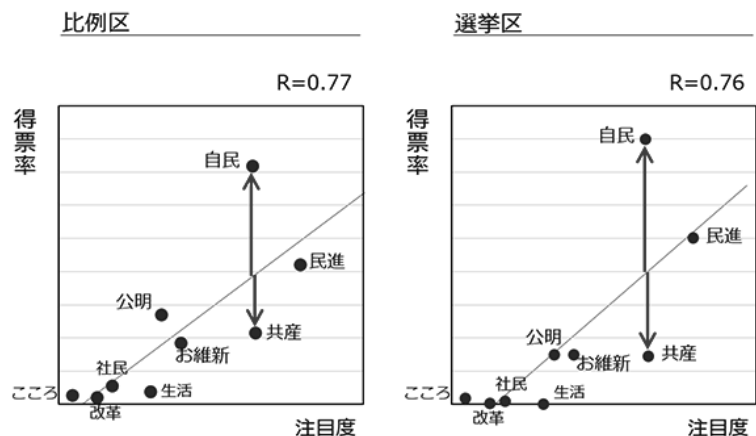
2. 前提条件と実際の結果の比較検証

過去の分析結果から導き出された6つの前提条件に対して検証を行いました。

前提 1 : 公示日から投票日前日までのネット上の注目度は、政党の得票数に直接的に相関する

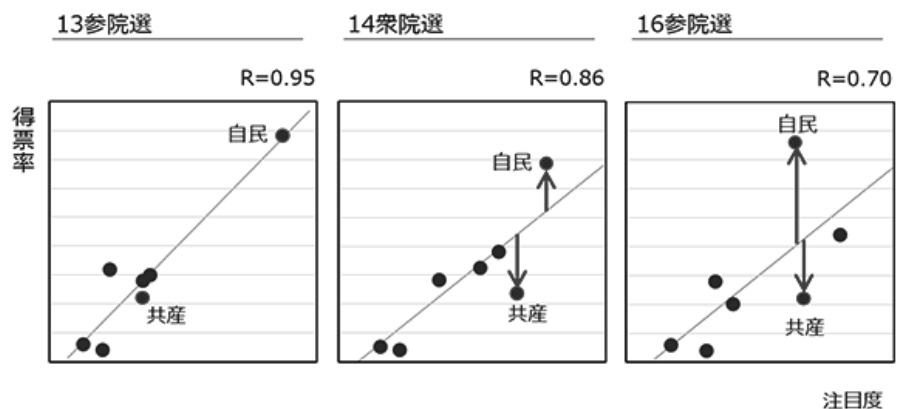
議席数予測において、もっとも重要な要素なのがこの前提1の条件です。第一回目のビッグデータレポートで公開したこの条件が今回の選挙にも当てはまっているのかを確認しました。その結果、図3に見るとおり、自民党と共産党が相関を表す回帰直線から大きく乖離していることがわかりました。

図3. 2016 参院選の注目度と得票率の関係



(資料:「Yahoo!検索」データ, 選挙結果データ)

図4-A. 比例区における注目度と得票率の関係の推移



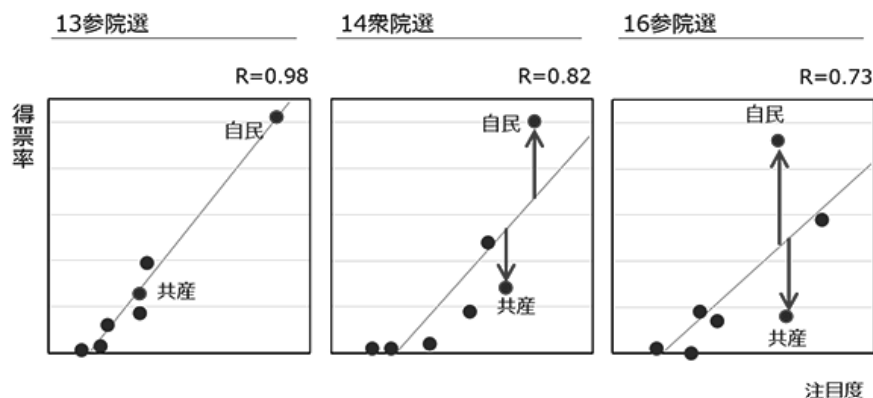
(資料:「Yahoo!検索」データ, 選挙結果データ)

本来は 2013 年参院選のグラフのようにすべてのプロットが直線上か誤差のほとんどない近い位置に並ぶという結果になるはずでしたが、比例区・選挙区とも

に自民党と共産党が大きく乖離しました。

その傾向は前回の2014年衆院選の時にも若干見られていましたが、今回の2016年参院選ではさらに大きく広がっており、この前提条件とのズレが予測を困難にしたことがわかりました。

