

第5回評価分科会 議事録

1 日 時 令和2年1月27日（月）10:00～12:16

2 場 所 総務省第二庁舎6階特別会議室

3 出席者

【委員】

椿 広計（分科会会長）、岩下 真理（分科会会長代理）

【臨時委員】

久我 尚子、山本 渉、美添 泰人

【専門委員】

神林 龍

【審議協力者】

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）、経済産業省大臣官房調査統計グループ統計企画室参事官補佐、国土交通省総合政策局情報政策課課長補佐、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長

【説明者】

内閣府：経済社会総合研究所研究官室上席主任研究官

総務省：統計研究研修所統計研修研究官ほか

財務省：財務総合政策研究所調査統計部調査統計部長、国税庁長官官房企画課課長補佐ほか

国土交通省：総合政策局情報政策課交通経済統計調査室課長補佐ほか

【事務局（総務省）】

山田統計企画管理官

統計委員会担当室：櫻川室長、栗原次長、鈴木次長、上田政策統括官付参事官、福田補佐、増成補佐

4 議 事

（1）分科会長の互選、分科会長代理の指名について

（2）当面の検討の進め方について

（3）精度検査報告書※提言に対応した取組について（法人企業統計調査、民間給与実態統計調査、造船造機統計調査、自動車輸送統計調査及び建築着工統計調査の補正調査）

(4) 諸外国における欠測値補完及び総務省による各府省の統計作成支援について

(5) 欠測値への対応に関する内閣府の研究成果について

※平成28年度統計法施行状況に関する審議結果報告書(統計精度検査関連分)(平成30年3月統計委員会)

○櫻川総務省統計委員会担当室長 皆様、おはようございます。若干定刻より早いですけれども、皆様おそろいになりましたので、ただ今から第5回評価分科会を開催させていただきます。

私は、本分科会の事務局を務めます総務省統計委員会担当室長の櫻川です。分科会長を選任いただくまでの間、議事の進行を務めさせていただきますので、よろしくお願ひします。

評価分科会は、統計改革推進会議の最終取りまとめにおいて、統計委員会の必置機関として自律性・中立性を確保した評価チームを設置することとされたことを受けて、平成30年に統計委員会令の改正により設置されました。

また、資料1のとおり、10月14日付で内閣総理大臣から評価分科会に属すべき委員、臨時委員及び専門委員の指名が行われています。

さらに、オブザーバーとして関係府省、日本銀行、地方公共団体から東京都と神奈川県の方にも御参加いただくこととしています。

続きまして、分科会の会長を御選任いただきたいと存じます。これについては、統計委員会令第1条第4項により委員の互選によることとされています。どなたか御推薦などはございますか。

岩下委員、どうぞ。

○岩下委員 僭越ではございますが、統計技術、そして、調査実務に精通なさっています椿委員を力強く御推薦したいと思ひます。

○櫻川総務省統計委員会担当室長 ただ今椿委員を分科会長に推薦する御意見ありましたが、皆様、いかがでございましょうか。

(「異議なし」の声あり)

○櫻川総務省統計委員会担当室長 御異議がないようですので、椿委員に分科会の会長をお願いしたいと存じます。

それでは、恐縮ではございますが、椿委員には分科会長席、こちらになりますけれども、お移りいただきまして、就任の御挨拶をいただければと存じます。

○椿分科会長 ただ今分科会長に選任いただきました椿と申します。

統計の理論とか実務ということに関して、少しの経験はあるところですが、こういう話を今後、徐々に政府全体の統計の中で実現していく、ちょっとずついい方向にしていくことについては、いろいろな方々の広い意見の集約のもとで実現できるのではないかと考えていますので、御審議、その他もろもろの意見について積極的に御協力いただければと思ひます。

私、統計委員会は久しぶりでして、なかなか雰囲気とパターン、昔のことになれてしま

っているのではないかと思いますので、是非忌憚なく、運営に関しても御意見頂戴できればと思います。どうぞよろしくお願い申し上げます。

○櫻川総務省統計委員会担当室長 それでは、以後の進行は椿分科会長にお願いしたいと存じます。

椿分科会長、お願いします。

○椿分科会長 どうもありがとうございます。

それでは、本日用意されている資料につきまして、事務局から簡単に御説明いただければと思います。よろしく申し上げます。

○福田総務省統計委員会担当室長補佐 では、お手元の資料について確認させていただきます。

まず、評価分科会に属すべき委員、臨時委員及び専門委員についての資料が資料 1。

それから、当面の評価分科会の検討の進め方（案）についての資料が資料 2。

精度検査報告書において、平成 30 年度までに実施すべきとされた事項についての関係府省の取組の現状（財務省、国土交通省関係）についての資料が資料 3。

法人企業統計調査の欠測値補完についての資料が資料 4。

民間給与実態統計調査における欠測値補完等についての資料が資料 5。

造船造機統計調査における調査対象事業所の整理についての資料が資料 6。

自動車輸送統計調査における欠測値補完に関する取組等についての資料が資料 7。

建築着工統計調査補正調査の見直しについての資料が資料 8。

諸外国における欠測値補完についての資料が資料 9。

各府省の統計作成支援のための業務相談窓口についての資料が資料 10。

「欠測値補完に関する調査研究」についての資料が資料 11 です。

その他、参考資料 1 として、統計委員会令。

参考資料 2 として、統計委員会運営規則。

参考資料 3 として、欠測値への対応に関する各府省研究成果。

参考資料 4 として、欠測値への対応に関する各府省等職員研究発表の紹介。

参考資料 5 として、平成 28 年度統計法施行状況に関する審議結果報告書（統計精度検査関連分）。

参考資料 6 として、令和元年度統計委員会評価分科会審議結果報告書（第 1 回～第 4 回審議分）。

参考資料 7 として、第 4 回評価分科会議事概要を準備しています。

資料は以上です。

○椿分科会長 資料の過不足ございませんでしょうか。よろしいでしょうか。

それでは、次に、分科会長代理の指名を行いたいと思います。

参考資料の 1 を御覧いただければと思います。

統計委員会令第 1 条第 6 項におきまして、「分科会長に事故があるときは、分科会に属する委員のうちから分科会長があらかじめ指名する者が、その職務を代理する」とされています。

