

## 世帯調査の方法と調査世帯の性質

—世帯構成，年収，学歴に関する比較—\*<sup>1</sup>

佐野 晋平\*<sup>2</sup>

多田 隼士\*<sup>3</sup>

山本 学\*<sup>4</sup>

### 要 約

本稿は，調査方法の異なる家計関連統計を比較し，各統計の特徴と利用上の注意点を明らかにした。二人以上世帯については，調査された世帯属性の平均には大きな違いはなかった。ただし，世帯属性の分布には調査方法の違いにより差が見られ，国民生活基礎調査は全国消費実態調査，家計調査より平均的ではない家計を多く捉えている可能性がある。一方で，単身世帯については，国民生活基礎調査で割合が低くなった。代替世帯を取らず比推定も用いない国民生活基礎調査では，単身世帯が過小に把握されていると考えられる。

学歴に関しては，統計間で概ね近い分布を示すものの，国民生活基礎調査と国勢調査に含まれる「学歴不詳」者が比較を困難にしている。国勢調査のコホート情報と労働力調査の学歴分布を用いた推計によると，学歴不詳者は小中高卒グループで多く発生していることが明らかとなった。学歴不詳者を集計から除くことはバイアスを生む可能性が高いと考えられる。

キーワード：世帯調査，非標本誤差

JEL Classification：C80, C83

## I. はじめに

家計を対象とした調査は，社会経済の実態把握や学術研究において重要な役割を果たす。経済厚生 of 重要な決定要因である所得・消費・貯

蓄などの情報は，家計収支を調査することで把握できる。特に，所得格差の指標のように家計間の非対称性を分析するには，個々の家計を調

\*<sup>1</sup> 本稿の作成にあたり，国民生活基礎調査の個票を利用した。また宇南山卓総括主任研究官（財務省）より貴重な助言を得た。関係者各位に厚く御礼申し上げる。なお，ありうる誤りはすべて筆者の責にある。本稿の意見は筆者らが所属する組織の公式見解を示すものではない。

\*<sup>2</sup> 千葉大学法政経学部准教授

\*<sup>3</sup> 財務省財務総合政策研究所研究部研究員

\*<sup>4</sup> 財務省財務総合政策研究所研究部研究員

査するミクロデータが不可欠である。家計収支以外にも、労働力状態を調査することで失業の状況が把握でき、教育に関する統計によって一国の人的資本水準が計測できる。

家計を対象とした公的統計は複数存在するが、調査方法や調査項目の違いによりそれらの結果が異なることがある。たとえば、所得格差の指標を示すジニ係数は用いる統計よりその数値が異なることが指摘されている（橋木1998、舟岡2001、大竹2005、小塩2010、内閣府2010など）。一般に、公的調査の調査方法は悉皆調査と標本調査に大別されるが、標本調査のうちでも調査手順は異なる。

ほとんどの調査は、層化多段抽出と呼ばれる方法で、国勢調査区から調査地区を無作為に抽出するが、その詳細な手順は統計によって異なる。特に、厚生労働省が調査・公表している国民生活基礎調査は、調査対象地区のすべての世帯を調査する「集落抽出」の手法を採用している。それに対し、総務省が調査・公表する家計調査や全国消費実態調査などは、抽出された調査地区の中からさらに調査世帯を抽出して調査している。また、家計調査や全国消費実態調査は調査結果を集計する際に、世帯属性の偏りを補正するように個々の標本に重みをつける比推定と呼ばれる手法を採用している。こうした異なる調査手順は、理論的な性質としてはほぼ同等であるが、調査対象が調査を拒否した場合には実務上の対応次第で統計的な性質に差がでる。

本稿では、家計を対象とした異なる統計同士を比較することで、各統計の性質を明らかにする。特に、世帯構成、所得、学歴などの世帯属性に関する主要な変数を比較する。統計間の差異を比較することで、調査方法による結果の違

いを、それぞれの調査方法の特徴と関連付けることで、統計利用上の注意点を明らかにした。

比較の結果、二人以上世帯については、調査された世帯属性の平均には大きな違いはなかった。これは、調査方法によらず、日本の主要な官庁統計がおおむね日本の代表的な世帯を調査できていることを示唆する。ただし、世帯属性の分布には調査方法の違いにより差が見られ、国民生活基礎調査は全国消費実態調査、家計調査より平均的ではない家計を多く捉えている可能性がある。

一方で、単身世帯については、国民生活基礎調査で割合が低くなった。単身世帯は調査拒否の可能性が高いことは知られており（新田、2004）、代替世帯を取らず比推定も用いない国民生活基礎調査では、単身世帯が過小に把握されていると考えられる。

学歴に関しては、統計間で概ね近い分布を示すものの、国民生活基礎調査と国勢調査に含まれる「学歴不詳」者が比較を困難にしている。国勢調査のコホート情報と労働力調査の学歴分布を用いた推計によると、学歴不詳者は小中高卒グループで多く発生していることが明らかとなった。学歴不詳者を集計から除くことはバイアスを生む可能性が高いと考えられる。

本稿の構成は以下の通りである。第Ⅱ節では、世帯に対する基幹統計である国民生活基礎調査、全国消費実態調査、家計調査の3統計の各統計の調査概要や調査手順を概観する。第Ⅲ節では、世帯人員や世帯主年齢などの世帯属性に注目し、各統計を比較する。第Ⅳ節では、世帯ベースで学歴の情報を得られる国民生活基礎調査や国勢調査などの比較をする。第Ⅴ節はまとめである。

## Ⅱ．家計に対する調査の概要

### Ⅱ－１．国民生活基礎調査

国民生活基礎調査は「保健、医療、福祉、年金、所得等国民生活の基礎的事項を調査し、厚生労働行政の企画及び運営に必要な基礎資料を得るとともに、各種調査の調査客体を抽出するための親標本を設定することを目的」とした調査で、厚生労働省により実施されている。3年おきに実施され詳細な事項まで調べる大規模調査と、その間の2年間に実施される小規模調査からなる。本稿では大規模調査のみ焦点を当てる。

大規模調査は、調査事項ごとに「世帯票」「健康票」「所得票」「貯蓄票」「介護票」5つの調査票に分かれている。厚生労働省から都道府県を通じ、「世帯票」「健康票」「介護票」については保健所に、「所得票」「貯蓄票」については福祉事務所にそれぞれ調査委託がなされる。調査対象地区の保健所、福祉事務所の調査員は各世帯に調査票を配布し、後日調査員が配布する方法をとる。ただし、貯蓄票は密封回収、健康票・所得票はやむをえない場合のみ密封回収をしている。

調査対象は、原則として全ての世帯であるが、単身赴任者、遊学中の者、社会福祉施設の入所者、長期入院者、収監中の者などは調査対象外である。また、調査対象地区内のすべての世帯が調査対象となるが、調査対象そのものは世帯ではなく個人であり、調査対象となるかどうか個人単位で決まる。

調査世帯の抽出方法は、全国から調査地区を

無作為に抽出し、その地区のすべての世帯を調査対象とする集落抽出である。「世帯票」「健康票」については全国約300,000世帯が調査対象となり、「所得票」「貯蓄票」「介護票」の調査対象世帯は、「世帯票」「健康票」の対象世帯の中から再抽出される。サンプルサイズは、「所得票」「貯蓄票」が全国約40,000世帯、「介護票」は介護保険法の要介護者及び要支援者約7,000人である<sup>1)</sup>。

調査対象世帯の抽出方法は、「世帯票」「健康票」の場合、国勢調査の調査区<sup>2)</sup>から、都道府県・政令指定都市および地区の基本情報<sup>3)</sup>によって層化し、層化無作為抽出で調査地区が決定される。集落抽出であり、その調査区内のすべての世帯が調査対象となる。

一方、「介護票」「所得票」「貯蓄票」の調査対象世帯は、「世帯票」「健康票」の調査対象世帯の中から抽出される。「所得票」「貯蓄票」については、「世帯票」「健康票」の調査対象になった国勢調査の調査区をさらに地理的条件等で分割した単位区<sup>4)</sup>を設定し、層化無作為抽出によって単位区を再抽出する。調査対象単位区が決まると、その中のすべての世帯が調査対象となる。「介護票」の調査対象世帯の抽出方法は「世帯票」「健康票」の調査対象地区から、再び層化無作為抽出で調査区が抽出され、調査区内の悉皆調査がされる。

このように国民生活基礎調査のすべての調査票について集落抽出が採用されており、調査地

1) 調査票ごとに対象世帯数が異なるのは後述する抽出方法の違いに起因する。

2) 1調査区は概ね50世帯になるように分割されている。この調査区のうち後置番号が「1」及び「8」の調査区が対象となる。後置番号とは調査区の属性を表す番号であり、「1」は一般調査区、「8」は概ね50人以上の単身者が居住している寄宿舍・寮が存在している地区である。

3) 例えば人口集中地区であるかどうか、農漁村や工業地帯であるかどうかなどを指す。

4) 1単位区は約25世帯である。設定された単位区のうち国勢調査の調査区の後置番号が「1」の単位区が対象となる。

区ごとに悉皆調査されるため、調査拒否世帯が発生しても代替世帯を調査することはできない。

## Ⅱ－２．家計調査

家計調査は「国民生活における家計収支の実態を把握」を目的とした調査で、総務省により実施されている。同調査は毎月実施され、全国約9,000世帯（うち単身世帯は約800世帯）を対象とする。他の2調査と比較すると小規模な調査であるが、調査頻度が高いことが特徴である。

調査票は「世帯票」（または「準世帯調査票」<sup>5)</sup>）、「家計簿」、「年間収入調査票」、「貯蓄等調査票」の4つであり、調査世帯は全ての調査票に回答する。「世帯票」については調査開始前に調査員が質問して調査する。「家計簿」は、二人以上世帯は6か月間、単身世帯は3か月間継続して調査を行う。調査世帯は、二人以上世帯では毎月1/6ずつ、単身世帯では毎月1/3ずつ入れ替わるように設計されている。「年間収入調査票」は調査開始月を含む過去1年間の収入について、記入開始後1か月目の後半に記入を行う。「貯蓄等調査票」は二人以上世帯のみ実施し、記入開始後3か月目の前半に記入を行う。

