固定電話と携帯電話を対象とした電話調査の導入と推定値の評価

福田 昌史

Introduction and Evaluation of Dual-frame RDD Surveys for Landline and Cell-phone Users

Masafumi FUKUDA

The Yomiuri Shimbun introduced dual-frame RDD surveys of landline and cell-phone users for monthly public-opinion surveys. To improve the survey methods, this study investigated experimental dual-frame RDD surveys. The results show that cell-phone surveys are a feasible method for use in Japan, improve the coverage of telephone surveys, effectively reach young respondents, and produce data of adequate quality. The unequal weighting effects (UWEs) of the dual-frame surveys were also examined to measure increases in variance. Increases in variance due to weighting were small, and the UWE was better than in single-frame landline surveys.

Key words: RDD, dual frame survey, unequal weighting effect

キーワード: 固定電話調査, 携帯電話調査, デュアルフレーム調査, 不等加重効果

1. 序

読売新聞社は2016年4月, 電話方式で行う全国世論調査の対象を, 固定電話のみから固定電話と携帯電話の両方に調査する方法に切り替えた。

報道機関が行う電話調査は, 対象の番号を無作為に生成するRDD (Random Digit Dialing) と呼ばれる方式で行われており, これまでは固定電話のみを対象としていた。

RDDは面接方式や郵送方式などと比べて短期間で準備・実施でき, 電話帳に掲載されていない番号にも調査できるという利点があった。ただ近年は, 自宅に固定電話がなく携帯電話しか持たない人, いわゆる「携帯限定層」が若年層を中心に増えており, 固定電話のみを対象とした調査のカバレッジ低下が懸念されていた。

また, 固定電話調査は対象者が在宅していることが必要となるため, 調査期間中に複数回電話をかけてもほとんど自宅にいない人への調査が難しい。

海外では, 携帯限定層の割合が高い国を中心に携帯電話への調査が行われ, 固定電話調査と組み合わせる

手法が広く使われている。これをふまえ, 読売新聞社は2回の共同調査と2回の独自調査の計4回の実験調査を行ったのち, 固定電話と携帯電話の併用方式を導入した。本稿は, 携帯電話調査の導入の背景を概観し, 実験調査のデータを使って調査の精度を検証する。具体的には, 併用方式で使う集計ウェイトが推定値の分散に与える影響を評価する。

2. 固定電話調査を巡る状況

2.1. 固定電話離れ

携帯電話の普及に伴って固定電話を持たない人が徐々に増えている状況は, 報道各社が面接方式や郵送方式の世論調査で捉えてきた。読売新聞社が実施した郵送調査の結果をみると, 2011年に9%だった携帯限定層の割合は, 16年には15%前後まで増えた(図1)。若いほど携帯限定層の割合が高く, 16年1~2月に行った2回の調査を平均すると, 20歳代は31%, 30歳代39%, 40歳代16%などで, 30歳代~50歳代はいずれも5年間で2倍以上に膨らんだ。(図2)。

携帯限定層の増加は日本に限った現象ではない。

米国では, 公的機関であるCDC(疾病管理予防センター)が定期的な大規模な調査を行っており, 2003年の上半期時点では成人の3%が携帯限定層(Blumberg &

図 1. 携帯限定層の割合

図 2. 年代別携帯限定層の割合の変化

Luke, 2007) だったが, 15 年下半期には 48% まで膨らんだ (Blumberg & Luke, 2016). 25~29 歳の携帯限定層の割合は 73% と非常に高く, 65 歳以上は 21% で, 若い世代での割合が高いという日本と似た傾向がみられる。

欧州で公表されている 28 か国の統計 (European Commission, 2016) をみると, 固定電話を持たず携帯電話だけを持つ世帯の割合は, フィンランドの 87% が最も高く, チェコ 84%, リトアニア 71% など, 5 割を超える国は 12 か国ある。主要国をみると, イタリア 43%, イギリス 22%, フランス 19%, ドイツ 15% などと差があるが, 多くの国で増加傾向がみられる。

2.2. 固定電話調査への影響

国内で携帯限定層が増えてきている中で, 固定電話調査の精度に及ぼす影響が検証されてきた。福田 (2010) は, 2010 年 9 月に行った面接調査の結果から, 携帯限

定層の割合は 5% で, 固定電話を持つ人と携帯限定層の間に極端な意識の差がないことを示した。川崎・宮本 (2011) は, 2011 年 1~2 月に行った郵送調査をまとめ, 携帯限定層の割合は 9% で, 携帯限定層を含む全体の調査結果と携帯限定層を除いて集計した結果にほとんど違いがないことを示し, 固定電話のみを対象とした調査への影響は小さいことを報告した。また, 齋藤 (2013) は, 2012 年の郵送調査で「固定電話がない」と答えたのは全体の 12% で, 固定電話がある人となない人では内閣支持にほとんど差がないこと, ない人はある人より無党派層の割合がやや多めであることを示した。

