

Google Surveys と有権者名簿抽出ネット調査

—朝日新聞社の新しい試み—

Searching for Alternatives: Google Surveys and an Online Survey Using Probability Samples

齋藤 恭之

Yasuyuki Saito

1. Google Surveys (GS) について
2. 東京都知事選, 新潟県知事選でのGS設計
3. GS 結果分析
4. GS まとめと問題点
5. 郵送ネット調査の目的と運用
6. 郵送ネット調査の結果分析
7. 郵送ネット調査まとめと問題点
8. おわりに

〈要旨〉

朝日新聞社は、2016年から全国世論調査で固定電話と携帯電話のデュアルフレームRDD調査を開始した。しかし、携帯電話番号には地域情報がないため、地域調査や地域を限定した選挙調査は固定電話のみを対象としている。そのため、携帯はあるが固定電話はない「携帯限定層」には、固定電話による選挙調査ではアプローチできない。「携帯限定層」が増え、カバレッジエラーが増えると調査精度が悪くなる。そこで今後の選挙調査をどうするかという観点から、新しい試みを行った。一つはGoogle Surveys (GS) の利用である。GSではGoogleの検索エンジンを使って検索した人にランダムに最大10問までの短い質問をポップアップとして表示し、回答してもらう。ここでは、2016年の東京都知事選と新潟県知事選での各3回行ったGS実験について論じる。もう一つは、有権者名簿から抽出した対象者に手紙を送って、ネットで回答してもらう調査の実験である。ここでは、2016年の衆院東京10区の補選での実験調査について論じる。

In 2016, Asahi Shimbun started the dual-frame RDD survey of landline and cell phones for the nationwide public opinion surveys. However, because cell phone numbers do not indicate regional information, we have to conduct the regional pre-election poll on landline phones only. Although, we can call people who have both cell and landline phones, we cannot approach those with cell phones only. As the number of people who only have cell phones is increasing, the coverage error of landline phone surveys also increases and their precision decreases. Therefore, we have attempted to find alternative survey methods for future pre-election polls. One alternative is Google Surveys (GS), whereby people who use Google's search engine are randomly invited to answer a short questionnaire that consists of up to 10 questions, which appears in a pop-up window. Here, I will discuss our testing of GS, three times each, for both the Tokyo Metropolitan Governor and Niigata Prefectural Governor elections in 2016. Another option is an online survey of people who are randomly chosen from the voter list; letters are sent to those selected inviting them to answer a questionnaire on the website. Here, I will discuss our testing of such an online pre-election poll for the Lower House supplemental election in Tokyo 10th district in 2016.

1. Google Surveys (GS) について

Google は2012年にGSを始め、2015年には日本でも開始した。Googleは、GSをウェブサイト運営者やユーザー（調査対象者）と調査依頼者双方にメリットのある「エコシステム」と位置づけている。

ウェブサイトの運営者がGoogleの「サイト運営者向けネットワーク」に入ると、Google検索で自社サイトの有料コンテンツなどにアクセスしてきたユーザーに対して、閲覧前にアンケートの依頼ができる。サイト運営者は調査依頼者から収益を得られる。一方、ユーザーにとってはアンケートの回答で有料コンテンツを無料で見られる。

GSはGoogle独自のアルゴリズムにより、ユーザーにPCやスマホに最適化された調査画面をポップアップ表示する。ユーザーは調査に回答することによって、有料コンテンツを見ることが出来る。Googleは、回答者の閲覧履歴からユーザー属性を推定している（Cookieから性別と年齢を、IPアドレスから場所を推定）。