家計調査も原則として全世帯を対象としているが、住み込みのいる世帯、社会福祉施設、長期入院者、自衛隊の官舎内居住者、学生の単身世帯は調査対象外である。また国民生活基礎調査とは単身赴任者の扱いが異なり、単身赴任者世帯と単身赴任者を送り出した世帯は2つの別の世帯として調査される。さらに、これも国民生活基礎調査と異なり世帯単位での抽出となるため、調査対象外となる場合は、世帯単位で調査対象外となる。

調査対象世帯の抽出方法は層化3段抽出であり、第一段階で調査対象市町村を、第二段階で

調査単位区を、第三段階で世帯を抽出する。第一段階では168市町村を全国から選定する。第二段階では、選定された市町村内から国勢調査の調査区を単位として調査単位区を抽出する。第三段階では、二人以上世帯は選定された単位区で農林漁家、勤労、勤労以外の世帯グループに分けて1単位区から計6世帯選定する。単身世帯は二人以上世帯と同様の単位区より無作為に1世帯を選定する。なお、単身世帯の抽出では一般単位区の選定と同様に寮・寄宿舎を11単位区程度抽出し、1単位区から1つの寮・寄宿舎を選定し、そこから6世帯を抽出する。また、調査拒否世帯が発生した場合には代替標本を調査する。

さらに、単身世帯の調査世帯数は調査区の世帯分布と独立に決まっているため、世帯属性を補正するように、比推定と呼ばれる手法で世帯分布を調整している。具体的には労働力調査の世帯分布結果を基に地方・世帯人員・年齢階級などの構成を補正している。すなわち、世帯人員の分布や年齢分布などは労働力調査の分布と一致するように集計されているのである。この比推定によって、調査拒否が発生したとしても少なくとも一定の世帯属性についてはある程度偏りを補正することができる<sup>6)</sup>。

## Ⅱ－３．全国消費実態調査

全国消費実態調査は「家計の構造を「所得」、「消費」、「資産」の3つの側面から総合的に把握することを目的」とした調査で、総務省により5年おきに実施されている。同種の調査である家計調査で得られない詳細な結果を得るため、調査対象は約57,000世帯（うち単身世帯は約4,400世帯）と大きい。

調査票は「家計簿A（9、10月）」、「家計簿B（11月）」、「耐久財等調査票」、「年収・貯蓄等

5) 準世帯とは原則として1) 間借り・下宿などの単身者 2) 会社などの独身寮の単身者 3) 寮・寄宿舎の学生・生徒 4) 病院・療養所の入院者 5) 社会施設の入所者 6) 自衛隊官舎内居住者 7) 矯正施設の入所者 8) その他、を指す。

6) たとえば、松井（2005）を参照。

調査票」,「世帯票」の5つから成り立っている。「家計簿」は、二人以上世帯は9～11月の3か月間、単身世帯は10～11月の2か月間調査を行う。年収及び貯蓄等調査票の年収、貯蓄、借入金等の事項については、「年収・貯蓄等調査票」に11月末日現在の状況を記入し、世帯に関する事項については、「世帯票」に二人以上世帯では9月1日現在、単身世帯については10月1日現在の状況を記入する。

調査対象世帯は、家計調査とほぼ同じで、住み込みのいる世帯、社会福祉施設、長期入院者や学生の単身世帯などは調査対象外とされている。単身赴任者についても家計調査と同様に単身世帯に含まれる。また、これも家計調査と同様に世帯単位での抽出となるため、調査対象外となる場合は、世帯単位で調査対象外となる。

調査世帯の抽出方法は家計調査と同様に多段階抽出であるが、二人以上世帯と単身世帯で異なる。二人以上世帯ではまず調査する市町村を選定する。市についてはすべての市を対象とするが、町村については地理的配置、産業別特徴などを考慮して選定される。この調査対象に選定された市町村に調査単位区<sup>7)</sup>を設定し、それぞれ12世帯を抽出する。単身世帯では二人以上世帯の際に抽出されたそれぞれの調査単位区から系統抽出によって計4,400世帯抽出する。ここで、国民生活基礎調査の調査対象の抽出の際に実施したような寮・寄宿舎の地区は除外も区別も行わない。ただし、平成16年以前の調査では寮・寄宿舎の単身者を区別し、全国の30人以上居住する寮・寄宿舎を居住する者の所属する産業により7層に区分し、100か所(約600世帯)を抽出していた。

調査拒否世帯が発生した場合には、代替標本を調査する<sup>8)</sup>。こうして得られた標本を用いて、比推定と呼ばれる手法で補正し、集計を行って

いるのも家計調査と同様である。

## Ⅱ－4. 各統計の調査方法の比較

表1は前述した国民生活基礎調査、家計調査、全国消費実態調査の概要をまとめたものである。本稿では調査方法の違いにより、家計を対象とした統計間に違いがあるかを問題にしているので、各調査の特徴を概観した後、調査方法の違いにより生じる調査拒否への対応に焦点を当てて比較を行う。

家計調査は毎月調査されているため、1回の調査のサンプルサイズは小さいものの、季節性を考慮した上での、調査時の家計の収入、支出が把握できる。この点で他の2つの統計と異なる。ただし、家計簿への記入が必要であるため、調査世帯の負担が重いと言える。これは全国消費実態調査も同様である。

全国消費実態調査は5年に1度の大規模調査であるため、サンプルサイズは大きい、二人以上で9月～11月の3か月間、単身世帯では10月～11月の2か月間のみの調査であり、ボーナス期を含まない。そのため、他の2つの統計と異なり、季節性の影響を受けた偏りは避けられない。また家計調査と同様に、家計簿への記入が必要であるため、調査世帯の負担が重い。

国民生活基礎調査は、3年に1度の大規模調査に関しては十分なサンプルサイズが確保でき、所得や課税状況などは昨年度の調査になるものの、所得や課税がプリコード方式を採用しているため自由記入方式の家計簿方式に比べて、記入漏れは少ないと考えられる。また、「世帯票」「健康票」「貯蓄票」「介護票」は実施年度の状況に関する調査であるのに対し、「所得票」は住民税を除くと昨年度の所得や課税状況に関する調査である。つまり、他の2統計が調査時点の所得や税・保険料の支出を回答して

7) 1調査区単位とは国勢調査の近接2調査区に相当する。

8) 調査世帯として抽出された世帯で、三ヶ月間継続しての調査が不能となった場合は、代替の世帯を抽出し、残りの期間のみを調査している。なおⅢ節の集計においてはこれらすべての世帯を用いている。継続調査が不可能になったために二ヶ月のみ調査されている世帯については2/3世帯として含めている。

表1 家計を対象とした公的統計の比較

|            | 国民生活基礎調査                                                                                 | 全国消費実態調査                                            | 家計調査                                                |
|------------|------------------------------------------------------------------------------------------|-----------------------------------------------------|-----------------------------------------------------|
| 目的         | 福祉分野及び所得等国民生活の実態把握                                                                       | 「所得」、「消費」、「資産」の3つの側面から国民生活の実態把握                     | 家計収支の実態把握                                           |
| 間隔         | 3年に1回<br>(大規模調査)                                                                         | 5年に1回                                               | 毎月                                                  |
| 調査期間       | ・二人以上、単身世帯共に過去一年間のデータを6月(世帯、健康、介護)と7月(所得、貯蓄)に分けて調査                                       | ・二人以上世帯は(9～11月)3か月間<br>・単身世帯は(10,11月)2か月間           | ・二人以上世帯は連続する6か月間<br>・単身世帯は連続する3か月間                  |
| 調査規模       | ・「世帯票」「健康票」は全国約300,000世帯<br>・「所得票」「貯蓄票」は全国約40,000世帯<br>・「介護票」は介護保険法の要介護者及び要支援者を全国約7,000人 | 調査対象は約57,000世帯(うち単身世帯は約4,400世帯)                     | 全国約9,000世帯(うち単身世帯は約800世帯)                           |
| 調査票        | 世帯票、健康票、所得票、貯蓄票、介護票                                                                      | 家計簿A(9,10月)、家計簿B(11月)、耐久財等調査票、貯蓄等調査票、世帯票            | 家計簿、世帯票、準調査世帯票、年間収入調査票、貯蓄等調査票                       |
| 調査範囲       | 全国から人口や産業を層化基準とし層化無作為抽出                                                                  | 全ての市を調査対象に含んだ上で残りの町村を地理や産業を考慮して抽出                   | 政令市、県庁所在地は必ず含んだ上で残りを層化抽出                            |
| 調査対象外      | 単身赴任者、遊学中の者、社会府福祉施設、長期入院者、収監中の者など                                                        | 住み込みのいる世帯、社会府福祉施設、長期入院者、自衛隊の官舎内居住者や学生の単身世帯          | 住み込みのいる世帯、社会府福祉施設、長期入院者、自衛隊の官舎内居住者や学生の単身世帯          |
| 調査データ特徴    | 所得・課税(住民税を除く)が昨年度分                                                                       | 二人以上世帯で9～11月分の調査の為、ボーナス期を含んでいない                     | 年間収入調査票は調査開始月を含む過去一年分                               |
| 単身赴任者の調査状況 | 単身赴任者は除く                                                                                 | 単身赴任者は単身世帯に含み、二人以上世帯では単身赴任者のみを除いた他の世帯員の世帯として調査対象とする | 単身赴任者は単身世帯に含み、二人以上世帯では単身赴任者のみを除いた他の世帯員の世帯として調査対象とする |
| 勤労世帯       | 会社役員等は含める                                                                                | 会社役員等は勤労世帯に含まない                                     | 会社役員等は勤労世帯に含まない                                     |
| 調査拒否世帯への対応 | 補充は行わない                                                                                  | 補充する                                                | 補充する                                                |

