特集論文

「消費生活に関するパネル調査」新規対象の回収状況と回答傾向

坂口 尚文
（公益財団法人 家計経済研究所 次席研究員）

「消費生活に関するパネル調査」の第21回調査で追加したコーホートEについて、第22回調査での回収率とウエイト補正の可能性を検証した。コーホートEの回収率は86%で他のコーホートと比して低い。一方で、人口統計学的な変数で条件付けた場合、既存コーホートに比べて、回収率に大きな偏りがあるわけではなかった。また、摩耗した標本について、カテゴリー変数の回答割合に大きな偏りは確認できなかった。そのような中においても、人口統計学的情報を補助変数としたウエイト補正は、さらに偏りを縮小させる効果がある。回収標本の有配偶の所得分布は、若干、低い方にずれていた。対象者の調査協力理由が「謝礼」がもらえるかどうかの情報を組み込み、適切なweighting classを作成することで、所得分布の偏りの解消が確認できた。

1. はじめに

「消費生活に関するパネル調査（JPSC）」では、2013年に実施した第21回調査において1984年～1989年生まれの対象者を新たに追加した。JPSCでは、ほぼ5年おきに、後生まれの出生コーホートを年齢の隙間なく調査に加えてきた。追跡コーホートは順次アルファベットをふって識別しており、第21回調査での新規対象群はコーホートEと呼ぶ。今回報告を取りまとめる第22回調査は、コーホートEについて2回目の追跡調査で、初回の回答とあわせて対象の回答の変化を捉えることが可能となった。

一方、パネル調査では調査を続けていく上で対象者の調査からの脱落は避けられない。経年の標本摩耗のスピードは、近年に追加した後生の出生コーホートほど、概して大きい傾向にある。追跡2回目の調査での脱落率は、コーホートEの値がこれまで追加してきた5つの追跡コーホート中で最も大きい結果となった。対象の脱落の多さは、回収標本の回答割合に大きな歪みをもたらす可能性がある[1)]。そこで、本稿では対象者の調査からの脱落が多かったコーホートEについて、第22回調査での回収結果を精査し、バイアスの有無の検証およびウエイト値の付与による補正の可能性を探る。

あわせて、本稿では対象者の「調査への協力理由」が脱落に与える影響[2)]と、ウエイト作成に当該情報を組み込むことの効果を検証する。「調査への協力理由」はコーホートEの追加に合わせて、JPSCの調査項目に取り込んだ質問である。謝礼や社会のためなど、なぜJPSCの調査依頼を引き受けたかを問うている。コーホートEでは初回調査時より当該質問を尋ねたことになるため、対象者全員の回答が揃い、どのような回答をした対象者が脱落しやすいかどうかを検証できる。

本稿の構成は、第2節でコーホートEの回収状況を他のコーホートの比較で精査し、第3節で「調

査の協力理由」について初回調査での回答割合を提示する。第4節は補正ウエイトの作成と作成手法ごとのウエイト系列の比較を、第5節では補正効果を検証する。最後に第6節で、まとめと考察を述べる。

2. 第22回調査における コーホートEの回収状況

2013年の第21回調査から調査に追加したコーホートEは、1985年10月～1989年10月生まれの女性、648名の女性が対象である。対象の抽出は層化二段抽出によるもので、全国を地域（8ブロック）と都市規模（区市郡）で層化し、第一次標本単位（PSU）として「国勢調査」の調査区を抽出する。各PSUからの対象者の抽出は、年齢、配偶状態、世帯構成を考慮した割り当てを行い、4～9名の予備対象を配して目的数に達するまでアタックを行う。このような調査の設計上、コーホートEの初回抽出標本は自動加重標本になる。標本の抽出方法は、これまでのコーホートの抽出方法をほぼ踏襲しているため、回収、回答状況についてコーホート間での比較が可能である。図表-1において、コーホートEとJPSCが保有しているすべての追跡コーホートについて、それぞれの回収状況を比較してみる。

まず、初回応諾率の結果をみてみる。前述したように、JPSCでは標本抽出の際に、予備の対象を配して目的とする対象の数が揃うまで依頼を行う。つまり、名簿から抽出した対象に調査の依頼を断られたら、同等の属性を持つ次の対象候補に依頼を行っている。初回応諾率は、調査の応諾者数（＝各コーホートの初回標本サイズ）を依頼の総数で割ったものである。標本サイズに応諾率の逆数をかけた値が、初回標本を揃えるまでに要した数となる。コーホートEの応諾率は27.3%で、5つのコーホートの中で最低の値をとっている。

応諾率の下段には、それぞれのコーホートについて、2回目の調査での回収率を提示した。回収率は、標本全体での回収率と対象者の属性別の回収率を掲載している。標本全体の回収率について見てみると、コーホートEの値は86.0%である。これは5つのコーホート中で最も低い値である。回収率が最も高いコーホートAとは8.8%ポイントほど開きがある。標本全体の回収率はコーホートAからBにかけて大きく低下し、その後上下するものの趨勢として低下傾向にあるといえよう。コーホートCも86.6%で同様に低い回収となったコーホートであるが、両コーホートは初回応諾率も相対的に低い点で共通している。

ただ、回収率の低下が、回収標本から得られる推定値の偏りの増大を必ずしも意味するわけではない。ある特定の層に偏って脱落が発生していないかを精査することが、より重要な作業である。そこで、初回の標本抽出に用いた情報である配偶状態と地域ブロックについて、群別の回収率をまとめてみた。また、対象者の時間不変の属性については、どの群で未回収が観測されるかを明確に捉えることができる。学歴は、調査対象である24歳以降の女性にとって、ほぼ不変の属性といえる。図表-1には学歴別の回収率も掲載した。