調査依頼者は、回答数や各選択肢のパーセントなど調査の結果をGSのアンケート結果ページで随時確認できる。

ローデータも随時ダウンロードでき、ローデータには各ユーザーの回答データ以外に、推定の性別、推定の年代などが含まれる（詳細はGSのウェブサイトを参照）。

なお、この報告は、齋藤恭之、長崎緑子作成、Googleサーベイを使用して実施（2016年7月～10月）した調査結果を使っている。

図表1. ポップアップ画面の例（PC）



図表2. ポップアップ画面の例（スマホ）



2. 東京都知事選、新潟知事選でのGS設計

①日程と有効回答（日程はいずれも2016年）

東京都知事選：投開票日は7月31日

1回目 7月15日～17日（1008件）

2回目 7月23日～25日（1010件）

3回目 7月29日～31日（1004件）

新潟知事選：投開票日は10月16日

1回目 9月28日～10月3日（1000件）

2回目 10月6日～13日（1000件）

3回目 10月12日～17日（1000件）

②質問構成

東京・新潟ともに

1, 3回目：1問だけの構成「投票したい候補者」

2回目：7問構成

①関心度 ②投票したい候補者 ③投票に行くか ④（付帯質問）東京：前回知事選での投票候補者、新潟：原発運転再開への賛否 ⑤政党支持 ⑥性別 ⑦年代

3. GS 結果分析

3-1. 性別構成の分析

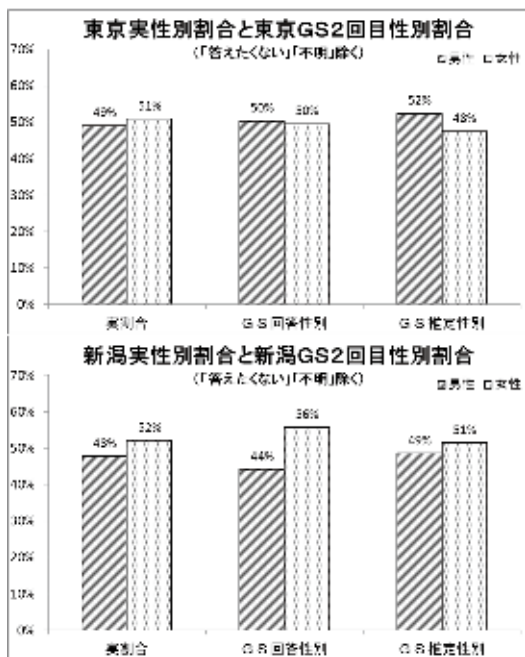
Googleで検索をしてあるサイトを閲覧すると、GoogleはCookie（Webブラウザ内に蓄積される情報）を使ってその情報を取得している。収集した回答者の閲覧履歴に基づいて、性別や年代などのユーザー属性情報を推定している。

今回の2回目のGSで性別質問を回答しなかった人、もしくはGSの推定性別で「不明」だった人を

除くと、GS の男女比は、東京は回答による性別、推定による性別は共に5割前後で国勢調査から求められる実際の割合と大きな違いはなかった。新潟は回答性別で女性が多かったが、推定性別は実際の割合に近かった。

GS の回答性別と推定性別は東京、新潟共に7~8割一致した。

図表3. 質問での回答性別とGS推定性別



図表4. 回答性別とGS推定性別の一致率

東京都GS (2回目)		GS推定性別		
		男性	女性	不明
GS回答性別	男性	80%	6%	14%
	女性	10%	72%	18%
DK/NA		32%	40%	28%

新潟県GS (2回目)		GS推定性別		
		男性	女性	不明
GS回答性別	男性	81%	4%	15%
	女性	11%	75%	14%
DK/NA		36%	40%	24%

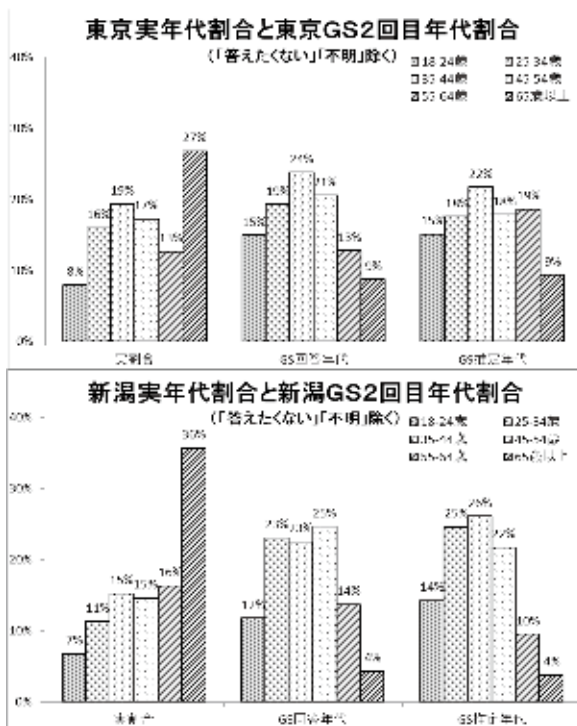
3-2. 年代構成の分析

有権者全体での65歳以上の構成比は東京27%、新潟36%である。しかし、2回目のGSの質問で年代を回答しなかった人、もしくは推定年代が「不明」だった人を除いた65歳以上の構成比は、回答年代、推定年代ともに、東京では9%、新潟では4%とかなり少なかった。また、実構成比では新潟の