いることと異なるため、回答内容に1年程度のずれが生じ、比較時には調整が必要になる。

各統計の調査方法の違いが最も大きく出る部分は、調査拒否への対応と比推定の有無である。国民生活基礎調査は比推定をする設計となっていないため、世帯属性の分布を調整するようなメカニズムは含まれていない。そのため、一定の属性を持った世帯で系統的に調査拒否が発生すると、その属性を持った世帯の分布が過小になってしまう。

それに対して家計調査や全国消費実態調査では比推定が採用されているため、少なくとも世帯人員や年齢分布に関しては事後的に補正がなされている。もちろん、系統的な属性を持つ世帯が調査拒否している場合、比推定では完全には偏りを除去できず、サンプリングバイアスがかかる可能性はある。例えば、回答者が相対的に学歴の高い者に偏った場合、所得水準や消費水準が上方にバイアスを持つ。

### Ⅲ. 各統計の世帯属性の比較

国民生活基礎調査、家計調査、全国消費実態調査の世帯構成、年齢、年収の項目を比較する。なお、年収についての比較を念頭に置くため、国民生活基礎調査は「所得票」を用い、年収調査年を揃える為に、国民生活基礎調査は2010年を、家計調査、全国消費実態調査は2009年を用いる。それぞれ個票データに基づ

き集計を行い、乗率ウェイトを用い復元している。必要に応じて、労働力調査の個票集計、家計消費状況調査と国勢調査の公表データの結果を用いている。出来るだけ統計間の条件を揃えた集計を行い、平均値だけではなく、分布も含めて統計間の比較を行う。

表2 世帯属性の比較

パネルA：総世帯(2009～2010)

|          | 国民生活<br>(2010) | 全国消費<br>(2009) | 家計調査<br>(2009) | 家計消費状況<br>(2009) | 国勢調査<br>(2010) | 労働力調査<br>(2009) | 国民生活Ⅰ<br>(2010) | 国民生活Ⅱ<br>(2010) |
|----------|----------------|----------------|----------------|------------------|----------------|-----------------|-----------------|-----------------|
| 集計世帯数    | 26,115         | 55,089         | 8,531          | 19,612           | —              | —               | 25,937          | 25,937          |
| 世帯人員     | (人)            | 2.65           | 2.48           | 2.49             | 2.50           | 2.46            | 2.66            | 2.66            |
| 18歳未満人員  | (人)            | 0.41           | 0.44           | —                | —              | 0.39            | —               | 0.42            |
| 65歳以上人員  | (人)            | 0.71           | 0.59           | —                | —              | 0.53            | —               | 0.72            |
| うち無職者人員  | (人)            | —              | 0.46           | —                | —              | —               | —               | —               |
| 有業人員     | (人)            | 1.29           | 1.17           | 1.13             | 1.21           | 1.15            | 1.20            | 1.30            |
| 世帯主の年齢   | (歳)            | 59.1           | 55.5           | 55.7             | 57.9           | —               | —               | 59.3            |
| 世帯主の性別 男 | (%)            | 79.9           | 77.7           | —                | —              | 76.0            | —               | 80.0            |
| 女        | (%)            | 20.1           | 22.3           | —                | —              | 24.0            | —               | 20.0            |
| 年間収入     | (千円)           | 5,496          | 5,532          | 5,350            | —              | —               | 5,525           | 5,525           |

パネルB：単身世帯(2009～2010)

|          | 国民生活<br>(2010) | 全国消費<br>(2009) | 家計調査<br>(2009) | 家計消費状況<br>(2009) | 国勢調査<br>(2010) | 労働力調査<br>(2009) | 国民生活Ⅰ<br>(2010) |
|----------|----------------|----------------|----------------|------------------|----------------|-----------------|-----------------|
| 集計世帯数    | 5,546          | 4,253          | 703            | 1,677            | —              | —               | 5,368           |
| 世帯人員     | (人)            | 1.00           | 1.00           | 1.00             | 1.00           | 1.00            | 1.00            |
| 18歳未満人員  | (人)            | 0.00           | 0.00           | —                | —              | 0.00            | —               |
| 65歳以上人員  | (人)            | 0.51           | 0.41           | 0.41             | —              | 0.29            | 0.40            |
| うち無職者人員  | (人)            | —              | 0.34           | —                | —              | 0.35            | —               |
| 有業人員     | (人)            | 0.47           | 0.56           | 0.54             | 0.56           | 0.48            | 0.47            |
| 世帯主の年齢   | (歳)            | 61.1           | 55.8           | 55.5             | 55.7           | —               | —               |
| 世帯主の性別 男 | (%)            | 43.5           | 46.8           | 47.5             | —              | 52.5            | 48.3            |
| 女        | (%)            | 56.5           | 53.2           | 52.5             | —              | 47.5            | 51.7            |
| 年間収入     | (千円)           | 2,629          | 3,274          | 3,100            | —              | —               | 2,679           |

パネルC：二人以上世帯(2009～2010)

|          | 国民生活<br>(2010) | 全国消費<br>(2009) | 家計調査<br>(2009) | 家計消費状況<br>(2009) | 国勢調査<br>(2010) | 労働力調査<br>(2009) | 国民生活Ⅱ<br>(2010) |
|----------|----------------|----------------|----------------|------------------|----------------|-----------------|-----------------|
| 集計世帯数    | 20,569         | 50,836         | 7,828          | 17,934           | —              | —               | 20,569          |
| 世帯人員     | (人)            | 3.11           | 3.10           | 3.11             | 3.12           | 3.12            | 3.11            |
| 18歳未満人員  | (人)            | 0.53           | 0.63           | —                | —              | 0.58            | —               |
| 65歳以上人員  | (人)            | 0.77           | 0.66           | —                | —              | 0.65            | —               |
| うち無職者人員  | (人)            | —              | 0.51           | —                | —              | —               | —               |
| 有業人員     | (人)            | 1.52           | 1.43           | 1.37             | 1.49           | 1.47            | 1.52            |
| 世帯主の年齢   | (歳)            | 58.5           | 55.4           | 55.8             | 58.8           | —               | —               |
| 世帯主の性別 男 | (%)            | 90.0           | 90.7           | —                | —              | 87.2            | —               |
| 女        | (%)            | 10.0           | 9.3            | —                | —              | 12.8            | —               |
| 年間収入     | (千円)           | 6,295          | 6,482          | 6,290            | —              | —               | 6,295           |

(注1) 国民生活、国民生活Ⅰの世帯主は各世帯が世帯主と回答した者、国民生活Ⅱの世帯主は最大所得者である。

ただし、ともに所得票のみを集計している。

(注2) 国民生活Ⅰはサンプルから無職学生を除いてある。

(出所) 国民生活基礎調査は個票より、全国消費実態調査、家計調査、家計消費状況調査、国勢調査、労働力調査は公表データより、筆者作成。

表3 世帯人員別世帯割合（総世帯）

| 世帯人員      | 国民生活(2010) | 全国消費(2009) | 家計調査(2009) | 国勢調査(2010) |
|-----------|------------|------------|------------|------------|
| 1人        | 21.8%      | 29.6%      | 29.5%      | 32.4%      |
| 2人        | 31.5%      | 28.3%      | 28.1%      | 27.3%      |
| 3人        | 21.1%      | 18.2%      | 18.3%      | 18.2%      |
| 4人        | 16.0%      | 15.6%      | 15.6%      | 14.4%      |
| 5人以上      | 9.6%       | 8.3%       | 8.5%       | 7.8%       |
| 平均世帯人員(人) | 2.65       | 2.48       | 2.49       | 2.42       |

### Ⅲ－１．世帯構成の比較

家計に関する各統計の世帯属性を比較したものが表2である。まず、総世帯（パネルA）を比較すると、国民生活基礎調査は他の統計に比べ、世帯人員が多く、高齢者の人数も多い。同調査の世帯人員は2.65人で、他の統計よりも0.15～0.2人多い。また世帯主の年齢は59.1歳で、他の統計より4歳高い。高齢なのは世帯主だけではなく、65歳以上人員を比較しても0.1人多くなっているため、世帯員にも高齢者は多いと言える。

一方、国民生活基礎調査以外の統計は比較的一致している。世帯人員は、2.45～2.5人である。また世帯主の年齢は家計消費状況調査が若干高いものの、それ以外は55歳代である。世帯人員の年齢構成については、18歳未満人員が0.4人、65歳以上人員が0.5～0.6人であり、ほぼ同質のサンプルとなっている。

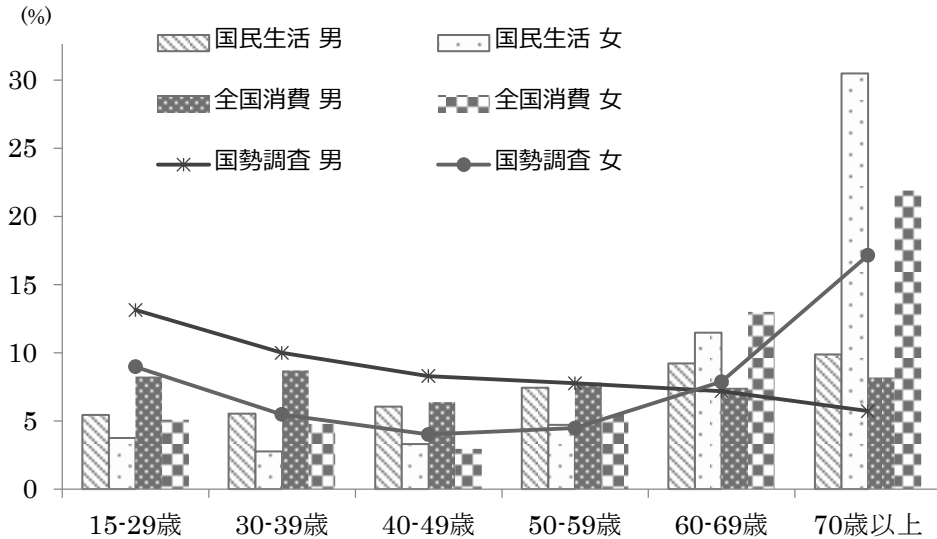
なぜ国民生活基礎調査には上記のような違いがあるのだろうか。総世帯の世帯人員別世帯割合を確認すると（表3）、国民生活基礎調査の単身世帯割合は21.8%であるのに対して、他統計では30%前後と、10%ポイントの差がある。この単身世帯の割合の低さが総世帯における世帯属性に影響しているのではないかと想定される。そこで、表2において単身世帯（パネルB）と二人以上世帯（パネルC）に分けて比較を行う。世帯人員について見ると、二人以上世帯では3.1人と等しい。単身世帯の世帯人員は定義上1人であるから、総世帯で国民生活基礎調査の世帯人員が多かったのは、単身世帯の割合が10%ポイント過小であったためである。