そこで、私から分科会長の代理を指名したいと思います。

私といたしましては、岩下委員に分科会長代理を強くお願いしたいと思いますけれども、岩下委員、よろしいでしょうか。

○岩下委員 了解いたしました。よろしくお願いします。

○樫分科会長 では、恐縮ですが、よろしくお願いします。

それでは、議事に入りますけれども、最初に、今後の検討の進め方ということで、資料2に基づきまして、これも事務局から御説明いただければと思います。

栗原次長、よろしくお願いします。

○栗原総務省統計委員会担当室次長 それでは、事務局から説明させていただきます。資料2の方を御覧いただければと思います。

当面の評価分科会の検討の進め方ということで、進め方について御確認させていただきたいと思います。

1番、当面の検討課題のところですが、(1)といたしまして平成28年度統計法施行状況に関する審議結果報告書(統計精度検査関連分)(平成30年3月統計委員会)。以下、精度検査報告書と申し上げますけれども、そちらにおきまして、30年度まで実施すべきとされた事項、具体的には事業所・企業を対象とします基幹統計調査における欠測値への対応につきまして、関係府省の取組を聴取して、統計技術の観点から評価を行うというのが1点です。

それから、2点目、こちらは欠測値への対応に関します各府省、それぞれ独自の研究などの成果の共有化ということで、こちらは審議というより、そういうものを御紹介いただくというものです。

2番のスケジュールの方、本日が第5回でございまして、精度検査報告書のフォローアップとして財務省の統計と国土交通省の統計を取り上げます。

それから、各府省の研究成果の共有の1回目といたしまして、記載のものについて、今回、御紹介いただきます。

続きまして、第6回が2月19日ですが、フォローアップの2回目としまして、厚生労働省の統計と、それから、各府省の取組の共有化の2回目ということで、記載のとおりものを予定してございます。

そして、第7回、3月25日で、今回の審議の結果の取りまとめと、来年度検討テーマの検討ということで予定させていただいているところです。

簡単ではございますが、以上です。

○樫分科会長 どうもありがとうございました。

ただ今の御説明につきまして、何か質問等あれば、よろしくお願いします。いかがでしょうか。特によろしいでしょうか。

それでは、当面の評価分科会の検討の進め方につきましては、資料2のとおりということで御了解いただいたということにします。ありがとうございました。

それでは、次の議事に入りたいと思います。

先ほど御了承いただきました当面の評価分科会の検討の進め方ということに従いまして、

まず、精度検査報告書のフォローアップを行いたいと思います。

まず、法人企業統計調査の取組につきまして、事務局及び財務省様から説明を聴取したいと思います。それでは、お願いします。

○栗原総務省統計委員会担当室次長 それでは、資料の3番の方を御覧いただければと思います。

こちらは、本日の審議いただきます財務省、それから、国土交通省の統計の関係につきまして整理したものです。

1枚めくっていただきまして、まず、財務省の法人企業統計調査です。

左から3列目のところに、精度検査報告書におきましてこういう指摘がありますという記載が書いてございます。

それから、その右側の欄が取組の現状ということで、それぞれの省庁の方で取組を行った結果の概要がまとめてございます。

まず、私の方から報告書におきます指摘事項を簡単に御紹介した上で、各府省の方から取組の状況につきまして詳しい説明ということで、資料4以下の該当する資料を使って説明していただくというスタイルで考えていますので、よろしくをお願いします。

まず、法人企業統計調査ですが、こちらは、一部非回答に対しまして0値補完を行っている統計調査ということになっています。

売上等の主要項目の記載があることを前提に、最終的に一部未回答の状態となっており、発生している数十社程度に対しまして、当該一部非回答の項目に0値補完をしているということで、集計上0として処理をしているということです。

当該対応につきましては、過小推計につながる懸念もあるということから、第Ⅲ期の基本計画におきまして指摘されています欠測値の補完方法の改善方法の検討に、その0値補完の検証も含めることが必要であるというような指摘がなされているところです。

これについての説明の方、それでは、財務省の方からよろしくをお願いします。

○巴財務省財務総合政策研究所調査統計部調査統計部長 よろしくをお願いします。財務省です。

それでは、資料4、法人企業統計調査の欠測値補完について御説明させていただきます。

1ページおめぐりいただきますと、調査の概要がございまして、法人企業統計調査ですが、四半期別調査と年次別調査という2本に分かれておりまして、いずれも我が国の営利法人等の企業活動の実態を明らかにすることなどを目的として実施されているものです。

まず、四半期別調査ですが、国内に本社のある資本金1,000万円以上の営利法人等から約3万2,000社を標本抽出しています。

一方、年次別調査につきましては、国内に本社のある全営利法人等から約3万6,000社を標本抽出しています。

標本の抽出方法につきましては、中ほどにありますとおり、資本金階層別、業種別に層化し、金融業、保険業を除く全産業については、資本金5億円以上は全数抽出、資本金5億円未満につきましては、各階層ごとに等確率系統抽出により標本を抽出しております。

調査事項ですが、こちらは、お手数ですが、資料の一番裏側の調査票の様式というのが、四半期別のがついています。それで御紹介したいのですが、大きく3つのパートに分かれています。ちょっと字が小さくて恐縮ですが、まず、1つ目が、調査票の左側から真ん中ぐらいいかけまして、資産、負債、固定資産の増減など、貸借対照表、BSの項目を調査しておる部分が1つございます。