方が高齢者が多いのに、GSでは、東京の方が高齢者の比率が高くなっている。

また、GSの推定年代と回答年代の一致率は4割~7割とばらついている。

図表5. 質問での回答年代とGS推定年代



図表6. 回答年代とGS推定年代の一致率

東京都調査 2回目	GS回答年代	GS推定年代						
		18-24歳	25-34歳	35-44歳	45-54歳	55-64歳	65歳以上	不明
18-24歳	18-24歳	56%	9%	1%	2%	0%	0%	31%
25-34歳	25-34歳	11%	43%	11%	3%	2%	1%	30%
35-44歳	35-44歳	2%	11%	54%	9%	5%	2%	17%
45-54歳	45-54歳	1%	5%	11%	47%	18%	3%	16%
55-64歳	55-64歳	2%	4%	4%	13%	62%	6%	10%
65歳以上	65歳以上	0%	1%	1%	3%	16%	70%	8%
DK/NA		12%	12%	10%	9%	12%	2%	43%

新潟県調査 2回目	GS回答年代	GS推定年代						
		18-24歳	25-34歳	35-44歳	45-54歳	55-64歳	65歳以上	不明
18-24歳	18-24歳	70%	9%	2%	3%	0%	0%	16%
25-34歳	25-34歳	5%	71%	7%	2%	1%	1%	12%
35-44歳	35-44歳	3%	15%	53%	11%	1%	1%	18%
45-54歳	45-54歳	3%	3%	23%	48%	5%	1%	17%
55-64歳	55-64歳	2%	2%	11%	22%	40%	7%	16%
65歳以上	65歳以上	3%	3%	5%	0%	24%	42%	24%
DK/NA		16%	8%	23%	15%	5%	1%	33%

3-3. 地域精度

Googleは検索したユーザーのPCやスマホのIPアドレスを認識しており、そのIPアドレスから地域を推定し、調査当該地域だと特定されたユーザーにGS独自のアルゴリズムで質問のポップアップ画面を表示する。

しかし、IPアドレスの地域特定には誤差があり、調査対象である東京都、新潟県以外のユーザーに対してポップアップ画面が表示される可能性がある。そこで、「投票したい候補者」の質問には、東京の2回目の調査では「東京都に住んでいない」という選択肢を加え、東京の1、3回目、および新潟調査では「18歳未満または東京都（新潟県）に住んでいない」という選択肢を入れた。

その結果、調査対象地域ではなかった割合は、東京では15%以上だが新潟では9%以下と、新潟の方が推定地域の精度は高くなっている。この理由として、近隣県から東京に通勤する人が職場などでGSの回答をしている可能性が考えられる。

図表 7. GSの地域推定外れ度合い

「投票したい候補者」質問			
東京GS	1回目	2回目	3回目
(18歳未満または)東京都に住んでいない	15%	16%	16%
新潟GS	1回目	2回目	3回目
18歳未満または新潟県に住んでいない	9%	6%	8%

4. GS まとめと問題点

- ①結果の比較 … 3回目の調査支持率と実得票率では東京、新潟ともに大きな違いはなかった。
- ②調査コストが安価 … 複数回行い、一定期間での支持の上がり下がりを見ることが出来る。
- ③対象者選定のアルゴリズムの不明確さ … どのようにポップアップを出しているのか。代表性は担保できるのか。また、回答率が不明。
- ④GS チームによる質問文のチェックがある … 性別や年代などの個人情報に関係する質問には「答えたくない」などの選択肢の追加を要求される場合がある。
- ⑤調査開始時刻が決まっていない … GSのチェックで問題ないと判断されてから調査が開始される。調査依頼者は調査開始時刻を指定できない。
- ⑥調査対象地域ではない対象者が含まれる … オプションで地域特定用のフィルタリング質問を付加し、地域を限定することは可能。
- ⑦調査期間が一定ではない … 有効回答約1000件を獲得するのに、東京では3日、新潟では1週間前後かかった。