これらを含む, 携帯限定層に関する先行研究は, 概して次のような共通の結論を得ている。

- (1) 携帯限定層は徐々に増えており, 最近是有権者全体の 15% 前後を占める。ただ, 全体から見れば大きな集団ではない
- (2) 若年層で携帯限定層の割合が高い
- (3) 政治意識に関して, 携帯限定層と固定電話保有層の間に大きな差はない
- (4) 携帯限定層と固定電話保有層の間に意識差がある場合, その差は, 各集団の年齢構成の違いによって生じている
- (5) 上記 (1) と (3) の理由から, 固定電話調査への影響は小さい

携帯限定層の影響だけでなく, 藤木 (2014) は, 固定電話調査の結果に影響しうる集団をさらに拡張して検証を行った。2014 年 1~2 月に実施した郵送調査で, 電話の保有状況に加えて電話調査が行われている週末 3 日間 (金, 土, 日) の在宅状況を尋ね, 「接触不可能層」を, 「固定電話を持っていない人, もしくは常に外出している人」として定義した。調査の結果, 接触不可能層は 15% を占めていたが, 接触不可能層の意見はそれ以外の人の意見と傾向が変わらないため, 固定電話調査への影響はほとんどなかった。ただ, 週末の金曜日午後から日曜日夜にかけての 13 の時間帯について, 半分 (7 つ) 以上「外出していた」と答えた人の割合は, 20 歳代, 30 歳代で 7 割前後を占めて他の年代より高く, 短期間の調査では自宅にいる若者への接触が難しいことが分かった。

2.3. 共同実験調査

固定電話の不完全なカバレッジを補完するため, 海外では携帯電話への調査が行われていることから, 読売新聞社を含む報道関係 6 社 (朝日新聞, NHK, 共同

通信、日経リサーチ、毎日新聞、読売新聞）と日本世論調査協会は、2014年9月と10月の2回、携帯電話への実験調査を行い、国内での携帯電話調査の実施可能性を探った（携帯RDD研究会, 2015）。

実験調査では、総務省が「電話番号指定状況」として公開している携帯電話番号の上6桁に、無作為な5桁の数字を組み合わせて、調査対象の携帯電話番号11桁を生成し調査対象を選んだ。

この調査では、回答者の4人に1人が固定電話を持っていないと答え、携帯限定層の回答を確実に集められることが分かった。また、20歳代、30歳代からの回答を固定電話調査より多く得られることが分かった。これらのメリットがある一方、男性の割合が女性より高く（2回目の調査では男63：女37）、男女比が偏るというデメリットも明らかになった。同様の傾向は、その後に行われた朝日新聞による実験調査（Saito & Lavrakas, 2015）や、NHKによる実験調査（小野寺・塚本, 2015）でもみられた。

政治意識に関する質問では、固定電話調査などと同様の結果が得られた。また、回答の質に関しては、

- 外出中だと周囲への気遣いや落ち着いて答えられないことによる無回答が増える懸念があったが、外出中の対象者と自宅で答えた対象者の回答の間に統計的に有意な差はなかった
- 通話音質の悪さによる無回答の増加も懸念されたため、音質による回答差を比較したが、統計的に有意な差はなかった

など、対象者の回答環境による回答の差はみられなかった。

2.4. 電話に出ない人

携帯電話への共同実験調査で女性の割合が低かった要因やその影響を探るため、読売新聞社は15年1～2月に行った郵送調査で検証した。

知らない電話番号や、番号非通知の相手から携帯電話に着信があった場合、その電話に出るかどうかを尋ね、「いつも出る」、「ときどき出る」、「あまり出ない」、「全く出ない」、「携帯電話はない」の5つの選択肢の中から、回答を1つ選んでもらった。その結果、「全く出ない」との回答は男性より女性に多く、実験調査の回収状況と矛盾しない結果が得られた。ただ、出る人と出ない人の回答傾向がほぼ同じであることから、電話に出る人の回答の偏りにはほとんどつながっていないことが分かった（福田, 2015）。