2つ目の部分は、一番右側の方の真ん中ぐらいい、4ポツで何々中の損益というところがございまして、売上高、損益と、損益計算書、PL項目を調べる部分がございまして。

それから、もう1つ、3つ目のパートが、その下、右の一番下です。5ポツで人件費と書いてあるところございまして、給与、賞与など、人件費項目を調査してございまして。

今回、議題となっています一部非回答項目への0値補完というのは、この5ポツの一番右下の人件費の部分にだけ適用になっているということ、まず、御説明したいと思います。

済みません、また資料を戻していただきまして、2ページを御覧ください。

精度検査報告書に関する取組状況の概要ですが、回答値が0である場合と欠測値の場合のデータにつきまして、現在、両者を区分しておりませんが、今後、必要な予算措置、それから、システム改修を行った後、両者を区分して把握することとします。

それから、2点目、また、学識経験者等を交えました研究会におきまして、一部非回答項目、これは先ほど申しましたとおり、人件費のところですが、人件費に対する0値補完について検証を行いましたところ、0値補完による影響は軽微であるとの結論に至りました。後ほど詳細に説明させていただきます。

3ページを御覧ください。フローチャートが出ていますが、これは回答値が0というものと欠測値の区別について事務処理をフローであらわしたものです。

まず、調査票の審査段階におきまして、調査項目に0との回答、または未回答がある場合、他の項目との整合性があるかないかをまず見ます。

整合性がある場合、例えば0として扱ってもBSとかPLの計算が合う場合がございます。こうした場合は、0という処理が通常の正しい0という形で行われます。

他方、その合計値が合わなくなるような場合がございます。そういう場合や、それから、あとは人件費、PLとかBLから独立している項目につきましては、企業側に照会を行います。企業側の照会に対する回答の結果、数値が0として正しい場合は、これもまた正しい0として扱われるということです。

他方、企業側から照会に対する回答が得られなかった場合は0として補完処理、集計上は0として処理されてしまうということです。

ただし、この表は一部欠けているところがありまして、BS、PL項目につきまして、整合性がなくて、かつ、企業側から回答がない場合は0値補完ではなくて、集計外という処理をします。済みません、その部分はこのフローチャートから省略されておりまして非常に分かりづらくなっていますので、この場で補足させていただきます。申しわけございません。

今回の検証ですが、BS、PL項目は整合性がとれているという、それを前提として、

独立項目である人件費が企業側から回答が得られずに0と補完処理されている点、調査期ごとに件数は異なりますが、おおむね数十社程度ですが、こちらについて他の補完方法と比較することにより、過小推計となっていないかにつき、検証を行ったというところです。

4ページが検証方法ですが、現行の0値補完、これをAとしていますが、それ以外のものとして2パターンを実施してございます。

平均値型のC、それから、対数回帰型のEというものをを用いて補完を行っています。

Cの平均値型ですが、一部未回答法人と業種、それから、資本金規模が同じで、かつ、記入がある提出法人の平均値を利用して補完を行うとした場合がCです。

対数回帰型のEですが、一部未回答法人と製造業、非製造業の産業区分が同じ提出法人の売上高の対数を説明変数とする単回帰式による補完を行った。これがEのバージョンです。

これらの補完方法につきまして、未提出法人の欠測値補完の際、補完された提出法人を用いる場合をバージョン1、それから、用いないものをバージョン2として、それぞれ2通りで3種類ずつ、計6パターンを試算してみました。それが、次のページ、5ページです。

5ページは、現行の補完方法、現行がA2という、黒い実線が現行の補完方法ですが、これと他の補完方法との推計値の比較を行ってみました。

御覧になってお分かりになるとおり、役員給与グラフの方はほぼ同じような動きになっています。

従業員給与グラフ、これは多少推計値の差はございますが、ほぼ同じような伸びになっています。

ちなみに人件費の項目はこれ以外もあるのですが、代表的な、一番使われるのがこの2本ですので、この2本を出させていただきました。

それから、6ページですが、今、5ページで使った推計値を伸び率で比較したものでございまして、こちらにつきましては、役員給与、それから、従業員給与ともに、補完方法による差はほとんど見られないというところです。

7ページ、結論ですが、以上を踏まえまして、有識者も交えました研究会における検証の結果ですが、1つ目として人件費項目の未回答法人数は数十社程度であり、それほど大きくないこと。

2つ目として現行の補完方法でも伸び率はゆがみなく推計できていること。

以上のことから、現行方式の人件費項目に対する0値補完については、特段の問題はないという結論に至ったものです。

次に、せっかくの機会ですので、8ページ以降で、御参考までに、平成30年度統計法施行状況報告において報告いたしました欠測値補完のこれまでの検討、それから、実施状況について若干御紹介させていただきます。

8ページの方が結論の概要ですが、結論といたしましては、年次別調査の審査事務等にEDINET情報を一層活用すること。

それから、2つ目は、欠測企業の補完に当たっては、過去データを利用することが有効。

ただし、対象とする過去データの範囲について引き続き研究が必要なことを報告してごさいます。

それぞれ若干補足させていただきます。9ページを御覧ください。

9ページは、EDINET情報が各省の統計でどんなふうに使われているのかというのを調べたものですが、いずれの統計調査もEDINET情報等は審査事務における計数照会での活用が主というものでございました。

10ページは、EDINET情報を、もう一步踏み込んで、我々のデータベースに流し込めないかというのを検討したのですが、まず、課題として、一番上にごさいますとおり、私どもの法人企業統計の対象は法人の単体決算ですが、EDINETの四半期報告は連結情報のみということで、なかなか使えないということ。

それから、2つ目にごさいますように、全ての調査項目に関する情報は取得できなかったというようなことが分かりまして、システム的にデータを取り込むということについてはまだまだ研究が必要だという結論に至ったところです。

他方で、データの流し込み、取り込みは無理でも、EDINET情報をこれまで以上に審査に活用できないかということで、EDINET情報を閲覧するための市販ソフトというものがございまして、それが使えないかということを検討してみました。