配偶状態別の回収率について、コーホートEは有配偶と無配偶のいずれも全コーホートでほぼ最低の値となっている。有配偶と無配偶の回収率を比較すると、いずれのコーホートでも無配偶の回収率が有配偶の回収率よりも低くなっている。有配偶と無配偶の回収率のポイント差もコーホートEではわずかながら大きくなっているが、両配偶状態の回収率を相対比で比較するとコーホートEの値が必ずしも大きいわけではない。無配偶の回収率は常に有配偶より低いが、無配偶の方に偏った脱落の傾向が観測されてはいないといえる。

地域ブロック別の回収率は、コーホートEの値はいずれも80%台である。コーホートEは全体の回収率が86%であるから、回収率が顕著に高い地域もなければ、低い地域もないといえる。他方、コーホートA以外では、回収率の高い地域と低い地域のポイント差がコーホートEよりも大きく、地域ブロック間での回収率の散らばりが観測できる。ただし、全コーホートを通じて、特定の地域ブロックが常に回収率が高いわけでもなく、また低いわけでもない。回収率の差異を地域ブロック

図表-1 コーホート別の回収率

	コーホートA	コーホートB	コーホートC	コーホートD	コーホートE
出生年	1958 ～ 1969 年	1969 ～ 1973 年	1973 年～ 1979 年	1979 年～ 1984 年	1984 年～ 1989 年
サンプルサイズ	1,500	500	836	636	648
初回応諾率	41.4%	34.3%	28.4%	30.5%	27.3%
回収率（2 回目調査）					
全体	94.8%	88.4%	86.6%	90.1%	86.0%
	0.6	1.4	1.2	1.2	1.4
配偶状態別					
有配偶	95.7%	90.0%	88.3%	93.1%	88.6%
	0.6	2.1	1.7	1.7	2.2
無配偶	93.0%	87.3%	85.4%	88.5%	84.8%
	1.1	1.9	1.6	1.6	1.7
地域ブロック別					
北海道	95.8%	63.6%	87.5%	100.0%	85.7%
	2.4	10.3	5.8	0.0	6.6
東北	89.8%	90.9%	82.7%	91.3%	88.1%
	2.9	5.0	5.2	4.2	5.0
関東	94.4%	92.4%	83.6%	89.4%	87.9%
	1.0	2.0	2.1	2.1	2.1
中部	97.8%	88.9%	89.4%	88.6%	88.5%
	0.9	3.3	2.5	3.0	3.0
近畿	94.6%	81.2%	87.9%	91.3%	80.8%
	1.5	4.2	2.7	2.8	3.9
中国	94.0%	91.3%	88.9%	97.1%	88.9%
	2.6	5.9	4.7	2.8	5.2
四国	93.8%	92.9%	91.7%	88.9%	83.3%
	3.5	6.9	5.6	7.4	8.8
九州	94.9%	91.8%	89.5%	85.5%	81.6%
	1.8	3.9	3.3	4.0	4.4
学歴別					
中学・高校	94.7%	87.4%	85.8%	90.2%	85.2%
	0.8	2.2	2.0	2.0	2.6
専門・専修	96.1%	89.7%	84.5%	88.5%	86.5%
	1.1	3.4	2.9	2.9	3.3
短大	94.4%	90.1%	88.7%	92.9%	85.9%
	1.3	2.8	2.3	2.8	3.6
大学・大学院	93.9%	87.5%	87.5%	89.8%	86.3%
	1.8	3.5	2.4	2.1	2.1
R- 指標	0.94	0.84	0.92	0.87	0.92

注: 下段は標準誤差

の特性に帰着することは、JPSCの標本からは難しい。地域ブロック間で回収率がばらつく要因は、標本サイズが30を切る地域ブロックも多くあり、推定値の精度が悪いこと。さらに標本サイズが小さいために、一つの地域ブロックから抽出する調査区（PSU）の数が少なく、抽出された特定の調査区がその属する地域ブロックを過大に代表しているなど、JPSCの調査設計に起因するところが大きいように感じる。PSU内・間の回収率の傾向は今後の課題としたい。

学歴別の回収率については、コーホートEはどの学歴においても85、86%の回収率である。全体の回収率とほぼ同じ値であり、学歴間で回収率の差異はない。他のコーホートについても学歴間で回収率に大きな差はなく、概ね3%ポイントの範囲におさまっている。

これら属性別の回収率の違いを総合的に評価する指標として、R-指標（Representativity-Indicator）がある（Bethlehem et al. 2009）。R-指標は、対象者個々の回収確率の標準偏差について、それを2倍したものを1から引いた値として定義される。指標の値は基本的に0から1の範囲をとり、値が1に近ければ回収標本の偏りは小さく、0に近ければ偏りは大きい。図表-1で、R-指標を計算する回収確率は、人口統計学的属性（地域ブロック、区市郡、年齢、配偶状態）の情報と学歴を補助変数としてロジスティック回帰を行い推計した。応答変数は各コーホートとも2回目の調査での回収・未回収の状況である。指標の作り方から、上記の補助変数で標本を交差分類したとき、それらのセル間で回収確率に違いがないときには標準偏差の値が小さくなり、R-指標は1に近くなる。コーホートEの指標の値は0.92であり、コーホートEよりも回収率が高かったコーホートBやコーホートDに比べて指標の値は1に近くなっている。すなわち、コーホートEは回収率自体は低くなったものの、人口統計学的な属性や学歴といった観点から見れば、属性間で回収率の偏りが相対的に大きくなかったといえる。