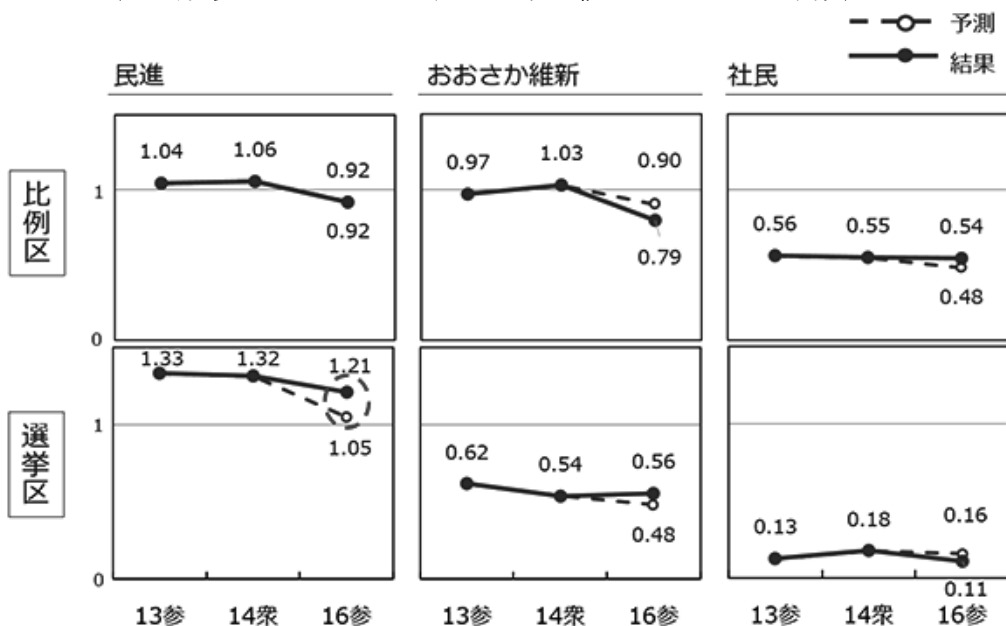
図4-B. 選挙区における注目度と得票率の関係の推移



(資料:「Yahoo!検索」データ, 選挙結果データ)

前提2: ネット上の注目度が政党の得票にどの程度つながるかは、一部の例外を除き党ごとに一定である

図5. 自民、共産以外の主要政党のつながりやすさの予測と結果
(主要政党のつながりやすさの平均値=1.0とした指数)



(資料:「Yahoo!検索」データ, 選挙結果データ)

一部の例外政党を除き、各政党のネット注目度がどの程度得票数につながるかは一定であるという前提条件は果たしてどうでしょうか。

図5に見るとおり、自民党、共産党、公明党を除いた主要政党はほぼ一定であり、大きなズレは見られませんでした。そのため、つながりやすさの予測値と選挙結果の間にも大

前提3：前提2で表したつながりやすさは、自民党が上昇傾向に、共産党が下降傾向にある

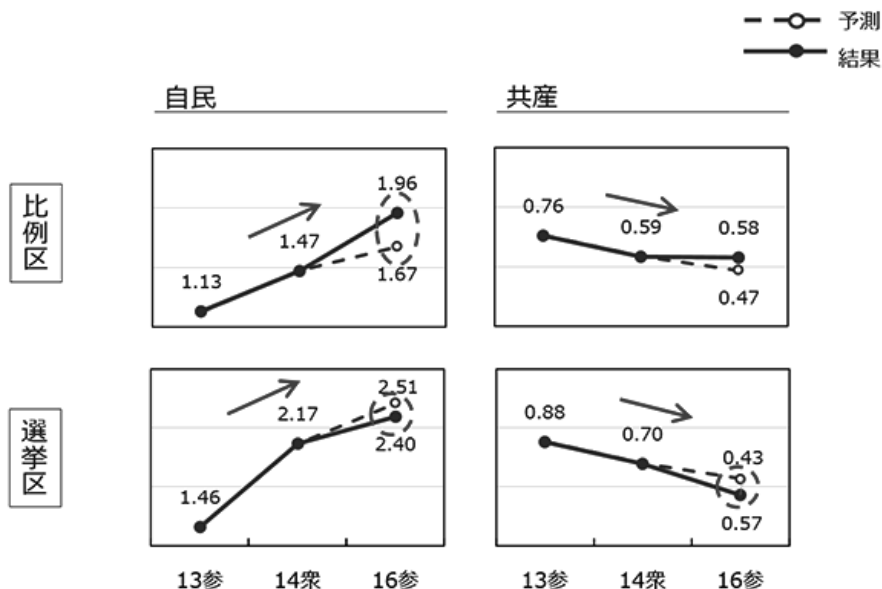
きな相違はありませんでした。ただし、民進党の選挙区でのつながりやすさは下がる予想をしていましたが結果は横ばいとなり、これが選挙区における民進党の議席数を過少に予測した一因となりました。

