

2018 年家計調査結果における変動調整値の推定方法

柴田 卓也[†]Estimation of discontinuity-adjusted figures on the Family Income and Expenditure Survey results
in 2018

SHIBATA Takuya

総務省統計局の実施する家計調査は、全国の世帯から層化三段抽出法により約9,000世帯を抽出し、毎日の収支について家計簿を記入してもらうことで、家計に関する詳細な情報を得ることを目的とした、統計法における基幹統計調査の一つである。2018年からの家計調査の実施においては、家計簿の様式を従来のものから一新し、世帯の記入負担の軽減を図ると共に、より詳細な調査結果を得られるようにした。一方で、調査で用いる調査票の変更は、少なからず調査結果に影響を与え、従来の結果との間に時系列的な断層を生ずる可能性が見込まれた。そこで総務省統計局では、調査票変更が調査結果に与える影響を推定し、その影響を除いた前年比を「変動調整値」として2018年1月分から公表している。本稿では変動調整値の推定方法及び推定結果について解説する。

キーワード： Rubinの因果モデル、因果効果、傾向スコア、ロジスティック回帰

The Family Income and Expenditure Survey (FIES), which is conducted by Statistics Bureau, Ministry of Internal Affairs and Communications, is the one of fundamental statistics in the Statistics Act and obtains the detail of household income and expenditures from household account books of about 9,000 households which are selected based on a three-stage stratified sampling method from all households in Japan. In the FIES, new form of account books has been introduced since January 2018. It would be expected to make keeping account books easier and entry contents more precise. However, the change of survey forms may cause a gap for the time series of survey results. By the Statistics Bureau, the effect by changes of survey forms is estimated and the ratios of survey results year over year without the effect, named “discontinuity-adjusted figures”, are released after the results of January 2018. In this paper, the details of the estimation and the results of discontinuity-adjusted figures are shown.

Key words: Rubin’s causal model, causal effect, propensity score, logistic regression

[†] 総務省統計局統計調査部消費統計課

1. 変動調整値導入の経緯

総務省統計局の実施する家計調査は、全国の世帯から約 9,000 世帯を抽出し、毎日の収支について家計簿を記入してもらうことにより、家計に関する詳細な情報を得ることを目的とした、統計法における基幹統計調査の一つである。その家計簿を記入してもらうという調査方法の関係上、家計調査は以前から、調査客体の回答負担の大きい調査と言われており、そのため調査に協力する世帯が、専業主婦のいる世帯などに偏っているという指摘があった（宇南山（2011））。また、自動車やリフォーム工事費用など、高額で購入頻度の低い財やサービスの支出については、日常で発生する少額かつ高頻度の支出に比べ、家計簿への記入がされにくいのではないかという懸念もあった。このような状況を受け、2018年1月から、家計調査では調査負担の軽減と結果精度の向上を図り、オンライン家計簿の導入及び家計簿様式の変更を行った。このうち家計簿様式¹の主な変更点は以下のとおりである。

- ① 「口座自動振替による支払」欄の改修
 - ・ 項目の追加、細分化
 - ・ 「クレジット・掛買い・月賦」欄を追加
- ② 「口座への入金（給与・年金等）」欄の新設
- ③ 日々の収支の記入欄の拡充
 - ・ 1ページあたりの行数を増加
 - ・ 現金以外による購入の支払方法を見直し
- ④ 高額で購入頻度の低い財やサービスについて、記入確認リストの導入

これらは、従来の家計簿と比べ、収支をより記入しやすくするための変更であるため、今まで記入されにくかった収支の記入が増えることで収支の平均値が増加するなど、様式の変更が集計結果に影響を与える可能性がある。このため、総務省統計局では、過去の調査結果との比較を行うために、家計簿変更が結果に及ぼす影響を推定する方法を検討し、その影響を除いた調査結果の前年比（又は前年同月比、前年同期比）を「変動調整値」として2018年1月調査分から公表している。その変動調整値の推定方法及び実際の推定結果について、以下に解説する。ただし本稿における意見は筆者個人のものであり、所属する組織を代表するものではない。

本稿の構成を以下に述べる。2章では推定手法の概論として、2018年家計調査へのRubinの因果モデル（Holland（1986））の導入と、傾向スコアを利用した因果効果の推定方法について述べる。3章では公表されている変動調整値の推定に利用した実際の手順を紹介し、4章でその推定結果について述べる。

2. 推定手法

2.1 反実仮想による因果効果の考え方

一般的に、調査方法の違いが調査結果に及ぼす影響（以下、本稿では因果効果と呼ぶ）を調べるためには、母集団から同じ方法で無作為抽出した2つのグループに対し、異なる調査方法を適用した場合の結果を比較する方法が考えられる。この場合、無作為抽出が適切に行われていれば、関心のある変数に対する共変量（それ自身の変化が、関心のある変数に影響を及ぼす量）の

¹ 詳細は総務省統計局 Web サイト内「家計調査の調査票」<http://www.stat.go.jp/data/kakei/form.html> を参照。

分布も同一になっていると考えられるので、2グループの調査結果の差を因果効果とみなすことができる。

2018年の家計調査においても、同様の考え方にに基づき、全国の調査対象世帯を2つのグループに分け、一方では新しい家計簿（家計簿A）、他方では従来の家計簿（家計簿B）を使用することにより、家計の収支金額に対する因果効果を調査できるようにした。グループ分けは家計調査の調査単位区を単位とし、以下の規則（丸数字の順に優先）に従って、各グループに含まれる世帯数が同一となるように設定した。

- ① 政令市及び県庁所在市は、市ごとに家計簿Aと家計簿Bが半々となるように振り分け
- ② 地方ごと（可能であれば都道府県ごと）に半々
- ③ 都市階級ごとに半々
- ④ 開始月ごと（可能であれば開始月の地方ごと）に半々

しかし、家計簿ABの振り分けは、調査実施上の制約から、

- ・ 調査世帯ごとではなく、調査単位区ごとの振り分け
- ・ 2018年標本改正における新調査市町村では全て家計簿Aを使用
- ・ 単身世帯のうち寮・寄宿舍世帯は全て家計簿Aを使用