その結果、複数法人のデータを一括してエクセルで閲覧できるということが分かりまして、今現在我々がしている、企業のホームページを当たって、一件一件PDFで確認していることをしていますが、それと比べて作業時間を短縮できるのではないかという結論に至りました。

結果として、そのためにEDINET閲覧ソフトを年次別調査の審査、照会事務の効率化に活用し、調査結果の精度向上に取り組むことといたしました。

これが、EDINET情報の活用です。

11ページですが、これは今回の検証対象となっています現行の補完方法について示したものでございまして、詳細については割愛しますが、欠測法人と同業種で、なおかつその資本金規模の近い10社の計数を基にした補完がなされるような設計となっているというのが現在の補完方法です。

12ページです。12ページは、学識経験者による研究会において、欠測値補完方法について更なる研究を行ったものです。

学識経験者の皆様からは、欠測値の補完について、現状の補完方法をベースとしつつも、欠測企業の過去の提出データを利用することも有効ではないかという意見が出されまして、シミュレーションを行って検討いたしました。

具体的には、提出があった法人のデータを基に、提出法人の真値を擬似的に欠測させて、その欠測部分を現行の補完方法で埋めた方法と、当該法人の過去データを利用して埋めた方法と、どちらが真値との差が小さいのか大きいのかというのを比較したところです。

それが12ページの左側の図ですが、こちら見ますと、過去データを利用した補完方法の方が真値との絶対差が小さい法人が多くて、その有効性が認められたということです。

この資料では設備投資というのを例に出してございまして、他の調査項目につきまして

も過去データの有効性というのは確認できたところです。

他方、過去データを持ってくるときに、対象とする過去データの範囲をどこまでにするかという点がございまして、12ページの右の図では、検討段階の図を置いてございますが、例えば年次別調査ですと、スタートから下に矢印があって、前年のデータを探します。前年がない場合は前々年のデータを探して、それがなければ現状の補完方法でやろうというものです。

それから、四半期別の方は、同様に前年同期比のデータを探して、それがなければ前々年同期比のデータ、それがなければ、直近の、前期のデータを使ったらどうか。それがなければ今の補完方法でやろうという案となっておりますが、使用するデータの範囲につきましては、皆様からいろいろ御意見が出て、正直、決め手がないところでございまして、引き続き研究を続けているというところです。

以上、駆け足ですが、最近の動きも含めて欠測値の補完方法について御説明させていただきました。

○樫分科会長 どうもありがとうございました。様々な対応を研究されているということが非常によく分かりました。

ただ今の御説明につきまして、何か御質問あるいは御意見などがあれば、是非よろしくお願いします。

神林専門委員、よろしく申し上げます。

○神林専門委員 質問が2点ございます。1点目は、グラフの読み方なのですが、6ページの伸び率というのは、5ページの全産業の平均値の伸び率。それで、伸び率の平均ではないですねというのが1つです。

もう1つは、2ページのサマリーなのですが、今の最後のところ、0値補完による影響は軽微であるとしてあるのですが、その0値補完による影響というのは何に対する影響なのかがちょっと分からないので、質問したいと思います。2点です。

○樫分科会長 御回答よろしく申し上げます。

○佐藤財務省財務総合政策研究所調査統計部調査統計部電子計算システム課長 6ページの伸び率ですけれども、5ページでお示しした推計値についての前年同期比を表したものです。

0値補完による影響は、全体推計値に対する影響です。

○神林専門委員 特定の推計値ではなくて、統計全体の推計値ですか。

○佐藤財務省財務総合政策研究所調査統計部調査統計部電子計算システム課長 そうです。母集団推計値ということになります。

○神林専門委員 ありがとうございます。

○藤原財務省財務総合政策研究所調査統計部調査統計部調査統計課長 補足しますと、今回の課題が先ほど事務局から御説明があったように、0値補完をすることが過小推計につながっているのではないかというものでございまして、それに対する影響というのを見させていただきました。

○美添臨時委員 あまり時間もないので、余計なことは言いませんが、私はこの統計に、

かれこれ 50 年ぐらい関わってしまして、学識は怪しいけれども、経験者として参加しています。

神林専門委員の御質問はもつともで、マイクロデータを公表する場合には個別の欠測補完が大事ですけれども、今回求められているのは欠測値が全体の推計に与える影響の評価なので、推計値に関しては実験する前から誤差が少ないことは分かっていたけれども、何年もかけて丁寧な分析をやった結果、軽微であると、ここに示したような結果になっています。

世の中で実際に使われているのは前年同期比が多いので、それで見れば更に誤差は少ないことが読み取れますが、それでも改善の余地はあると学識経験者たちが言っていて、事務局はまだやる気になっているというのが今の報告だと思います。

○樫分科会長 どうもありがとうございます。

ほかに御質問。久我臨時委員、よろしくお願いします。

○久我臨時委員 ニッセイ基礎研究所の久我です。

1点とても簡単な質問なのですがけれども、3 ページでチャートを説明いただいたのですが、今回、結論として、0 値補完をしても非常に軽微なので問題がないということは理解しました。

推計なども理解したのですがけれども、数十社程度なので、軽微なので問題ないところだと思うのです。三万数千社のサンプルに対して数十社なので、0.1%ぐらいなのでということだと思うのですが、そもそも回答のない対象数が毎回同じ数十社程度なのかということと、ある程度、こういう、無回答が増えると軽微ではなくなると思うのですが、その場合の閾値など、例えば無回答が何%を超えたらどうか、そういったものがあるのかということをお教えいただければと思います。

○巴財務省財務総合政策研究所調査統計部調査統計部長 まず、最初の方です。未回答法人というのは何かこう傾向的なものがあるのかというお尋ねかと思いますが、これは私もその研究会の中で先生方から御質問が出たのですが、結論としては、数十社のうちで幾ばくかのところは継続的に人件費の部分は出してくれないという企業がございます。それに加えて、あとは本当にアドホックに答えてくれなかったり、くれたりとかということもあるので、一概に傾向を言えないところがございます。