自民党と選挙区の共産党については、図6に見るとおり、今回の選挙でも前提条件通りの傾向が確

認されました。しかし、自民党の比例区においては結果は予測以上に上昇をし、選挙区では予測ほどは伸びない結果となりました。選挙区の共産党についても同様に、下降傾向を盛り込んだ予想をしましたが、結果は予想以上に下がりました。

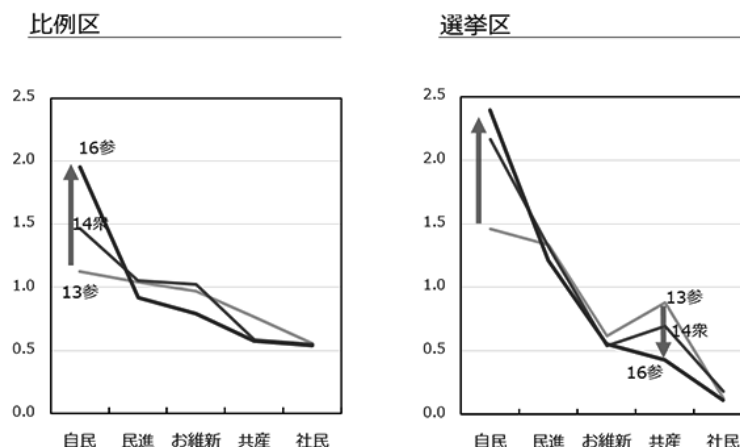
結果、図7のように自民党の上昇傾向、共産党の下降傾向は予測に盛り込んでいたものの、その上昇・下降度合いについてズレが生じた結果となりました。

図6. 自民と共産のつながりやすさの予測と結果
(主要政党のつながりやすさの平均値=1.0とした指数)



(資料：「Yahoo!検索」データ、選挙結果データ)

図7. 政党別にみた注目度の得票へのつながりやすさ
(主要政党のつながりやすさの平均値=1.0とした指数)



前提4：公明党の得票率はネット上の注目度に影響されず、一定幅の中で周期的に変動する

前回の2014年衆院選の振り返りレポートで発見したこの前提は、今回の選挙にも当てはまったのかも見てみました。

図8のとおり、比例区の予測値15.3%に対して結果は13.5%となり、周期性からのズレが発生しましたがゆらぎの範囲内には収まったため、サンプル数が増えていくほど精度が高まりそうです。

図9のとおり、選挙区は予測値4.7%に対して結果7.5%と大きく乖離し、ゆらぎの範囲からも外れる結果となりました。しかし、選挙区に関しては今回新たな発見がありました。

参院選の選挙区では、公明党の候補者数と得票率の間に高い相関($R=0.99$)があることがわかりました。これはつまり、選挙区で候補者を擁立すればその擁立人数に比例して選挙区の総得票数に占める公明党の得票率も増えるという公明党の強い組織力を示しています。プロットした点もほぼ一直線上に並んでおり、興味深い発見といえます。

図8. 公明党の得票率の予測と結果：比例区 (単位：%)

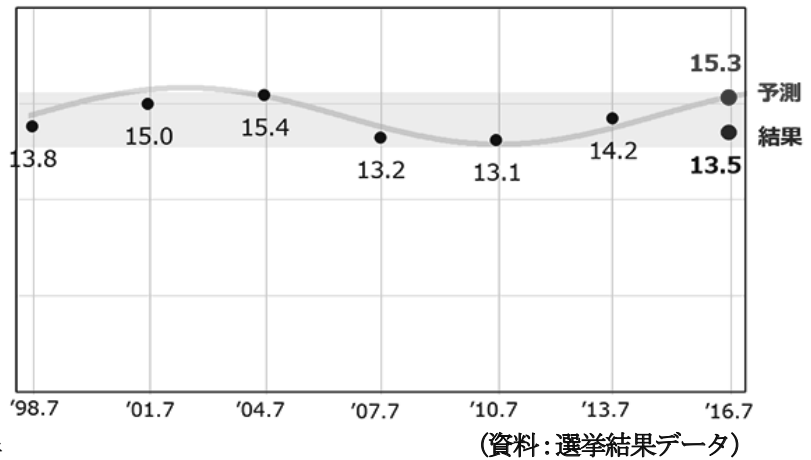


図9. 公明党の得票率の予測と結果：選挙区 (単位：%)

