【資料１－２】

全国家計構造調査（旧全消）「年平均推定値」の推定方法について公表に向けて

2022年3月1日

慶應義塾大学大学院経済学研究科
／理化学研究所AIPセンター
松永将志・二荒麟
慶應義塾大学　経済学部／理化学研究所AIPセンター
星野崇宏

前回の「まとめと今後の課題」

【まとめ】全国家計構造調査の年次集計に家計調査を利用する場合、過去の家計調査と全消の乖離の原因を理解する必要

“同じ月の比較で全体として全消の方が大”

①標本の違い　　　収入　家計＜全消　⇒　全消の消費大

②季節性　　　　　全消の対象月は↓

③調査継続による効果　　　家計調査は6か月継続のため↓

本研究では①～③を考慮する方法を提案し合理的な値を推計

前回との違いは「共変量選択」「標準誤差」「検定」の実施

【課題】

(1)まだ一部費目(光熱費)のみ説明不能⇒2019年調査への適用？

(2)各費目の地域ごとの集計値の算出

(3)最終的に何を除去するのか？何を目標にするのか？についての議論が必要⇒何を考慮した“年間支出”か？

今後の方向性:研究及び公表

【研究として】

- 調査継続効果のさらなる詳細な分析とその理解
- 2019年からのネット調査世帯（調査継続効果が低下すると期待）と通常調査の比較

【公表のための方向性】

参考系列として公開する際には以下の通りとする

①明示的には「調査継続効果」を考慮しない年平均推定値の算出

＊計算の容易さと説明の問題

＊家計調査と家計構造調査での調査継続効果が同質でない可能性があるため

②10大費目×地域 10大費目×都道府県

または一部中分類費目（家賃など）×地域　　等

粒度を細かくできるかの検討

期待される効果

家計調査だけでなく家計構造調査のデータを利用することで

①10大費目×地域　10大費目×都道府県

　など粒度を細かくした年平均の算出

②標準誤差を減らすことが可能

今回の新しい方法では

ないが、家計調査と

家計構造調査の融合に

よって右図のように

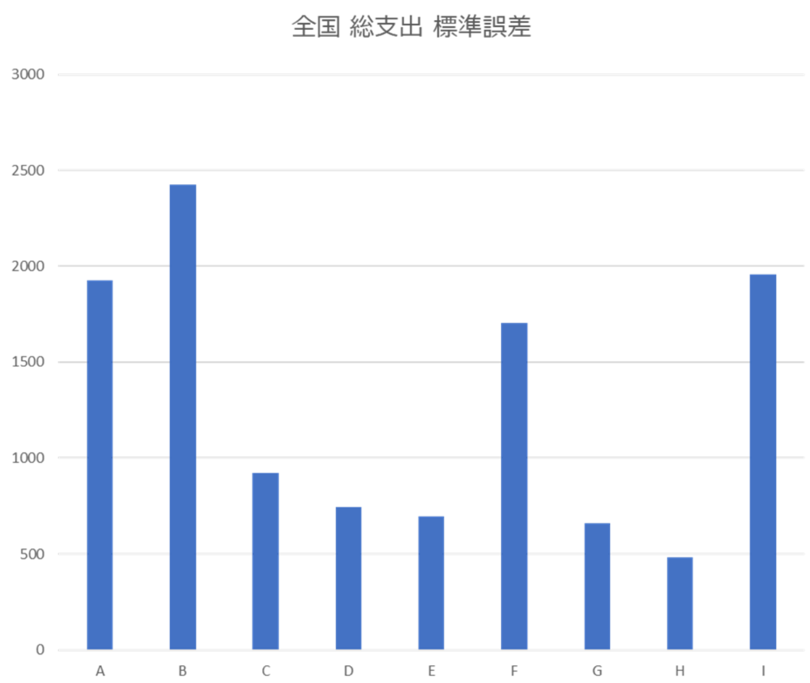
標準誤差を小さくすることが

可能になる

年平均の推定・標準誤差(2019年・総支出)

全国

(A)年推定 (調査継続効果平均)①
(B)年推定 (調査継続効果最小)②
(C)調査継続効果1,2か月目での
　年推定(回帰利用)③
(D)調査継続効果1,2か月目での
年推定(平均利用)④
(E)年平均 (家計)
(F)9,10 月平均 (家計)
(G)9,10 月平均 (構造)
(H)年平均 (合計＝家計と構造混合)
(I)年平均
　(家計 + 構造の効果
＝
　E＋（G－F））