○藤原財務省財務総合政策研究所調査統計部調査統計部調査統計課長 社数的なものとしては、今回検証している期間、5 年分くらい実施しているのですが、大体同じぐらいの社数で、数十社というところで推移しています。

先ほどの御質問の閾値についてですが、今時点では大きな変動がなかったものですから、正直、決めたものはございません。

ですので、御指摘のように、今後増えていくとかがあれば、そういったことを考えていかなければならないのではないかと考えているところです。

○樫分科会長 よろしいでしょうか。

ほかいかがでしょうか。山本臨時先生。

○山本臨時委員 1点だけ。電通大の山本です。お世話になります。

5 ページの右側のグラフについて少しお伺いしたいのですが、これまで0 値補完の影響は軽微であるということの結論で、ところが、調査の概要としましては、月例経済報告や経済財政政策立案の基礎資料で、四半期別GDP 速報など、国民経済計算の策定に使われているという前提でのお伺いなのですが、0.1%の未回答に対して、今、御提示いただいた補完のどれかが正しいのだとすると、およそ0.5 から1 %ぐらいの揺れが数字に生じているのですが、このことは本調査の影響ではなくて、利活用されている経済指標とか、GDP 等には影響は軽微なのでしょうか。

多分、軽微とおっしゃると、そのまま軽微であるとして使われそうなので、利活用されている人たちにとって軽微かどうかという点で、5 ページの右側のグラフ、もちろん全て補完したものですので、これらのどれかが正しいということではなく、このあたりの範囲に収まっているだろうというグラフだとは存じ上げての質問です。よろしくお願いします。

○美添臨時委員 GDP 推計にこれは使いません。設備投資だけです。

○山本臨時委員 ああ、そうですか。では、ほとんど関係ない。

○美添臨時委員 GDP には関係ない。ただ、誤差はあるのです。

ついでに余計なことを言うと、後ろの方にある真値との差を見ればすぐに分かりますが、山本先生がよく御存じのとおり、欠測値補完は何を実施してもうまくいかない。要するに左下に点が集中しています。

いろいろな手法を試してみても、改善はわずかで、この程度しか精度は上がらない。

ほかの省庁でもそうだと思います。ここまでよくできている分析は、ほかにはあり得ない。

○山本臨時委員 単に従業員給与グラフの方が役員給与に比べて縦軸が1 桁高かったもので、それで、割合としては軽微なのでしょうけれども、額としてはそれなりの額だと思いついて、ちょっとお伺いした次第です。

特に政府や行政等で利活用されている範囲で影響が軽微でしたら問題ないように思います。ありがとうございます。

○樺分科会長 財務省様の方で補足することはございますか。

○巴財務省財務総合政策研究所調査統計部調査統計部長 特段ございません。

○樺分科会長 分かりました。どうもありがとうございます。

ほかいかがでしょうか。よろしいでしょうか。

私の方で少し感じたところというのは、もともと0 値補完というのがある方向性への偏りが出るということに関する危惧があった。

今回、これを検討していただいたということで、その影響は軽微であるが、そのバイアスを補正するための方法論については検討を進めている、研究を進めている。ある種そういうメッセージが財務省様から一般に出ることが非常に、この欠測値の補完というのは、ある意味で存在しないデータに対してある値を作るということになっていますので、そのプロセス上、そういうことをきちんと実施しているということを示すことがまず第一義として大切であったということで、財務省様がやられたことは非常に評価できるのではないかというのが私自身の感想です。

それから、もう1つ非常に重要なことは、精度の問題というよりは、回答の値が0であるということとインプットして0になった、あるいは、欠測であるということ、今後、明確に識別いただく方針が出ているということです。

これが非常に重要な回答ではなかったかというふうに考えます。もちろんシステム上の整備という問題はあるようですけれども、これはマイクロデータの利活用においても必ずしもこれまでインプットしたものとそうでないものを識別するということが各府省でばらついていたのではないかという危惧があるところで、その意味ではこういうことを少し前向きに検討していただいたということは評価できるのではないかというふうに考えておるのです。

財務省の対応自体に対して今後、何か追加をお願いしたいというような意見がもしあれば、ここでお聞きしますが、よろしいでしょうか。

特にその種の御意見ないようでしたら、私どもとしては、今後も、ある種こういうメッセージを世の中に出すということが非常に重要だろうとは思っておりますけれども、今回、財務省で実施していただいたことの対応というのは特に問題はなくて、妥当であるというふうに評価したいと思いますけれども、委員の皆様方、よろしいでしょうか。

どうもありがとうございました。

財務省、どうもありがとうございました。

それでは、次に進めたいと思います。次は民間給与実態統計調査の取組ということで、これも事務局及び財務省からの説明を聴取したいと思います。よろしくをお願いします。

○栗原総務省統計委員会担当室次長 それでは、次に民間給与実態調査です。

民間給与実態調査におきましては、源泉徴収義務者に対しまして、所属する給与所得者から抽出された給与所得者に関する情報の報告を求めています。

この給与所得者の情報におきまして、一部非回答が発生した調査票の場合、当該源泉徴収義務者から得られました情報全体を集計対象外とした上で、ウェイト調整を行っている。

調査対象となる源泉徴収義務者の0.6%に相当しているということでございまして、当該対応は、一部非回答に対する1つの対処方法ではありますが、単一補完を行うことで、集計対象外とした情報を利用できる余地もあると考えられる。このため、30年度に有識者の意見を聞く等によりまして、改善の余地の有無について判断を行うことが必要である。判断の結果、改善の余地がある場合、行政記録情報による欠測値の対応も含め、その対処について引き続き検討を進めることが必要であるという指摘がなされています。