3. デュアルフレーム調査とウェート

3.1. 携帯電話調査への期待とその課題

携帯限定層の増加や、外出が多い若者の調査困難は今後も続くことが見込まれる。固定電話調査の結果に及ぼす影響が今のところ小さいとはいえ、増加によって固定電話で調査できる集団が徐々に小さくなり、高齢者が占める割合は高くなっていく。その結果、固定電話のみを対象とした調査では世代間で意識差のある項目での偏りが大きくなるおそれがある。

一方、2016年6月に選挙権年齢が18歳以上に引き下げられ、若者の意識を把握することの重要性は増している。読売新聞社は、若年層からの回答を得るのに有利な携帯電話調査を固定電話調査と組み合わせることで、電話調査のカバレッジを上げるとともに、サンプルに占める若者の割合を向上させることができると考えた。

ただ、携帯電話への調査は、共同実験によって実施可能であることが分かったものの、実際の導入には固定と携帯2種類の調査からどのように1つの推定値を算出するか、そして、その方法によって安定した推定値を得られるのかが課題だった。以降では、これらの課題を検討する。

3.2. オーバーラップ方式とスクリーニング方式

固定電話と携帯電話のように、2つの抽出枠から対象者を選び出す調査は、デュアルフレーム調査（Dual-frame survey）と呼ばれる。電話方式のデュアルフレーム調査には、集団の重なり部分、つまり固定電話と携帯電話の両方を持っている集団（図3のab）について、(1) 固定電話調査だけから回答者を選ぶ方式（スクリーニング方式）と、(2) 固定電話調査と携帯電話調査の両方から回答者を選ぶ方式（オーバーラップ方式）の2つが存在する（AAPOR Cell Phone Task Force, 2010）。この2つの方式のうち、読売新聞社はオーバーラップ方式を採用した。スクリーニング方式は、固定電話を持っている相手につながった場合、その相手には調査せず通話を終了させるので、無駄な発信が増え調査員の負担につながることや、コストが割高になる恐れがあるため不採用とした。以下では、オーバーラップ方式を前提に議論を進める。

3.3. デュアルフレーム調査のウェート

米国で行われている固定・携帯のデュアルフレーム調

図3. フレームの重なりと調査方式

査は、最終的な推定値を得るためのウェートを計算するために、複数のステップを踏んでいる (e.g. Kennedy, 2012). 以下に処理の概要を示す。

Step1 集団ごとの推定のための weighting

デュアルフレーム調査において、図3のように抽出枠AとBがあり、AとBに重なりがある場合、母集団は、Aだけに含まれる集団aと、AとBの両方に含まれる集団ab、Bだけに含まれる集団bの3つの集団に分けられる。

調査変数 y の母集団総計 Y を推定する場合、集団a, ab, bの3集団ごとの y の総計の推定値を使う。オーバーラップ方式の場合、集団abについては、フレームAからのサンプルによる推定値と、フレームBからの推定値の2種類が得られる。

電話調査の場合、この段階の推定で対象が不等確率抽出であることを考慮する必要がある。固定電話調査では、つながった世帯に何人の有権者がいるか、電話番号がいくつあるかで、対象者が選ばれる確率が異なる。また、携帯電話調査でも携帯電話を何台持っているかで選ばれる確率が異なる。これを調整するため、選ばれる確率の逆数に比例したウェートを付加する。この際、過大なウェートの発生を抑えるため、世帯内有権者数や電話番号数、携帯電話数に上限値指定 (capping) が施される (Massay & Botman, 1988)。

さらに、a, ab, b各集団内での男女比や年代構成などの属性の分布 (母集団における真の割合) が既知であれば、それに合わせて無回答の補正を施す。

Step2 統合パラメーターを使ったデータの統合

調査変数 y の母集団総計 Y は、

$$\hat{Y} = \hat{Y}_a^A + \theta \hat{Y}_{ab}^A + (1 - \theta) \hat{Y}_{ab}^B + \hat{Y}_b^B \quad (1)$$

で推定できる (Hartley, 1962). ここで右辺の $\hat{Y}$ に付いている上付き添え字A, Bは、サンプルの抽出枠を表し、下付き添え字は Y の総計値を推定している集団a, ab, bを表している。 θ は、統合パラメーター (composite parameter, mixing parameter) と呼び、

$0 \leq \theta \leq 1$ の値をとる。実用性の面から、 $\theta = 0.5$ がよく使われ、推奨されているとの報告がある (e.g. Lohr, 2011; Brick, 2015). この値は、0.5以外にも Effective sample sizeに基づいて算出する方法 (Frankel et al, 2007) などがあり、バリエーションが存在する。

Step3 無回答を補正するための調整

Step1, 2でのウェートを課したサンプルの属性の構成比が、母集団での構成比に合うように、ウェート値を調整する。方法は、事後層化やレイキングなど複数ある (e.g. Kalton & Flores-Cervantes, 2003).

