

全国消費実態調査の世帯ウエイト

市村英彦 川口大司 鳥谷部貴大 深井太洋

東京大学経済学研究科 政策評価研究教育センター

2019年3月28日

1. ウエイトの作成方法

2. 全国消費実態調査と国勢調査の時点調整

3. 基本属性の分布の確認

4. 世帯所得・消費分布

全国消費実態調査のセル

セルを定義するために用いられている変数は

- 世帯主年齢：35歳未満、35～59歳、60歳以上（単身世帯のみ）
- 性別：男性、女性（単身世帯のみ）
- 世帯人数：1人、2人、3人、4人、5人以上
- 地域：二人以上世帯：10地域、単身世帯：6地域

世帯主の就業状態、より細かい年齢階級、都道府県において回収率が異なる可能性

新しく提案するセル

セルを定義するために用いる変数は

- 世帯主の年齢：25歳未満、25～34歳、...、75～84歳、85歳以上
- 性別：男性、女性（単身世帯のみ）
- 世帯人数：1人、2人、3人、4人、5人以上
- 就業状況：勤労世帯、無職世帯、その他
- 地域：都道府県

サンプルサイズの問題から、これらすべての変数の同時分布を国勢調査に合わせることは難しい

→ 同時分布ではなく、複数の周辺分布を合わせる

セルの定義

以下の5つの周辺分布が国勢調査に合うようにウエイトを作成する

0層目 調整済み調整係数（次のスライドで説明）

1層目 世帯人数×性別×世帯主の年齢

2層目 単身世帯×就業状況

3層目 都道府県×就業状況

4層目 単身世帯×都道府県

5層目 世帯主の年齢×就業状況

調整済み調整係数

調整済調整係数の定義

$$\text{調整済調整係数} = \begin{cases} \frac{\text{ある市の世帯数}}{\text{調査予定世帯数}} \cdot \frac{\text{調査予定世帯数}}{\text{回収世帯数}} & (\text{回収率 } 50\% \text{以上}) \\ \frac{\text{ある市の世帯数}}{\text{調査予定世帯数}} \cdot 2 & (\text{回収率 } 50\% \text{未満}) \end{cases}$$

$$\text{集計用乗率} = \text{補正係数} \times \text{調整済調整係数}$$

もし調査への参加がランダム（かつ回収率が50%以上）ならば調整済調整係数をウェイトとして用いることで人口分布を再現できる

しかし、特定の世帯タイプが調査へ参加しにくい可能性があるためこの偏りを補正係数を用いて修正する

ウエイトの作成方法

データ：2014年全国消費実態調査

2005と2010年の国勢調査を用いて予測した2014年
時点の世帯分布（予測方法については後述）¹

Step 0 調整済み調整係数をベースウエイト (ω_0) として用いる

Step 1 ウエイト ω を用いて1層目の変数で定義される人口分布が
国勢調査と合うようにウエイトを更新する。つまり、各セルで

$$\omega_1 = \frac{\text{国勢調査の世帯数}}{\text{全国消費実態調査の世帯数}} \times \omega_0$$

(分母は ω_0 で重みづけしたもの)

¹一般世帯、単身かつ通学中の世帯を除く。ウエイトに使う変数のいずれにも欠損値のない世帯のみを用いた。

ウエイトの作成方法

Step j ウエイト ω_{j-1} を用いて j 層目の変数で定義される人口分布が国勢調査と合うようにウエイトを更新する。つまり、各セルで

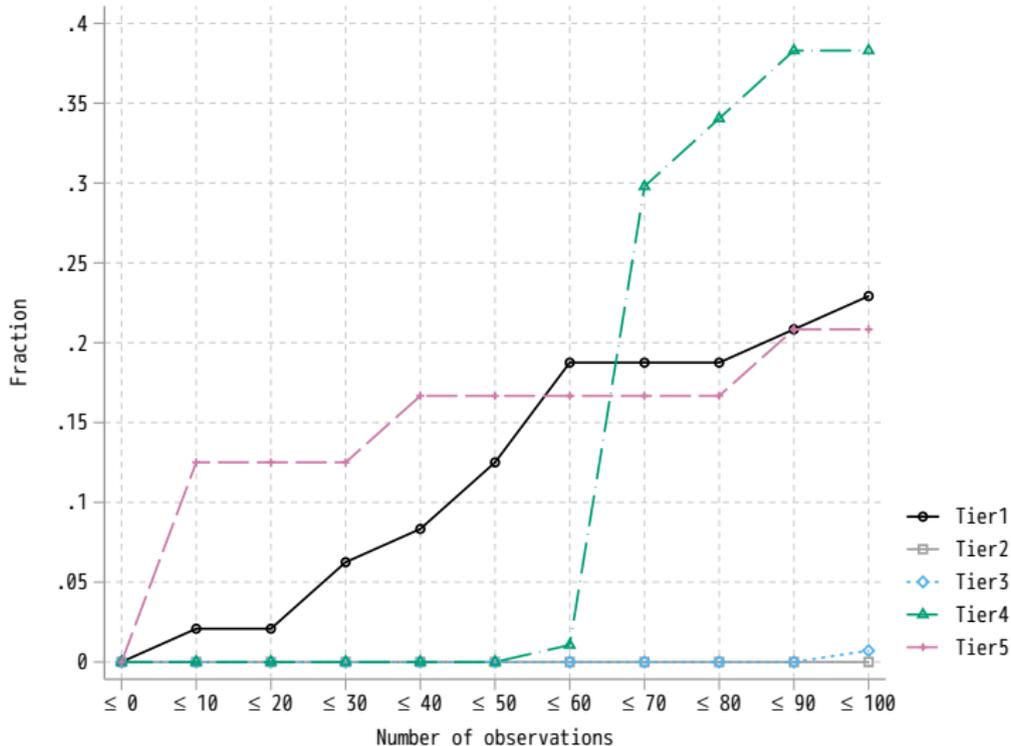
$$\omega_j = \frac{\text{国勢調査の世帯数}}{\text{全国消費実態調査の世帯数}} \times \omega_{j-1} \quad (\text{ただし、}\omega_0 = \omega_J)$$

(分母は ω_{j-1} で重みづけしたもの)