として行ったため、2018年の家計簿別の調査世帯の属性の分布に、結果として偏りが生じている可能性があった。家計簿別の調査世帯分布が同等のものとみなせない場合、各調査グループにおける世帯属性の分布の違いが収支金額の分布に影響を及ぼすと考えられるため、2グループの調査結果の差には、因果効果のほかに世帯分布の違いによる影響が含まれる。

2018年の家計調査における、調査グループと使用した家計簿の関係を図1にまとめた²。調査結果として得られるデータは、図中の「(1)家計簿Aの使用グループにおける家計簿Aによる調査結果」及び「(4)家計簿Bの使用グループにおける家計簿Bによる調査結果」の2つのみである。

「(2)家計簿Bの使用グループにおける家計簿Aによる調査結果」及び「(3)家計簿Aの使用グループにおける家計簿Bによる調査結果」は、現実的に得ることができない。

	家計簿Aの使用グループ	家計簿Bの使用グループ
家計簿Aによる調査結果	(1)調査結果あり	(2)調査不可能
家計簿Bによる調査結果	(3)調査不可能	(4)調査結果あり

図1 調査グループと使用した家計簿の関係

そこで今回の分析では、家計簿の種類と調査世帯の関係にRubinの因果モデルを導入し、家計簿変更による因果効果の推定を行った。これは、(2)や(3)の結果について、本来存在するが測定されていない状態（欠測）であるという、現実にはありえない状況を考え（反実仮定）、(1)及び(2)

² 図の形式は星野（2009、22ページ）など、傾向スコアを用いた解析に関する文献にあるものを参考にしている。

と(3)及び(4)の結果の差の期待値（Rubin の因果効果、Rubin（1974））を推定する方法である。各調査世帯についての家計簿 A による調査結果を y_1 、家計簿 B による調査結果を y_0 と書くと、因果効果は $E(y_1 - y_0)$ と表すことができる。

2.2 傾向スコアを用いた因果効果の推定

Rubin の因果効果の期待値を推定する方法として、家計簿 AB の使用グループへの割り付けが完全は無作為に行われていれば、各グループにおける平均の差（(1)の平均－(4)の平均）を求めればよい。しかし 2.1 節で述べたように、2018 年の家計調査における 2 グループへの割り付けは完全は無作為に行われているわけではないため、家計収支に影響を与えるような世帯属性の分布が、2 グループの間で異なっている可能性がある。この場合、2 グループの世帯属性分布の違いによる影響を除去するための操作（共変量調整）を行うことで、因果効果を推定することができる³。

	家計簿Aの使用グループ	家計簿Bの使用グループ
家計簿Aによる調査結果	(1)調査結果あり	(2)欠測
家計簿Bによる調査結果	(3)欠測	(4)調査結果あり
世帯票等による調査結果	世帯属性(世帯人員, 年収等)	

図2 Rubin の因果モデルによる調査結果の考え方

共変量調整の方法としては、マッチング、層別解析、回帰モデルを用いる方法等、様々なものがあるが、本研究においては、因果効果だけではなく、2017 年以前（家計簿 B のみ使用）又は 2019 年以降（家計簿 A のみ使用）の家計調査の公表値と比較するための、家計簿 A 又は B を使用した場合の集計結果の推定が必要である。そこで、共変量を調整し、家計簿 A 又は家計簿 B を使用した場合の調査結果を推定できる手法として、傾向スコアを用いた逆確率重み付け法（IPW: Inverse Probability Weighting）を用いた⁴。

傾向スコアとは、Rosenbaum and Rubin (1983) において提唱された、バランシングスコアと呼ばれる量の一種である。バランシングスコアとは、それを条件付けることによって、2つのグループへの割り付けと共変量の分布が独立になるような共変量の関数である（傾向スコアの具体的な推定方法については次節以降で説明している）。

ここでの分析において、傾向スコアは家計調査の調査世帯が家計簿 A の使用グループに割り当てられる確率として定義される。このとき、傾向スコアの推定値 e_i を用いて計算される、家計簿 A

³ 厳密には、2 グループへの割り付けが、共変量である世帯属性のみに依存し、結果変数である収支金額に依存しないという条件（強く無視できる割当て条件）が成立していることが必要（後述）。

⁴ マッチング等の手法に関する問題点については、星野（2009、2.6 節）を参照。

を使用した世帯についての以下の推定値 $\hat{E}(y_1)$ (IPW 推定量) は、家計簿の割当てによらない y_1 の期待値 (周辺期待値) $E(y_1)$ の一致推定量となる (星野 (2009、69-70 ページ))。

$$\hat{E}(y_1) = \sum_i \frac{z_i w_i y_i}{e_i} / \sum_i \frac{z_i w_i}{e_i} \quad (1)$$

ここで、 i は調査世帯、 z_i は家計簿 A を使用した世帯で 1、家計簿 B を使用した世帯で 0 となる割り付け変数、 w_i は家計調査の集計用乗率、 y_i は収支金額を表す。同様に、家計簿 B を使用した世帯についての以下の推定値 $\hat{E}(y_0)$ は、家計簿の割当てによらない y_0 の期待値 $E(y_0)$ の一致推定量となる。

$$\hat{E}(y_0) = \sum_i \frac{(1 - z_i) w_i y_i}{1 - e_i} / \sum_i \frac{(1 - z_i) w_i}{1 - e_i} \quad (2)$$

上記の推定値の差が因果効果の推定値となる。

3. 実際の計算手順

3.1 傾向スコアの推定

IPW 法による因果効果推定のためには、まず家計調査の調査世帯に対する傾向スコアを推定する必要がある。傾向スコアは各サンプルが 2 グループのどちらかに割り付けられる確率を表すため、2 グループのどちらに属するかを表す割り付け変数を目的変数とした、ロジスティック回帰分析やプロビット回帰分析によって推定する方法が一般的である⁵。2018 年の家計調査における変動調整値の推定では、使用した家計簿の違いにより世帯を 2 グループに分け、家計簿 A を使用した世帯で 1、家計簿 B を使用した世帯で 0 をとる割り付け変数を目的変数とし、各世帯の世帯属性 (世帯人員や世帯主の年齢など) を説明変数としたロジスティック回帰により、傾向スコアの推定を行った。推定に使用したデータは、二人以上の世帯及び単身世帯のそれぞれについて、2018 年の 1 月から 12 月までのデータを一つにまとめた 1 年分のデータである⁶。