ただ、この調整によって過大な値や過小な値をとるウェートが発生すると、不等加重効果が大きくなり、調査誤差の増加につながる。これを避ける方法として、極端に小さいウェートや極端に大きいウェートを、境界として設定した値に置きかえる方法 (トリミング) がある。上下数%ずつをトリミングしたり、あらかじめ決めた値 (0.2未満と5.0超など) をトリミングしたりする手法が使われる。

4. 固定・携帯併用方式調査の設計と精度の検証

4.1. 不等加重効果

固定・携帯併用方式の電話調査を設計するにあたっては、調査の誤差 (推定値の分散) を小さく抑えることが重要である。ここでは、デザイン効果 (Kish, 1965) をその基準として考える。デザイン効果は、推定値の分散と単純無作為サンプルを仮定したときの推定値の分散の比をとったものである。

さらに、ここでは、Levine & Harter (2015) と同様に、デザイン効果と不等加重効果 (UWE= Unequal weighting effect) (Kish, 1965, 1992)

$$UWE = \frac{n \sum_{i=1}^n w_i^2}{(\sum_{i=1}^n w_i)^2}, \quad (2)$$

がほぼ同等であると仮定し、これを小さく抑えるよう設計を考える。ここで n はサンプルサイズ、 w_i は、回答者 i ($i = 1, \dots, n$) に適用するウェートを表す。上式はウェートと調査変数が無相関であることを仮定しているが、これについては4.7節で検証する。

4.2. 電話利用状況の把握

固定電話と携帯電話のデュアルフレーム調査を設計し適切なウェートを計算するには、母集団における携帯限定層の割合など、各集団の大きさを把握する必要

表 1. 郵送調査による電話利用形態別の割合 (%)

調査実施	両方	固定のみ	携帯のみ	DK
2015 年 8~9 月	73	10	15	2
16 年 1~2 月	72	9	17	2
16 年 1~2 月	76	8	14	2
平均 (除 DK)	75.2	9.2	15.6	—

表 2. 携帯電話実験調査と固定電話調査の実施概要

実施	固定電話調査		携帯電話調査	
	対象世帯	回答数	応答者	回答数
2015 年 11 月	1734	1060	1297	607
16 年 2 月	1743	1060	1294	555

がある。国内には米国のような公的統計が存在しないため、読売新聞社は定期的に実施している郵送調査でこれを調べている。調査では、(1)「あなたのご家庭に、固定電話はありますか、ありませんか」との質問と、(2)「あなたは、携帯電話やスマートフォンを使っていますか、使っていませんか」との質問の両方を尋ね、携帯限定層など各集団が母集団に占める割合を推定している。

さらに、電話調査において対象者がどの集団に属するかを分類する必要があるため、携帯電話調査で上記 (1) の質問、固定電話調査で (2) の質問を尋ね、3 種類の調査で使う質問を統一させた。この質問を使った 2015 年以降の郵送調査で、上記 (1) と (2) の質問への回答を組み合わせて推定した各集団の割合を表 1 に示した。

4.3. 独自実験調査

読売新聞社独自の携帯電話への実験調査は、2015 年 11 月と 16 年 2 月の 2 回、固定電話への正規の調査と同時に実施した (表 2)。実施に当たっては、共同実験調査で得た知見を踏まえ、調査員が話す内容や複数回発信のスケジュール、発信結果の分類法など、独自の方法を構築した。携帯電話調査と固定電話調査の回答傾向が近い点や、携帯では男性の回答が多めである点などは、共同実験調査と同様の結果が得られた。

実験調査の結果、4.2 で挙げた電話利用に関する質問については、表 3 の結果が得られた。

4.4. フレームごとのサンプルサイズ

デュアルフレーム調査において、それぞれのフレー

表 3. 固定電話調査と携帯電話実験調査での電話利用に関する質問への回答 (%)

(固定電話調査) 携帯電話を			
実施	使っている	使っていない	DK
2015 年 11 月	82	17	1
16 年 2 月	83	15	2
平均 (除 DK)	83.8	16.2	—

(携帯電話調査) 固定電話が			
実施	ある	ない	DK
2015 年 11 月	76	24	0
16 年 2 月	73	25	2
平均 (除 DK)	75.3	24.7	—