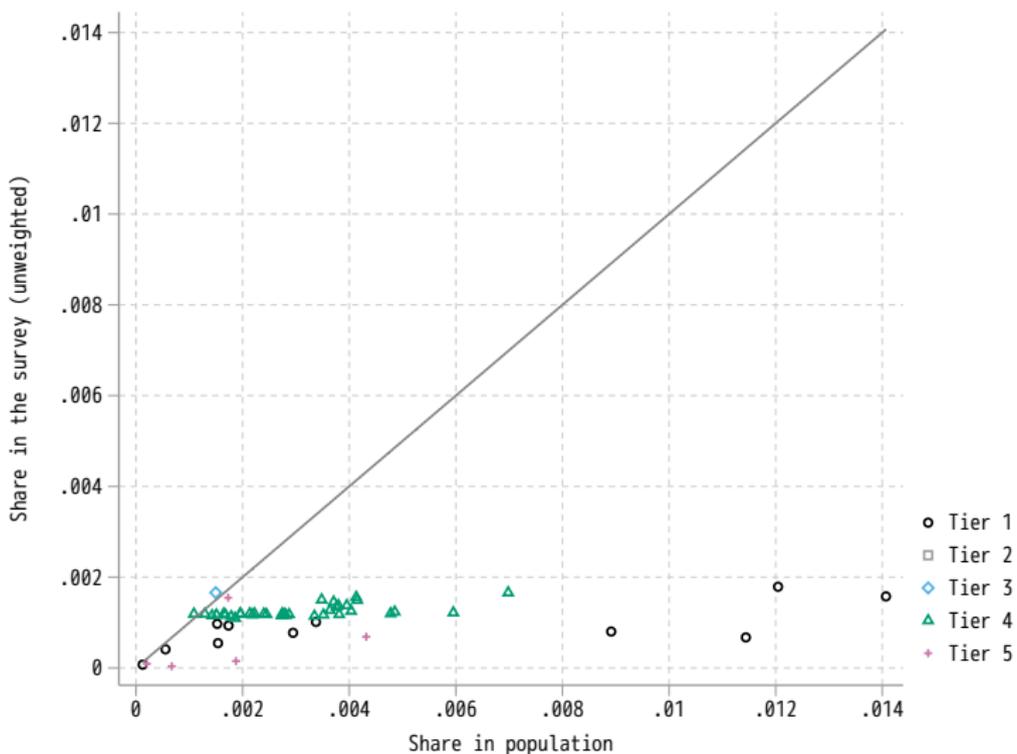
すべての層で周辺分布が国勢調査と一致するまで Steps 1- J を繰り返す
収束判定は

$$\|f_{j,\text{NSFIE}}(\omega) - f_{j,\text{Census}}\|_2 = \sqrt{\sum_{c \in \mathcal{C}_j} (f_{j,\text{NSFIE}}(\omega)(c) - f_{j,\text{Census}}(c))^2} < 10^{-8}$$

観測数の少ないセル：各層の総セル数に対する割合



観測数の少ないセル（100世帯以下）



観測数の少ないセル

1, 4, 5層で観測数の少ないセルがある

- 1層と4層の観測数の少ないセル
→ 調査の設計上、単身世帯が過少サンプルされているため
- 5層の観測数の少ないセル
→ 24歳以下で無職世帯・その他世帯、85歳以上で勤労世帯

ただし、上記のセルは国勢調査でも比較的小さく、特定のセルに過度に大きなウエイトを割り振ることは起きていないと考えられる

1. ウエイトの作成方法

2. 全国消費実態調査と国勢調査の時点調整

3. 基本属性の分布の確認

4. 世帯所得・消費分布

全国消費実態調査と国勢調査の時点調整

全国消費実態調査は2014年だが、ウェイト作成に使用できる直前の
国勢調査は2010年であるため調査時点に4年間のずれがある

→ この調査時点のずれを調整する必要がある

2005年と2010年の国勢調査から2015年の人口分布を予測し
その結果を用いて2014年全国消費実態調査のウェイトを計算する

世帯分布の予測方法

Simple prediction

各セル (a, g, f, e, p) について、セルの人数 $N_{agfep,t}$ の変化率が
2005~2010と2010~2015で同じだと仮定

$$\hat{N}_{agfep,2015}^s = \frac{N_{agfep,2010}}{N_{agfep,2005}} \cdot N_{agfep,2010}$$

- a : 年齢階級 (5歳刻み)
- g : 性別 (単身世帯のみ)
- f : 世帯人数
- e : 世帯主の就業状況
- p : 都道府県

世帯分布の予測方法

Regression-based prediction

2005年と2010年の国勢調査を用いて、グループごとに

$$N_{agfep,2010} = f_j(N_{agfep,2005}, N_{a-5,gfep,2005}) + \varepsilon_{agfep,t}$$

$$\hat{N}_{agfep,2015}^r = \hat{f}_j(N_{agfep,2010}, N_{a-5,gfep,2010})$$

(推定には2次の Taylor series を使用)

グループの定義

単身かどうか、10地域、世帯主年齢が65歳以上かどうか

(世帯主の就業状況や都道府県を考慮しても結果は変わらない)

世帯分布の予測方法

Hybrid prediction

Regression-based prediction では0以下の予測値や“非現実的”な予測値が出る場合があるので、それを Simple prediction で修正

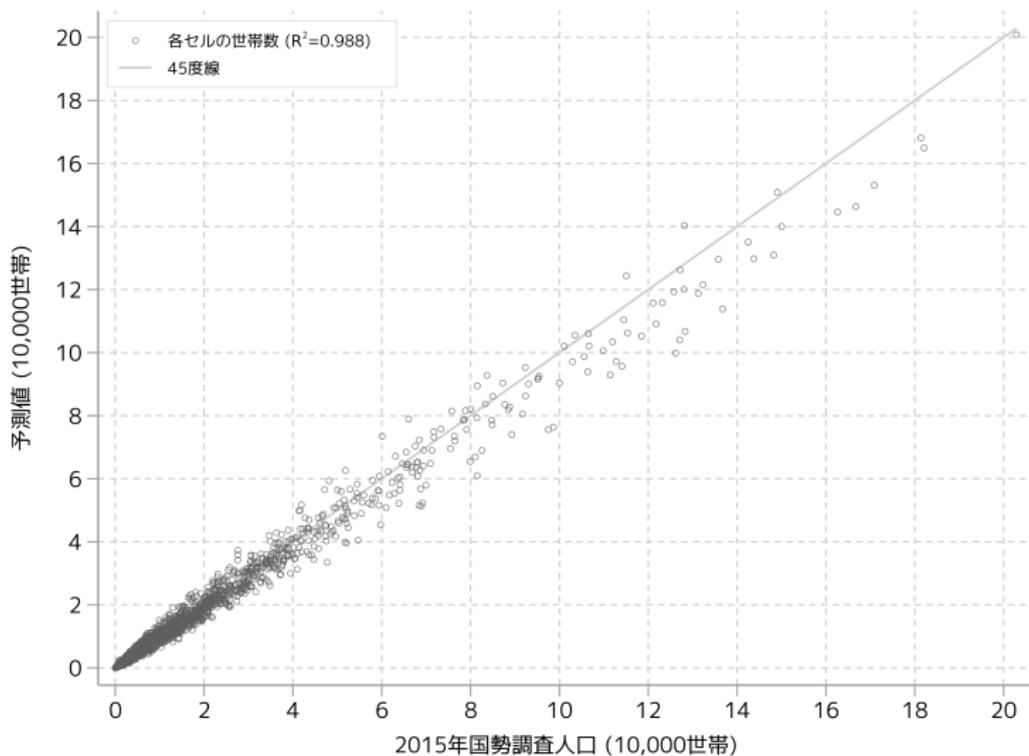
$$\hat{N}_{agfep,2015}^h = \begin{cases} \hat{N}_{agfep,2015}^r & \text{if } \hat{N}_{agfep,2015}^r > 0 \text{ and } \left| \frac{\hat{N}_{agfep,2015}^r - N_{agfep,2010}}{N_{agfep,2010}} \right| \leq 0.5 \\ \hat{N}_{agfep,2015}^s & \text{otherwise} \end{cases}$$