3.「調査への協力理由」の回答結果

JPSCではコーホートEの追加に合わせ、対象者の調査に対する態度を測る質問を収録した。

追加した調査項目は、「記入にかかった時間」、「調査への協力理由」、「調査票をどのように記入したか」の3つで、いずれも調査対象者に記入してもらっている。JPSCの調査単位は個人（女性）だが、調査事項は世帯調査の側面も強く、他の世帯員の労働時間や所得など対象者が正確に把握していないケースも多いと考えられる。「調査票をどのように記入したか」は、そのようなケースを想定し、対象者が回答を記入するにあたって、自分一人ですべての質問を回答したか、あるいは家族に尋ねたり、一緒に回答したかを調査したものである。追加したこれら3つの調査項目は、調査協力に対する動機の違いや調査に対する取り組み方の差異を測定している。調査の指向性と呼べるもので、調査への継続参加を予測する指標として活用できる可能性がある（土屋 2006）。本稿ではこれらの項目の中から、「調査への協力理由」の回答結果について取り上げる。本節ではコーホートEの回答結果について俯瞰し、次節では補正ウエイト作成時に情報を組み込むことの効果を計量する。

「調査への協力理由」の質問で提示した選択肢は、「回答することが社会のためになる気がするので」、「家計や自分の生活を振り返る機会になると思うから」、「回答したら謝礼をもらえるから」、「特に断る理由もないから」、「その他」の5つであり、対象者にはこの中からいずれか一つを選ぶように指示した。図表-2に提示した値は、コーホートEにとっての初回調査である第21回調査での回答割合である。

図表-2の最上段の行の値は、コーホートEの標本全体での回答割合を示したものである。最も割合が大きい回答は、「特に断る理由もないから」であり、全体の49.7%にあたる。半数近くの対象者は明確な理由や意思を持って調査に参加していないことになるが、他方で残りの半数は調査の参加にあたって何らかの積極的な理由を見出している