ムにどの程度のサンプルサイズを割り当てるのがよいかは、最も重要な問題の 1 つといえる。

Lohr & Brick (2014) や Levine & Harter (2015) は、固定電話と携帯電話それぞれの調査費用を考慮したサンプルサイズの割り当てについて論じたが、ここでは適切な割り当ての目安を知るため、データの統合段階で生じるウェートの不等加重効果を試算した。具体的には、仮に固定電話調査と携帯電話調査の両方を行い、それぞれ得られたサンプルの構成比が表 3 の平均値だった場合に、2 つのサンプルを表 1 の平均値の構成比になるよう統合するためのウェート (3.3 節の Step1~2 にまでのウェートに相当) を算出し、不等加重効果を計算した。

図 4 は、試算した不等加重効果 (垂直軸) を、携帯電話サンプルの割合 (水平軸) を変化させてプロットしたものである。統合パラメーター θ は、0.5 を使用して計算した。

この図からは、携帯電話調査サンプルの割合が 0.4~0.7 の間であれば、誤差を増大させることなく推定値を得られることが分かる。この結果から、携帯電話調査のサンプルの割合は、0.5 を基準に設計することとした。ただし、調査の性質上、各調査で得られる有効回答数は変動するので、厳密に同数になることはない。ただ、図 4 からは、0.5 から多少変動しても、誤差の増大には影響しないことが分かる。

4.5. 固定、携帯電話調査サンプルの統合

サンプルの統合のためには、電話の使用状況は必須の情報となる。ただ、これらの質問には一定数の無回答 (0~2%) が発生するため、補完処理を行う。具体

図 4. デュアルフレーム調査での携帯電話調査サンプルの割合と不等加重効果

的には、固定・携帯電話を保有・使用している確率を性別・年代ごとに推定し、この確率に基づいて乱数を発生させ値を代入した。

統合パラメーター θ は 0.5 とした。前述の通り、デュアルフレーム調査が先行している米国でよく使われていることや、両フレームから得る回答者数の比を 1 : 1 とする設計とも相性が良いのが理由である。

3.3 節で論じた、デュアルフレーム調査のウェートを、実験調査のデータに適用し、不等加重効果を検証するため、15 年 11 月の実験調査のサンプルから 525 件を無作為に抽出し、同時に実施した固定電話調査のサンプルから 525 件を無作為に抽出し、初めに Step1～2 のウェートを施して両サンプルを統合した¹⁾。Step1 では、抽出確率の逆数に比例したウェートを付けたが、各集団における属性分布の真の割合が不明であるため無回答の補正は省略した。Step2 では、電話保有に関する 3 つの集団の割合が表 1 で示した平均値になるよう統合した。

図 5 は、統合後のウェートの分布を表したものである。全体の 98% が 0.4～2.0 の間に分布しており、極端に大きな値が発生していない。この段階での最小値は 0.22、最大値は 2.34、不等加重効果は 1.12 で、これによる調査の誤差幅の拡大率はその平方根として得られ、値は 1.06 だった。

4.6. 無回答の補正

ここでは、次の段階として Step3 のウェートを検討する。無回答の補正は、性別と年代 (6 分類)、地域 (全

図 5. Step1～2 のウェートの分布

国の都道府県を 11 地域に分類したもの) の 3 変数の分布が目標母集団での分布と同じになるよう施した。具体的には、4.5 節で計算したウェート (3.3 節の Step2 までに相当するウェート) に対して、キャリブレーション (Deville & Särndal, 1992) を適用した。

キャリブレーションは、補正する変数の分布を目標母集団での分布に合わせつつ、適用する前と後でのウェートの変化を最小にする手法で、その変化量を測定する方法として、3 種類の距離関数

(1) 線形関数

$$\sum_s w_i G(w_i, w_i^c) = \sum_s w_i \frac{(w_i^c - w_i)^2}{2}$$

(2) 乗法関数

$$\sum_s w_i G(w_i, w_i^c) = \sum_s w_i \left\{ \frac{w_i^c}{w_i} \log \frac{w_i^c}{w_i} - \frac{w_i^c}{w_i} + 1 \right\}$$

(3) ロジット関数

$$\begin{aligned} \sum_s w_i G(w_i, w_i^c) &= \sum_s w_i \left[\frac{1}{A} \left\{ \left(\frac{w_i^c}{w_i} - L \right) \log \frac{w_i^c/w_i - L}{1 - L} \right. \right. \\ &\quad \left. \left. + \left(U - \frac{w_i^c}{w_i} \right) \log \frac{U - w_i^c/w_i}{U - 1} \right\} \right] \end{aligned}$$

が使われている (詳しくは、土屋, 2009)。ここで、 s はサンプル、 w_i は、回答者 i に対するキャリブレーション前のウェート、 w_i^c は、キャリブレーション後のウェートである。線形関数を使うことで、結果的に一般化回帰推定量、乗法関数を使うことでレイキング推定量を求めるウェートが得られる。また、ロジット関数を使用すると、パラメーター U と L を決めることでキャリブレーション前後のウェートの比の範囲を調整