ロジスティック回帰の説明変数の選択基準は以下のとおり。

- ① 家計調査の調査事項における世帯属性について、収支の調査結果に依存しているものを外す。
- ② その上で、消費支出と相関のある変数をできるだけ多く選択する。
- ③ ②で選択した世帯属性変数のそれぞれについて、家計簿使用グループの割り付け変数を目的変数としたロジスティック回帰を行い、疑似決定係数が小さい (他の変数を使用した場合に比べ 10 分の 1 程度) 変数を外す。
- ④ 残った世帯属性変数全てを説明変数の候補とし、後述 (3.2 節) の条件に最もよく当てはまる組合せ (交互作用項を含む) を、傾向スコアの推定モデルとして選択する。なお説明

⁵ ただし最近では、ランダムフォレストのような機械学習的手法を用いた推定も行われている。

⁶ ただし 2018 年 1 月分の公表から 11 月分の公表までの間は、1 年間分のデータが存在しないため、暫定的に異なるデータで推定を行った。1 月分から 3 月分までは各月のデータのみを使用し、4 月分から 11 月分までは、1 月からその時点の最新月までの結果を合わせたデータを用いた。また単身世帯については、各月のデータが少ない関係上、二人以上の世帯と合わせた総世帯のデータについて傾向スコアの推定を行い、そのうちの単身世帯のデータを抜き出して変動調整値の計算を行った。その後、2018 年 12 月分の公表に合わせ、1 月分から 11 月分の結果について、ここで述べた方法により遡及改定を行った。

変数選択には変数増加法を用いた。

3.2 強く無視できる割当て条件の確認

3.1 節の方法で推定した傾向スコアが適切に推定されているかどうかは、「強く無視できる割当て」条件の成立により確かめることができる。強く無視できる割当て条件とは、星野（2009、43 ページ）によると『割当てはあくまで共変量にのみ依存し、結果変数には依存しない』という仮定が成り立つことであり、傾向スコアを用いた共変量調整により因果効果を推定する場合に成り立っている必要がある。強く無視できる割当て条件の成立を直接検証することは難しいため、2018 年の家計調査における変動調整値の推定では、間接的な検証方法として以下の方法を用いた。

- ① 傾向スコア推定モデルにおける説明変数の組合せについて、当てはまりのよいものを選択する（疑似決定係数、尤度及び c 統計量を評価基準として使用）。
- ② 傾向スコアによる共変量の調整を行い、各共変量について 2 グループの分布がより近くなる推定結果を選択する（分布の近さについての評価基準は後述）。

3.3 IPW 推定量の算出

3.2 節の方法により確定した傾向スコアの推定モデルを使用し、実際のデータについて傾向スコアの推定及び IPW 推定量の算出を行った。IPW 推定量を算出した項目は以下のとおり（括弧内は算出した世帯区分を表す）。

- 消費支出（全ての世帯、勤労者世帯、勤労者以外の世帯及び無職世帯）
- 世帯主収入のうち定期収入（勤労者世帯）
- 世帯主収入のうち臨時収入・賞与（勤労者世帯）
- 世帯主の配偶者の収入（勤労者世帯及び無職世帯）
- 他の世帯員の収入（勤労者世帯及び無職世帯）
- 勤め先収入を除く実収入（勤労者世帯及び無職世帯）
- 非消費支出（勤労者世帯及び無職世帯）

ただし、勤労者以外の世帯及び無職世帯についての推定は、二人以上の世帯のみ行った。

3.4 変動調整値の算出

2018 年及び 2019 年結果における対前年比（月別及び四半期結果の場合は対前年同月比及び対前年同期比）には、経済・社会的変動のほか、家計簿変更による影響が含まれる。このうち家計簿変更による影響を推定し、取り除くことで算出した対前年比を「変動調整値」と呼ぶ。変動調整値の計算においては、2.2 節で説明した、2 つの IPW 推定量の差を因果効果とする方法ではなく、IPW 推定量から家計調査の公表値に対する増加影響分を推定し、調整する方法を採用した。二人以上の世帯及び単身世帯における変動調整値の具体的な計算手順を以下に示す。

二人以上の世帯

2017 年と比較した 2018 年の対前年比を算出する場合（旧家計簿基準）：

- ① 3.3 節に挙げた各収支項目について、(2)式に従い、2018 年の調査において家計簿 B を使用した世帯の IPW 推定量（以下、B 推定値）を、各月ごとに算出。
- ② ①で算出した B 推定値と、2018 年の各月の全世界帯についての単純な集計結果（家計簿 A

使用世帯と家計簿 B 使用世帯をまとめて、家計調査の集計用乗率で加重平均した結果で、家計調査の公表値と等しい。以下、原数値)を比較して、原数値の方が大きい場合(家計簿変更により収支に増加影響が見られる場合)は B 推定値、B 推定値の方が大きい場合(家計簿変更による収支への増加影響が見られない場合)は原数値を、「2017 年結果と比較可能な数値」とする。また 2017 年と比較可能な数値と原数値の差を算出し、「調整額」と呼ぶこととする。すなわち、

$$\begin{aligned} \text{(調整額)} &= \text{(2017 年と比較可能な数値)} - \text{(原数値)} \\ &= \begin{cases} \text{(B 推定値)} - \text{(原数値)} & : \text{(B 推定値)} < \text{(原数値)} \text{ の場合} \\ 0 & : \text{(原数値)} \leq \text{(B 推定値)} \text{ の場合} \end{cases} \end{aligned}$$

このとき、調整額は 0 以下の値をとる。

- ③ 3.3 節に挙げた収支項目の内訳については、各月の原数値の構成比を用いて調整額を按分し、原数値から按分した調整額の絶対値を差し引くことで「2017 年と比較可能な数値」を作成する。また上位項目については、下位項目の「2017 年と比較可能な数値」を足し上げることで作成する。
- ④ ②で決定した 2017 年と比較可能な数値を使用し、各月の対前年同月比を算出。
- ⑤ 四半期及び年平均の結果については、各月の 2017 年と比較可能な数値を用いて算出。