国民生活基礎調査で単身世帯の割合が低い理由を明らかにするため、単身世帯（パネルB）を比較すると、国民生活基礎調査は他の統計に比べ、高齢者と女性が多く、年間収入が低いことが見て取れる。世帯主年齢は61.1歳と他の統計より5歳高く、65歳以上の者が5割を占める。また女性の占める割合は56.5%となっている。2010年の国勢調査における女性の割合（47.5%）が真であると考えれば、かなり女性側に偏っていることになる。年間収入は約260万円であり、他の統計よりも50～60万円低い。

これらの違いは、単身世帯の定義の違い、サンプルの偏りから生じている可能性が考えられる。まず、単身世帯の定義の違いの可能性を検討してみよう。国民生活基礎調査と家計調査・全国消費実態調査には、前者には「学生」が含まれるが後二者には含まれず、前者には「住み込み又はまかない付きの寮・寄宿舎に住む単身世帯」は含まれないが後二者には含まれるという違いがある。そこで、国民生活基礎調査の個票を用い、学生を除いて世帯属性を再度、比較したものが表2の国民生活Ⅰである。学生を除いてもほとんど変わらないことが分かる。また全国消費実態調査で、単身世帯のうち寮・寄宿舎に住む世帯の割合を確認すると1%弱に過ぎない。したがって寮・寄宿舎に住む世帯を含むか含まないかは、結果にほとんど影響しないと考えられる。以上から、これらは定義の違いから生じたものではない。

次に、サンプルの偏りの可能性を検討してみよう。ここでは「悉皆」調査である国勢調査と比較を行う。一般に、悉皆調査である国勢調査

図1 単身世帯の年齢階層別・性別の分布



の方が、標本調査である国民生活基礎調査、全国消費実態調査よりも、真の値に近いと考えられる。単身世帯について年齢階層別・性別の分布を確認したものが図1である。国勢調査と比較すると、国民生活基礎調査、全国消費実態調査の双方で、50歳未満の占める割合が少なく、60歳以上の占める割合が多い。国勢調査との乖離幅に注目して、国民生活基礎調査と全国消費実態調査を比較すると、男女ともにほとんどの年齢階層で、国民生活基礎調査の方が全国消費実態調査よりも乖離幅は大きい。特に70歳以上の女性に注目すると、70歳以上の女性が占める割合は国勢調査では約17%なのに対して、国民生活基礎調査で約31%と、2倍近い開きがある。このように単身世帯のサンプルには明らかに偏りがあり、若年者が少なく、高齢者特に70歳以上の女性が多くなっている。一般に、若年単身世帯の多くは仕事についており、比較的時間に余裕のある高齢単身世帯よりも調査への協力は得られにくいと考えられる。また高齢単身世帯には女性が多い。そのため、国民生活基礎調査でも全国消費実態調査でも単身世帯の中での高齢女性の占める割合が高くなっているのだろう。

しかし、若年単身世帯の協力が得られにくいという事情はどの統計調査でも変わらないはずである。にもかかわらず、統計間で差異が生じているのは、標本抽出および集計方法の違いによると考えられる。多段階抽出法で世帯を抽出単位としている家計調査および全国消費実態調査では、代替標本の取得が可能で、しかも比推定を採用することで補正が可能であるのに対して、集落抽出法を採用する国民生活基礎調査では、これらの方法をとることができない。そのため、若年単身世帯で調査拒否が発生すると国民生活基礎調査では単身世帯が総世帯に占める割合が低くなり、単身世帯の中でも高齢女性への偏りが見られると考えられる。

### Ⅲ－2. 二人以上世帯の比較

これまで見た通り、国民生活基礎調査とその他の統計では標本抽出方法の違いにより、単身世帯および単身世帯を含む総世帯では、隔たりが大きいことが分かった。そこでここからは二人以上世帯に絞って比較を行う。二人以上世帯で比較すると、世帯人員を始めとして概ね一致しているが、65歳以上人員、世帯主年齢、有業人員に若干の差異が残っている。国民生活基

礎調査の65歳以上人員は0.77人で、他の統計より0.1人多い。世帯主年齢については国民生活基礎調査と家計消費状況調査では58歳代となっており、他の統計より3歳高くなっている。有業人員は家計調査で1.37人と他の統計より0.1人過小である。

世帯主年齢の違いは世帯主の定義に起因する可能性がある。全国消費実態調査・家計調査・家計消費状況調査では世帯主を最大所得者と定義しているが、国民生活基礎調査では各世帯が世帯主と回答した者が世帯主である（舟岡2001）。そこで国民生活基礎調査が世帯員個別の年間収入を調査しており、個票では最大所得者に定義しなおすことが可能である点を利用して、世帯主を最大所得者として再計算したのが表2における国民生活Ⅱである。世帯主年齢は55.2歳となり、他の統計と同水準となった。以上から国民生活基礎調査の世帯主が他の統計に比べ高齢であったのは、世帯主の定義の違いによるものであると言える。ただ世帯主の定義を変えることで世帯主年齢は一致したが、代わりに男性世帯主の割合が81.3%と10%近く下がってしまった。その理由は今後の検討課題である。

このように、65歳以上人員、有業人員に若干の違いはあるものの、二人以上世帯に限定し定義の違いを調整すれば、平均値では概ね一致していると言える。しかしマイクロデータを用いて分析する意義があるのは、平均だけではなく分布の情報が得られるからである。そこで、ここからは各統計の世帯属性の分布を見る。国民生活基礎調査、全国消費実態調査、家計調査、国勢調査の世帯属性について分布を確認したものが表4である。なお、国民生活基礎調査の世帯主は、最大所得者に直してある。

世帯属性のうち、世帯主年齢、世帯人員、世帯類型の分布については国勢調査と比較が可能である。そこでまず、国民生活基礎調査、全国消費実態調査、家計調査と国勢調査を比較する。世帯主年齢や世帯人員では、国勢調査と大きな違いはない。全国消費実態調査、家計調査

で大きな違いが無いのは、ある意味当然である。なぜならば、これらの統計では集計の際に比推定を採用しており、地方別、世帯人員別の割合が調整されているからである。一方、国民生活基礎調査ではそのような調整はなされていないが、概ね一致している。このことから二人以上世帯に限って言えば、比推定の有無により大きな差は生じないと言える。

ただし比推定で調整していない部分には違いがある。世帯類型を比較すると、どの統計でも夫婦のみの世帯、夫婦と未婚の子供のいる世帯の和で大部分を占めるが、国民生活基礎調査では両者の和は70%強であり国勢調査にはほぼ一致するのに対して、全国消費実態調査、家計調査では80%弱と10%近く高くなっている。国民生活基礎調査も国勢調査と完全に一致している訳ではなく、片親と未婚の子のいる世帯の占める割合は（全国消費実態調査、家計調査と同様に）低く、逆に三世帯世帯の占める割合は高い。

続いて国勢調査とは比較できない、有業人員、従業上の地位、年間収入について比較する。有業人員では、国民生活基礎調査は有業人員3人以上の世帯の割合が高いのに対し、全国消費実態調査、家計調査は有業人員1人の世帯の割合が高い。従業上の地位では、国民生活基礎調査は他の2統計と比べて自営業者の占める割合が高くなっている。年間収入を見ると、年間収入300万円未満の世帯の割合は、国民生活基礎調査では20%強であるのに対して、全国消費実態調査、家計調査では12~13%であり、国民生活基礎調査の方が年間収入の低い者の占める割合が高い。ただし、国民生活基礎調査のサンプルが年間収入の低い者に偏っているかという点を決してそうではなく、年間収入1500万円以上の世帯の割合は4.2%で、1%程度ではあるが全国消費実態調査、家計調査よりも高い。これは、山口（2014）の、全国消費実態調査は国民生活基礎調査と比較して、所得水準分布の中央が厚く両端が小さくなる特徴があるという指摘と整合的であり、所得だけではなくそ

表4 二人以上世帯の分布

単位：％

| 項目     |               | 国民生活 | 全国消費 | 家計調査 | 国勢調査 |
|--------|---------------|------|------|------|------|
| 世帯主年齢  | 29歳以下         | 3.7  | 2.4  | 2.1  | 3.7  |
|        | 30-39歳        | 15.6 | 15.0 | 14.5 | 14.9 |
|        | 40-49歳        | 19.3 | 18.7 | 19.3 | 17.8 |
|        | 50-59歳        | 19.5 | 21.1 | 20.1 | 19.1 |
|        | 60-69歳        | 21.7 | 24.5 | 24.2 | 22.7 |
|        | 70歳以上         | 20.2 | 18.4 | 19.8 | 21.7 |
| 世帯人員別  | 2人            | 40.3 | 40.3 | 39.9 | 40.3 |
|        | 3人            | 27.0 | 25.8 | 25.9 | 26.9 |
|        | 4人            | 20.4 | 22.1 | 22.2 | 21.3 |
|        | 5人以上          | 12.3 | 11.8 | 12.0 | 11.5 |
| 有業人員別  | 0人            | 17.2 | 18.3 | 19.4 | －    |
|        | 1人            | 32.7 | 34.7 | 37.0 | －    |
|        | 2人            | 35.2 | 35.7 | 33.5 | －    |
|        | 3人以上          | 15.0 | 11.3 | 10.1 | －    |
| 従業上の地位 | 雇用者世帯の比率      | 53.6 | 56.7 | 53.9 | －    |
|        | 自営業者世帯の比率     | 21.6 | 17.7 | 18.3 | －    |
|        | 無業世帯の比率       | 24.8 | 25.6 | 27.8 | －    |
| 世帯類型   | 夫婦のみの世帯       | 32.0 | 34.2 | 34.0 | 29.3 |
|        | 夫婦と未婚の子供のいる世帯 | 39.3 | 44.2 | 45.9 | 41.3 |
|        | 片親と未婚の子のいる世帯  | 8.0  | 7.0  | 5.7  | 12.9 |
|        | 三世代世帯         | 11.2 | 6.8  | －    | 7.0  |
|        | その他・不詳        | 9.5  | 7.8  | －    | 9.5  |
| 年間収入   | 199万円以下       | 9.5  | 4.4  | 2.2  | －    |
|        | 200-299万円     | 11.6 | 8.4  | 10.1 | －    |
|        | 300-399万円     | 13.8 | 15.1 | 17.3 | －    |
|        | 400-499万円     | 12.6 | 14.5 | 15.3 | －    |
|        | 500-599万円     | 10.9 | 12.4 | 12.7 | －    |
|        | 600-799万円     | 16.1 | 19.3 | 18.6 | －    |
|        | 800-999万円     | 10.6 | 11.8 | 10.7 | －    |
|        | 1000-1499万円   | 10.7 | 10.8 | 10.1 | －    |
|        | 1500万円以上      | 4.2  | 3.4  | 3.0  | －    |