2014年の世帯分布の予測

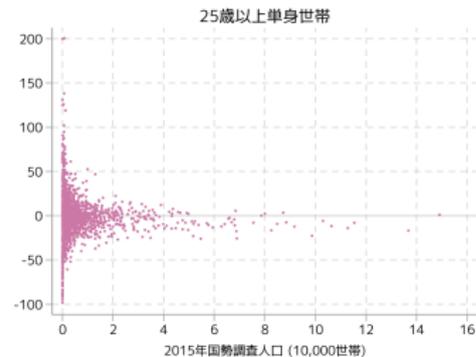
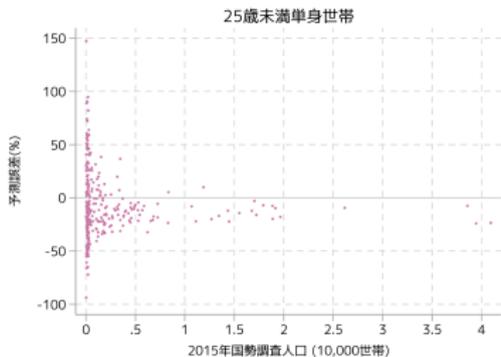
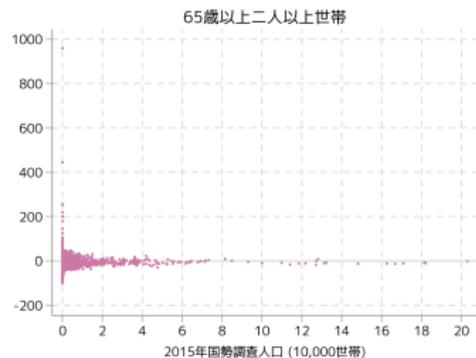
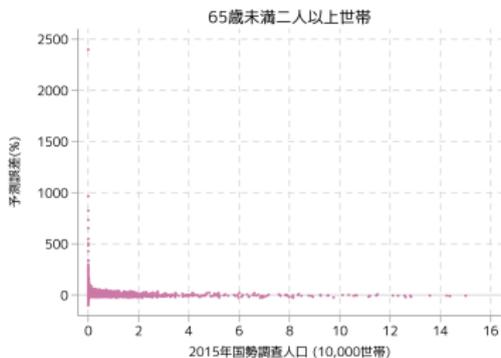
2010年の世帯分布と $\hat{N}_{agfep,2015}^h$ の線形補間によって、2014年の世帯分布の予測する

$$\hat{N}_{agfep,2014} = 0.2 \cdot N_{agfep,2010} + 0.8 \cdot \hat{N}_{agfep,2015}^h$$

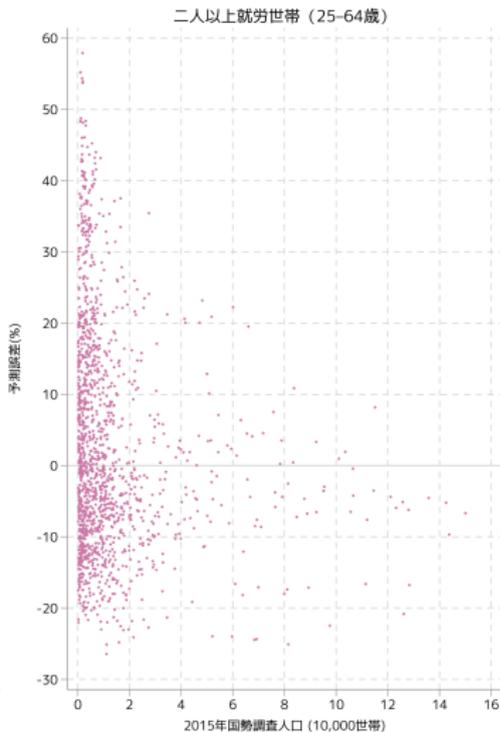
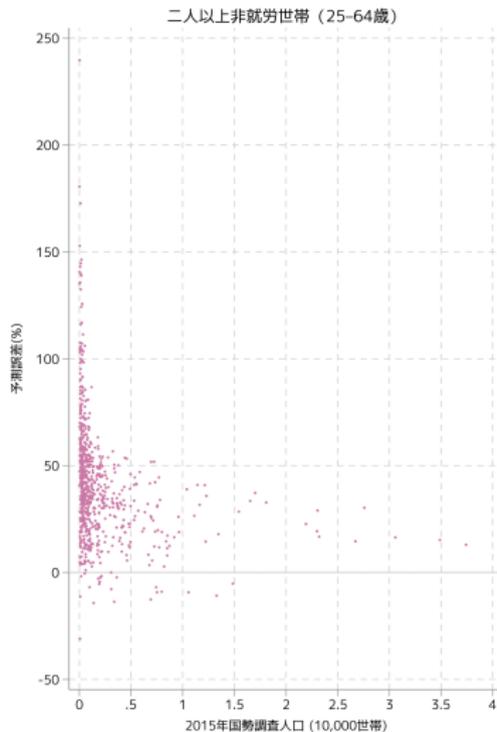
2015年の予測値と国勢調査の値の比較