5. 郵送ネット調査の目的と運用

5-1. 調査の目的

2016年10月に衆議院東京10区で補選が行われた。有権者名簿から確率抽出した2000人の対象者に対し、案内の手紙を送り、ネットで回答してもらうという実験調査（以降、「郵送ネット調査」）を行った。

調査の目的は、以下の通り。

- ・有権者名簿から抽出した対象者に対し、郵送でネット調査を依頼する手法が可能かどうか検証する。
- ・回収率がどのくらいかを計る。
- ・事前謝礼や予告はがきの文章、案内の手紙の紙質によって回収率がどう違うか調べる。
- ・回答者の性別、年代などの属性が、有権者名簿の属性とどの程度合っているか調べる。
- ・回答結果とRDD調査の結果や選挙結果と比較し、同じ傾向なのか違うのか確認する。
- ・回答デバイスの種類や謝礼の種類に何らかの傾向があるのか調べる。

5-2. 調査の運用方法

まず、衆議院東京10区の地域（豊島区および練馬区の一部）の有権者名簿に対して、層化無作為2段抽出法で2000名の調査対象者を抽出した。抽出時の対象者の情報は、氏名、住所、性別、年齢（生年月日）。

調査の運用だが、まず対象者一人ひとりにインターネットの調査画面とそのURL、QRコードを用意。パソコンやスマホのブラウザにURLを入力したり、スマホでQRコードを読み取ったりすることでその人の調査画面を開いてもらう。

図表 8. URLとQRコードのイメージ



調査開始約1週間前に対象者に予告ハガキを送付。各人のURLとQRコード、問い合わせ用のフリーダイヤルの電話番号を記載した案内の手紙は、調査開始1日前に送付した。調査画面は告示日にオープンし、投開票日前日にクローズ。調査期間は12日間だった。

問い合わせ用のフリーダイヤルに電話してきた人で、インターネットにアクセスできる環境がなく、回答を希望する人には、電話でこちらから質問を読み上げ回答してもらった。回答の謝礼はQUOカード、図書カード、Amazonギフト券（共に500円分）から一つを選んでもらい、後日郵送した。主な調査スケジュールは以下の通り（日付はすべて2016年）

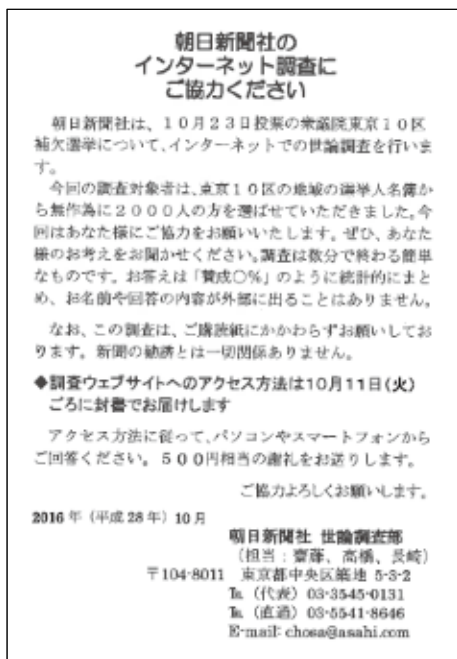
日付	曜日	作業内容
9月13日	火	有権者名簿からの抽出開始
9月16日	金	有権者名簿からの抽出終了
10月3日	月	予告はがき発送
10月10日	月・祝	調査案内の手紙発送
10月11日	火	告示日・ネット調査画面オープン
10月15日	土	未回答者への督促手紙発送
10月22日	土	ネット調査画面クローズ
10月23日	日	投票開票日

分析のため対象者を次の3つの条件（①～③）を変えてランダムに500人ずつ4つのグループに分類した。①予告はがきの文面が「文章形式」と「箇条書き」。②案内・督促の手紙の紙質が「再生紙」と「上質紙」。③事前謝礼「あり」と「なし」。「あり」のグループには案内の手紙にボールペンを同封した。