れ以外の属性についても、国民生活基礎調査は全国消費実態調査、家計調査より平均的ではない家計を多く捉えている可能性がある。

国民生活基礎調査の分布が、全国消費実態調査、家計調査の分布と異なる要因として考えられるのが、調査方法の違いである。前述の通り、国民生活基礎調査は調査対象地区のすべての世帯を調査する「集落抽出」の手法を採用しているのに対し、全国消費実態調査、家計調査は、抽出された調査地区の中からさらに調査世帯を抽出して調査している。理論的な性質としてはほぼ同等のはずであるが、調査拒否の実務上の対応の違いにより差が生じている可能性がある。

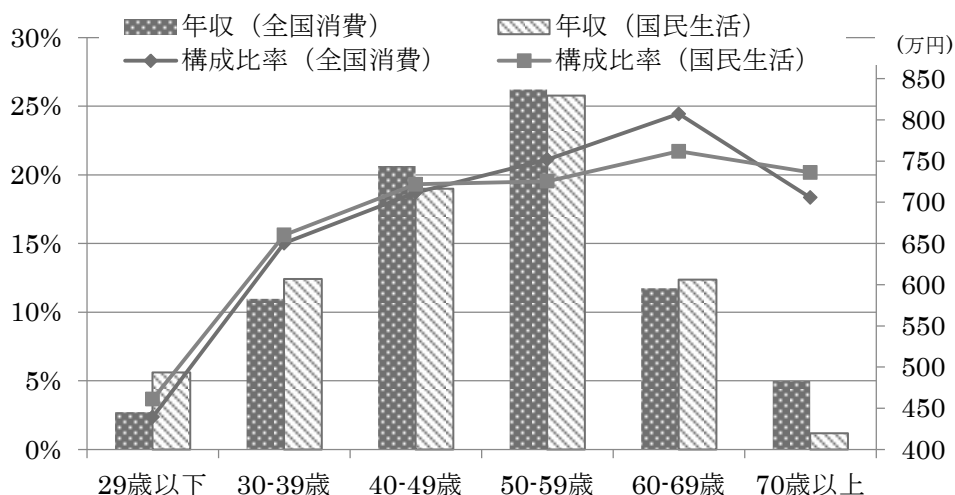
もう一点考えられるのが、調査負担の違いである。国民生活基礎調査はあらかじめ記入欄が設けられた「ブリコード方式」を採用してお

り、調査時に1回記入すればよいが、全国消費実態調査、家計調査はそれぞれ3ヶ月間、6ヶ月間継続して家計簿をつけることが要求されるため、調査客体の負担が重く、調査拒否につながりやすい。このような調査方法の違い、調査負担の違いから、分布の差が生じていると考えられる。

### Ⅲ－３．年収の比較

三統計で比較可能な経済変数である、世帯の年間収入について比較したものが表2の各パネルの一番下の行である。すでに見たように、単身世帯には、大きなサンプルの違いがあるため、ここではパネルCに注目する。二人以上世帯では家計調査と国民生活基礎調査は約630万円とほぼ一致するが、全国消費実態調査では約650万円で20万円程度高くなっている。これら

図2 二人以上世帯の各年齢階層の年収及び構成比率（全国消費vs 国民生活）



は3%程度の差であり、決して大きな差ではないが、ここでは簡単にこの差が生じている要因について考察する。

国民生活基礎調査と全国消費実態調査では、前述の通り、世帯主年齢階層で比較すると、前者に若年者や高齢者が多く、後者に「50-69歳」の層が多かった。日本では年齢による年収差が大きく、一般に50代前後で年収がピークを迎えることを考えると、世帯主年齢階層における構成割合は世帯年収にも影響を与えと考えられる。そこで年齢階層別の構成割合を左軸に、年齢階層別の平均年収を右軸に表したものが図2である。国民生活基礎調査では年収の高い「50-69歳」の割合が全国消費実態調査より低く、年収の低い「70歳以上」の割合が全国消費実態調査より高い。そこで両者の構成割合を同じにすると、全体の差の20万円のうち、10万円程度が説明できる。

国民生活基礎調査と全国消費実態調査の年収差の残り10万円は、70歳以上の平均年収の差より生じている。国民生活基礎調査の世帯主年齢70歳以上の平均世帯年収は約420万円で、全国消費実態調査の平均世帯年収約480万円より、60万円低い。そこで世帯主年齢が70歳以上の世帯について、表5で調査項目別にどの収入で差がついているのを確認すると、公的年金・恩

給、企業年金・個人年金等、財産所得で、差の約77%が説明できる。

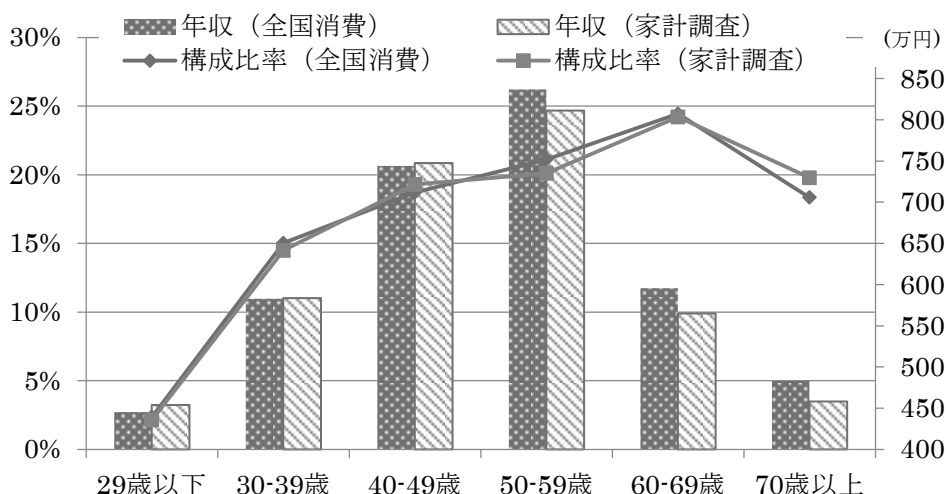
このうち財産所得の差約20%については両統計の調査方法の違いから説明可能である。宇南山（2011）は家計消費状況調査と家計調査を比較することで「プリコード方式」の方が「アフターコード方式」よりも特定品目の記入漏れが少ない可能性をしているが、同様のことが所得についても言える。例えば、国民生活基礎調査における「財産所得」は全国消費実態調査の「家賃・地代の年間収入」と「利子・配当金」の合計とほぼ同じになると考えられるが、全国消費実態調査は国民生活基礎調査よりも13万円多くなっている。つまり単に「財産所得」を聞いている国民生活基礎調査よりも、「家賃・地代の年間収入」「利子・配当金」を別々に確認している全国消費実態調査の方が、財産所得についてはよりよく反映している可能性がある。

ところで家計調査と全国消費実態調査の間にも20万円の差が存在していた。世帯主年齢階層における構成割合に若干の差異はあるものの、大きな差はない。そうであるにも関わらず、なぜ世帯年収で約20万円の差が生じるのかが問題になる。上記と同様に、家計調査と全国消費実態調査の年齢階層別の構成割合を左軸

表5 世帯主年齢70歳以上の世帯の収入の比較（調査項目別）

| 全国消費(2009)   |       | 国民生活(2010)  |       | 単位:万円 |
|--------------|-------|-------------|-------|-------|
|              |       |             |       | 差     |
| 勤め先からの年間収入   | 78.9  | 雇用者所得       | 75.2  | 3.7   |
| 農林漁業収入       | 5.5   | 農耕・畜産所得     | 3.4   | 2.1   |
| 農林漁業以外の事業収入  | 17.3  | 事業所得        | 14.4  | 2.9   |
| 内職などの年間収入    | 2.1   | 家内労働所得      | 0.8   | 1.3   |
| 公的年金・恩給      | 281.1 | 公的年金・恩給     | 269.2 | 11.9  |
| 家賃・地代の年間収入   | 38.7  | 財産所得        | 32.7  | 13.0  |
| 利子・配当金       | 7.0   |             |       |       |
| 企業年金・個人年金受取金 | 37.1  | 企業年金・個人年金等  | 15.2  | 21.8  |
| 親族などからの仕送り金  | 2.6   | 仕送り         | 0.5   | 2.1   |
| その他の年間収入     | 7.1   | 雇用保険        | 0.8   | -1.0  |
|              |       | 児童手当等       | 0.2   |       |
|              |       | その他の社会保障給付金 | 1.5   |       |
|              |       | その他の所得      | 5.6   |       |
| 現物消費の年間見積り額  | 3.3   | (該当するものなし)  |       | 3.3   |
| 合計           | 480.7 | 合計          | 419.6 | 61.1  |