2015年の予測値と国勢調査の値の比較



2015年の予測値と国勢調査の値の比較



2015年の予測値と国勢調査の値の比較

各セルにおける世帯数は、全体としては実際の値から大きく外れることなく予測することができる

ただし、世帯属性によって、過大もしくは過少予測する傾向

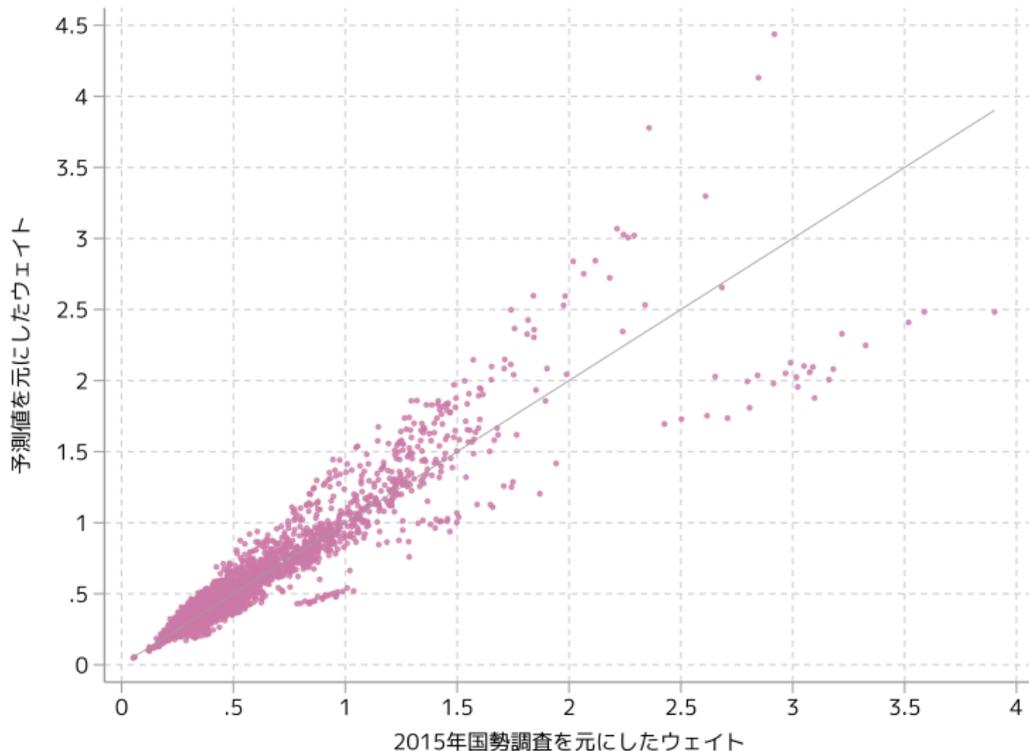
世帯構成

- 二人以上世帯を過大に予測する傾向
- 単身世帯（特に25歳未満）を過少に予測する傾向

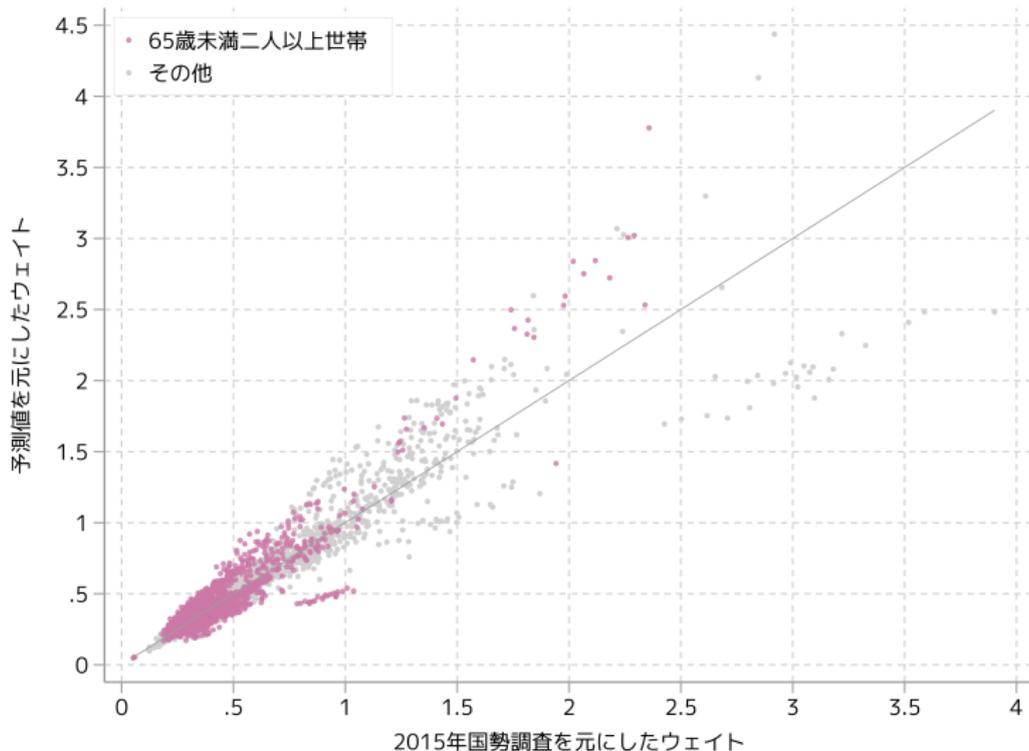
就業状況

- 非就労世帯を過大に予測する傾向
- 就労世帯を過少に予測する傾向

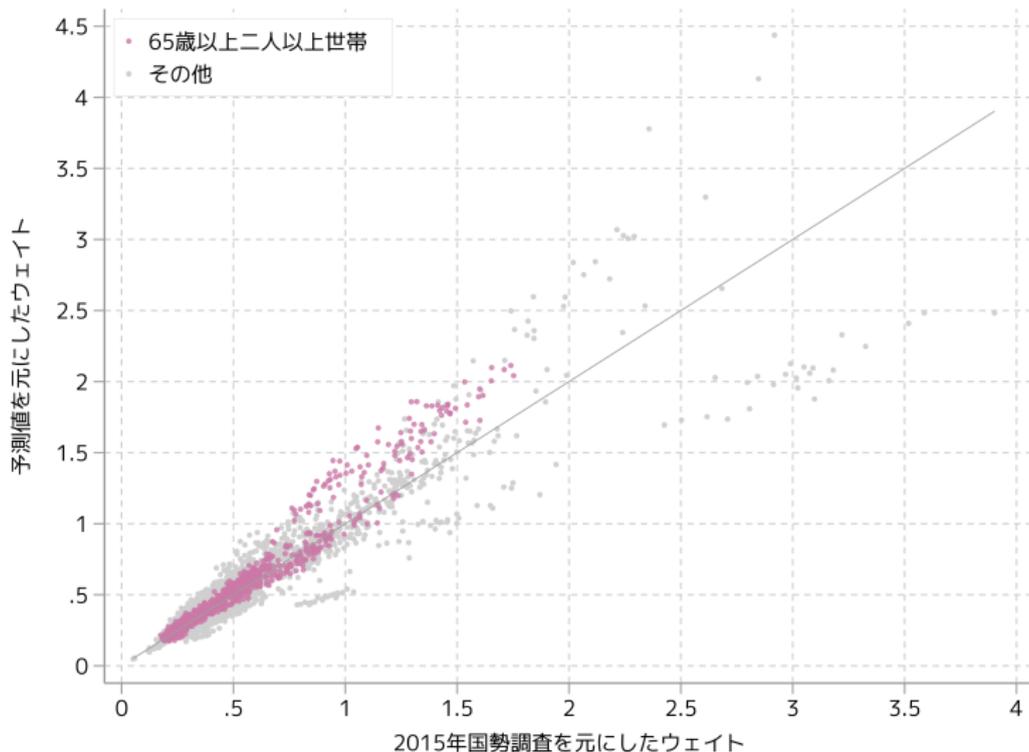
