

第2回 評価分科会 議事概要

1 日 時 平成31年3月12日（火）13:30～14:50

2 場 所 総務省第2庁舎6階特別会議室

3 出席者

【委員】

西郷 浩（分科会長）、北村 行伸（分科会長代理）

【臨時委員】

岩下 真理、神林 龍、美添 泰人

【審議協力者】

総務省統計局統計調査部経済統計課調査官

財務省大臣官房総合政策課経済政策分析官

文部科学省総合教育政策局調査企画課課長補佐

厚生労働省政策統括官付参事官付統計企画調整室室長補佐

農林水産省大臣官房統計部企画管理官補佐（統計調整班担当）

国土交通省総合政策局情報政策課建設経済統計調査室室長

日本銀行調査統計局経済統計課統計総務グループ

東京都総務局統計部調整課課長代理

神奈川県統計センター企画分析課課長

【事務局（総務省）】

横山大臣官房審議官

統計委員会担当室：櫻川室長、阿南次長、上田次長、肥後次長、福田補佐

4 議 事

- （1）建設工事統計調査（施工調査）の欠測値対応の見直しについて
- （2）来年度以降のテーマについて
- （3）その他

5 議事概要

- （1）建設工事統計調査（施工調査）の欠測値対応の見直しについて

国土交通省より、資料1「建設工事施工統計調査における欠測値補完の見直しについて（案）」に基づき、説明が行われた。

行政情報活用による回収率向上検討の必要性等いくつか指摘があったが、基本的な方向性については了解を得た。

委員等からの主な発言等は以下のとおり。

- ・立派な分析なので、国交省の中でこのような技術を是非後任の方にも継承して欲しい。
 - ・平成2年度非回答業者のフォロー調査の際、どういう根拠で欠測値は調査結果にほとんど影響を与えないという結論を出したのか分かるか。当時は影響が無く今回は影響があるとなったのが、調査の仕方の違いによるのか、世の中の実態が変わったのかは大事な問題。更に学ぶべきところがあるかもしれない。今後のためにもきちんと調べておけば蓄積になる。後で調べて教えて欲しい。
- 平成2年度の実態調査は、非回答業者に実績があるか無いかのみを問い合わせているものと思われる。大部過去のことで手元に資料が無いが、確認して見る。
- ・資料 p.4 で、経済センサスの回答とアンマッチが生じている要因は何か。建設許可番号の記入が無く、名称、所在地でもマッチできなかったためか。
- センサス回答で建設許可番号を記載してもらうことになっていたが、記載が無かったり、誤記があったりしたためのものである。
- ・回収率向上は本当に大事だが、報告者負担の大きい調査項目を削除するのはいけないと思う。行政資料の活用がポイント。行政資料であるものはそれを活用する、無いものでどうしても欲しいものは、負担を御願ひするしかない。問題を切り分けれないといけない。行政情報の活用をここに書いて欲しい、回収率の向上も図れるし、必要な情報もとれる。
- 行政情報活用検討の必要性は御指摘の通り。建設業許可手続きの際にいくつかの経営に関するデータが業者から提出される、それを活用して報告者の負担を減らすことができるという御指摘については、検討しているが、許可の多くは知事許可なので都道府県から情報を入手する必要があるという問題がある。省内許可部局と相談しながら検討を進めたい。
- ・ウエイトの修正の仕方について少々お聞きしたい。7ページの表をみるとセルを切ったときに、それぞれのセルでは差の検定をして平均値の差が回答事業所と非回答事業所で分かれる、違うという箇所が散見される。この表全体をみて、回答事業所と非回答事業所の差がないと判断をして、すべてのセルにおいて回答事業所の分布と非回答事業所の分布の平均値の差がない、ウエイトだけを修正するという判断をなさったと思う。セルを切っているのに、平均値が違う・かい離していると判断されたセルに関しては、バイアスの修正もするという選択肢があったと思うのだけれども、それはなぜしなかったのか。
- 復元をした時の数字(p.11)の規模感が合っているというところがあったので、この点についてこれ以上細かな検証はしていない。

- ・検定値（推定値）の $W \cdot X$ の X の部分が、各セルの平均の工事高になると思う。これが回答事業所だけのデータを使って X を復元しているけれど、これが例えば総合工事事業の500万から1000万の法人であれば、（非回答事業所と回答事業所のどちらがどうなっているかわからないのだが）平均値に差があると（国土交通省は）判断されている。この差の部分を X で修正することはできると思うのだが、それはどうしてやらなかったのか。全体でバイアスがないと判断できるなら、このウエイトだけを修正すればよい。けれども、バイアスがあると判断できる場所もあると思うので、これはロジックがうまく通っていないところもあると思う。
 - ・影響はP値だけで見てはいけない。全体の推定量でどれだけ出るかということ。この検定は層がちょっと細か過ぎる。これだけ細かいP値を出しても統計学的には意味が無い。今回は簡便な方法を使って補完をただけでも、p.11のように大幅な改善が図られている。今後の検討課題にさせていただきだとは思いますが、思うけれども。
- 確かに、層を細かく分けていくとサンプル数が少なくなってしまう。これから検討していく際に御指摘いただいた点については考慮していきたい。
- ・未回答がでたときに、こういう形でセル別にグループを分けて各セルのなかで未回答がどれくらいあるのか、バイアスがあるのかを検証してそれを復元するという方法と、そうは考えずにばらっと全サンプルでどれくらい平均的に未回答のバイアスがあるのかという検証の仕方の2つある。どちらかというとな後の方が簡便で、ある程度仮定をおけば一致すると思う。他の統計への応用ということを考えると、どちらがよいのかということを検討してもらえるとよいのではないかな。
 - ・5 ページ当たりで非回答について分析されているのは貴重だが、一番大事なのは、非回答事業者の属性がいったいどうだったのか、規模と売上高、そのデータも持っているはずなので、もう少し分析していただくと、今の指摘にも対応して、今後のブラッシュアップにも通じると思う。
 - ・未回答のバイアスを考えるときに、回答事業所と非回答事業所のサンプルを全部集めてきて、単に regression（回帰分析）のみで回答・未回答がどれくらい平均の差がでるのか。コントロール変数を作って、それだけで平均的なバイアスだけを計算して、あとは修正するとなれば、計算手順としてはかなり楽になると思う。それが、セル別にバイアスをいちいち計算して、積み上げていくのと、どれくらい最終的に結果が違ってくるのかということを出せば、とても役に立つと思う。他の統計で、ここまでフォローアップして、ある意味全数がバックグラウンドでわかっている統計は、あまりないのではないかな。復元の方法・未回答の扱いのやり方によって、どの程度違ってくるのかということが、ある程度わかると、他の統計にも役に立つのではないかな。