2018 年と比較した 2019 年の対前年比を算出する場合(新家計簿基準)：

- ① 各収支項目について、(1)式に従い、2018 年の調査において家計簿 A を使用した世帯の IPW 推定量(以下、A 推定値)を、各月ごとに算出。
- ② ①で算出した A 推定値と原数値を比較して、原数値の方が小さい場合(家計簿変更により収支に増加影響が見られる場合)は A 推定値、A 推定値の方が小さい場合(家計簿変更による収支への増加影響が見られない場合)は原数値を、「2019 年と比較可能な数値」とする。また 2019 年と比較可能な数値と原数値の差を算出し、「調整額」と呼ぶこととする。すなわち、

$$\begin{aligned} \text{(調整額)} &= \text{(2019 年と比較可能な数値)} - \text{(原数値)} \\ &= \begin{cases} \text{(A 推定値)} - \text{(原数値)} & : \text{(原数値)} < \text{(A 推定値)} \text{ の場合} \\ 0 & : \text{(A 推定値)} \leq \text{(原数値)} \text{ の場合} \end{cases} \end{aligned}$$

このとき、調整額は 0 以上の値をとる。

- ③ 3.3 節に挙げた収支項目の内訳については、各月の原数値の構成比を用いて調整額を按分し、原数値に按分した調整額を足し上げることで「2019 年と比較可能な数値」を作成する。また上位項目については、下位項目の「2019 年と比較可能な数値」を足し上げることで作成する。
- ④ ②で決定した 2019 年と比較可能な数値を使用し、各月の対前年同月比を算出。
- ⑤ 四半期及び年平均の結果については、各月の 2019 年と比較可能な数値を用いて算出。

単身世帯

二人以上の世帯に比べサンプルサイズが小さいため、まず(1)式又は(2)式による各月に対しての計算結果を単純平均することで、家計簿 A 又は家計簿 B を使用した世帯、それぞれの IPW 推定量の年平均値を算出した。その上で、単身世帯の各月の推定値を以下のように求めた。

- ① 年平均の集計値と推定値の差を「調整額」とする。
- ② 二人以上の世帯における年平均の調整額と各月の調整額の比率を用いて、単身世帯における各月の調整額を以下のとおり求める。また、各月の調整額と集計値を足すことで、2017年又は2019年と比較可能な数値を求める。

$$\left(\text{単身の各月調整額} \right) = \left(\text{単身の年平均調整額} \right) \cdot \frac{\text{二人以上の各月調整額}}{\text{二人以上の年平均調整額}} \quad (3)$$

- ③ 以降は二人以上の世帯における③以降と同様。

単身世帯の各月の調整額を求めるために二人以上の世帯の調整額を利用した理由の一つとして、単身世帯のサンプルサイズが小さいため、そのまま各月の調整額を求めると誤差が大きいと考えられる点がある。また消費支出や実収入には月ごとの季節性が存在することにより、調整額の大きさも月ごとにある程度決まった変動があると考えられるため、二人以上の世帯の結果の持つ季節性と単身の結果の季節性が同じという仮定を置き、単身の各月の調整額を求めた。

上記手順において、一律に推定値を用いて対前年比の計算を行わないことは、以下の理由による。新しい家計簿への様式変更は、従来の家計簿より記入者の負担を軽減すると共に、より詳細な記入をしやすくするためのものであることから、家計簿 A を使用した場合の収支項目に関する集計結果は、家計簿 B を使用した場合の集計結果と比べて大きくなることが予想され、また実際の集計結果でも家計簿 A 使用グループにおける集計結果の方が、家計簿 B 使用グループに比べて大きい結果を得ていた。そのため、家計簿 A の導入による増加影響が見られないのは、推定値及び原数値に含まれる誤差の影響と考え、対前年比を過大に見積もることがないように、原数値を使用して対前年比の計算を行った。

4. 推定結果

前述の手順に沿った推定結果について、手順ごとに以下に述べる。

4.1 傾向スコアの推定モデル⁷

計算は二人以上の世帯及び単身世帯について別々に行った。家計調査の2018年1月から12月までの世帯別データをまとめ、世帯属性を説明変数としてロジスティック回帰を行った(表1、2)。この際、世帯の年間収入や世帯主の年齢など、連続変数は適当な区分で階級化し、全ての変数をカテゴリ変数としてダミー変数による回帰を行った。また目的変数は、家計簿 A 使用グループで1、家計簿 B 使用グループで0となる割り付け変数を使用した。

⁷ ただし2018年1月分から11月分までは、暫定的に各月で異なるデータを使用して推定したため、傾向スコアの推定にもここに挙げたものと異なるモデルを使用した。

なお、単身世帯における傾向スコアの推定には、寮・寄宿舎に居住する世帯を除いたデータを使用した。これは、寮・寄宿舎の単身世帯は全て家計簿 A の使用グループに割り付けられていることが確定しており、家計簿 B を使用した場合の結果推定が不可能なためである。

ロジスティック回帰分析の計算は、全て統計解析環境 R (バージョン 3.4.4) の glm 関数を以下のように使用した (AB は割付変数、data は世帯別に AB 及び説明変数の値が入った data.frame)。

```
result <- glm(formula = AB ~ ., data, family = binomial(link = "logit"))
```

表1 傾向スコアの推定に使用した変数一覧 (二人以上の世帯)