¹⁾ 読売新聞社が行う毎月の電話調査は、平均的に 1050 人程度の回答を得ている。

表 4. 線形関数と乗法関数を用いたキャリブレーションの結果

	線形関数	乗法関数
最小	0.20	0.20
最大	3.31	3.31
不等加重効果	1.21	1.21

表 5. ロジット関数を用いたキャリブレーションの結果

(L, U)	(0.5, 1.7)	(0.6, 1.7)	(0.7, 1.7)
最小	0.21	0.21	0.20
最大	3.31	3.38	3.51
不等加重効果	1.22	1.22	1.22

(L, U)	(0.5, 1.6)	(0.6, 1.6)	(0.7, 1.6)
最小	0.21	0.21	0.20
最大	3.24	3.30	3.39
不等加重効果	1.22	1.22	1.22

(L, U)	(0.5, 1.5)	(0.6, 1.5)	(0.7, 1.5)
最小	0.22	0.21	0.20
最大	3.26	3.28	3.35
不等加重効果	1.22	1.22	1.23

(L, U)	(0.5, 1.4)	(0.6, 1.4)	(0.7, 1.4)
最小	0.21	0.21	0.20
最大	3.24	3.25	3.28
不等加重効果	1.22	1.22	1.24

することができ、 L は、 w_i^c/w_i の最小値を、 U はその最大値を指定する。 A は $(U - L) / \{(U - 1)(1 - L)\}$ である。

最適な距離関数を検討するため、4.5 節で統合したデータに対し、まず線形関数と乗法関数を用いてキャリブレーションを適用した。表 4 は、キャリブレーション適用後のウェートの最小値と最大値、不等加重効果を示しており、同じ結果になった。

不等加重効果は 1.21、その平方根は 1.10 で、キャリブレーション適用前の 1.06 と比べて誤差幅の増加がほとんどないことが分かる。

次にロジット関数を検討した。線形関数や乗法関数を使った時の w_i^c/w_i の最小値が 0.46、最大値が 1.85 だったため、 L を 0.46 より大きい値、 U を 1.85 より小さい値を使って、不等加重効果を小さくできるか試行した。その結果を、表 5 に示した。

この結果をみると、 L や U を動かして w_i^c/w_i の範囲を狭めても、 w_i^c の範囲を狭めたり、不等加重効果を小さくしたりする効果がないことが分かる。16 年 2 月

図 6. 最終的なウェートの分布

の実験調査のデータでも、線形関数と乗法関数の不等加重効果が 1.25 で、ロジット関数を使って L と U を変化した場合でも 1.25～1.26 となり、同様の結果が得られた。

キャリブレーション手法に関する以上の試行から、距離関数として乗法関数を採用することとした。もともと属性構成比のバランスが良く、ウェートが極端な値をとらないため、ロジスティック関数でウェートの範囲の拡大を抑えることや、Step3 におけるウェートのトリミングが不要で、得られる推定量は、広く使われているレイキング推定量と同じであることが選択の理由である。また、乗法関数はロジスティック関数において $L = 0, U = \infty$ としたものであるので、偶発的に極端なウェートが発生した場合には、トリミングに頼らずロジスティック関数に切り替えて対応しやすいという利点もある。

15 年 11 月のデータに対して乗法関数を用いてキャリブレーションを施した後のウェートの分布を、図 6 に示した。範囲は表 4 に示した通りで、0.4 から 2 の間に 91% のウェート値が収まっている。

4.7. ウェートと調査項目の相関関係

Spencer(2000) は、ウェートと調査項目に相関関係がある場合の不等加重効果を近似的に導出し、Kish (1965, 1992) が示した (2) 式と異なることを示した。以下では、調査項目とウェートの相関の有無と、相関が不等加重効果に及ぼす影響を検討する。

まず 4.6 節で求めた 15 年 11 月調査のウェートと、調査項目の相関関係を見るため、調査での各質問の選択肢 (53 項目) を一つずつダミー変数化し、ウェートとの相関係数を計算した。その結果、53 個の相関係数のうち、37 個が $-0.05 \sim 0.05$ の間に収まり、その範囲

図 7. ブートストラップ法によるデザイン効果の推定

外は 16 個だった。そのうち、8 個で無相関の仮説が棄却された (有意水準 5%)。

さらにこれらの相関の、デザイン効果への影響をみるため、53 項目の標本平均の標準誤差をブートストラップ法 (リサンプル回数 = 1000) によって推定し、デザイン効果を推定した。ブートストラップの過程では、リサンプルする度にウェートを再計算した。