これに対する説明の方をよろしくをお願いします。

○宮本国税庁長官官房企画課課長補佐 国税庁です。どうぞよろしくをお願いします。

資料の5番を使って御説明させていただきます。

まず、1ページをおめくりいただきまして、民間給与実態統計調査の概要です。

こちらの調査は、民間給与の実態を明らかにし、租税に関する制度及び税務行政の運営に必要な基礎的事項を明らかにすることを目的とする調査として、調査事項としましては、左下に記載がございまして、源泉徴収義務者に関する事項というのが事業者の基礎的な情報で、資本金ですとか年間給与の支払い総額ですとかを聞いています。

あと、もう1つの方が給与所得者に関する事項ということで、こちらがいわゆる従業員の情報でございまして、給与の総額ですとか、あとは、税法上の各種控除の情報などを調査しています。

調査方法といたしましては、2段階抽出という形になっておりまして、まず、事業者の方を、企業規模別、給与所得者の、従業員者の数の方で区分いたしまして、そこから無作為に抽出し、その後、標本事業所において給与所得者を、層別に一定の抽出率が決まっていますので、そちらに基づいて一定の数、従業員の情報をお書きいただいている形になります。

調査の周期としては年に1回で、今まさに調査中といったところです。

1ページおめくりいただきますと、実際にお送りしています調査票の現物がこちらになりまして、紙の場合ですが、左側の方が事業所の基礎的な情報を御回答いただいているもので、右側が給与所得者の情報という形になります。

1ページおめくりいただきまして、こちらが御指摘いただいた一部無回答に関する現状のフローといった形になっています。

まず、調査対象となりました事業所の方で、民間給与実態統計調査事務局という、こちらは民間に委託しておるところですが、そちらに調査票をまず御提出いただきます。

そちらで回答の内容を審査いたしまして、疑義照会等で不明の部分を解明いたしまして再提出いただく形になりますが、それでも修正し切れなかった場合は、⑥番に書いてあるとおり、欠測のある回答については、まず、エラーフラグを立てた上で納品いただいて、こちらで、国税庁で集計しておるのですが、そちらを全て削除した上で集計しておるのが現状です。

したがいまして、結論として、何らかの欠測のある回答については全て無効回答といった処理となっています。

また1ページおめくりいただきまして、こちらが実際の欠測値の発生割合になります。

こちらは、まず、左側に書いてあります源泉徴収義務者用の調査票に関する欠測が、先ほど御指摘いただいたところで、28年分と言いますと、約0.68%という形になっています。

こちらの欠測が発生しますと、当該事業所の回答を全て無効回答としますので、例えばそちらの事業所が20人お書きいただくとした場合、20人分の給与所得者の情報が全て無効回答というふうな扱いになっておるところです。

右側に書いています給与所得者用の調査票の情報というものにつきましては、各従業員の年齢ですとか性別ですとかそういった情報の欠測です。

こちらは0.2%から0.3%程度の発生といったところです。

その欠測値の中身につきましては(2)番のところに記載がございしますが、一番多いのは給与対象者の人数不足です。

また、3番目の層番号誤りというのも、こちらは結論としては、書いている人数が合わないということですので同じところです。

これらの誤りにつきましては、源泉徴収義務者用の調査票データのエラーという形になりますので、当該事業所の回答が全て無効回答という形になっています。

層別で申し上げますと、第1層の一番小さな事業所になりますけれども、こちらは毎年、無作為抽出でほとんど当たることがないので、このような方のエラーが多いというところ
です。

また、先ほど申し上げた人数不足の数が多い原因といたしましては、口頭の電話だけの
確認では補完できずに、再提出などを御依頼しなければいけないので、ちょっと御協力を
得られないケースが多いといったところも1つあるところです。

1 ページをおめくりいただきまして、ここから欠測値の補完の検討に入っております
が、まず1点目としては、いわゆる行政記録情報による補完が可能かどうかというところ
ですが、こちらの民間給与実態統計調査でとっています給与所得者の給料ですとか税法上
の控除などの情報は、所得税法上、年末調整されている方で給与が500万円以下の方につ
きましては、税務署に提出の義務がございませんので、まず、我々としてデータを持って
いないというのが1点です。

また、民給の記載上、氏名を書かなくても社内の会社員の番号、社員番号などで御提出
いただけますので、そもそも誰の情報かというのが、まず我々は分からないというところ
もございます。

また、先ほど申し上げたとおり、毎年無作為抽出の層に比較的欠測値が多いところもご
ざいますので、過去データというものもありません。

そういったところで、国税庁としましては、行政記録情報からの補完は難しいだろうと
いったところになります。

続きまして、1 ページおめくりいただきまして、こちらが有識者の先生と御相談させて
いただきまして検討した、いわゆる統計学上の処理といったところです。

横浜市立大学の土屋先生に御相談させていただいたところ、まず前提として、民給の欠
測値の件数は比較的少なく、調査結果に与える影響はそれほど大きくないだろうという
ところではあるのですが、一方、何らかの項目が抜けている回答につきましても、有効回答
としていくべきであって、何かしらは対応した方がいいのではないかという御指摘をいた
だいています。

先ほどから申し上げているとおり、まず、件数として最大となっています記入対象者の
人数不足ですとか層番号の誤りとかいったところは、やはり捨ててしまっているデータも
大きいところもございますので、対応を検討しておるところです。

具体的な対応としましては、本来10人書くべきところを8名しか書いていないような場
合には、その8を有効回答として、会社のサイズに膨らまして処理をすれば構わないとい
うことを御指摘いただきまして、単一補完の方法で申し上げますと、層化平均値代入法と
同じような結論になるところです。

一方で、年齢等の基礎情報の記載漏れの個別項目の欠測につきましては、件数が少ない
というところと影響も軽微というところで、個別の対応より、先ほど申し上げました、一
部欠測のある方のデータを消した上で、その会社の人数に膨らませるという処理も考えら
れるだろうというところであるのですが、こちらとしてはまだ結論が出ていないところ
ではございます。