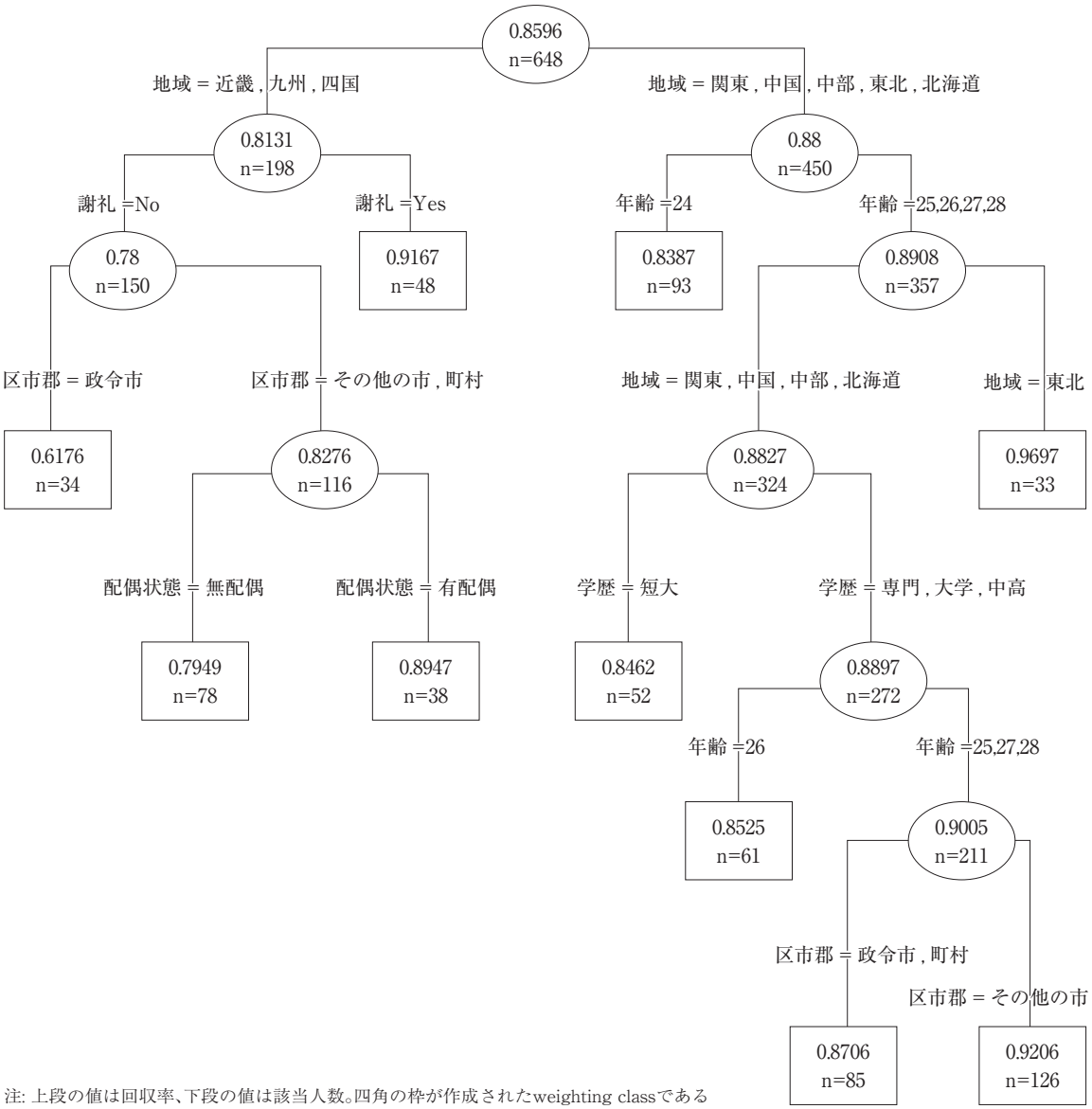
図表-2「調査への協力理由」の回答割合

	社会のため	振り返る機会	謝礼がもらえる	断る理由がない	その他	無回答	n
全体	11.3%	9.6%	25.0%	49.7%	3.7%	0.8%	648
回収別							
回収	10.8%	9.3%	26.4%	49.6%	3.2%	0.7%	557
未回収	14.3%	11.0%	16.5%	50.5%	6.6%	1.1%	91
	x-squared = 4.4767, df = 3, p-value = 0.2144						
配偶状態別							
有配偶	9.5%	9.5%	23.9%	52.7%	3.5%	1.0%	201
無配偶	12.1%	9.6%	25.5%	48.3%	3.8%	0.7%	447
	x-squared = 1.4995, df = 3, p-value = 0.6824						
地域ブロック別							
北海道	21.4%	7.1%	21.4%	50.0%	0.0%	0.0%	28
東北	9.5%	7.1%	11.9%	69.0%	2.4%	0.0%	42
関東	10.8%	13.0%	29.4%	43.7%	2.6%	0.4%	231
中部	8.8%	5.3%	23.0%	57.5%	5.3%	0.0%	113
近畿	12.5%	6.7%	28.8%	46.2%	1.9%	3.8%	104
中国	13.9%	13.9%	25.0%	44.4%	2.8%	0.0%	36
四国	11.1%	11.1%	16.7%	50.0%	11.1%	0.0%	18
九州	10.5%	9.2%	19.7%	52.6%	7.9%	0.0%	76
	x-squared = 25.165, df = 21, p-value = 0.2401						
学歴別							
中学・高校	9.3%	7.1%	26.9%	52.2%	3.3%	1.1%	182
専門・専修	7.7%	10.6%	25.0%	51.0%	4.8%	1.0%	104
短大・高専	9.8%	12.0%	20.7%	54.3%	2.2%	1.1%	92
大学・大学院	14.4%	10.0%	25.2%	45.9%	4.1%	0.4%	270
	x-squared = 8.4218, df = 9, p-value = 0.4923						

と捉えることもできる。特定の協力理由で回答割合が大きいのは「回答したら謝礼をもらえるから」であり、全体の4分の1にあたる25.0%が選択している。「回答することが社会のためになる気がするので」と「家計や自分の生活を振り返る機会になると思うから」の回答割合は、両者ともほぼ1割前後で、「社会のため」が11.3%、「振り返る機会」が9.6%となっている。

回収別の値は、コーホートEにとって2回目の調査である第22回調査の調査票が回収か、未回収かで分類したものである。回収群と未回収群での独立性の検定を行うと、得られたp値は0.21となり、両群で回答傾向に有意な差があるとまではいえない[3)]。何らかの積極的理由があるからといって調査への継続参加が保証されるわけではなく、逆に積極的な理由がないからといって調査から顕著に脱落するわけでもないことになる。ただし、「回答したら謝礼をもらえるから」の割合については、回収で26.4%、未回収で16.5%と10%ポイントほどの差がある。図表への掲載は割愛するが、検定の設定を変えて、協力理由が「謝礼」か否かで検定を行うと、p値は0.06となり、10%水準で評価すれば両群の回答に差があるという結果が得られる。「謝礼」を協力理由にしていたかどうかは、そ

図表-3 決定木による推計結果



注: 上段の値は回収率、下段の値は該当人数。四角の枠が作成されたweighting classである

の後の回収確率に影響を与えている。

属性別の回答割合の結果については、おしなべて各群での独立性が支持される結果が出ており、回答傾向に明確な差異があるとまでは言えない。地域別の回答割合は各群で若干ばらつきがあるものの、前述したように各地域ブロックのサイズはばらつきがあり、大きさが30に満たない地域もいくつかある。この結果を地域性の違いと直截的に解釈するのは差し控えた方がよい。学歴別の協力理由については、「回答することが社会のためになる気がするので」という回答が、大学・大学院で14.4%と他の学歴層より若干、値が大きくなっている。今後のデータの蓄積や他コーホートとの比較から、今後、学歴と協力理由については何らか

図表-4 ウエイト値の分布

	最小値	25パーセンタイル値	中央値	平均値	75パーセンタイル値	最大値	不等加重効果
「謝礼」含む							
レイキング	0.98	1.12	1.17	1.16	1.21	1.37	1.004
傾向スコア	1.04	1.11	1.15	1.16	1.20	1.47	1.004
傾向スコア（クラス分け）	1.07	1.11	1.15	1.16	1.21	1.32	1.003
決定木	1.03	1.09	1.15	1.16	1.19	1.62	1.009
「謝礼」含まない							
レイキング	1.05	1.12	1.16	1.16	1.20	1.33	1.002
傾向スコア	1.07	1.12	1.15	1.16	1.19	1.38	1.003
傾向スコア（クラス分け）	1.10	1.12	1.15	1.16	1.20	1.30	1.002
決定木	1.03	1.09	1.15	1.16	1.19	1.37	1.004

の示唆が引き出せる可能性はあるかもしれない。

4. 補正ウエイトの作成

本節では、コーホートEについて回収標本へのウエイト値の作成、およびウエイト値の付与によるバイアス補正効果を扱う。ウエイトの作成に用いる補助変数は、基本的に人口統計学的属性のみを用いる。人口統計学的属性は、いずれのコーホートについても全員の値が分かり、また時間に対してほぼ不変であることから、ウエイトの作成において必須の項目となる。本稿では、JPSCにおけるウエイト補正の効果を測る端緒として、人口統計学的属性に焦点を絞り、その補正効果をまず明らかにする。さらに、本稿でのもう一つの主眼である「調査への協力理由」が与える補正効果を明確にするため、用いる補助変数を制限し、協力理由を補助変数に含めるかどうかで効果がどの程度変動するかを計測する意図もある。なお、「調査への協力理由」の補助変数への導入は、回収と未回収とで有意な回答差が認められた、「回答したら謝礼をもらえるから」を選択したかどうかにのみ着目する。