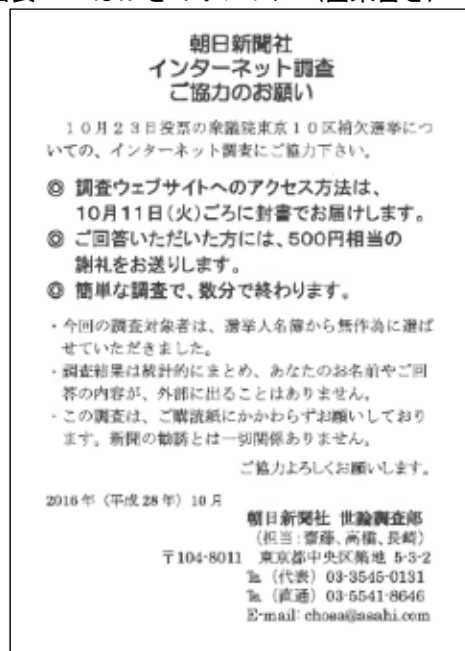
図表9. 各グループとその内容

	①予告はがき	②手紙	③事前謝礼
グループA	文章形式	再生紙	なし
グループB	箇条書き	再生紙	なし
グループC	箇条書き	上質紙	なし
グループD	箇条書き	上質紙	あり

図表10. はがきのサンプル（文章形式）



図表11. はがきのサンプル（箇条書き）



質問の内容は次の通り。

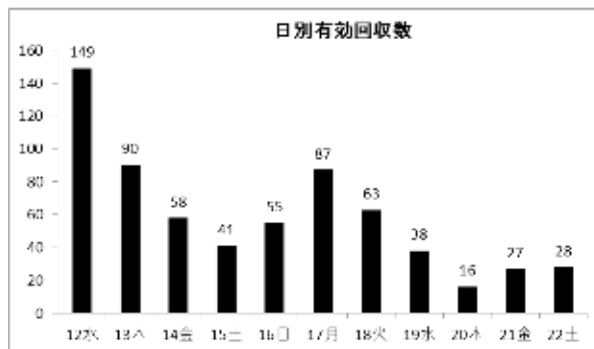
- ①関心度 ②投票する人を決めたか ③誰に投票したいか ④投票に行くか ⑤2016年7月の都知事選で誰に投票したか ⑥政党支持 ⑦内閣支持 ⑧普段どの程度投票に行っているか ⑨性別 ⑩年代 ⑪職業 ⑫学歴 ⑬回答デバイス ⑭URL か QR コードか ⑮謝礼の品。

6. 郵送ネット調査の結果分析

6-1. 回収数と回収率の分析

回収数は652件。回収率は33%。回収数を日別に見ると、調査開始の次の日の10月12日（水）が最も多く、149件で、その後、徐々に減少。督促の手紙の後は若干増加した。

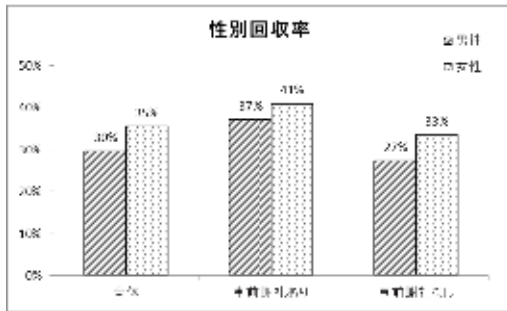
図表12. 日別有効回収数



また、グループ別の回収率は、Aが31%、Bが29%、Cが31%、Dが39%と事前謝礼の効果が、

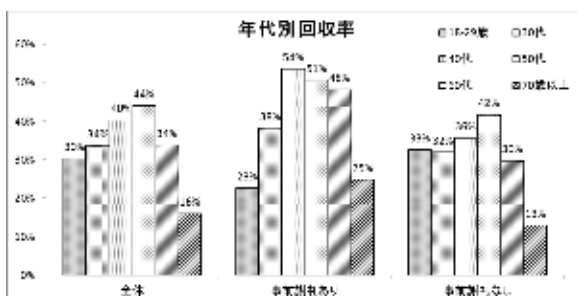
予告はがきの文章や手紙の紙質による効果より有意に大きかった。性別の回収率は、男性が30%で女性が35%と女性の方が若干回収率が高かった。事前謝礼ありの方が、男性、女性とも事前謝礼なしの場合より高かった。