後ほど説明をさせていただきますが、国税庁としては、この土屋先生の御指摘に従ったシステム対応を検討しています。

また1ページおめくりいただきまして、こちらは28年分のデータに基づいて統計精度等を試算してみたところではあるのですが、システム改修前でございます、注で書かせていただいておりますが、記載不足の事業所の回答につきまして、当該事業所の給与所得者からランダムに抽出して、1人補完するような形ですので、復元方法としては若干違うのですけれども、マッチング代入法のようなところで一旦できる範囲で試算したものです。

結論としては、精度といたしましては0.01%程度の影響で、平均給与については0.1%程度の影響でございました。

こちらは、この年、たまたまといったらあれですけれども、大きな事業所で欠測があったというところもあろうかとは思いますが。

また、土屋先生の議論の前提といたしまして、欠測がランダムに発生しているというところが必要というところではあるのですけれども、正確なデータをおとりできなくて恐縮なのですが、コールセンターを担当しています委託事業者等に聴取したところ、記載対象者の過不足や層番号の誤りというのは、比較的計算間違いですとか、そういった単純な誤りが多いということですので、我々としては、何らかのこのような形で欠測値補完をしてもまず問題ないだろうと考えています。

また1ページおめくりいただきまして、こちらは本年、予定していますシステム改修についての御説明です。

従来の復元処理の方法につきましては、当然抽出率の逆数を掛けて復元しておったところなのですが、令和元年分の調査より、回答者負担の軽減のために第7層の記載対象者に上限を設けたことがございまして、その結果、単純に抽出率の逆数だけでは復元ができなくなりましたので、それに併せまして、2段目に書いてあるとおり、新たな復元処理を現在、開発中です。

復元方法としましては、最初に申し上げました、源泉徴収義務者の調査票の方に、実際の会社の雇われている人数をお書きいただく欄がございますので、そちらの人数で各会社ごとに復元をするといった処理の方法に変更する予定です。

欠測値の処理につきましては、それに併せまして、過不足があったとしても、その記載いただいた人数で、あとは源泉徴収義務者用の調査票に御記載いただいた、その会社の従業員数の方で引き伸ばすといった形の処理を予定しています。

最後のところに少し書いてございますが、また、年齢等の個別項目の欠測の対応につきましては、引き続き検討といったところが現状です。

1ページおめくりいただきまして、こちらはその他の対応といったところではございますが、そもそもそういった誤りをなくすための工夫といったところではございます。

こちらに記載のあるものにつきましては、令和元年分調査より源泉徴収票等の情報を自動で転記するツールの提供を開始します。

具体的には、こちらのツールを使いますと、会社の持っている、いわゆる社員の方の給与台帳などのデータをCSVではき出して、それで、自動的に、何人を記載すべきかとい

うところも含め、必要な人数だけ抽出して項目が自動転記されるシステムになっています。

ただ、民間給与実態統計調査独自の数値というものがございまして、例えば勤続年数ですとか、あとは性別ですとか、そういった一部の情報につきましては、自動で転記する元データがございませんので、そういったところだけ御記載いただく形で、約 32 項目中約 20 項目、自動で転記する形になっています。

あと、最後のページですが、その他の簡単な対応でございまして、記載対象者の先ほどの誤りにおいて、記載対象者の人数につきましても、その部分だけ自動計算するツールなども 2 年ほど前から提供していますし、あと、その他簡単な督促等の改善などもこちらに記載させていただいています。

簡単ではございますが、以上となります。

○椿分科会長 どうもありがとうございます。

ただ今の御説明につきましても、何か御質問あるいは御意見などありましたら、よろしくをお願いします。いかがでしょうか。

では、美添臨時委員。

○美添臨時委員 議題が多いので、あまり余計なこと言いたくないのですが、今回の回答はほとんど完全な答えだと思うんですが、調査の設計上、もう少し改善できる点は幾つかありまして、例えば提案では各社ごとの乗率を掛けて復元したものを合計するので、もちろんそのやり方はありますけれども、各層での全社の合計と、給与の合計とで、割り算が比推定の形になるのを、個別比推定でやるか、合計比推定でやるかで、これで推定量の分散に差が出るはずです。

土屋先生に相談されているということなので、その辺も確認されたのかと思いますけれども、せっかくシステムを改善されるのであれば、推定量の性質まで検証して、この機会に更なる改善を目指していただけるとよいかと思います。

○椿分科会長 どうもありがとうございます。欠測値補完に関する御検討をいただきましたけれども、復元について非常に精力的に実施していただいたということで、今のような話があると思いますけれども、これは適宜御検討いただくということでよろしいでしょうか。

○宮本国税庁長官官房企画課課長補佐 そうですね。

○椿分科会長 あと、もともと復元の方がかなり重要であるということについて、ウェイト調整を主として実施していらしたというふうに経緯を見ていたわけですがけれども、今回、逆に欠測値を補完する話を新たに導入するということに、欠測値補完したものと、ウェイトの問題というのは、何か検討されているということはございますか。

○宮本国税庁長官官房企画課課長補佐 現状につきましては、回収率の逆数を計算したところでウェイト調整いたしていますので、欠測値を補完できた部分は無回答のウェイトから外れると思います。

まだ具体的な対応をしておるわけではございません。

○椿分科会長 どうもありがとうございます。欠測値を入れたものはウェイトをかけずそこは完全にそのまま足すというだけであると。

分かりました。そういうアプローチでよろしいでしょうか。

いかがでしょうか。これも実は各府省によって対応が違うということは承知していますが、それでも。

○美添臨時委員 時間があつたらゆっくり伺いたいのですけれども、そもそも国税庁の調査ですから、回収率が非常に高いのです。それは本当に恵まれている立場だと思うのですが、この調査の設計そのものを検討してほしい。せっかく土屋先生を入れているのだから、更なる改善が可能です。各事業所に抽出する給与所得者の数を指示しているようだけれども、その後の作業手順は多分調査の手引きにありますよね。系統抽出だと思います。