図3 二人以上世帯の各年齢階層の年収及び構成比率（全国消費vs家計調査）



に、年齢階層別の平均年収を右軸に表したものが図3である。これをみると、家計調査と全国消費実態調査では構成割合にはほとんど差がなく、また世帯主年齢49歳以下の世帯では平均世帯年収にも大きな差はないが、平均年齢50歳以上の層ではどの階層でも約30万円、全国消費実態調査の方が家計調査よりも高い。このように差が生じる一因としては、家計調査と全国消費実態調査の調査票の違いが考えられる。収入主体別には、家計調査では「世帯主」「他

の世帯員」の収入を確認しているのに対し、全国消費実態調査では「世帯主」「世帯主の配偶者」「他の世帯員（65歳以上）」「他の世帯員（65歳未満）」の収入を確認している。また収入の種類別には、家計調査が8種類の収入を確認しているのに対し、全国消費実態調査では11種類の収入を確認しており、後方で確認している「家賃・地代の年間収入」「企業年金・個人年金受取金」「利子・配当金」「親族などからの仕送り金」は、前者の「その他の年間収入」に含ま

れると考えられる。このような調査票の違いが、勤め先からの年間収入以外の収入が増え始

める50歳以上の階層での収入の差を生んでいると推測される。

## Ⅳ．学歴の比較

### Ⅳ－１．世帯調査における「学歴」の利用可能性

家計に関する分析で、教育に関する分析が近年注目されている。教育水準は、家計の労働者としての質を決定する。また、教育水準は、社会階層の決定要因としても重要であり、消費や余暇にも大きな影響を与える。しかし、各家計の教育水準の指標としての学歴について、日本の統計はまだ十分に整備されていない。そこで、ここでは現在利用可能な学歴に関する情報をまとめることで、教育水準に関する調査の課題を明らかにする。

わが国の公的統計から学歴情報を得る方法は3つある。1つ目は国勢調査をはじめとする家計に対する調査で、調査対象者および対象者の世帯人員の在学状況や最終学歴を直接尋ねる方法である。2つ目は賃金構造基本調査などの事業所に対する調査で、調査対象事業所側が抽出した労働者の最終学歴を尋ねる方法、あるいは対象事業所の新規採用者の学歴別人数を尋ねる方法である。3つ目は学校基本調査などの学校に対する調査で、在学者および卒業者の人数を尋ねる方法である。

それぞれの方法には利点および欠点がある。家計への調査の利点は、調査対象者本人あるいはその世帯構成人に直接尋ねている点で最も簡便である。加えて、調査時点の状況を把握でき、個人を集計することで地域別、マクロの集計が可能である。他の調査項目と組み合わせることで、教育のリターンの推計といった分析が可能となる（佐野・安井2009）。一方で、調査対象者が正しく回答しないことや中退などの扱いによる測定誤差や、系統的な属性を持つ個人が学歴情報に回答しない非標本誤差の問題があ

りえる。

事業所への調査の利点は、個人の学歴を尋ねている意味では家計側と同様である。加えて、その調査の目的から雇用・賃金情報が豊富であるため、ミンサー型賃金関数の推計などの利用に適している（川口2011）。一方、調査時点で雇用されている労働者を対象としており、調査対象の労働者の抽出は事業所の賃金台帳を基に実施されるため、ランダムサンプリングの仮定が満たされているとは限らない。加えて、事業所情報は豊富であるが、対象労働者の家族構成といった家計の情報を得ることが出来ない。

学校への調査の利点は、悉皆調査であることから在学者数・卒業者数をほぼ正確に把握でき、地域・マクロ集計も容易な点である。ただし、年齢別の教育水準を集計することは必ずしも容易ではない。集計入学・卒業年と年齢が対応している義務教育段階の人数は容易に捕捉できるが、入学・卒業年が必ずしも一様ではない高等教育段階の年齢別の人数の捕捉は困難だからだ。また、そもそも個人に対する調査ではないため個人・家計に関する情報を得ることができない。

学歴情報を把握し利用する際、どの調査方法が適しているかは、その分析対象に依存する。たとえば、賃金に関する詳細な情報と教育達成の関係を計測する賃金関数の推計は事業所調査が優れているが、就業選択行動を考慮した推計を実施する際には世帯属性が重要となるため家計を対象とした調査が優れている。一国全体の教育水準を計測する場合、家計による調査をマクロレベルで集計することで対応できるが（Barro and Lee 2000）、測定誤差の問題を考慮

すれば過去からの学校調査に関する情報を積み上げることで対応できる。(Godo 2011)

本節では、家計側の調査である2010年の国民生活基礎調査、国勢調査および労働力調査(年平均)の公表データから得られる学歴情報の比較を行う<sup>9)</sup>。比較に際しては、以下の2点に注意が必要である。1点目は、調査間における定義の差異である。各調査は主問で在学の有無を尋ね、枝間で学歴区分を尋ねている。学歴区分はそれぞれ、国民生活基礎調査は小学・中学、高校・旧制中、専門学校、短大・高専、大学、大学院の6区分、国勢調査は小学・中学、高校・旧制中、短大・高専、大学・大学院の4区分、労働力調査は小中高、短大高専、大学大学院の3区分であるが、専門学校をどこに含めるかは調査により異なる。国民生活基礎調査は調査票に専門学校の項目を準備しており、回答者自身の判断により専門学校と記入するが、国勢調査と労働力調査は入学資格と修業年限により区分される<sup>10)</sup>。専門学校の定義の違いにより、各調査に差異が生じる可能性がある。

2点目は、「不詳」の問題である。前述したように、家計の調査は特定の属性を持つ個人が調査あるいは特定の調査項目に回答しない可能性がある。性別、世帯人員は他の情報により把握可能だとしても、婚姻状況、就業状況や学歴などは調査されない限り把握できない。さらに、調査に回答したとしても、特定の設問に非回答である場合もある。これらはすべて「不詳」として扱われる。学歴情報は個人にとってセンシティブな情報であるため、調査自体に回答するものの学歴の設問に非回答である「学歴不詳者」が無視できないほど存在する可能性があり、そのことが統計間の比較を困難とする可

能性がある。

## IV-2. 各調査の学歴の構成

家計に対する各調査から得ることが出来る学歴情報の特徴を検討しよう。表6はそれぞれ男女別の国民生活基礎調査(パネルA)、国勢調査<sup>11)</sup>(パネルB)、労働力調査(パネルC)の学歴分布を比較したものである。なお、各統計の学歴区分と合わせ<sup>12)</sup>、最終学歴が大きく変化しないと考えられる25歳以上に限定し、集計を行った<sup>13)</sup>。

各パネルの列(1)によると、総数では各調査とも概ね近い値であるが、国民生活基礎調査における卒業者数は男女とも約500万人過小となっている。これは国民生活基礎調査で学歴の設問に回答していないという意味での「学歴不詳」者が約10%と多いことに起因すると考えられる。

各統計における最終学歴の分布を示したものが、列(2)である。国民生活基礎調査と国勢調査を比べてみよう。総数に占めるシェアは統計間で近い傾向であるが、人数は統計間で乖離している。小学・中学卒業者数は、国民生活基礎調査は国勢調査と比べ男女とも約50万人過小である。国民生活基礎調査の高卒は男性で約100万、女性で約130万人国勢調査より少ない。短大・高専グループは、専門学校の定義の違いもあり、国民生活基礎調査は国勢調査と比べ男性で約80万人、女性で約160万人過大である。男女の大学・大学院卒の人数は両統計でほぼ一致する。ただし、「学歴不詳者」の割合が国民生活基礎調査で約10%、国勢調査で約13%と共に無視できないほど大きく、彼らがどの学歴グループにカウントされるかで結果は変わりうる

9) 家計に対する構造調査である『就業構造基本調査』もその比較候補として考えられるが、調査実施年が2007年と2012年であるため、利用を断念した。

10) たとえば、高卒であることを入学要件とし、修業年限が4年以上であれば大卒・大学院卒に、高卒が入学資格だが修業年限が2年以上4年未満なら短大・高専に区分される。