補正ウエイトの作成には、レイキングによる方法、傾向スコアを利用した方法、決定木による方法の3つを用いた[4)]。レイキングは、事後層化を複数の補助変数に一般化したものである。本来的には、レイキングをはじめとするキャリブレーションによるウエイト作成は、母集団情報として外部情報、例えば「国勢調査」等の大規模調査を用いて回収標本のずれを補正できることに利点がある。ただ今回は、後述するように初回抽出の標本からのずれを検証するため、外部情報ではなく初回抽出標本の情報を用いる。同じ補助変数を用いるため、次の傾向スコアによる方法と同様のウエイト系列が得られる可能性が高い。ここでは、他の作成方法が大きく外れたウエイト値を算出していないか、メルクマールとして活用する。

傾向スコアによる方法は以下の手続きをとる。対象の属性（補助変数の値）で条件付けたとき、各対象がそれぞれ回収できる確率を傾向スコアと呼ぶ（Rosenbaum and Rubin 1983）。ウエイト付けは傾向スコアの逆数を用いて行う。真の回収確率は分からないため、本稿ではロジスティック回帰で推定する。傾向スコアを用いる方法の欠点として、過大なウエイト値が算出される点がある（Little and Rubin 2002）。過大な値を付与することを避ける簡易的な方法は、得られたウエイト値を数値の大小で複数の階級に分け、その代表値で置き換えることである（Valliant et al. 2013）。ここでは10個の階級に分け、その中央値を代表値として扱った結果も併記する。また、ロジスティッ

図表-5 ウエイト間の相関

「謝礼」含む

	レイキング	傾向スコア	傾向スコア（クラス分け）	決定木
レイキング	1.00	0.97	0.96	0.56
傾向スコア		1.00	0.97	0.65
傾向スコア（クラス分け）			1.00	0.62
決定木				1.00

「謝礼」含まない

	レイキング	傾向スコア	傾向スコア（クラス分け）	決定木
レイキング	1.00	0.98	0.97	0.62
傾向スコア		1.00	0.98	0.60
傾向スコア（クラス分け）			1.00	0.60
決定木				1.00

図表-6 初回サンプルとの距離D

	D	比率
ウエイトなし	0.47099	1
「謝礼」含む		
レイキング	0.40103	0.85
傾向スコア	0.40710	0.86
傾向スコア（クラス分け）	0.40586	0.86
決定木	0.44983	0.96
「謝礼」含まない		
レイキング	0.41634	0.88
傾向スコア	0.41837	0.89
傾向スコア（クラス分け）	0.42117	0.89
決定木	0.50067	1.06

ク回帰を行う際に共変量に交差項を用いた推計も考えられるが、今回は用いていない。組み合わせ数が多くなることと、意味のある組み合わせを自動的に抽出したものが、次の決定木による方法ととらえることもできるからである。

決定木は探索アルゴリズムの一種で、回収確率の違いに基づいて補助変数を自動的に分類していくものである。ここでは、classification regression tree（CART）アルゴリズムを採用している。分類されたクラスの回収確率の逆数がウエイト値となる。分類されるクラスの最低サイズは、標本サイズの約5%にあたる30に設定した。「謝礼」と人口統計学的属性を補助変数として、決定木により識別した各クラスの回収確率を図表-3に提示している。

図表-4には上記の方法で作成した各ウエイト値の分布を示した。上段は補助変数に人口統計学的属性に「謝礼」の要因を加えて作成したウエイト、下段が人口統計学的属性だけを用いたものである。表の右端にある不等加重効果（unequal weighting effect）は、関心のある変数と補正ウエイトとが無相関であることを想定したとき、ウエイトを導入することで推定量の分散が何倍になるかを示す指標である（Kish 1992）。不等加重効果は定義式を変形することにより、ウエイトの変動係数を2乗したものに1を加えたものと等しくなる。

まず、ウエイト値の平均について見ると、どの作成法でも、「謝礼」の要因を含める場合と含めない場合で、平均値が1.16と同じ値を示している。ウエイト値の最小値については1前後でどの作成方法でもほぼ違いがない。最大値は「謝礼」の要素を含めた場合に作成方法間でやや違いがあり、決定木による方法が1.62で大