変数名	区分	分類又は階級	備考
地方	10	北海道、東北、関東、北陸、東海、近畿、中国、四国、九州、沖縄	
世帯の年間収入階級	10	269万円未満、269～327、327～384、384～449、449～520、520～603、603～700、700～822、822～1026、1026万円以上	2017 年家計調査結果における十分位階級の階級値を使用。
世帯人員	4	2、3、4、5人以上	
世帯主の年齢階級	7	29歳以下、30～39、40～49、50～59、60～69、70～79、80歳以上	
世帯主の職業	6	労務作業、民間職員、官公職員、個人営業、無職、その他	その他には、法人経営者、自由業者、その他が含まれる。
世帯主の産業	12	「建設業」、「製造業」、「情報通信業」、「運輸業、郵便業」、「卸売業、小売業」、「金融業、保険業」、「学術研究、専門・技術サービス業」、「宿泊業、飲食サービス業」、「教育、学習支援業」、「医療、福祉」、「他のサービス業」、「公務」、「その他」	他のサービス業には「生活関連サービス業、娯楽費」、「複合サービス事業」及び「サービス業(他に分類されないもの)」が含まれる。 その他には「農業、林業」、「漁業」、「鉱業、採石業、砂利採取業」、「電気・ガス・熱供給・水道業」、「不動産業、物品賃貸業」、所属産業のないもの及び無職が含まれる。
世帯主の勤め先の企業規模	8	1～9人、10～29、30～99、100～299、300～499、500～999、1000人以上、その他	その他には、不詳、記入なし及び官公庁が含まれる。
就業人員	4	0、1、2、3人以上	
住居の所有関係	6	持家(一戸建て)、持家(その他)、民営の賃貸住宅(借間を含む)、公営の賃貸住宅、都市再生機構・公	

		社等の賃貸住宅、給与住宅（社宅・公務員住宅など）	
住居の居住室数	9	1、2、3、4、5、6、7、8、9室以上	1室は不詳を含む
住居の延床面積	5	79平方メートル未満、79～100、100～120、120～149、149以上	2017年家計調査結果における五分位階級の階級値を使用。
世帯主の就業状態	2	就業（正規）、その他	その他には、就業（正規以外）、非就業、不詳が含まれる。
世帯内の就学者数	4	0、1、2、3人以上	

表2 傾向スコアの推定に使用した変数一覧（単身世帯）

変数名	区分	分類又は階級	備考
都市階級	4	大都市、中都市、小都市 A、小都市 B・町村	
地方	10	北海道、東北、関東、北陸、東海、近畿、中国、四国、九州、沖縄	
年間収入	10	107万円未満、107～143、143～171、171～197、197～224、224～253、253～295、295～360、360～480、480万円以上	2017年家計調査結果における十分位階級の階級値を使用。
年齢	7	29歳以下、30～39、40～49、50～59、60～69、70～79、80歳以上	
職業	6	労務作業、民間職員、官公職員、個人営業、無職、その他	その他には、法人経営者、自由業者、その他が含まれる。
産業	12	「建設業」、「製造業」、「情報通信業」、「運輸業、郵便業」、「卸売業、小売業」、「金融業、保険業」、「学術研究、専門・技術サービス業」、「宿泊業、飲食サービス業」、「教育、学習支援業」、「医療、福祉」、「他のサービス業」、「公務」、「その他」	他のサービス業には「生活関連サービス業、娯楽費」、「複合サービス事業」及び「サービス業(他に分類されないもの)」が含まれる。 その他には「農業、林業」、「漁業」、「鉱業、採石業、砂利採取業」、「電気・ガス・熱供給・水道業」、「不動産業、物品賃貸業」、所属産業のないもの及び無職が含まれる。
勤め先の企業規模	8	1～9人、10～29、30～99、100～299、300～499、500～999、1000人以上、その他	その他には、不詳、記入なし及び官公庁が含まれる。

住居の所有関係	6	持家（一戸建て）、持家（その他）、民営の賃貸住宅（借間を含む）、公営の賃貸住宅、都市再生機構・公社等の賃貸住宅、給与住宅（社宅・公務員住宅など）	
住居の居住室数	9	1、2、3、4、5、6、7、8、9室以上	1室は不詳を含む
住居の延床面積	5	42平方メートル未満、42～72、72～100、100～132、132以上	2017年家計調査結果における五分位階級の階級値を使用。
就業状態	2	就業（正規）、その他	その他には、就業（正規以外）、非就業、不詳が含まれる。
性別	2	男、女	
年収×年齢		交互作用項	
年収×職業		交互作用項	

4.2 強く無視できる割当て条件の確認

家計簿 AB の割り付けが強く無視できる割当て条件を満たしていることを、4.1 節に挙げた傾向スコアの推定モデル及び傾向スコアの推定値を用いて、間接的に確認する。確認の方法は星野（2009、4.3 節）に詳しいが、ここでは以下の2つの点を確認する。

(1) 共変量によって割り付けが説明されていることの確認

傾向スコアの推定モデルのフィットが良ければ、共変量だけで割り付けを説明できていることの確認となるため、推定モデルの疑似決定係数（McFadden（1974）、以下 McFadden の疑似決定係数と呼ぶ）及び c 統計量（後述）の確認を行った。傾向スコアの推定に使用した2つのモデルについて、疑似決定係数及び c 統計量の値を表3に示す。

表3 傾向スコア推定モデルのフィットの良さを表す指標

	McFadden の疑似決定係数	c 統計量
二人以上の世帯	0.0046	0.55
単身世帯	0.035	0.62

McFadden の疑似決定係数 R_{MC}^2 は以下の(4)式で与えられる。

$$R_{MC}^2 = 1 - \frac{\ln L_1}{\ln L_0} \quad (4)$$

ここで、 L_1 は推定に使用したロジットモデルの尤度、 L_0 は全てのパラメータを 0 と仮定した時の尤度である。これを見ると、今回の推定に使用したモデルは、割り付け変数の分布に対する説明力がほとんどないことが分かるが、これは 2.1 節で説明した、使用する家計簿に関する割り付けが、おおむね無作為に近い形でできていたことを示しており、また後述するように、今回のモデ

ルで推定した傾向スコアによる調整では、2つの家計簿の使用グループ間の属性分布の差を少なくすることができていた。

また、c 統計量⁸はロジスティック回帰結果について ROC (Receiver Operating Characteristic) 曲線⁹のグラフを書いたときに、グラフの下部の面積が全体の面積に占める割合として定義でき、AUC (Area Under Curve) とも呼ばれる量である。今回の計算結果による ROC 曲線を図3及び4に示す(横軸が家計簿 B 使用グループの世帯を、間違っ家計簿 A 使用グループと判定した割合、縦軸が家計簿 A 使用グループの世帯を、正しく家計簿 A 使用グループと判定した割合)。