予測値と国勢調査の値ベースのウェイトの比較



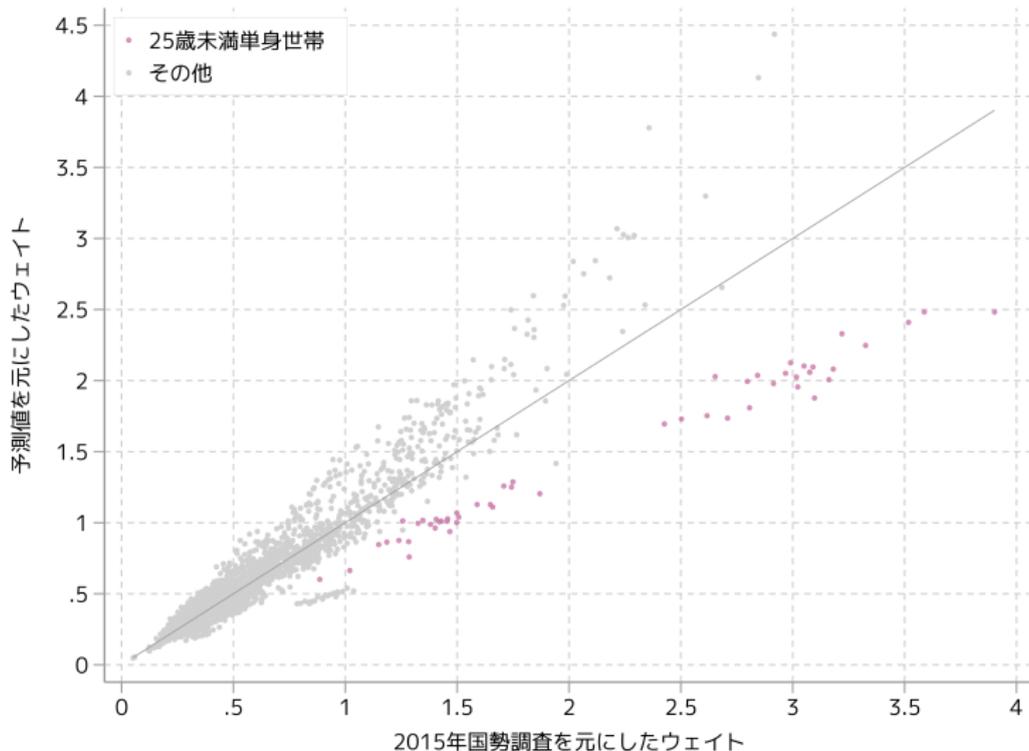
予測値と国勢調査の値ベースのウェイトの比較



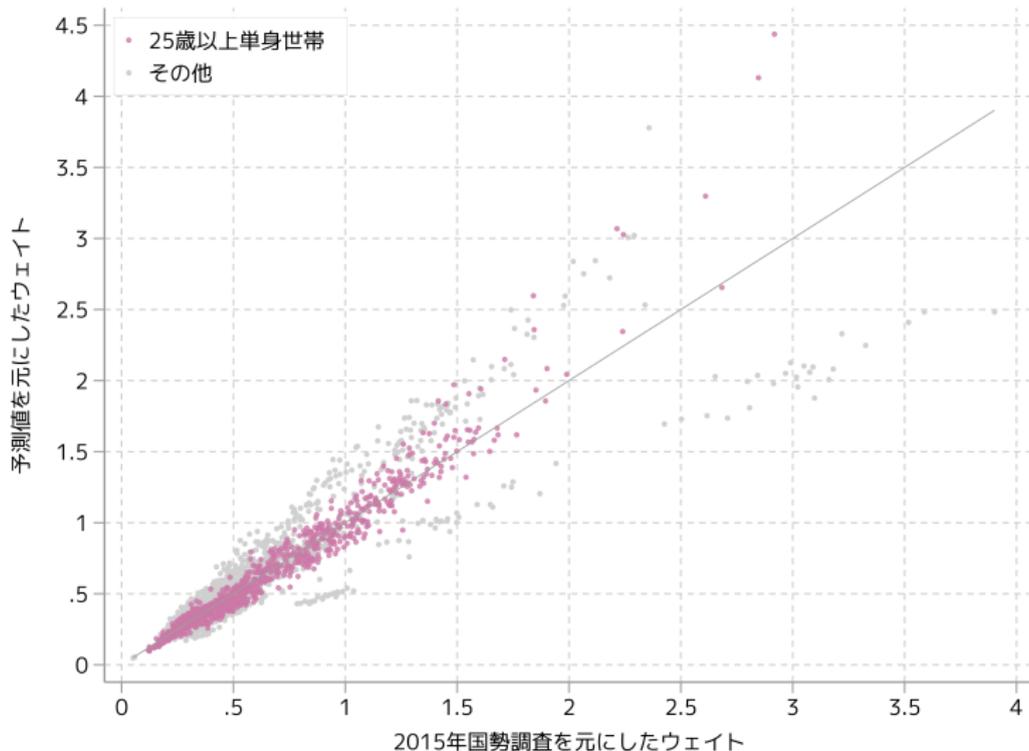
予測値と国勢調査の値ベースのウェイトの比較



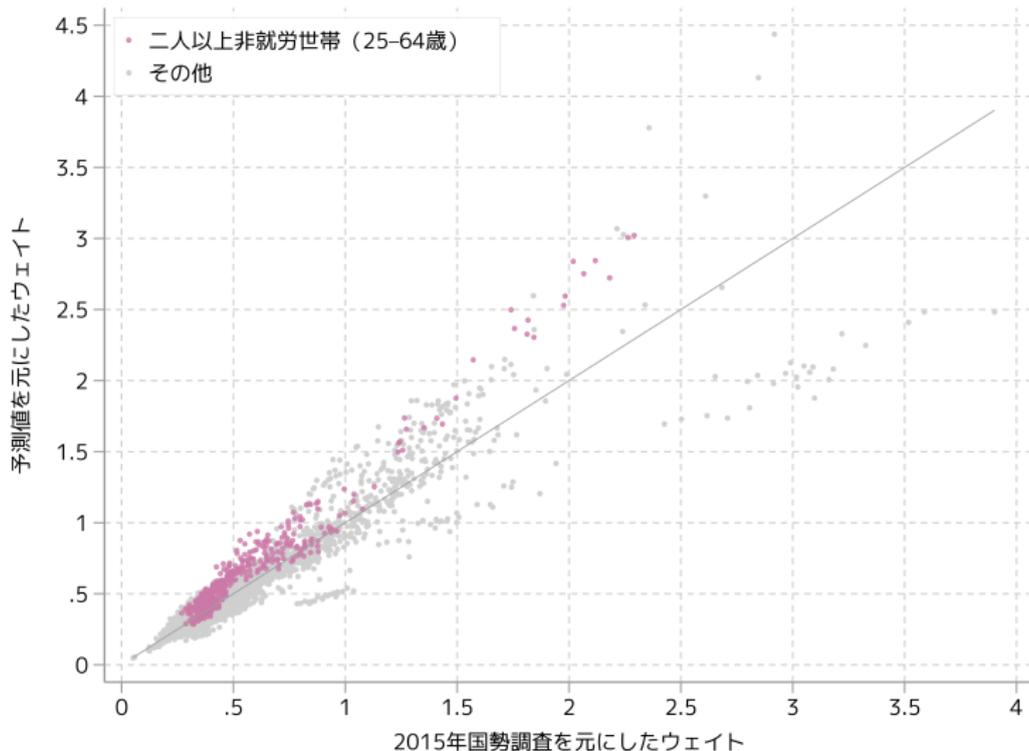
予測値と国勢調査の値ベースのウェイトの比較



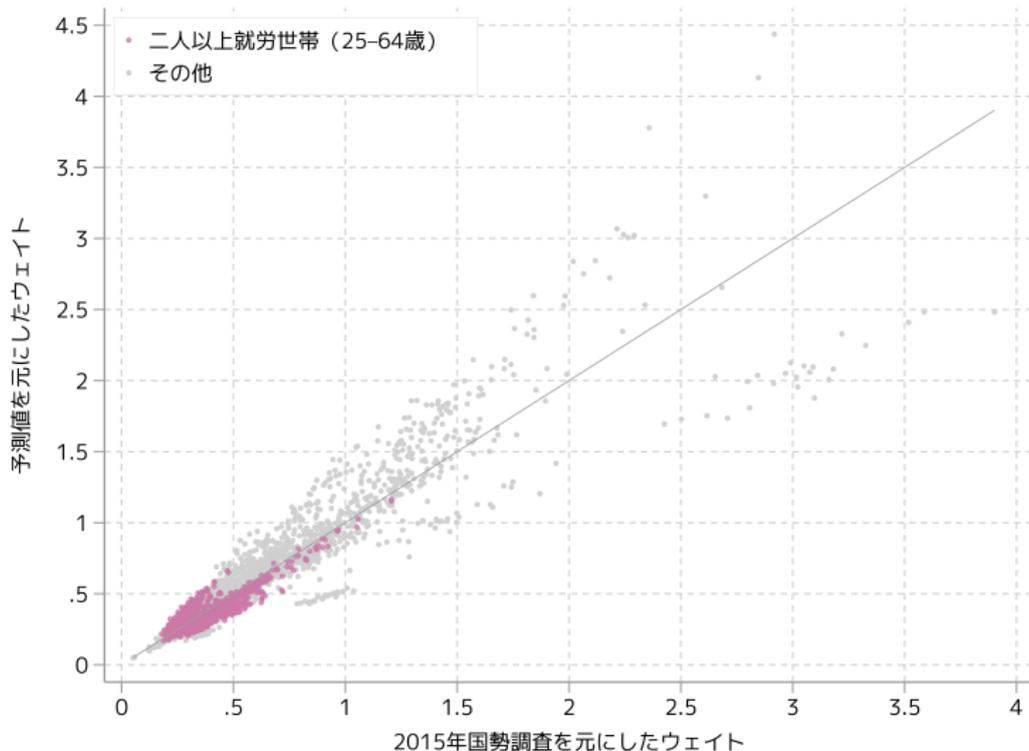
予測値と国勢調査の値ベースのウェイトの比較



予測値と国勢調査の値ベースのウェイトの比較



予測値と国勢調査の値ベースのウェイトの比較



予測値と国勢調査の値ベースのウェイトの比較