図 7 は、53 個の標本平均 (選択肢が選ばれた割合: 水平軸) と、デザイン効果の推定値の平方根 (垂直軸) をプロットしたもので、ウェートとの相関係数が -0.05 より小さいものは黒丸印、 0.05 より大きいものは $\times$ 印、 $-0.05 \sim 0.05$ の範囲内のもは白丸印をつけた。

全体的には、(2) 式で算出した不等加重効果 1.21 の平方根である 1.1 (破線) の付近に点が分布している。また、ウェートと弱い負の相関を持つ項目の黒丸は、全体の下側に分布し、弱い正の相関をもつ項目の $\times$ 印は全体の上側に分布する傾向がある。とはいえ、調査の誤差幅が大きくなる平均値 0.5 付近では、点は 1.1 と同じかそれより下に分布しており、ウェートと調査項目の間の相関による誤差幅への影響は非常に小さいことが分かる。

4.8. 固定電話単独調査からの改善

固定、携帯電話の併用調査について、固定電話のみの調査から精度がどの程度改善したかを検証するため、15 年 11 月と 16 年 2 月の固定電話調査のサンプルに 3.3 節の Step1 と Step3 の処理を施したウェートを試算した (Step2 はデータ統合プロセスのため適用不要)。条件をそろえるため、併用方式と同様に Step3 で性別、年代、地域を補正対象の変数とし、乗法関数を使ったキャリブレーションを適用した。その結果、15 年 11

図 8. 年代別のウェート分布 (15 年 11 月調査)

月と 16 年 2 月とも不等加重効果は 1.82 で、固定と携帯の併用方式の 1.2 台より大きかった。

図 8 に 15 年 11 月調査の固定単独調査 (左) と固定、携帯併用調査 (右) のウェートの分布を、年代別にボックスプロットで示した。垂直軸はウェートの値を、水平軸は各年代を表している。固定単独の場合、若年層の割合が低いいためウェートの値が大きく、その範囲も大きい。一方、併用方式の場合は取得したサンプルの属性バランスが良いため、特定の年代のウェートが極端な値を取ることはなく、範囲も各年代で大きく変わらないことが見てとれる。以上から、併用方式を採用することで調査の誤差幅を改善できたことが分かる。

5. 結 び

読売新聞社は、固定電話と携帯電話の併用調査の導入にあたり、実験調査などを通じていくつもの側面から点検を行った。

まず、携帯電話への調査は、相手がさまざまな状況下にあることが想定されたが、一定の回答数が得られることが分かった。これによって携帯限定層に調査できていなかったというカバレッジの問題が改善することを確認した。また、携帯電話調査のデータの質については、外出しているかどうかや、通話音質の良し悪しなど、固定電話調査では問題とならない要因による影響がないことを確認した。

このほか、「電話に出ない人」の意見を吸い上げていないという、電話調査だけでは検証しきれない懸念については、郵送調査によって、その影響がほとんどな

いことを確認した。

さらに、調査の誤差については、サンプルの統合や無回答の調整のためのウェイトによる誤差の拡大は、調査結果の解釈に影響しない程度に十分小さいことが分かった。携帯電話調査では、回答者に占める女性の割合が男性より低かったが、性別を含めて補正しても誤差幅への影響が十分小さいことが分かった。

ウェイトは、多くの調査項目との間で無相関だったが、厳密にゼロであるとはいえない項目もあった。ただ、誤差幅には影響していなかった。

これらの検証から、固定電話調査と携帯電話調査の併用方式への切り替えは、内閣支持のような、短期間で変化しやすい意識を捉える方法としては、これまでの固定電話のみを対象とする方法と比べて、現在取り得る方法として、より良い方法に発展させることができたと考えられる。

とはいえ、調査の質を維持するには、継続的な取り組みが必要となる。携帯限定層の割合を定期的に調べ、固定電話サンプル、携帯電話サンプルの割り当てサイズが適切かどうかを検証することや、調査ごとにウェイトの分布や、不等加重効果の大きさが問題ないかどうかのチェックが不可欠といえる。