- ・ 欠測値補完については、経産省は大規模な調査と分析を長い間行っている。また、財務省も法人企業統計について長らく検討を行っている。そういった情報をまとめて、各省で情報を共有できるような体制を作るためには、この場がふさわしいのではないか。各省で今までいくつかの方法を適用して改善できたけれど、まだ改善できていない問題点が残っている、このような技術的情報を共有することが大事だと思う。この場をそういう目的にも使っていただきたい。
- ・ 未回答事業者の実績の有無を、調査票の工夫により見分けることができないか。調査技術としては、ゼロだから書かないという場合と、ゼロではないのだが忙しかったりして書かないという場合とは、なるべく調査票の段階で区別できた方がいい。後から見分けるというのではなくて、実績が無い場合には、実績無しで回答してくださいという質問項目を調査票の最初に設け、実績が無い場合でも調査票を返してもらえるようにする方が望ましい。実績がゼロで回答しない場合と、実績はあるが何らかの理由で無回答という場合は、できるだけ区別ができるように調査票を作っていくことが基本。検討の余地があれば検討してもらいたい。
- ・ p.3 で、平成 2 年度実績と平成 23 年度実績を見比べると、実績無しで回答無しの事業者数はあまり変わらず、平成 2 年度において回答していた 1 万 2 千ほどの会社が、23 には実績有りで回答無しに動いているようなイメージ。前であれば答えていた事業者が平成 23 年度になって答えていないという状況の様に見える。何らかの形で実績が無い場合は無しと回答してもらえるようにする仕組みが必要だと思う。
- ・ 平成 2 年度と 23 年度の間で相当廃業があると思うが、どのくらいのスケールであるのか。
 - 廃業は多い年で平成 14 年度 4 万 2 千業者、15 年 1 万 4600、16 年 1 万 4400、17 年度はまた増えて 4 万 482 と年によって増減があるが、そのような規模で出ている。
- ・ 大きいところは固定されていて、規模の小さい事業者が入れ替わっているということか。
 - 廃業事業者の資本金階層別数というのはちょっと手元に数字が無いが、感覚的には規模の小さいところが廃業の割合が高まるといった感じである。
- ・ 調査票の工夫で、実績ゼロで無回答ということと、実績があるのだが何らかの理由で無回答というところが区別できるようにしておくのが基本。もし検討の余地があれば検討してもらいたい。
- ・ 全般的には評価の高い取組。今後の取組として、行政記録情報の活用等注文もあったが、おおよそここに書かれている通りで進めていただければ。特に、無回答の事業所の属性についてはもう少し詳しく調べていただいて、それに応じて、神林先生から御提案のあった回帰分析のようなことを考えたらいいか、

層別を細かくして対応した方がいいのか、非線形の状況が起きているかとか、そういったことも検証していただくと、この調査そのものへのプラスの影響もあるし、他の調査への情報提供にもなる、よろしく御願いたい。

(2) 来年度以降のテーマについて

西郷分科会長から、資料2「第1回評価分科会にて各委員より御提案いただいた検討テーマ案」に基づき、説明が行われた。

来年度以降の分科会における検討テーマ案について検討が行われ、「①欠測値への対応におけるスタンダード（モデル）の検討」及び「③ローテーション・サンプリングの有効性の検討」にまず着手すること、「②指数のあり方についての技術的な検討」については分科会長と事務局の間で進め方を整理することとなった。

委員等からの主な発言等は以下のとおり。

- ・事務局がどのくらい対応できるのかが問題だ。欠測値補完については、1983年に米国の研究者と実務家による議論を整理した3巻本（Incomplete Data in Sample Surveys, Volumes 1, 2, & 3）という基本的な文献があり、以前、私の研究会でも西郷先生たちとともに呼んだ経験がある。しかし、それを事務局に訳してもらうことを考えると対応が難しいと思う。もう少し負担を軽くして、これまで各省で検討されてきたことをまとめて紹介してもらうというのであればできるだろうが。
- ・ローテーション・サンプリングについては、最近、日本統計学会の特別セッションのために私が整理したものがあるので、紹介することはできる。
- ・指数については、対象が広いので、何をどう検討するのか絞り込みが必要。
- ・取り上げるべき議題は、まずは①（欠測値への対応におけるスタンダード（モデル）の検討）、次に③（ローテーション・サンプリングの有効性の検討）。②（指数のあり方についての技術的な検討）と④（回収率と精度の関係）についてどのように検討していくかを事務局と私で考えさせていただく。

(3) その他

- ・事務局から、資料3に基づき、説明が行われ、当面の評価分科会の検討の進め方について確認された。
- ・次回の評価分科会の場所と日程は改めて調整する旨、事務局から案内された。

以上