11) 日本人及び外国人を含めた集計である。日本人に限定しても結論は変わらない。

12) 国民生活基礎調査の専門学校は短大・高専に含めている。

13) 前節で見たように単身世帯を含めるかどうかは比較の上で重要となるが、本節では世帯人数で比較している。

表6 学歴分布の比較（上段：男性、下段：女性）

| パネルA<br>国民生活基礎調査 |            |            |            | パネルB<br>国勢調査 |            |            | パネルC<br>労働力調査 |            |
|------------------|------------|------------|------------|--------------|------------|------------|---------------|------------|
| 男性               | (1)        | (2)        | (3)        | (1)          | (2)        | (3)        | (1)           | (2)        |
| 総数               | 45,785,000 |            |            | 46,779,145   |            |            | 45,730,000    |            |
| 卒業生数             | 40,912,000 |            |            | 46,571,165   |            |            | 45,500,000    |            |
| シェア              | 89.4       |            |            | 99.6         |            |            | 99.5          |            |
| 小・中学校            |            | 6,888,000  | 23,936,000 |              | 7,170,602  | 25,096,747 |               | 28,120,000 |
| シェア              |            | 15.0       | 52.3       |              | 15.3       | 53.6       |               | 61.5       |
| 卒業生シェア           |            | 16.8       | 58.5       |              | 15.4       | 53.9       |               | 61.8       |
| 高校               |            | 17,048,000 |            |              | 17,926,145 |            |               |            |
| シェア              |            | 37.2       |            |              | 38.3       |            |               |            |
| 卒業生シェア           |            | 41.7       |            |              | 38.5       |            |               |            |
| 短大・高専            |            | 4,188,000  |            |              | 3,300,127  |            |               | 4,070,000  |
| シェア              |            | 9.1        |            |              | 7.1        |            |               | 8.9        |
| 卒業生シェア           |            | 10.2       |            |              | 7.1        |            |               | 8.9        |
| 大学・大学院           |            | 11,965,000 |            |              | 11,759,351 |            |               | 13,310,000 |
| シェア              |            | 26.1       |            |              | 25.1       |            |               | 29.1       |
| 卒業生シェア           |            | 29.2       |            |              | 25.3       |            |               | 29.3       |
| 卒業学校不詳           |            | 823,000    |            |              | 6,414,940  |            |               |            |
| シェア              |            | 1.8        |            |              | 13.7       |            |               |            |
| 卒業生シェア           |            | 2.0        |            |              | 13.8       |            |               |            |
| 在学者数             | 205,000    |            |            | 159,324      |            |            | 160,000       |            |
| シェア              | 0.4        |            |            | 0.3          |            |            | 0.3           |            |
| 未就学者数            | 43,000     |            |            | 43,676       |            |            | 70,000        |            |
| シェア              | 0.1        |            |            | 0.1          |            |            | 0.2           |            |
| 不詳               | 4,625,000  |            |            | 4,980        |            |            |               |            |
| シェア              | 10.1       |            |            | 0.0          |            |            |               |            |
| パネルA<br>国民生活基礎調査 |            |            |            | パネルB<br>国勢調査 |            |            | パネルC<br>労働力調査 |            |
| 女性               | (1)        | (2)        | (3)        | (1)          | (2)        | (3)        | (1)           | (2)        |
| 総数               | 50,824,000 |            |            | 51,008,550   |            |            | 49,640,000    |            |
| 卒業生数             | 45,139,000 |            |            | 50,800,343   |            |            | 49,510,000    |            |
| シェア              | 88.8       |            |            | 99.6         |            |            | 99.7          |            |
| 小・中学校            |            | 8,680,000  | 28,382,000 |              | 9,122,494  | 30,221,221 |               | 33,480,000 |
| シェア              |            | 17.1       | 55.8       |              | 17.9       | 59.2       |               | 67.4       |
| 卒業生シェア           |            | 19.2       | 62.9       |              | 18.0       | 59.5       |               | 67.6       |
| 高校               |            | 19,702,000 |            |              | 21,098,727 |            |               |            |
| シェア              |            | 38.8       |            |              | 41.4       |            |               |            |
| 卒業生シェア           |            | 43.6       |            |              | 41.5       |            |               |            |
| 短大・高専            |            | 10,697,000 |            |              | 9,001,378  |            |               | 10,370,000 |
| シェア              |            | 21.0       |            |              | 17.6       |            |               | 20.9       |
| 卒業生シェア           |            | 23.7       |            |              | 17.7       |            |               | 20.9       |
| 大学・大学院           |            | 5,103,000  |            |              | 5,154,984  |            |               | 5,660,000  |
| シェア              |            | 10.0       |            |              | 10.1       |            |               | 11.4       |
| 卒業生シェア           |            | 11.3       |            |              | 10.1       |            |               | 11.4       |
| 卒業学校不詳           |            | 957,000    |            |              | 6,422,760  |            |               |            |
| シェア              |            | 1.9        |            |              | 12.6       |            |               |            |
| 卒業生シェア           |            | 2.1        |            |              | 12.6       |            |               |            |
| 在学者数             | 165,000    |            |            | 129,419      |            |            | 50,000        |            |
| シェア              | 0.3        |            |            | 0.3          |            |            | 0.1           |            |
| 未就学者数            | 96,000     |            |            | 74,350       |            |            | 80,000        |            |
| シェア              | 0.2        |            |            | 0.1          |            |            | 0.2           |            |
| 不詳               | 5,424,000  |            |            | 4,438        |            |            |               |            |
| シェア              | 10.7       |            |            | 0.0          |            |            |               |            |

注：25歳以上で集計している。年齢不詳は除外している。国民生活基礎調査の専門学校は短大・高専に含んでいる。公表データでは労働力調査の「不詳」者数は集計されていないが、個票データによると、約2%である。

ため、統計間の比較を困難にしている。

国民生活基礎調査と国勢調査の学歴区分を労働力調査に合わせた列（3）で比較してみよう。労働力調査の学歴不詳者は公表データに表章されていないものの、個票データから計算した不詳者は約2%であり、学歴不詳の問題は深刻ではない可能性が高い。シェアと人数の両方から、国民生活基礎調査と国勢調査は労働力

調査と比べ、国民生活基礎調査の短大・専門を除き、どの学歴グループにおいても過小である。とりわけ、男女の小中高の人数およびシェアの乖離は大きく、約300万から500万人の乖離がある。ただし、国民生活基礎調査の短大・専門には専門学校が含まれているため人数が多く見えるが、専門学校卒人数は男性で約300万人、女性で460万人であることを考えると、短

大・高専に絞れば、労働力調査と比べ過小であると言える。すなわち、専門学校の定義の差を勘案したとしても、統計間における各最終学歴者の人数の差は残るが、その差は「学歴不詳者」が原因であると考えられる。

#### Ⅳ－３．学歴不詳問題

家計への調査はその調査の性質上、調査に回答しないという問題を内包している。加えて、調査には回答するが特定の項目に回答しない場合もある。それらはすべて「不詳」と分類されるが、特定の属性を持つ個人が「不詳」者になるとすれば、彼らを集計に含めるか否かで結果にバイアスをもたらす。他の統計調査の基礎となる国勢調査でさえも、「不詳」の問題がしばしば指摘されている（宇南山2013、山田2012、阿部2013、小池・山内2014）。国勢調査における「不詳」の問題は、他の方法で把握可能である性別と世帯人員を除く、年齢、婚姻、労働力状態、現在・過去の居住状況および学歴で生じやすい。とりわけ、学歴については深刻であり、表6によると人口に占める卒業学校不詳者のシェアは約13%であり、前回大調査（2000年）時点での約3.5%から大幅に増加している。国民生活基礎調査においても同様の問題は内包しており、表6によると、25歳以上に占める学歴不詳者の割合は男性で10.1%、女性で10.7%と無視できないほど大きい。

「不詳」をどのように処理するかはその構造をどのように仮定するかに依存する。不詳をその集計から除く場合は、不詳群が不詳以外のものと平均的には同じ性質を持つことを仮定することとなるが、その仮定が正しいとは限らない。たとえば、都市部のオートロックマンションに単身で居住者している個人を捕捉できないことが原因で不詳となる場合、彼らと不詳以外のものには系統的な差が発生する。学歴不詳者の属性は先験的に不明だが、どのような学歴グループが不詳者となりやすいかを他の統計を組

み合わせることで類推可能である。

どのような学歴グループが学歴不詳になりやすいかを2つの異なる方法を用いて推計しよう。1つの方法は前回大調査である2000年の国勢調査よりコホート情報による推計である。2000年の国勢調査での卒業者に占める卒業学校不詳者の割合は約3-4%であり、不詳の問題は深刻ではない。加えて、2時点の情報を用いることで、同一コホートの10年間の推移を検討することが出来る。別の方法は、『労働力調査』から得られる学歴分布を「真の学歴分布」と見なし、その情報から予測される学歴分布から推計する方法である。労働力調査は国勢調査における性別、年齢構成、地域を考慮し、その情報を復元できるようなランダムサンプリングを行っている。性別、年齢構成、地域に基づくランダムサンプリングから得られる学歴分布が、国勢調査における真の学歴分布を代表しているのならば、労働力調査の学歴分布の情報を用い他の統計調査の学歴分布を復元できる。いずれの方法も完全ではないが、異なる方法から得られた結果を比較することで、学歴不詳者の特徴を確認することが出来る。

コホート情報を用いた方法について説明しよう。たとえば、1970年生まれの高卒者は2000年に30歳であるが、10年後の40歳でも高卒である可能性が高い<sup>14)</sup>。2000年から2010年にかけての人口変化から予測できる人数と実際の人数を比較し、1970年生まれの高卒者が大幅に減少していれば、彼らが2010年時点で不詳とカウントされている可能性が高い。

いま調査年 $t$ 、年齢コホート $c$ 、人口 $n$ 、ある学歴グループの卒業生数を $e$ とする。2000年のコホート $c$ の人口を $n_c^{00}$ 、2010年のそれを $n_c^{10}$ とすると、コホート変化率は $r_c = n_c^{10}/n_c^{00}$ である。各学歴グループでコホート変化率が一定であると仮定すると、ある学歴グループの2000年から予測される2010年の卒業生数は $r_c \cdot e_c^{00}$ と推計でき、実際の卒業生数との乖離が

14) もちろん、この期間に大学等に進学することで、学歴自体が変化する可能性はある。

グループごとに計算できる。グループの乖離の和と2010年における不詳者から2000年から予測される不詳者の差は一致する。このような計算を、生年コホートごとに、性別、学歴別に行い、集計を行う。各学歴グループの予測値と実測値の乖離を人口で割ることで計算した値は、どの学歴グループが不詳になりやすいかを示す指標であると解釈できる。なお、推計値自体は正負いずれもとりのうが、正であれば予測値に対して実測値は過小であり、負であれば過大である。推計においては、最終学歴が大きく変化しないと考えられる、2000年に25-29歳で2010年に35-39歳となる1971-75年生まれコホートから、集計可能な2000年に70-74歳、2010年に80-84歳となる1926-30年生まれコホートまでに限定する。

労働力調査を用いた方法について説明しよう。労働力調査から得られる学歴分布が「真の分布」とであると仮定し、そのもとで国民生活基礎調査と国勢調査の総数から予測される分布と実際の分布を比較することで、不詳者がどのグループに存在するか検討する。

予測の方法は以下の通りである。労働力調査の総数を $E^{LF}$ とし、教育グループを $i$ 、年齢グループを $j$ とする。各グループのウェイトを $w_{ij}$ 、各グループの人数を $e_{ij}^{LF}$ とすると、

$$E^{LF} = \sum_j \sum_i w_{ij} e_{ij}^{LF}$$

となる。学歴区分<sup>15)</sup>と年齢区分を労働力調査のそれに合わせた上で、国民生活基礎調査と国勢調査をそれぞれ $S$ とし、各統計の総数 $E^S$ 、各グループの人数を $e_{ij}^S$ とすると、労働力調査の学歴ウェイトから予測される総数は

$$\widehat{E^S} = \sum_j \sum_i w_{ij} e_{ij}^S$$

と計算できる。予測値と実際の乖離 $\widehat{E^S} - E^S$

は学歴不詳者数であるが、各年齢グループで集計し、人口で割ることにより、どの学歴グループが不詳になりやすいかを示す指標を推計する。推計に際しては、コホートの時と比較するために、2010年時点で35歳以上に限定する<sup>16)</sup>。