図表-7 配偶状態別の所得分布とウエイト値による補正（税込み、単位：万円）

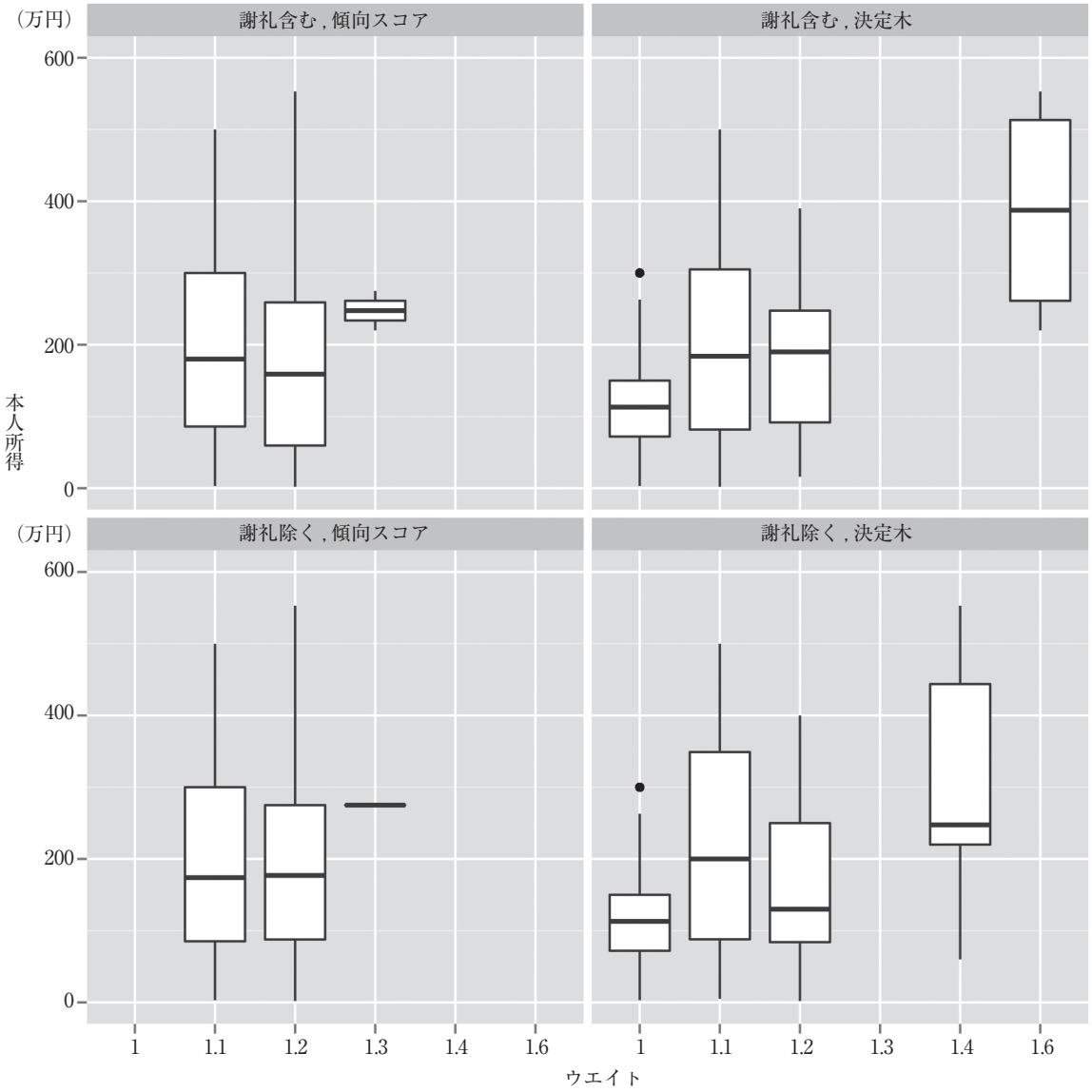
	有配偶（所得あり）			無配偶		
	10パーセンタイル	中央値	90パーセンタイル	10パーセンタイル	中央値	90パーセンタイル
初回サンプル	50.0	200.0	390.0	101.2	240.0	380.0
	12.9	26.9	16.4	4.7	7.7	12.8
回収標本						
ウエイトなし	46.0	180.0	382.0	100.7	240.0	381.2
	12.4	27.0	18.5	5.0	7.7	12.8
「謝礼」含む						
レイキング	44.6	181.2	381.5	100.3	240.0	380.0
	12.8	27.0	18.5	5.3	7.7	12.8
傾向スコア	45.0	181.8	381.4	100.1	240.0	380.0
	12.7	27.1	18.5	4.9	7.7	12.8
傾向スコア（クラス分け）	45.2	182.6	381.5	100.3	240.0	380.0
	12.6	27.1	18.5	5.1	7.7	12.8
決定木	47.1	200.0	388.0	100.0	240.0	380.0
	12.4	28.0	29.1	5.0	7.7	12.8
「謝礼」含まない						
レイキング	45.2	180.3	380.8	100.1	240.0	380.0
	12.6	26.8	18.4	4.8	7.7	12.8
傾向スコア	45.4	180.8	380.9	100.0	240.0	380.0
	12.5	26.9	18.4	4.9	7.7	12.8
傾向スコア（クラス分け）	45.5	181.6	381.2	100.2	240.0	380.0
	12.5	27.0	18.5	4.8	7.7	12.8
決定木	45.8	191.7	383.7	100.0	240.0	380.0
	12.3	27.2	18.6	5.1	7.7	12.8

注: 下段は標準誤差

きくなっている。「謝礼」の要素を含めない場合は、決定木による方法の値が下がり、最大値はいずれの作成方法でも大きな違いはなくなる。不等加重効果は、コーホートＥでは人口統計学的属性間に回収率の差があまりなかったことを反映して、全体的に極めて小さな値をとっている。「謝礼」要因を追加したことの影響は、不等加重効果でも最大値と同様の傾向がみられ、決定木による方法の値が変化する。まとめると「謝礼」を補助変数に加えることは、決定木ではウエイト値に影響を与え、他の方法ではそれほど大きな影響を与えない。

ついで図表-5は、作成方法の間の相関をみたものである。レイキングと傾向スコア法のウエイトの相関はほぼ1である。傾向スコアによる方法の欠点としてあげられる、過大なウエイト値も算出されていないため、ほぼ同等の系列が作成されている。得られた傾向スコアをそのまま用いたものとクラス分けしたものとの間の相関はほぼ1となっている。先の不等加重効果をみれば、傾向スコア法により作成したウエイト値の散らばりは小さく、区切り幅も小さなものになったため原ウエイトとクラス分けしたものとの差が広がらなかったと考えられる。ただし、これらの系列と決定木による系列の相関係数は0.5 ～ 0.6の値をとり、まったく

図表-8 weighting classごとの所得分布（有配偶女性、所得あり）



同一の値となっていない。傾向スコア法と決定木による方法では、同一の対象に対して異なる値を付与しているケースが散見されることを意味している。

5. 回収標本のバイアスとウエイト値による補正効果

第22回調査での未回収標本の当該調査での回答は分からないため、回収標本にどの程度のバイアスがあり、ウエイト値を付与することによってどの程度の補正効果があるかを評価することは、不