各セルの世帯数の予測と同様に、全体としては予測値を用いたウェイトは国勢調査の値を用いたウェイトから大きく外れることなく計算できている

ただし、世帯数の予測の乖離を反映して、ウェイトもセルの属性によって過大・過少評価される

世帯構成

- ・ 二人以上世帯の人口成長を過大に見積もる
- ・ 単身世帯（特に25歳未満）の人口成長を過少に見積もる

就業状況

- ・ 非就労世帯の人口成長を過大に見積もる
- ・ 就労世帯の人口成長を過少に見積もる

世帯属性による予測のずれの原因

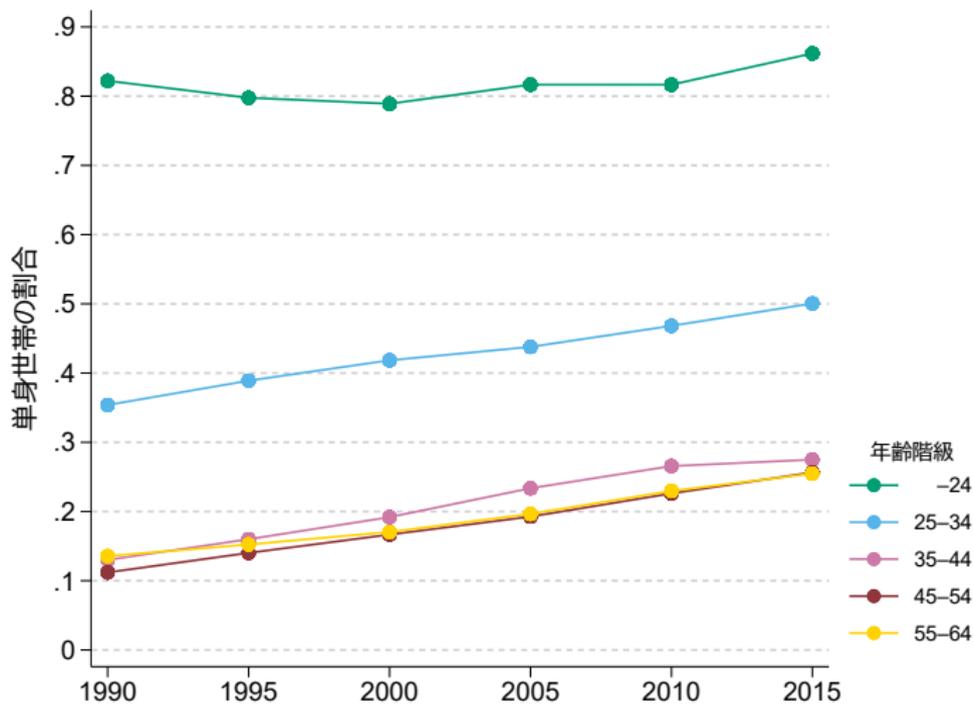
2005年と2010年の国勢調査から2015年の人口分布を予測
その結果を用いて2014年全国消費実態調査のウェイトを計算した