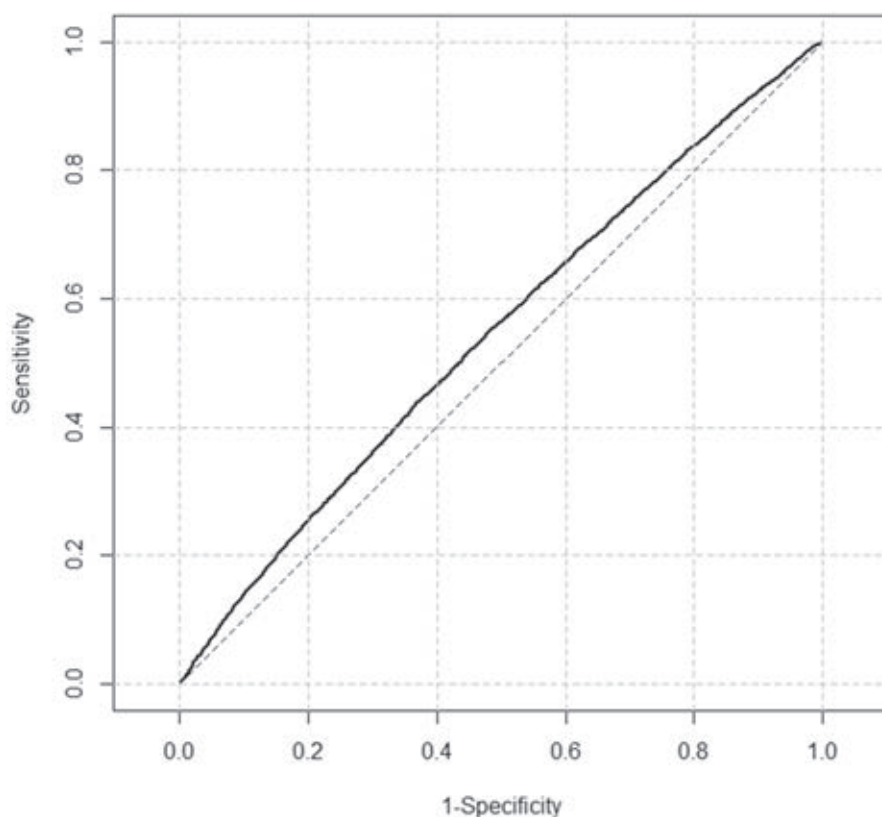


図3 二人以上の世帯についての傾向スコア推定値による ROC 曲線 (実線)

⁸ 詳しくは、例えば星野 (2009、125 ページ脚注) などを参照。

⁹ ROC 曲線とは、ロジスティック回帰分析の予測値を元にデータを 2 群 (例えば A 群と B 群) に分ける際に、本当は B 群のデータを A 群と判定する割合 (偽陽性率) を横軸、A 群のデータを正しく A 群と判定する割合 (真陽性率) を縦軸とし、2 群の判定に使用する基準 (その数値に対する予測値の大小関係で群の判定する基準で、カットオフとも呼ばれる) を 0 から 1 まで変化させたときの、偽陽性率と真陽性率で決まる点が描く曲線である。

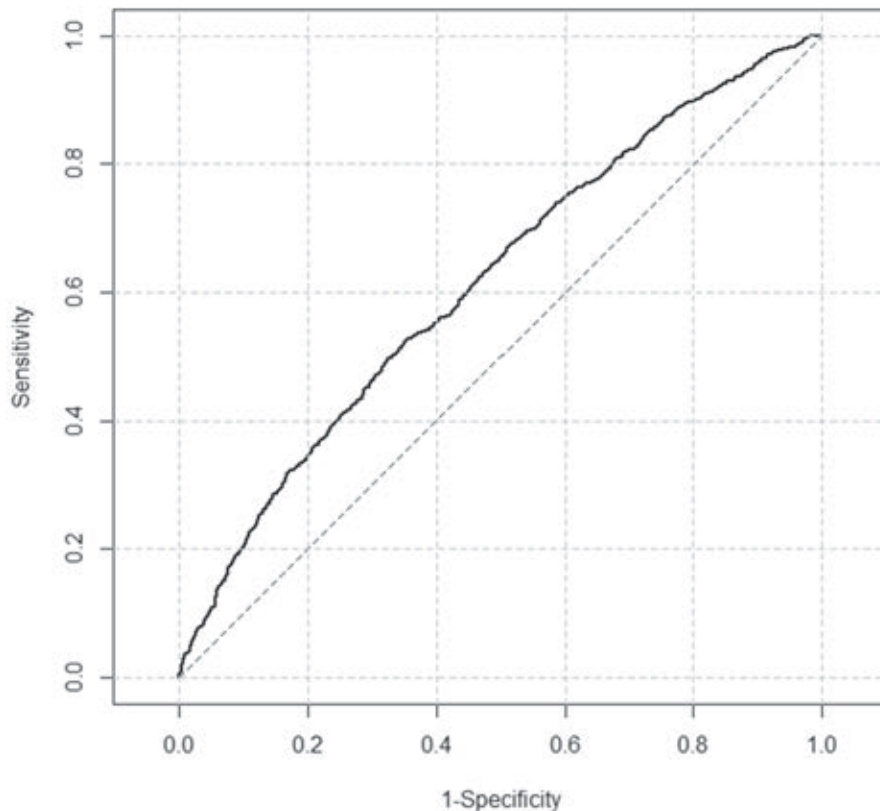


図4 単身世帯についての傾向スコア推定値によるROC曲線（実線）

傾向スコアによる割り付けの説明がうまくいっていることの見方は、一般的にc統計量が0.8以上とされている（星野, 岡田（2006））。一方、Brookhart, et.al.（2006）におけるシミュレーション結果では、c統計量が0.67程度であっても、従属変数に関連の強い共変量を選択することで、偏りの小さい調整が可能である場合があることも示されている。

以上により、今回の傾向スコア推定モデルは必ずしも割り付けをよく説明するものではないが、傾向スコアによる調整を行うことによって、収支項目について偏りの小さい結果を得ることができる可能性があることが分かる。

(2) 傾向スコアにより共変量の分布が近くなることの確認

傾向スコアの定義より、傾向スコアを条件付けたときに割り付けと共変量の分布は独立となる。このため、割り付けの2グループについて、それぞれ正しく推定された傾向スコアにより共変量分布の調整を行えば、同じ分布が得られるはずである。ここでは、推定した傾向スコアにより、家計簿別のグループにおける世帯属性の分布が近づくことを、以下の方法により確認した。