可能である。ここでは、一つのアプローチとして、初回調査である第21回調査で抽出したコーホートEの全対象（以降、初回サンプル）と第22回調査で回収できた標本（以降、回収標本）について、それぞれの第21回調査の回答を用いて比較することにする。つまり、初回サンプルの第21回調査の回答割合を母集団割合とみなし、ウエイト値による補正を施すのは回収標本の第21回調査での回答である。

評価には、択一式の選択肢の質問群と、対象者の本人所得（前年年間、税込み）を用いた。択一式の選択肢質問では、初回サンプルの回答割合（p_{kc}）と回収標本の回答割合（$\hat{p}_{kcw}$）について、土屋（2006）と同様、回答分布間の距離を下記の量で評価する。

$$D=100\times\frac{1}{Q}\sum_{k=1}^{Q}\left\{\frac{1}{C_k}\sum_{c=1}^{c_k}(p_{kc}-\hat{p}_{kcw})^2\right\}^{1/2}$$

は質問項目kのカテゴリー数、Qは質問項目数である。採用した質問項目は、「生活満足度」、「社会に対する満足度」、「本人の健康状態」、「生活程度」、「去年から今年の所得の増減」、「来年の所得の増減予想」、「家計簿記帳の頻度」、「就業状態」の8つで、いずれの質問も4から5件法である。

図表-6には、ウエイト値による補正を行った場合と行わない場合のDの値を記載した。比率は各ウエイト値によるDを、補正を行わなかった場合のDで割ったものである。補正を行わなかった場合のDは0.47であり、指標Dの定義を考えれば、初回サンプルと回収標本で回答割合はほぼ一致しているといってよい結果である。回収率が低下したとはいえ、標本の86%が得られていれば、初回サンプルと回収標本のずれは小さいことがうかがえる。そのような状況下で、ウエイト値による補正は、概ねDの値をさらに小さくする効果がある。「謝礼」の要因を補助変数に含めた方がDの縮小幅が大きく、決定木より傾向スコアの方法がDを縮小する傾向にある。「謝礼」を含めなかった決定木による補正は、補正を行わなかった場合よりDの値を増大させている。

次に所得の補正効果を見てみる（図表-7）。女性の場合は、有配偶か無配偶かで就業形態や労働時間の構成比が大きく異なり、また有配偶は就業しているかどうかで所得分布が大きく分断される。図表-7の結果は、配偶状態別、また所得が0でない対象だけを用いている。無配偶の所得については、いずれの作成方法の値も大きく異ならない。回収標本の値をそのまま用いても初回サンプルの値にほぼ一致している。回収標本の値にウエイト値を付与することで、副作用が出るわけでもない。「謝礼」の要因も補正値に影響は与えていない。一方で、有配偶の所得は、初回サンプルと回収標本の間で若干のずれがある。回収標本をそのまま用いた値は初回サンプルの値より低いものとなっており、有配偶では相対的に所得の高い層での脱落があったことがうかがえる。ウエイト値による補正は、傾向スコアによる方法では値がほとんど変化していない。むしろ、10パーセンタイルでは値が縮小され、偏りの大きさがわずかではあるものの増幅されている。決定木による方法では、初回サンプルと回収標本の間でのずれが大きかった中央値と90パーセンタイルの値が拡大されており、一定の補正効果が認められる。特に補助変数に「謝礼」を含めることにより、有配偶女性の所得についての補正効果が上昇していることも確認できる。ただし、決定木による方法の標準誤差は他のものに比べて大きくなっている。

図表-8は、有配偶女性の所得について、傾向スコア（クラス分け）と決定木に関して、付与したウエイト値を小数点1桁でまとめ、その層ごとの所得分布を示したものである。箱は第1四分位と第3四分位の範囲を示し、箱内の太線は中央値である。この図からは、ウエイト値が所得の補正に対してどのように機能したかが見てとれる。傾向スコアによる方法は、ウエイト値間の散らばりも小さく、各クラスの中央値には大きな差がない。ゆえにウエイト値の付与による補正効果は限定的なものになる。一方で、決定木による方法のグラフは低いウエイト値を割り振る層では低い所得帯になっており、高いウエイト値を割り振る層では高い所得帯となっているため、初回サンプルとのずれをうまく補正した結果が得られた。図表-8で

1.6と高いウエイト値を付与したクラスを図表-3で特定すると、近畿・四国・九州ブロックで、「謝礼」を参加継続理由にあげていない、政令市の対象者である。今回の傾向スコア法は交差項を考慮せず、同じ属性の値については一律の寄与を与えている。決定木による方法は、その不備を解消し、回収率と所得の回答傾向の違いをうまく層別している。

本節で検証したコーホートEのウエイト補正効果についてまとめると、次のようになる。初回サンプルと回収標本を比較するとさほど大きな違いはない結果が得られている。その前提の上で、ウエイト補正にはさらなる一定の効果がみられた。択一式の選択肢の質問群では、人口統計学的属性について、それらが寄与する脱落確率を計算し、ウエイト値を付与した方法で補正効果が高い（今回の傾向スコアによる方法）。コーホートEでは、この方法により作成したウエイト値の分散は決定木による方法に比して相対的に小さくなっていた。一方の本人所得では、近畿の政令市といった具合に、回収率の低い特定のクラスを決定木により抽出し、相応のウエイト値を付与する方法の補正効果が高かった。こちらの方法はウエイト値の分散が相対的に大きくなっている。選択肢質問のような値の範囲が限定された質問項目では対象者間の回答傾向が似るためウエイト値の分散が小さい補正が、所得のように散らばりが大きく出やすい項目についてはウエイト値の分散が大きい補正が功を奏するのかもしれない。今回の結果はあくまでコーホートEの事例研究にとどまる。質問やその回答の性質とウエイト値の作成方法の関係については、今後の課題としたい。