- 予測値を用いたウェイトは、全体的には国勢調査の値を用いたウェイトと近い値となっている
- 予測を外している原因は、主に世帯人数と就労からきている

世帯人数や就労状況の予測のずれがどこから発生しているかを
考える必要がある

- 単身世帯の増加を十分に予測できていない可能性
- 失業率の変化など、景気の変動を十分に予測できていない可能性

国勢調査の単身世帯割合の推移



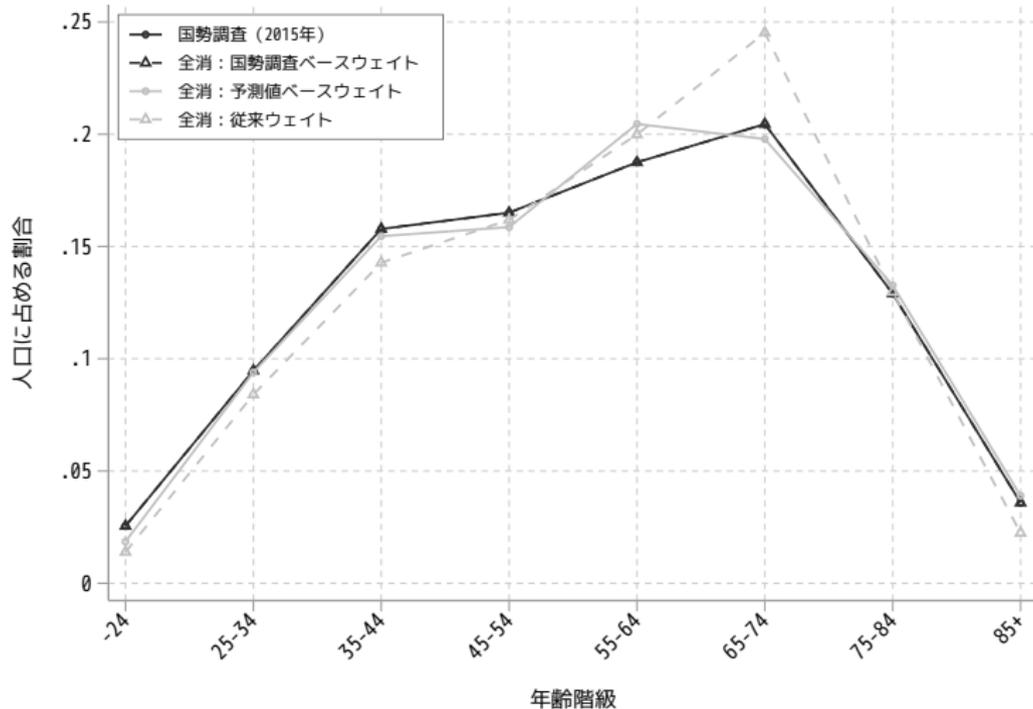
完全失業率の推移（労働力調査）



1. ウエイトの作成方法
2. 全国消費実態調査と国勢調査の時点調整
- 3. 基本属性の分布の確認**
4. 世帯所得・消費分布

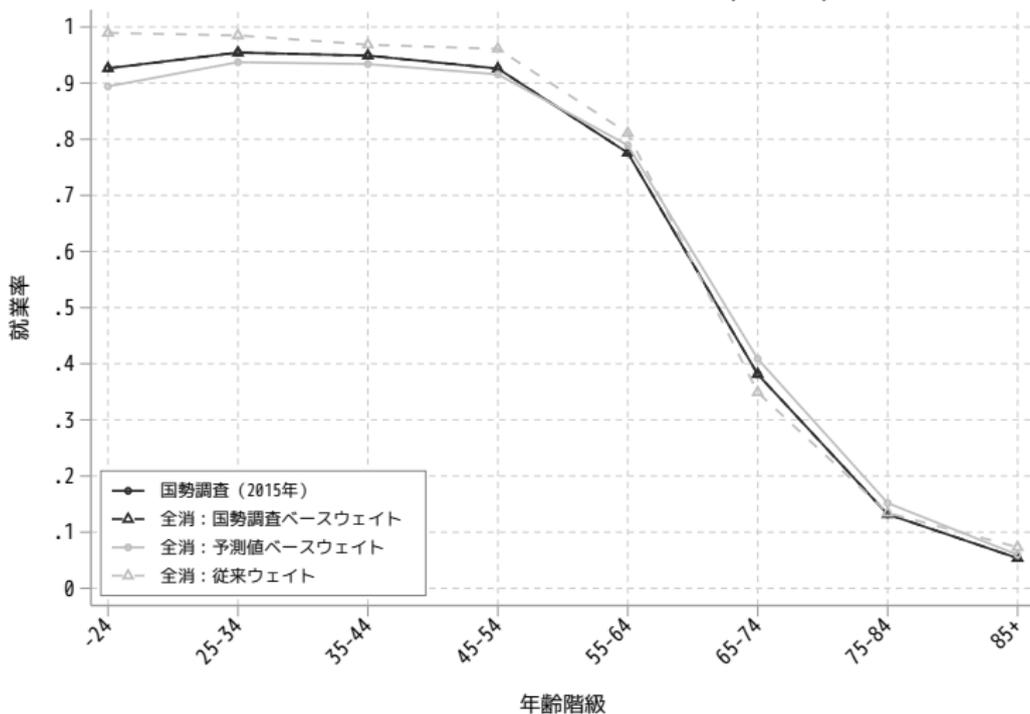
世帯主の年齢

国勢調査と全国消費実態調査の比較（2014年）



世帯主の年齢ごとの就業率

国勢調査と全国消費実態調査の比較（2014年）



世帯主の年齢・就業率

時点調整したウェイトを用いると国勢調査と比べて

- 世帯主の年齢が55～64歳の世帯が過大
- 若年・壮年世帯の就業率が過少

2015年の55～64歳の世帯の予測に、コホートサイズの異なる団塊世代の人口を使用していることや、不規則な景気変動を予測できていないことが原因の可能性

1. ウェイトの作成方法
2. 全国消費実態調査と国勢調査の時点調整
3. 基本属性の分布の確認
4. **世帯所得・消費分布**

年間世帯所得分布（1万円）

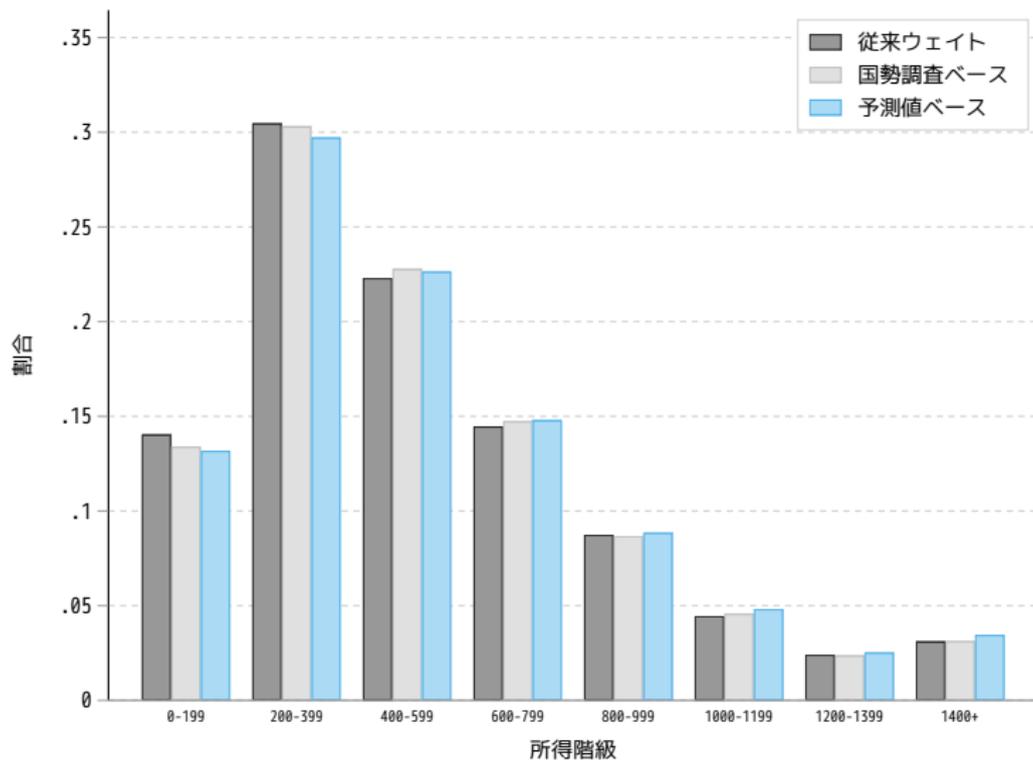


Table: 年間世帯所得の年齢プロファイル (1万円)