まず、傾向スコアの説明変数と同じく量的変数を適当に階級化し、世帯属性を質的変数として表した。そして各世帯属性について、全世帯数を10,000としたとき、各カテゴリに何世帯が含まれるか（一万分比）を求めた。

次に、世帯属性分布の近さ（一致度）を図るため、カテゴリ化した分布についての二乗平均群間差（Root Mean Square Difference, RMSD）を、以下のように定義した。

$$\text{RMSD} = \sqrt{\frac{1}{n} \sum_{i=1}^n (D_i^A - D_i^B)^2} \quad (5)$$

ここで、 n は世帯属性分布のカテゴリ数、 D_i^A 及び D_i^B はそれぞれ、家計簿 A 又は家計簿 B を使用したグループについて、第*i*カテゴリに属する一万分比の世帯数を表す。家計簿 A 及び B の使用グループについて、世帯属性分布に差がなければ、すなわち全ての*i*について $D_i^A = D_i^B$ が成立すれば、RMSD の値は 0 となる。このため、RMSD の値が小さいほど、グループ AB 間の分布差は小さいと判断することとした。なおこのように、傾向スコアを適切に推定できていることを確認するために、質的変数について 2 グループ間の分布を比較する方法は、MacKenzie, et.al. (2006)等を参考にした。

主な世帯属性についての RMSD の値を表 4、5 に示す。これらの RMSD は、2018 年 1 月から 12 月の各月について算出し、12 か月の値を単純平均したものである。また、年間収入や世帯主年齢などの連続変数について、階級化する際の階級値は表 1、2 と同じである。

表 4 傾向スコアによる調整前後での主な世帯属性分布に対する RMSD (二人以上の世帯)

変数名	調整前	調整後
地方	97.0	67.9
世帯の年間収入	113.3	88.8
世帯人員	84.2	106.9
世帯主の年齢	118.8	86.7
世帯主の職業	135.9	87.7
住居の所有関係	149.8	133.9

表 5 傾向スコアによる調整前後での主な世帯属性分布に対する RMSD (単身世帯)

変数名	調整前	調整後
地方	156.7	146.3
世帯の年間収入	181.4	98.0
世帯主の年齢	152.0	86.0
世帯主の職業	259.2	200.0
住居の所有関係	294.6	239.8

二人以上の世帯及び単身世帯での推定について、世帯人員を除いて調整後の方が RMSD の値が小さくなっていることが分かる。これにより、傾向スコアにより家計簿使用グループ間の分布の差をなくす方向へ調整できていると判断した。なお、この RMSD は傾向スコアの推定モデルの選択にも使用した。この場合、各世帯属性の分布について、RMSD をより小さくするモデルを良いモデルと判断した。

4.3 IPW 推定量及び変動調整値の算出

ここまでで推定した傾向スコアを用いて、(1)及び(2)式に従い計算することで、家計簿 A 又は家計簿 B を使用した場合の IPW 推定量 \hat{Y}^{IPW} を求めることができる。ここから、2017年又は2019年と比較可能な推定値 \hat{Y}^{est} を求める手順は以下のようになる。以下では、家計調査における集計値（公表値）を \hat{Y} と書く。なお、(1)及び(2)式に含まれる集計用乗率 w は、通常の家計調査における集計用乗率をそのまま用いた。

二人以上の世帯の場合：

2018年の各月結果における \hat{Y} と \hat{Y}^{IPW} を比較し、 $\hat{Y} > \hat{Y}^{IPW}$ のときは $\hat{Y}^{est} = \hat{Y}^{IPW}$ 、 $\hat{Y} \leq \hat{Y}^{IPW}$ のときは $\hat{Y}^{est} = \hat{Y}$ とした。また \hat{Y} と \hat{Y}^{est} の差 $\hat{Y}^{est} - \hat{Y}$ を調整額とした。

単身世帯の場合：

4.1節で求めた傾向スコアは寮・寄宿舍世帯を除いたデータで推定したため、IPW 推定量も寮・寄宿舍世帯を除いた世帯における推定値 \hat{Y}_{sub}^{IPW} となる。ここで、寮・寄宿舍世帯を除いた家計調査における集計値 \hat{Y}_{sub} と推定値 \hat{Y}_{sub}^{IPW} の差を、寮・寄宿舍世帯を除いた調整額 \hat{A}_{sub} とし、以下の式に従って、寮・寄宿舍世帯を含む全ての世帯の調整額 \hat{A} を求めた。

$\hat{A}_{sub} = \hat{Y}_{sub}^{IPW} - \hat{Y}_{sub}$ とするとき、

$$\hat{A} = \hat{A}_{sub} \cdot \frac{\hat{Y}}{\hat{Y}_{sub}} \quad (6)$$

これは、寮・寄宿舍世帯における集計値と調整額の比率が、寮・寄宿舍世帯を除く世帯における集計値と調整額の比率に等しいという仮定による。すなわちここでは、寮・寄宿舍世帯とそれ以外の世帯で、家計簿様式を変更したことによる収支金額の変動が同程度であると仮定している。以上で求めた調整額を使い、2017年又は2019年と比較可能な推定値 \hat{Y}^{est} を以下のように求める。

$$\hat{Y}^{est} = \hat{Y} + \hat{A} \quad (7)$$

実際に \hat{Y}^{est} を計算した結果を表5に示す。

表5 集計値及び家計簿変更による影響を調整した推定値の比較
(2018年平均、小数第一位で四捨五入、単位は円（年間収入を除く）)

	集計値 \hat{Y}	(参考) 家計簿 A のみ 集計値	(参考) 家計簿 B のみ 集計値	家計簿 A 推定値 \hat{Y}^{est}	家計簿 B 推定値 \hat{Y}^{est}
消費支出（二）	287,315	288,200	286,420	288,936	285,370
実収入（二勤）	558,718	574,388	542,580	579,571	537,243
年間収入（二） （共変量）	612万円	608万円	616万円	608万円	615万円
消費支出（単）	162,833	166,486	158,644	164,779	159,623
実収入（単勤）	330,867	362,705	288,966	356,869	300,969
年間収入（単） （共変量）	311万円	327万円	291万円	314万円	305万円