6. 考察

本稿では、2013年の第21回調査で追加したコーホートEの第22回調査での回収率、および回答傾向の偏りとウエイト値の付与による補正の可能性について検証してきた。コーホートEの回収率は86%でこれまでの追跡コーホート中で最低の値を記録した。一方で、人口統計学的な補助変数で条件付けた場合、既存コーホートに比べて、回収率に大きな偏りがあるわけではなかった。

コーホートEの初回サンプルと回収標本の比較では、カテゴリー変数の回答割合に大きな偏りは確認できなかった。そのような中において、人口統計学的情報を補助変数としたウエイト値の補正は、さらに偏りを縮小させる効果があった。また、有配偶の対象者の本人所得については、回収サンプルの分布の方が、若干、低い方にずれていた。傾向スコア法による所得への補正効果は限定的であるのに対し、決定木による方法は偏りを解消させる方向に働いていた。特に決定木では、調査への協力理由が「謝礼」がもらえるかどうかが地域情報と効果的に結びつき、回収率と目的変数である本人所得の双方に関連しているクラスを識別することに成功している。なお、ウエイト値の付与による推定値の分散増大の影響は、いずれの作成方法でも限定的であることが示されている。

ただ、今回のウエイト付与の結果が効果的であったからといって、経年摩耗が進んでいないコーホートEに対して常にウエイト補正を行うべきかと問われれば、筆者は積極的にくみしない。今回の推計では、weighting class間のウエイト値の散らばりも小さく、各classの平均や回答分布に差異がないものも多かった。86%と、まだある程度高い回収率を保っている状態では、ウエイト作成のコスト等と補正効果の影響をトレードオフにして考えるべきである。また、今回用いた「調査への協力理由」以外の、「記入時間」や「回答の方法」といった対象者の志向性を示す変数も、ウエイト値による補正の際に有用な情報となることが期待される。しかし、JPSCのコーホートは標本サイズが600前後と限られている。このような調査のメタ情報を多数捕捉することで、よりよい解決法が得られるかといえば、多すぎる情報はセルの細分化をもたらし、JPSCでは少数の対象者の値を過度に代表させることにもなる。むしろ大規模標本でない場合は、いかに脱落させないかが、とりわけ重要な課題である。志向性の変数について、事後的な補正の活用は当然のことではあるが、記入時間から回答者の負担を測るなど、対

象者の調査に対するモチベーションを上げ、標本の必要以上の摩耗を事前に防ぐ情報としても積極的に活用していくべきと考える。

最後に、本稿では取り上げなかったが、調査の指向性を示す変数は、時間不変の値ではない。第21回調査の回答と第22回調査の回答が変動している対象者も一定数いた。どのような回答の時系列を示す対象者が調査に残りやすいのか、また時系列の変動を補正の情報に組み込むにはどのようにすればよいかは、今後データの蓄積に応じて明らかにしていく必要があることを述べておく。

注

1）パネル調査のウエイト値による補正は、JPSCをはじめとして日本のパネル調査でもこれまで議論されている（坂本 2006; 石井・野崎 2014; 三輪 2014）。
2）世帯調査における、対象の性質と回収率の関係については、Groves and Couper（1998）に詳しい。また、土屋（2006）では、複数の調査の指向性変数について、回収標本のバイアス軽減効果を議論している。
3）検定にあたって、セルの該当数が少ない「その他」と「無回答」は、「特に断る理由がない」にまとめた。
4）ウエイトの作成方法については、例えば、Kalton and Flores-Cervantes（2003）、土屋（2009）などに詳しい。

文献

石井加代子・野崎華世，2014，「『慶應義塾家計パネル調査（KHPS）』と『日本家計パネル調査（JHPS）』におけるLongitudinal/Cross-sectionalウエイトおよびパネル統合ウエイトの作成」Panel Data Research Center at Keio University, Discussion Paper Series, DP2013-005.

坂本和靖，2006，「サンプル脱落に関する分析――「消費生活に関するパネル調査」を用いた脱落の規定要因と推計バイアスの検証」『日本労働研究雑誌』551: 55-70.

土屋隆裕，2006，「『調査への指向性』変数を用いた調査不能バイアス補正の試み――『日本人の国民性調査』データへの適用」『日本統計学会誌』36: 1-23.

――――，2009，『概説 標本調査法』朝倉書店.

三輪哲，2014，「NFRJ-08Panelにおけるウェイトによる脱落への対処」『家族社会学研究』26: 169-178.

Bethlehem, Jelke, Fannie Cobben, and Barry Schouten, 2009, “Indicators for the Representativeness of Survey Response,” *Survey Methodology*, 35: 101-113.

Groves, R. M. and M. P. Couper, 1998, *Nonresponse in Household Interview Surveys*, New York: John Wiley and Sons.

Kalton, G. and I. Flores-Cervantes, 2003, “Weighting Methods,” *Journal of Official Statistics*, 19: 81-97.

Kish, L., 1992, “Weighting for Unequal P_i,” *Journal of Official Statistics*, 8: 183-200.

Little, R. J. A. and D. B. Rubin, 2002, *Statistical Analysis with Missing Data*, 2nd ed. New Jersey: John Willey and Sons.

Rosenbaum, P. R. and D. B. Rubin, 1983, “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrica*, 70: 41-55.

Valliant, Richard, Jill A. Dever, and Frauke Kreuter, 2013, *Practical Tools for Designing and Weighting Survey Samples*, New York: Springer.

さかぐち・なおふみ　公益財団法人 家計経済研究所次席研究員。主な論文に「消費生活に関するパネル調査」における残存サンプルの配偶率について」（『季刊家計経済研究』100，2013）。労働経済学専攻。
（sakaguchi@kakeiken.or.jp）