謝 辞

査読いただいた匿名の2名の先生方から、論文の改善につながる数多くのご指摘をいただきました。深謝いたします。

参 考 文 献

- AAPOR Cell Phone Task Force (2010). New considerations for survey researchers when planning and conducting RDD telephone surveys in the U.S. with respondents reached via cell phone numbers, American Association for Public Opinion Research.
- Blumberg, S. J. and Luke, J. V. (2007). Wireless substitution: Early release of estimates based on data from the National Health Interview Survey, July – December 2006. National Center for Health Statistics, Retrieved from: <http://www.cdc.gov/nchs/nhis.htm>.
- Blumberg, S. J. and Luke, J. V. (2016). Wireless substitution: Early release of estimates from the National Health Interview Survey, July–December 2015. National Center for Health Statistics, Retrieved from: <http://www.cdc.gov/nchs/nhis.htm>.
- Brick, J. M. (2015). Design and weighting for dual frame survey, In AAPOR Webinar, Retrieved from <http://www.aapor.org/Education-Resources/Online-Education/Webinars.aspx>.
- Deville, J.C. and Särndal, C.E. (1992). Calibration estimators in survey sampling, *Journal of the American Statistical Association*, **87**, 376–382.
- European Commission (2016). E-communications and telecom single market, *Special Eurobarometer*, **438**, Retrieved from https://data.europa.eu/euodp/en/data/dataset/S2062_84_2_438.ENG.
- Frankel, M. R., Battaglia, M. P., Link, M. and Mokdad, A. H. (2007). Integrating cell phone numbers into Random Digit-Dialed (RDD) landline surveys, Proceedings of the Survey Research Methods Section, 3793–3800, American Statistical Association, Alexandria, VA.
- 藤木康裕 (2014). 週末外出する若者は調査を偏らせるか—RDD 接触不可能層の分析—, 政策と調査, **7**, 88–95.
- 福田昌史 (2010). 「携帯限定層」も数字に影響なし 世論調査「固定電話対象」は正確か, 2010年11月26日毎日新聞朝刊.
- 福田昌史 (2015). 電話に“出ない人”は調査を偏らせるか, 政策と調査, **9**, 19–26.
- Hartley, H. O. (1962). Multiple frame surveys, Proceedings of the Social Statistics Section, 203–206, American Statistical Association, Alexandria, VA.
- Kalton, G. and Flores-Cervantes, I. (2003). Weighting methods, *Journal of Official Statistics*, **19-2**, 81–97.
- 川崎英輝・宮本清史 (2011). 郵送調査の特性と課題—面接調査との比較から, 読売クオーターリー, 2011年夏号, 184–191.
- 携帯 RDD 研究会 (2015). 携帯電話 RDD 実験調査結果のまとめ, 日本世論調査協会, http://www.japor.or.jp/pdf/RDD_Report.pdf.
- Kennedy, C. (2012). Weighting Approaches for dual frame RDD surveys, In AAPOR Webinar, Retrieved from <http://www.aapor.org/Education-Resources/Online-Education/Webinars.aspx>.
- Kish, L. (1965). *Survey Sampling*, New York: John Wiley & Sons.
- Kish, L. (1992). Weighting for unequal, *Journal of Official Statistics*, **8**, 183–200.
- Levine, B. and Harter, R. (2015). Optimal allocation of cell-phone and land-line respondents in dual-frame surveys, *Public Opinion Quarterly*, **79-1**, 91–104.
- Lohr, S. L. (2011). Alternative survey sample designs: Sampling with multiple overlapping frames, *Survey Methodology*, **37-2**, 197–213.
- Lohr, S. L. and Brick, J. M. (2014). Allocation for dual frame telephone surveys with nonresponse, *Journal of Survey Statistics and Methodology*, **2**,

- 388–409.
- Massay, J. T. and Botman, S. L. (1988). Weighting adjustments for random digit dialed surveys, In R. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nichols II, & J. Waksberg (Eds.), *Telephone Survey methodology*, 143–160, New York, Wiley.
- 小野寺典子・塚本恭子 (2015). 携帯電話調査の実現可能性を探る～2015 年 3 月携帯電話実験調査から～, *放送と研究*, **2015-9**, 76–82.
- 齋藤恭之 (2013). 米国の「携帯世論調査」の現状と課題—米国世論調査協会の大会に参加して, *Journalism*, **2013-1**, 46–51.
- Saito, Y and Lavrakas, P. J. (2015). Testing the dual frame RDD surveying of the Japanese general population, Paper presented at 70th AAPOR annual conference.
- Spencer, B.D. (2000). An Approximate Design Effect for Unequal Weighting When Measurements May Correlate With Selection Probabilities, *Survey Methodology*, **26-2**, 137–138.
- 土屋隆裕 (2009). 概説標本調査法, 朝倉書店.
- (2016 年 5 月 12 日受理, 2016 年 10 月 19 日採択)
(この間審査 2 回・審査期間合計 89 日)