二人以上の世帯については、2グループの結果を比べたとき、集計値に比べて推定値の方が2グループの差が拡大する結果となっている。これは家計簿Aグループの方が平均の年間収入が低い
ため、傾向スコアによる調整がAの結果を高め、Bの結果を低める方向に作用したことによる。
一方、単身世帯については、集計値に比べて推定値の方が2グループの差が縮小しているが、こ
れは二人以上の世帯の場合と逆の理由による。ただしいずれにせよ、推定値の方が共変量の差は
縮小しており、各家計簿使用グループの集計値に比べて、世帯分布の偏りによる影響が小さい結
果が得られていると考えられる。

以上で求めた2018年における「家計簿変更による影響を調整した推定値」が、2017年又は
2019年と比較可能な結果であり、推定値を使用して求めた前年（同月、同期）比を「変動調整
値」と呼ぶ。2018年の前年同月比については、2018年に家計簿Bを使用した世帯のIPW推定量
（(2)式）と、2017年の公表値を使用する。また2019年の前年同月比については、2019年の公表
値と、2018年に家計簿Aを使用した世帯のIPW推定量（(1)式）を使用する。結果を表6に示す¹⁰。

表6 2018年平均における対前年比の比較（名目値、小数第二位で四捨五入、単位は%）

	集計値	変動調整値	ポイント差
消費支出（二）	1.5	0.8	-0.7
実収入（二勤）	4.7	0.6	-4.1
消費支出（単）	0.7	-1.2	-1.9
実収入（単勤）	4.2	-4.2	-8.4

表6で示した変動調整値の結果を見ると、消費支出に比べて実収入の方が、家計簿様式の変更
が調査結果に与える影響が大きかったことが分かる。これは、支出に影響を与えると考えられる
様式の変更点が日々の収支の記入欄の拡充のみであるのに対し、収入に影響を与えると考えられ
る様式の変更点が、日々の収支の記入欄の拡充に加え、「口座への入金（給与・年金等）」欄の新
設と、記入欄の大きな変更を伴っていたことに起因すると考えられる。従来の家計簿において
は、例えば毎月定期的に銀行口座に振り込まれる給与であっても、世帯員それぞれの分について
別々に、日々の収支記入欄に記入する必要があった。これに対し、新家計簿では口座への入金欄
を設けたことによって、記入負担が減少すると共に、より詳細な記入を可能としている。収入と
支出における家計簿変更の影響の違いは、以上のような理由によってもたらされたと考えられ
る。

また二人以上の世帯に比べて単身世帯の方が、消費支出、実収入共に家計簿様式の変更による
影響が大きかったことが分かる。これについては、まだはっきりとした原因が分からず、今後更
に分析が必要である。

¹⁰ 表6にない項目や、各月の結果に関する変動調整値は、統計局ホームページ「家計調査（家計収支編）時系列データ（二人以上の世帯）」<http://www.stat.go.jp/data/kakei/longtime/index.html>及び「家計調査（家計収支編）時系列データ（総世帯・単身世帯）」<http://www.stat.go.jp/data/kakei/longtime/soutan.html>を参照。

5. まとめ

本稿では、2種類の調査票を使用した統計調査において、調査票の違いが調査結果に対して与える影響を推定する手法を述べ、それを家計調査結果における時系列比較に適用した研究について解説した。2018年の家計調査では2種類の家計簿の使用グループの間で、世帯属性の分布に僅かな差異が見られたが、傾向スコアによる属性分布の調整を行った結果、世帯属性の分布の差を縮小し、前年と比較した場合の収支動向についてバイアスを減らした推定ができたと考えられる。これらの結果は総務省統計局により変動調整値として公表され、実社会で利用されている。

一方で、本稿においては傾向スコアの推定に伴う誤差については議論が及んでいないため、引き続き更なる検証が必要である。異なる調査票による調査結果の推定値について、本稿で述べたような点推定による評価ではなく信頼区間による評価を行うことで、調査票変更が調査結果に与える影響について、二人以上の世帯と単身世帯における差の検証など、より詳細な議論を行うことも可能になると考えられる。

謝辞

本研究の遂行及び本稿の執筆においては、総務省統計局統計調査部消費統計課の職員の方々から、様々な助言やコメントをいただいた。また2名の査読者の方には、原稿をととても丁寧に見ていただき、有用なコメントを数多くいただいた。ここに感謝の意を表したい。

参考文献

- [1] 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, 3-28.
- [2] 星野崇宏 (2009), 『調査観察データの統計科学 因果推論・選択バイアス・データ融合』, 岩波書店.
- [3] 星野崇宏, 岡田謙介 (2006), 『傾向スコアを用いた共変量調整による因果効果の推定と臨床医学・疫学・薬学・公衆衛生分野での応用について』, 保健医療科学, 55(3), 230-243.
- [4] Brookhart, M.A., Schneeweiss, S., Rothman, K.J., Glynn, R.J., Avorn, J., Stürmer, T. (2006), “Variable Selection for Propensity Score Models”, *American Journal of Epidemiology*, 163, 12, 1149-1156.
- [5] Holland, P.W. (1986), “Statistics and Causal Inference”, *Journal of the American Statistical Association*, Vol. 81, 396, 945-960.
- [6] MacKenzie, E.J., Rivara, F.P., Jurkovich, G.J., Nathens, A.B., Frey, K.P., Egleston, B.L., Salkever, D.S., Sharfstein, D.O. (2006), “A National Evaluation of the Effect of Trauma-Center Care on Mortality”, *The New England Journal of Medicine*, 354, 366-378.
- [7] McFadden, D. (1973), “Conditional Logit Analysis of Qualitative Choice Behavior”, in P. Zarembka (ed.) *FRONTIERS IN ECONOMETRICS*, 105-142, Academic Press: New York
- [8] Rosenbaum, P.R., Rubin, D.B. (1983), “The central role of the propensity score in observational studies for causal effects”, *Biometrika*, 70, 1, 41-55.

- [9] Rubin, D.B. (1974) , “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies”, *Journal of Educational Psychology*, Vol. 66, 5, 688-701.