図表 13. 性別回収率



回収率を年代別に見ると18~29歳30%、30代34%、40代40%、50代44%、60代34%、70歳以上16%と高齢層の回収率が低かった。事前謝礼ありと事前謝礼なしで比較すると、18~29歳は事前謝礼なしの方が回収率が高かったが、30代以上は事前謝礼ありの方が高かった。

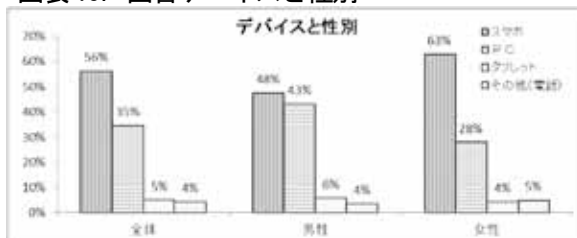
図表 14. 年代別回収率



6-2. 回答デバイスの分析

回答デバイスの割合は、スマホ56%、パソコン35%、タブレット5%、その他（電話）4%の順だった。性別では、女性の方が男性よりスマホで回答する人が多かった。

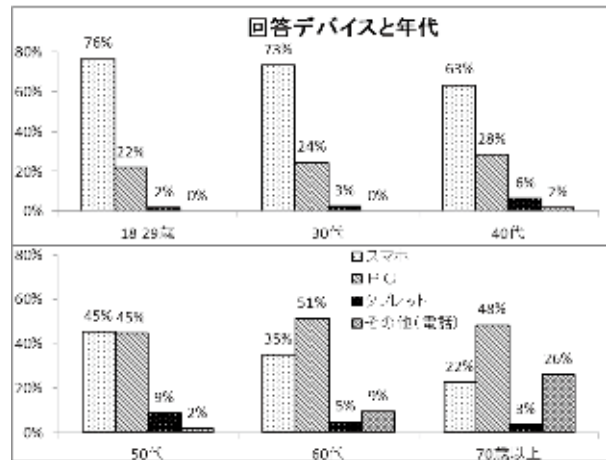
図表 15. 回答デバイスと性別



回答デバイスを年代別に見ると、若年層はスマホ、高齢層はパソコンで回答する傾向があり、70

歳以上は、4人に1人が電話だった。

図表 16. 回答デバイスと年代



6-3. URL と QR コードの分析

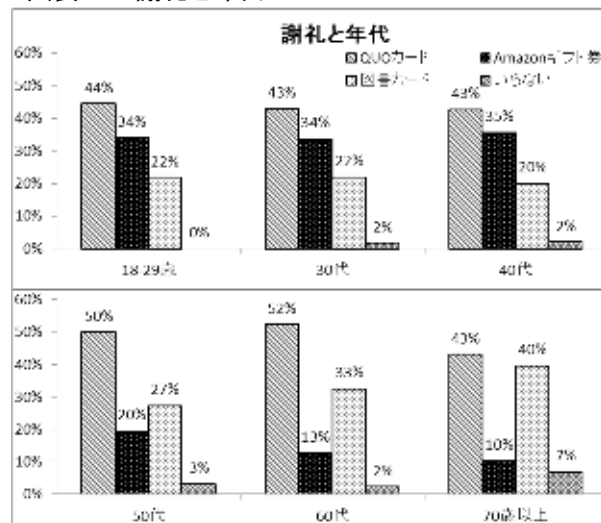
URL を使って質問のページにアクセスした人は39%で、QR コードの人は57%だった。パソコンで回答した人の99%はURL でアクセスし、スマホの95%がQR コードでアクセスしており、パソコンとURL、スマホとQR コードには強い相関があった。

6-4. 謝礼の品の分析

謝礼の品はQUO カード46%、Amazon ギフト券27%、図書カード25%、「いらない」が2%だった。

性別に見ると、男性と女性での謝礼の品の傾向に大きな違いはなかった。年代別では、全年代でQUO カードが最も多く、18歳~40代までAmazon ギフト券が2番目だが、50代以上は図書カードが2番目だった。