ウェイトの種類	従来ウェイト	国勢調査ベース	予測値ベース
全世帯	534	538	549
15-19	190	191	193
20-24	277	274	269
25-29	407	412	403
30-34	494	506	497
35-39	595	576	571
40-44	650	628	622
45-49	703	686	697
50-54	749	726	738
55-59	725	752	771
60-64	540	529	549
65-69	468	503	522
70-74	401	431	446
75-79	350	382	394
80-84	336	374	386
85-89	313	343	341
90+	328	367	379

年間世帯所得分布・年齢プロフィール

時点調整したウェイトを用いた平均年間世帯所得は2015年の国勢調査から作成したウェイトを用いたときと比較して2%高くなる

- 若年層では、所得は減少傾向
→ 非就労世帯数を過大に予測しているため
- 中年・壮年層では、所得は上昇傾向
→ 単身世帯数を過大に予測しているため

消費分布

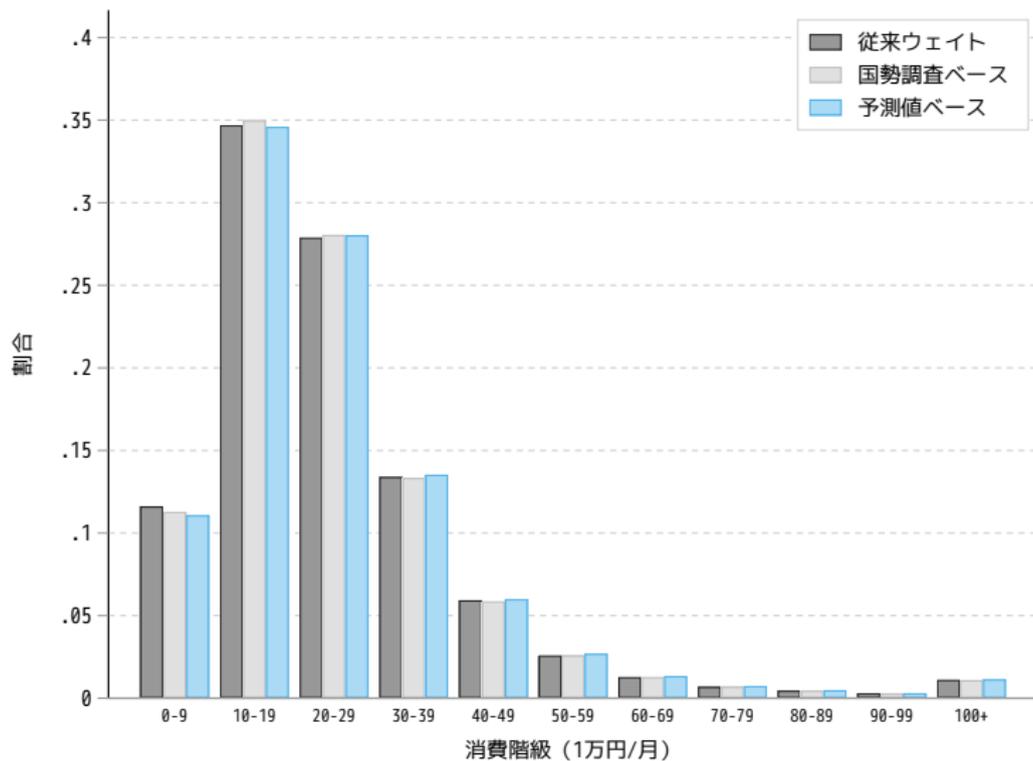


Table: 世帯消費の年齢プロファイル (1万円/月)

ウェイトの種類	従来ウェイト	国勢調査ベース	予測値ベース
全世帯	25.4	25.4	25.7
15-19	15.0	13.2	13.1
20-24	15.1	15.1	15.1
25-29	19.5	20.7	20.3
30-34	22.5	23.2	22.8
35-39	25.9	24.9	24.7
40-44	27.4	26.5	26.4
45-49	30.3	29.6	30.2
50-54	32.3	31.3	31.8
55-59	31.0	32.0	32.5
60-64	27.5	27.2	27.7
65-69	24.9	25.8	26.1
70-74	22.2	23.1	23.4
75-79	19.6	20.8	21.1
80-84	20.1	20.9	21.2
85-89	16.5	17.9	17.9
90+	16.7	18.0	18.1

月平均消費額・年齢プロフィール

時点調整したウェイトを用いても、平均消費額は大きく変化しない

- 所得の年齢プロフィールと同様、若年世帯で消費が減少
中年・壮年世帯で消費が増加という傾向がみられる
- 所得の変化と比べると消費の変化は小さい

住居の所有関係

ウェイトの種類	全国消費実態調査			国勢調査
	従来ウェイト	国勢調査ベース	予測値ベース	
持ち家	0.754	0.740	0.748	0.676
民営の賃貸住宅	0.161	0.174	0.170	0.219
公営の賃貸住宅	0.039	0.038	0.037	0.040
都市再生機構・公社等	0.013	0.013	0.012	0.015
社宅・公務員住宅	0.023	0.025	0.023	0.025
借間	0.006	0.007	0.007	0.009
寮・寄宿舍	0.003	0.003	0.003	0.017

共働き世帯の割合

以下を満たす世帯を共働き世帯と定義

- ・ 夫婦と17歳以下の世帯員のみで構成
- ・ 夫婦二人とも就業している

ウェイトの種類	全国消費実態調査				国勢調査
	ウェイトなし	従来ウェイト	国勢調査ベース	予測値ベース	
全世帯	0.221	0.160	0.166	0.159	0.165
二人以上世帯	0.235	0.233	0.231	0.219	0.231

高齢者世帯（夫婦ともに65歳以上の二人世帯）

ウェイトの種類	従来ウェイト	全国消費実態調査		国勢調査
		国勢調査ベース	予測値ベース	
割合	18.88	16.15	15.53	11.21
総支出	24.54	24.13	24.07	
食料	7.10	6.99	6.98	
住居	1.61	1.61	1.60	
光熱・水道	1.91	1.90	1.91	
家具・家事用品	0.97	0.97	0.97	
被服・履物	0.87	0.84	0.84	
保健医療	1.54	1.54	1.55	
交通・通信	3.03	2.93	2.91	
教育	0.00	0.00	0.00	
教養娯楽	2.98	2.89	2.88	
その他	4.52	4.45	4.43	
年間所得	434.45	424.14	426.82	

母子・父子世帯

ウェイトの種類	従来ウェイト	全国消費実態調査		国勢調査
		国勢調査ベース	予測値ベース	
割合	1.62	2.30	2.49	1.82
総支出	22.24	21.94	22.16	
食料	5.38	5.19	5.20	
住居	2.24	2.23	2.18	
光熱・水道	1.63	1.58	1.59	
家具・家事用品	0.63	0.63	0.63	
被服・履物	1.00	1.03	1.06	
保健医療	0.66	0.65	0.66	
交通・通信	3.33	3.34	3.41	
教育	2.04	1.96	2.01	
教養娯楽	1.96	1.92	1.96	
その他	3.38	3.42	3.45	
年間所得	387.74	370.05	370.94	