

総務大臣
武田良太殿

統計委員会委員長
北村行伸

諮問第147号の答申

労働力調査、就業構造基本調査、全国消費実態調査及び社会生活基本調査に係る匿名データの作成について

本委員会は、諮問第147号による労働力調査、就業構造基本調査、全国消費実態調査及び社会生活基本調査に係る匿名データの作成について審議した結果、下記のとおり結論を得たので、答申する。

記

1 計画の適否

令和3年1月27日付け諮問第147号「労働力調査、就業構造基本調査、全国消費実態調査及び社会生活基本調査に係る匿名データの作成について」（以下「本計画」という。）は、総務省統計研究研修所における検証結果を踏まえて審議した結果、以下のとおり、調査回答者の匿名性及び学術研究等における有用性が確保されるものと認められることから、本計画で、平成25年から令和元年までの労働力調査、平成24年及び平成29年の就業構造基本調査、平成21年及び平成26年の全国消費実態調査並びに平成23年及び平成28年の社会生活基本調査（以下「4調査」という。）の匿名データを作成することは、適当である。

2 理由等

(1) 上限値、下限値の設定等

本計画では、4調査における調査項目の上限値、下限値の設定等について、これまでに作成した労働力調査等の匿名データにおける匿名化措置の一部を変更して、以下のとおり措置することとしている。

ア 世帯員年齢のトップコーディング（一定の値を上限値とし、それを上回る場合に上限値以上でまとめる措置をいう。以下同じ。）は、85歳以上を90歳以上に変更

イ 就業構造基本調査における就職時期及び就業期間のトップコーディングは、世帯員年齢のトップコーディングの変更に合わせ、就職時期は70年前以上を75年前以上に、就業期間は70年間以上を75年間以上に変更

ウ 社会生活基本調査における末子年齢のリコーディング（分類の程度を粗いものにする措置をいう。以下同じ。）及びトップコーディングは、1～11歳の間でリコーディング及び12歳以上でトップコーディングを、1～9歳各歳をそのまま提供、10～17歳の間でリコーディング、18歳以上をトップコーディングに変更

エ 全国消費実態調査における住居等面積、年間収入、貯蓄現在高及び借入金残高のトップコーディング及びボトムコーディング（一定の値を下限値とし、それを下回る場合に下限値以下でまとめる措置をいう。）は、全国一律を地域区分ごとに変更
これらについては、データの分布状況の変化や他の統計調査の匿名データにおける匿名化措置を踏まえたものであり、有用性が高まることから、適当である。

(2)情報の削除

本計画では、社会生活基本調査の調査票Aにおいて、子どもの数の多い世帯及び父子世帯については、同調査の調査票Bにおける措置と同様に、地域区分ごとに出現率の少ない世帯を削除することとしている。

これについては、匿名性を確保する一方で、有用性の観点からは地域区分ごとに削除する措置が講じられており、適当である。

(3)新規追加調査項目の匿名化措置等

本計画では、4調査における新規追加調査項目について、以下のとおり措置することとしている。

ア 就業構造基本調査の匿名化措置

- ・卒業時期は、就学区分（小中高大など）と組み合わせることにより、世帯員の年齢を推測することが可能となるため、平成24年調査では昭和58年～平成24年の間で、平成29年調査では昭和63年～平成29年の間で、5年ごとにリコーディング
- ・雇用契約更新回数は、1～9回はそのまま提供、10回以上は、出現率が低くなるものがあるため、10～14回、15～19回でリコーディング、20回以上をトップコーディング
- ・平成24年調査における東日本大震災の仕事への影響に関する調査項目は、対象となる世帯が限られており、当該項目を複数の項目と組み合わせることにより、世帯を特定されるリスクが高まるため、当該項目は提供しない。

イ 全国消費実態調査の匿名化措置

- ・育児休業取得期間は、30か月まではそのまま（1か月単位で）提供、31か月以上は、出現率が低くなるものがあるため、トップコーディング
- ・平成26年調査における被災に関する調査項目は、対象となる世帯が限られており、当該項目を複数の項目と組み合わせることにより、世帯を特定されるリスクが高まるため、当該項目は提供しない。
- ・平成26年調査における建物の階数は、外観識別性が高い等のため、建物全体の階数は1～2階、3～5階、6～10階、11～14階でリコーディング、15階以上をトップコーディング、世帯が住んでいる階数は1～2階、3～5階、6～10階でリコーディング、11階以上をトップコーディング
- ・品目の購入地域は、項目の特性等を踏まえ、当該項目は提供しない。

ウ その他の新規追加調査項目については、項目の特性やデータの分布状況を踏まえ、そのまま提供することとしている。

ア及びイについては、匿名性を確保するために適当である。また、ウについては、匿名性が確保されるとともに有用性が高まることから、適当である。