図表 17. 謝礼と年代



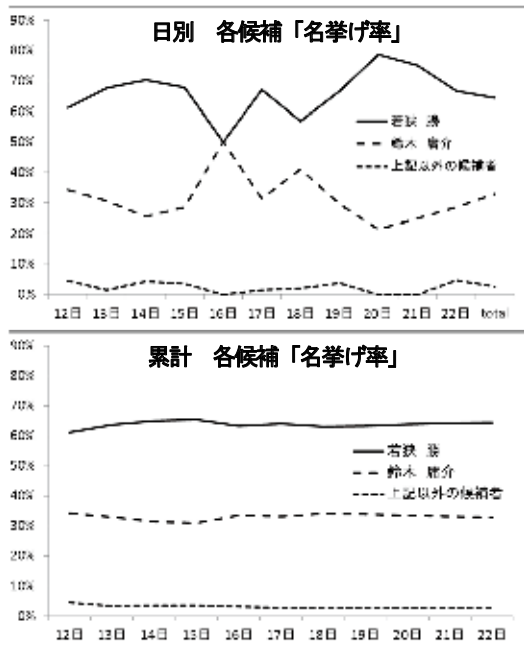
6-5. 投票したい候補者の分析

「投票したい候補者」の質問の選択肢は以下の通り。①若狭 勝（自民党）②鈴木 庸介（民進党）③上記以外の候補者（自由記入欄：任意）④わからない（①と②はランダムイズ）。

結果は、①若狭 49%、②鈴木 25%、③上記以外の候補者 2%、④わからない 25%だった。④を除いた「名挙げ率」は①若狭 64%で、②鈴木 33%。実得票率は若狭が 60%で、鈴木が 37%と、調査では若干、若狭に強めに出た。

日別の「名挙げ率」は、回収数の関係もあり、ばらつきがあるが、日別を累計すると、調査 2 日目までの結果と最終の結果はほぼ同じ数値になっており、調査 2 日目以降も最終日まで大きな変動はなかった。

図表 18. 別・累計各候補の「名挙げ率」



6-6. RDD 調査との比較

朝日新聞社は 2016 年 10 月 15、16 の両日に衆議院東京 10 区補選の電話情勢調査を行っている。コンピューターで無作為に作成した固定電話の番号に調査員が電話をかける RDD 方式で、衆議院東京 10 区の有権者を対象に調査した。有権者がいる世帯と判明した番号は 830 件、有効回答 506 人、回答率 61%だった。RDD 調査では、集計の際、電話回線数、世帯内の有権者人数、地域（区別）、性、年代の構成比に対するウェイト付けを行っている。

「関心度」「投票に行くか」「政党支持」質問に

ついて、RDD 調査と結果を比較してみた。

「関心度」「投票に行くか」には大きな違いはなかった。「政党支持」は、郵送ネットでは自民、民進が若干低く、無党派層が若干高かったが、RDD と比べて極端な違いはなかった。

図表 19. 郵送ネットと RDD の 回答比較

◆衆議院東京10区の補欠選挙関心度		
	郵送ネット	RDD
大いに関心がある	28%	26%
少しは関心がある	54%	56%
関心はない	18%	17%
その他（・答えない）	1%	1%

◆補選の投票に行くか		
	郵送ネット	RDD
必ず行く	63%	65%
できれば行きたい	32%	27%
行かない	5%	7%

※「その他（・答えない）」「わからない」を除く。「期日前投票を済ませた」は「必ず行く」に含めた。

◆政党支持		
	郵送ネット	RDD
自民党	29%	33%
民進党	7%	12%
共産党	6%	6%
その他の政党	7%	5%
無党派層	52%	44%

※5%以下の政党は「その他の政党」にまとめた。

6-7. 郵送ネット調査の質問と回答

（数字は%。小数点以下は四捨五入）

(1) あなたは、いま行われている衆議院東京 10 区の補欠選挙に、どの程度関心がありますか。

①大いに関心がある 28 ②少しは関心がある 54 ③関心はない 18 ④その他（自由記入欄：任意）1

(2) あなたは、東京 10 区の補欠選挙で投票する人を、もうお決めになりましたか。

①決めた 49 ②決めていない 50 ③その他（自由記入欄：任意）1

(3) 東京 10 区の補欠選挙で、仮にいま投票するとしたら、あなたに、だれに入れたいと思いますか。（①～②ランダムイズ）

①若狭 勝（自民党）49 ②鈴木 庸介（民進党）25 ③上記以外の候補者（自由記入欄：任意）2 ④わからない 25

(4) あなたは、今度の東京 10 区の補欠選挙には、投票に行きますか。

①必ず行く 56 ②できれば行きたい 29 ③行かない 5 ④期日前投票を済ませた 2 ⑤わからない 8

(5) 7月に行われた東京都知事選挙で、あなたは、だれに投票しましたか。(①～③ランダマイズ)