（旧）属性情報による回帰とTobit-type2 model

家計iの当年t月のデータ（またはそれ+1の対数）を家計に関する属性情報(子供の数や世帯収入等)x、季節ダミーD、調査継続ダミーEを用いて下記で表現

$$y_{it} = \gamma^T x_i + \sum_{m=1}^{12} \alpha_m D_{it} + \sum_{k=2}^{6} \beta_k E_{itk} + \varepsilon_{it}$$

但し$D_{it} = \begin{cases} 1 & t = m \\ 0 & otherwise \end{cases}$, $E_{itk} = \begin{cases} 1 & \text{家計}i\text{が}t\text{月に}k\text{か月目} \\ 0 & otherwise \end{cases}$

⇒回帰による回答者の違いの考慮

加えて

$$y_{it}^* = \lambda^T x_i + \epsilon_{it}$$

を背後に考え

教育費などゼロが多い項目のことも考慮してyではなくvが観測されるとする。

$$v_{it} = \begin{cases} y_{it} & if \;\; y_{it}^* \geq 0 \\ 0 & if \;\; y_{it}^* < 0 \end{cases}$$

（新）属性情報による回帰とTobit-type2 model

家計iの当年t月のデータ（またはそれ+1の対数）を家計に関する属性情報(子供の数や世帯収入等)x、季節ダミーD、調査継続ダミーEを用いて下記で表現

$$y_{it} = \gamma^T x_i + \sum_{m=1}^{12} \alpha_m D_{it} + \varepsilon_{it}$$

但し$D_{it} = \begin{cases} 1 & t = m \\ 0 & otherwise \end{cases}$

⇒回帰による回答者の違いの考慮

調査継続効果を考慮しない
Tobit-type2モデル

加えて

$$y_{it}^* = \lambda^T x_i + \epsilon_{it}$$

を背後に考え

教育費などゼロが多い項目のことも考慮してyではなくvが観測されるとする。

$$v_{it} = \begin{cases} y_{it} & if \;\; y_{it}^* \geq 0 \\ 0 & if \;\; y_{it}^* < 0 \end{cases}$$

（新）分析で出した指標

C. 「全国消費実態調査・家計構造調査を１年継続して調査した場合」の年平均

調査継続効果を考慮しない

$$\frac{N_{Z10}\bar{y}_{Z10}}{N}+\frac{N_{Z11}\bar{y}_{Z11}}{N}$$
$$+\sum_{m\neq 10,11}^{12}\frac{N_{Am}}{N}\frac{1}{\#(y_{Zi}>0)}\sum_{i:y_{Zi}>0}\left[\hat{\alpha}_m+\hat{\gamma}^T x_{Zi}+\hat{\sigma}\hat{\lambda}_{mills,Zi}\right]\frac{\#(y_{Zi}>0)}{N_Z}$$

＊ N_{Zm}, N_{Am}は全国消費実態調査・家計構造調査と家計調査のm月の人数,

＊ N_Z は全国消費実態調査・家計構造調査の人数,

＊ N_Aは家計調査の10,11月以外の月の人数,

＊ $N=N_Z+N_A$,

＊ $\bar{y}_{Zm}$は全国消費実態調査・家計構造調査のm月平均．ここで,

$$\frac{1}{\#(y_{Zi}>0)}\sum\nolimits_{i:y_{Zi}>0}\left[\hat{\alpha}_m+\hat{\gamma}^T x_{Zi}+\hat{\sigma}\hat{\lambda}_{mills,Zi}\right]\frac{\#(y_{Zi}>0)}{N_Z}$$

は$E[\alpha_m+\gamma^T x_{Zi}+\sigma\lambda_{mills}|y_{Zi}>0]P(y_{Zi}>0)+0*P(y_{Zi}=0)$の推定値

議論

（まとめ）

地域別にすると同期間で家計調査と家計構造調査では一定以上の乖離

⇒提案手法では平準化される

ただし一部の項目で上振れの傾向

→こちらについては原因究明中(Tobitの推定の問題か？)

（今後）

・より詳細な区分についての算出

・標準誤差の計算

⇒4月中には確定させたい