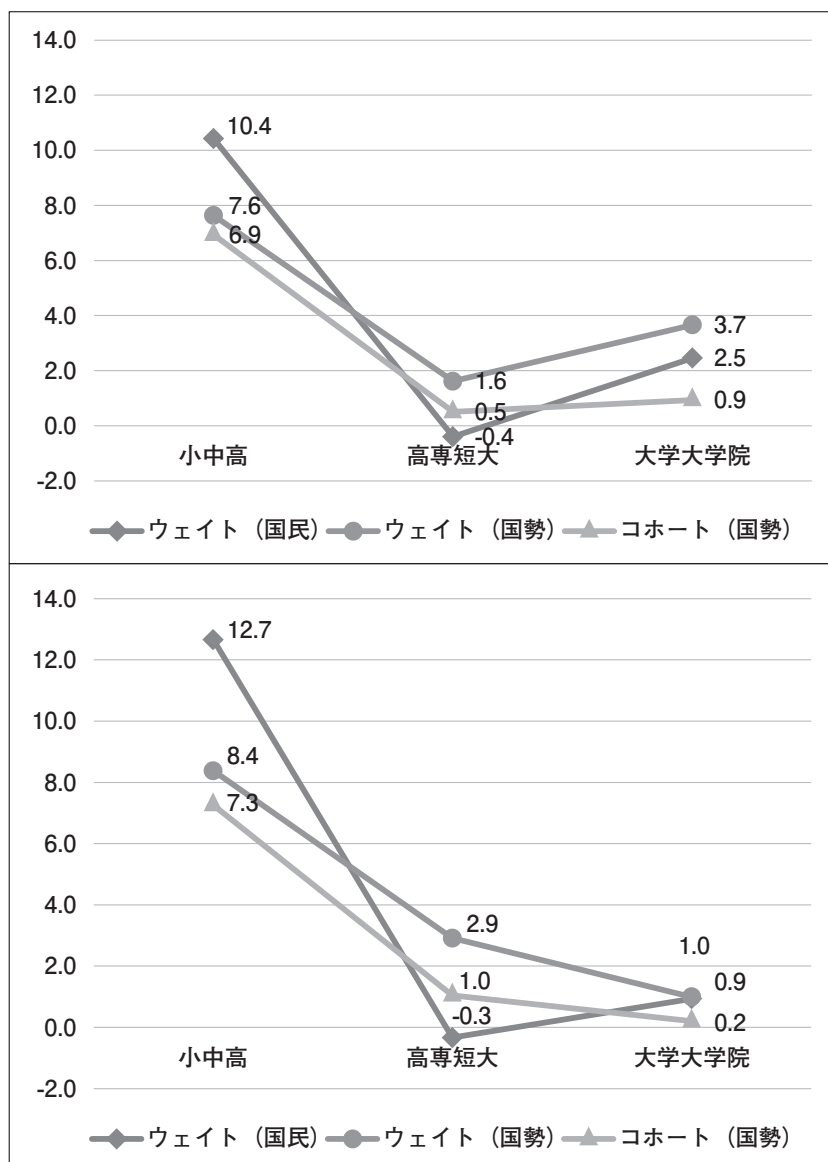
図4は国勢調査のコホート情報、労働力調査のウェイトを用い、総数に占める各学歴グループの予測値と実測値の誤差の割合（以下、不詳確率）を示したものである。上段の男性の結果によると、いずれの方法を用いたとしても、小中高のグループでの不詳確率が際だって高いことが確認できる。また、コホートによる推計と労働力調査のウェイトを用いた推計を比較すると、小中高卒の不詳確率は7.6と6.9%と共に高く、コホートでは高専短大グループと大卒大学院卒グループの割合は近いが、ウェイトを用いた推計では1.6と3.7と大卒の不詳割合が高い。ウェイトを用いた推計では、国民生活基礎調査と国勢調査で、不詳確率は小中高、大学大学院、高専短大の順となる傾向は同じである。なお、高専短大卒では、誤差が負の値であるが、これは国民生活基礎調査の同グループは過大であるためだと考えられる。

図4下段の女性の結果によると、それぞれの方法を用いたとしても、小中高のグループでの不詳確率が際だって高いことは男性の結果と同様である。ただし、高専短大と大学大学院グループに関しては、用いる統計による差異が確認される。国民生活基礎調査を用いると、大卒者が不詳者になりやすいが、国勢調査を用いると、高専短大グループが不詳者なりやすい。両統計に見られる差異は、シェアの高い高専・短大グループに含まれる専門学校の定義の違いである可能性がある。労働力調査の専門学校の定義は国勢調査と同じであるが、国民生活基礎調査と異なるため、両者の差異がこのグループに現れていると考えられる。

15) 高専・短大・専門学校の区分が必ずしも統計調査同士で共通していないため、このグループの乖離は大きくなる可能性が高い。

16) 25歳以上にしても傾向は変わらない。

図4 学歴別総人数に占める不詳発生割合（上段：男性、下段：女性）



注：2010年時点に35歳以上で集計している。ウェイトとは労働力調査の学歴分布を用いた推計を指し、コホートとは国勢調査のコホート情報を用いた推計結果を指す。

以上の結果より得られる示唆は2点ある。1点目は、「学歴不詳者」は低学歴層で発生する可能性が高いため、このグループを集計から外すことは、バイアスを生む可能性があることで

ある。2点目は、2000年の国勢調査や労働力調査は不詳者の割合が低いため、学歴の分布は信頼できることである。

## V. おわりに

本稿は、調査方法の異なる家計関連統計を比較し、各統計の特徴と利用上の注意点を明らかにした。二人以上世帯については、調査された世帯属性の平均には大きな違いはなかった。ただし、世帯属性の分布には調査方法の違いにより差が見られ、国民生活基礎調査は全国消費実態調査、家計調査より平均的ではない家計を多く捉えている可能性がある。一方で、単身世帯については、国民生活基礎調査で割合が低くなった。代替世帯を取らず比推定も用いない国民生活基礎調査では、単身世帯が過小に把握されていると考えられる。

学歴に関しては、統計間で概ね近い分布を示すものの、国民生活基礎調査と国勢調査に含まれる「学歴不詳」者が比較を困難にしている。国勢調査のコホート情報と労働力調査の学歴分布を用いた推計によると、学歴不詳者は小中高卒グループで多く発生していることが明らかとなった。学歴不詳者を集計から除くことはバイアスを生む可能性が高いと考えられる。

本稿より得られる含意をまとめると以下の通

りである。国民生活基礎調査と他の家計関連統計は二人以上世帯に関しては整合的であるため比較可能である。学歴に関しても、利用項目の定義により若干差が出るが、他の家計関連統計と比較しても整合的である。

集落抽出は平均的ではない家計を多く捉えることができる利点がある。しかし、比推定のような回答者の属性を補正するような仕組みを取り入れないと、集計結果が偏ってしまう可能性がある。ただし、比推定によってもその調査負担や記入方式により、系統的なサンプルの脱落が発生する可能性が存在することには留意が必要である。

学歴不詳問題からの示唆をさらに踏まえると、低学歴層の回答率がより低下する可能性があり、集落抽出、層化抽出双方で非標本誤差が深刻になる可能性がある。公的統計に関する理解を周知することは言うまでもないが、加えて全国消費実態調査などで調査されていない学歴を調査項目に含めるなど、統計間のクロスチェックの方法を充実することも一案だろう。

## 参 考 文 献

阿部隆（2013）「統計ウォッチング：人口統計  
国勢調査結果の「不詳数」に係わる諸問題」  
『統計』、2013年12月号、pp.51-54.  
宇南山卓（2011）「家計調査の課題と改善に向けて」『統計と日本経済』第1巻第1号、  
pp.3-28.  
宇南山卓（2013）「仕事と結婚の両立可能性と  
保育所：2010年国勢調査による検証」、RI-  
ETI Discussion Paper Series, 13-J-039  
大竹文雄（2005）『日本の不平等』日本経済新  
聞社

小塩隆士（2010）『再分配の厚生分析』日本評  
論社  
川口大司（2011）「ミンサー型賃金関数の日本  
の労働市場への応用」、阿部顕三、大垣昌夫、  
小川一夫、田渕隆俊 編『現代経済学の潮流  
2011』東洋経済新報社  
小池司朗・山内昌和（2014）「2010年の国勢調  
査における「不詳」の発生状況：5年前の居  
住地を中心に」『人口問題研究』No. 70-3、  
pp.325-338.  
佐野晋平・安井健悟（2009）「日本における教

- 育のリターンの推計」『国民経済雑誌』第200巻第5号, pp.71-86.
- 橋本俊詔 (1998)『日本の経済格差』岩波新書
- 内閣府 (2010)『経済財政白書』
- 新田功 (2004)「国民生活基礎調査における無回答データ等の影響を考慮した調査設計に関する研究」『厚生の指標』第51巻第11号, pp.24-28.
- 舟岡史雄 (2001)「日本の所得格差についての検討」『経済研究』第52巻第2号, pp.117-131.
- 松井博 (2005)『標本調査法入門―基礎から学ぶ, 標本調査の理論と実際―』日本統計協会
- 山口雅生 (2014)「所得格差拡大と高齢化の再検討」『経済研究』第65巻第1号, pp.86-93.
- 山田茂 (2012)「2010年国勢調査が把握した大都市地域の性別年齢別人口の精度に関する考察」『政経論叢』第158号, pp.35-67.
- Barro, R. and J. Lee (2000) “International Data on Educational Attainment Updates and Implications”, NBER Working Paper 7911
- Godo (2011) “Estimation of Average Years of Schooling for Japan, Korea and the United States”, PRIMCED Discussion Paper Series, No. 9