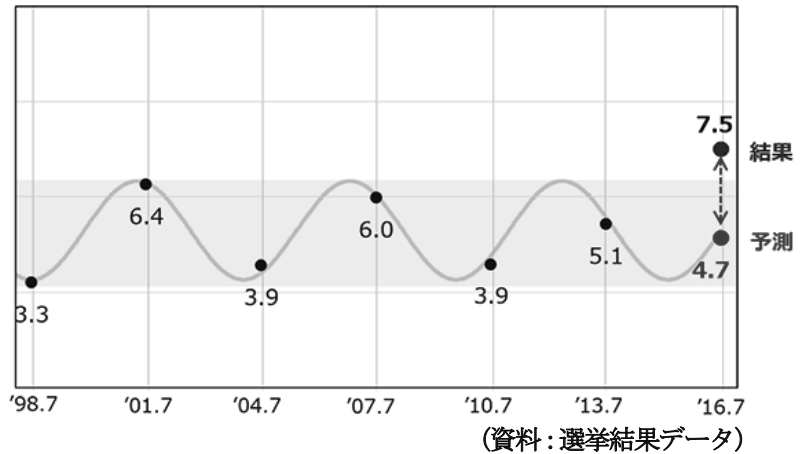
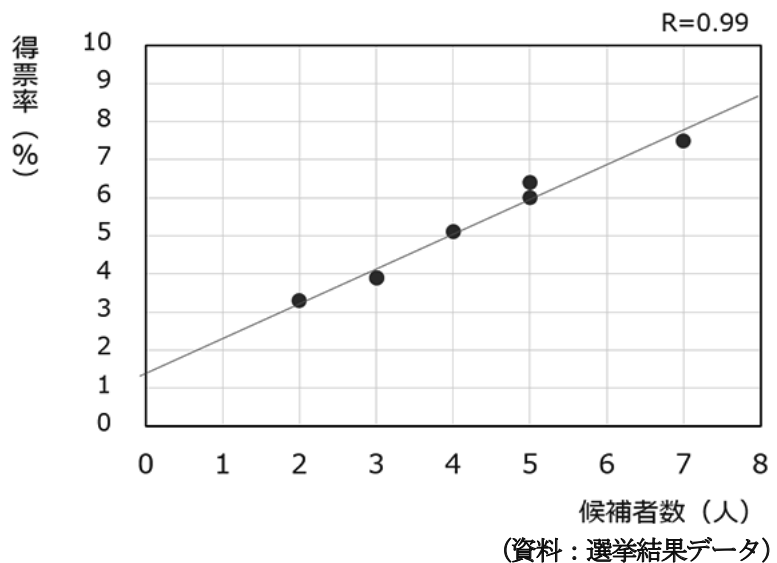


図10. 公明党の選挙区における候補者数と得票率の関係



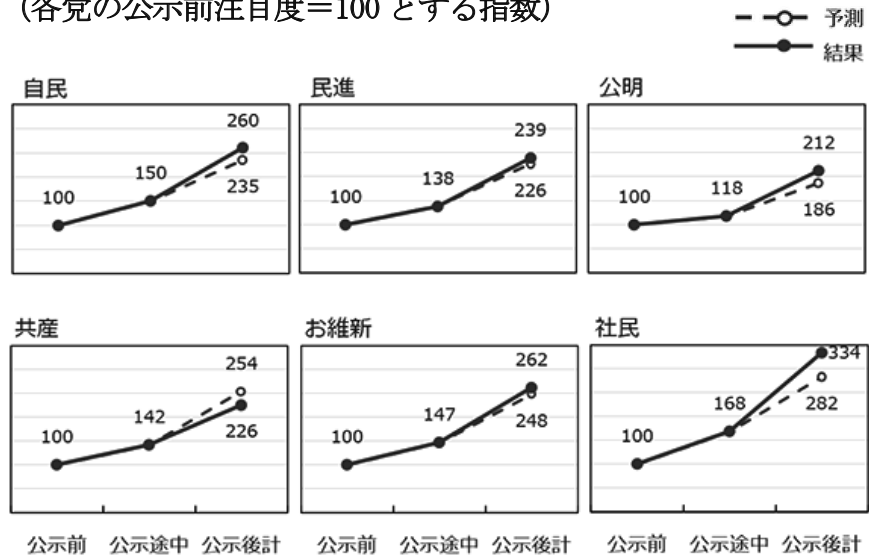
前提5：投票日までの注目度の盛り上がり度合いは党によって一定である

公示日以降の注目度の盛り上がり度合いは、2014年衆院選では予測とほぼ合致した条件です。今回の選挙でも合致したのかを検証しました。

チャートから見てもわかるとおり、この前提条件は今回の参院選でもほぼ合致しました。

全体の傾向としては、共産党以外は盛り上がりが予測よりも若干上回りましたが、共産党だけは逆に予測よりも少し下回っていました。これも選挙区における共産党の過剰予測の要因となりました。