①小池 百合子 54 ②増田 寛也 13 ③鳥越 俊太郎 12 ④上記以外の候補者(自由記入欄:任意) 5 ⑤投票に行かなかった 14 ⑥その他(自由記入欄:任意) 2

(6) あなたは、いま、どの政党を支持していますか。

①自民党 29 ②民進党 7 ③公明党 2 ④共産党 6 ⑤日本維新の会 2 ⑥社民党 1 ⑦生活の党 1 ⑧日本のこころ 0 ⑨上記以外の政党(自由記入:任意) 0 ⑩支持する政党はない 51 ⑪その他(自由記入欄:任意) 1

(7) あなたは、安倍内閣を支持しますか。支持しませんか。

①支持する 46 ②支持しない 41 ③その他(自由記入欄:任意) 13

(8) あなたは、ふだん、国の選挙や地方の選挙で、どの程度投票に行っていますか。

①必ず行っている 52 ②行くことの方が多い 32 ③行かないことが多い 12 ④ほとんど行かない 4 ⑤その他(自由記入欄:任意) 1

(9) あなたは、男性ですか。女性ですか。

①男性 43 ②女性 57

(10) あなたの年代を教えてください。

①18～19歳 2 ②20代 14 ③30代 18 ④40代 24 ⑤50代 20 ⑥60代 13 ⑦70代 7 ⑧80歳以上 2

(11) あなたのご職業は、以下のどれにあたりますか。

①事務職 17 ②専門・技術 17 ③管理職 9 ④製造 1 ⑤販売 4 ⑥サービス 8 ⑦自営 7 ⑧農林漁業 0 ⑨主婦・主夫 17 ⑩学生 5 ⑪無職 4 ⑫年金生活 6 ⑬その他(自由記入欄:任意) 5

(12) あなたの最終学歴を教えてください。(在学中の方は、いま通っている学校を選んでください)

①小・中学校(または旧制高等小学校) 1 ②高校(または旧制中学、高等女学校など) 19 ③専門学校 12 ④短大・高専 11 ⑤大学 49 ⑥大学院 7 ⑦その他(自由記入欄:任意) 0

(13) あなたは、この調査を以下のどちらで回答していますか。(電話回答は「その他」に分類)

①パソコン 35 ②スマートフォン 56 ③タブレット端末 5 ④その他(自由記入欄:任意) 4

(14) あなたは、この調査画面にアクセスするとき、調査用ホームページのアドレスを使いました

か。QRコードを使いましたか。(電話回答は「その他」に分類)

①調査用ホームページのアドレス 39 ②QRコード 57 ③その他(自由記入欄:任意) 4

(15) 調査に回答していただきまして、ありがとうございました。以下の①～③の謝礼の品の中から一つを11月以降にお送りしますので選んで下さい。(①～③ランダマイズ)

①QUOカード(500円分) 46 ②図書カード(500円分) 25 ③アマゾンギフト券(500円分) 27 ④いない 2

7. 郵送ネット調査まとめと問題点

①回収数 652 … RDD と同等の回収数を得ることができた。

②結果の比較 … 調査結果と実得票率との比較で大きな違いはなかった。

③高齢層での低回収率

④速報性への対応 … 12日間で33%の回収率。

⑤コストへの対応 … 抽出費用、郵送費用、事前・事後謝礼と相応のコストがかかる。

⑥カバレッジへの対応 … ネットの環境がない人の対応。

8. おわりに

GSも郵送ネット調査も「投票したい候補者」と実際の得票率で大きな違いはなかった。しかし、GSでは地域設定などの問題、郵送ネット調査では速報性やコストなどの問題がある。また、ともに高齢者の低回収率の問題もある。メディアの選挙調査としてすぐに実用化するのには難しいだろう。

(朝日新聞社・世論調査部)

参考リンク

Google Surveys 公式サイト
<https://g.co/surveys>

参考文献

Nagasaki, Saito, Synodinos (2017) Comparison Between the Google Survey and Landline RDD in Two Japanese Regional Elections (<https://www.aapor.org/getattachment/Conference-Events/Annual-Meeting/2017-Abstract-Book.pdf.aspx>)