図11. 公示後注目度の盛り上がり度の予測と結果
(各党の公示前注目度=100とする指数)

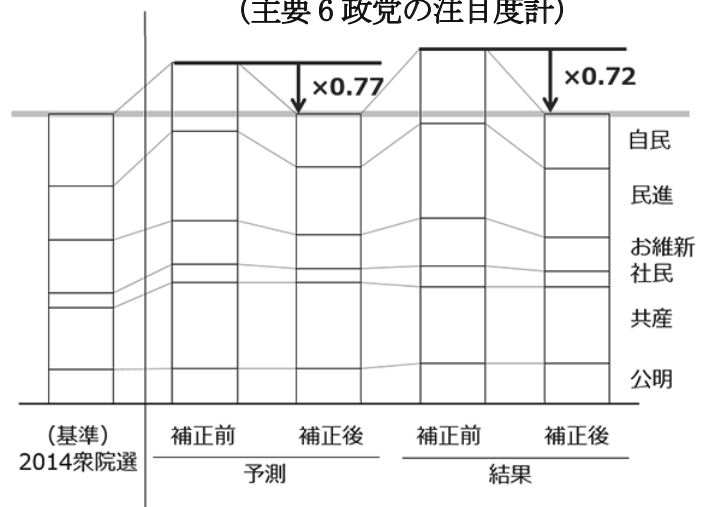


前提6：予測の基準としている選挙と比べて公示後の注目度にズレがあった場合は、基準としている選挙と同水準の注目度になるようベースラインを補正する必要がある

最後の前提はベースライン補正です。2014年衆院選の注目度を基準にして、ベースラインを補正して適用しました。

結果として、予測で実行した補正の値は実際に必要だった補正值と大きな乖離がなかったことがわかりました。

図12. ベースライン補正值の予測と結果
(主要6政党の注目度計)



3. 振り返りのまとめ

また、前提条件以外のところでも注目すべき点がありました。それは、選挙区の一入区の予測値と選挙結果との比較です。

図13に見るとおり、自民党が当選した一入区の一一致率は95%と高い値となりましたが、一方で野党連合勝利となった選挙区の一一致率は18%に留まりました。これが

(資料：「Yahoo!検索」データ)

今回の選挙区予測のズレの3分の2に該当します。また、一人区で野党連合が勝利した選挙区は得票数でも接戦が多かったこともわかっており、それも予測が難しかった一因でした。野党連合という形態に対する予測方法や、接戦が予想される場合の予測方法についての課題がみえてきました。

図 13. 2016.7 参院選選挙区予測の一致率

		選挙区数 (うち接戦)	議席数	一致率 (%)
一人区	自民党勝利	21 (3)	21	95
	野党連合勝利	11 (9)	11	18
複数人区		13 (9)	41	85

(接戦の定義：当落の得票率の差が10%未満)
(資料：「Yahoo!検索」データ，選挙結果データ)

以上が、第24回参院選の議席数予測レポートの振り返りになります。前提条件の振り返りでは、前提条件そのものがずれてしまっているものと合致しているものがあることがわかりました。その中でも主として前提3の変化に伴い、選挙予測の基礎となっている前提1にズレが生じていることが予測値に大きな影響を与える結果となっていることがわかりました。これは、総じて言えば自民安定政権での信頼度の向上、共産党の支持層の変化を数量的に示していると言えます。また、公明党の立候補者数と得票の関連は組織としての強さを示しています。

このように、Yahoo! JAPAN ビッグデータレポートではさらなる仮説検証を続ける中で、新たなデータ視点での知見を提供していければと考えております。

(ヤフー株式会社 チーフストラテジーオフィサー：CS0)

この抄録は「Yahoo! JAPAN ビッグデータレポート」に掲載されたものを整形して再録したものである。

〈参考リンク〉

Yahoo!JAPAN, 「Yahoo! JAPAN ビッグデータレポート (前編)」

<http://docs.yahoo.co.jp/info/bigdata/> , (July 12, 2016).

—————, 「Yahoo! JAPAN ビッグデータレポート (後編)」,

<http://docs.yahoo.co.jp/info/bigdata/> , (August 9, 2016).