職種、年齢、経験によって給与は大きく違うので、構成を考えると、更に精度が上がると予想できます。多分そこまでやらなくても従来はいい結果が出ていたのかと想像はしますけれども、改善の可能性は非常に大きいと思うのです。ということで期待しておきます。

○椿分科会長 先生、これは今回のフォローアップというよりは今後の期待ということでよろしいですか。

○美添臨時委員 せっかく改修しようというのに、これだけではもったいない。

○椿分科会長 そうですね。ありがとうございます。

ほかに何か御意見、御質問ありますでしょうか。山本臨時委員。

○山本臨時委員 最後の方に御紹介いただきました、ツールの御提供というのはすごくすばらしい取組だと思っておりまして、記入者負担の軽減といいますか、既存の、行政資料ではなく、調査客体側がお持ちのデータから加工できるというのはいいことだと思います。そのことについてちょっとお伺いしたいのですが、これを使うときに、例えば年齢が未記載だとエラーが出て提出できないということになるのか、それでも、エラーのまま提出できるようになっているのかという点と、これを導入したことによって、実際にエラーは減りつつあるのかという、その2点お伺いしたいのですが。

○宮本国税庁長官官房企画課課長補佐 まず、1点目ですが、このシステム上、先ほど申し上げたとおり、一部空欄で一旦転記されるのですけれども、そのままですと、御提出は一応できない形になっていますので、これをお使いいただくと、基本的には適正なデータが入るといったところかと思えます。

あと、現状、こちらの提供が本年から開始ですので、まだ成果は何も分からないという状態です。申しわけありません。

○山本臨時委員 あと、済みません、ついでに。利用率がどの程度になりそうかみたいな、ダウンロード数とかの見込み等はございますか。

○宮本国税庁長官官房企画課課長補佐 一応ツールもどうしても初年ですので、正確なことはあれなのですけれども、一応オンラインの利用自体は現在、37%程度ですので、それらをお使いの方がこちらを、比較的便利かと思えますので、お使いいただけるのではないかと期待しておるところです。

○山本臨時委員 ありがとうございます。

○椿分科会長 どうもありがとうございます。

いかがでしょうか。よろしいでしょうか。

それでは、特に美添臨時委員からは、標本設計に係ることの今後の期待はございましたけれども、今回、フォローアップということに関しての項目につきましては、国税庁様非常によくやっていただいたということかと存じますので、本日の回答も含めて、この分科会としては了承したいと思えます。よろしいでしょうか。

どうもありがとうございました。

それでは、引き続きまして、造船造機統計調査の取組につきまして、これも事務局、それから、国土交通省様から説明を聴取したいと思えます。よろしく申し上げます。

○栗原総務省統計委員会担当室次長 続きまして、造船造機統計調査です。

こちらの調査の方は、回収率が8割程度以下で、単一補完またはウェイト調整による欠測値への対応が行われていない調査における対応ということになってございます。

造船造機統計調査は、回収率約75から80%ですが、全部非回答に対して、単一補完やウェイト調整などの欠測値への対応が行われることなく、得られた回答のみで単純合算集計が行われており、調査結果が過小となっている懸念がある。一方、同調査は、届出名簿に基づく統計調査であることから、調査対象に廃業事業所等が含まれる可能性があり、当該廃業事業所が回収率を下げている可能性もある。このため、30年度に調査対象事業所の休業及び廃業状況を確認し、公表数値に係る捕捉状態等の検証を行うことが必要であるといったことが指摘されています。なお、届出名簿に基づき調査を引き続き実施する場合、定期的に休業及び廃業状況の確認が必要であるとされているところです。

これに対しまして、説明の方をよろしく申し上げます。

○樫分科会長 国土交通省様、よろしく申し上げます。

○土屋国土交通省総合政策局情報政策課交通経済統計調査室課長補佐 国土交通省の交通経済統計調査室で課長補佐をしています土屋と申します。よろしく申し上げます。

それでは、資料6の造船造機統計調査における欠測値への対応に関する資料について御説明します。1ページ目を御覧ください。

まず、造船造機統計調査の概要です。

目的といたしましては、我が国の造船及び造機の実態を明らかにするとともに、造船・船用工業関係行政の基礎資料を得ることとしています。

具体的には、例えば国民経済計算や、I I P、I O表、あと、日銀の企業物価指数だとか、そういうところで御活用いただいています。

調査の範囲です。造船調査につきましては、鋼製船舶、それと、FRP船だとかの鋼製以外の船舶の総トン数20トン以上などの製造設備を有している工場、造機につきましては、船用工業機関等の製造だとか修繕を行っている工場としています。

選定方法につきましては、両方とも全数調査です。

調査事項につきましては、造船調査につきましては、製造船舶であれば、トン数や船価など、修繕船舶であれば、隻数や工事金額などを調査しています。造機調査につきましては、製造高や、修繕高などを調査しています。

それと、調査系統ですが、調査方法につきましては、郵送及びオンライン調査で実施しておりまして、オンラインについては大体2割程度の利用実績ということになってござい

ます。

飛ばしまして、調査時期ですが、造船調査については毎月調査、造機調査については四半期調査という形で実施しておりまして、公表の形としては速報、月報、四半期報という形でそれぞれ公表しています。

2 ページ目を御覧ください。

平成 30 年 3 月の統計委員会で決定いただいています審議結果の報告書の内容です。①番のところで、回収率が 8 割以下で、単一補完またはウェイト調整による欠測値への対応が行われていない統計調査ということで、捕捉状態等の検証を行うことが必要であるという御指摘をいただいています。

その検証結果が、3 ページ以降になりますので、3 ページ目を御覧ください。

調査対象事業所につきましては、造船法の第 6 条に基づいて調査対象名簿というものを国土交通省の方で整備しておりまして、これをベースに、現地訪問であったり、電話による確認であったり、あと、経済センサスー活動調査の調査票情報とのマッチングを実施して、調査対象事業所の再整理というものを行っています。

下の方に行きまして、造船調査の結果です。

左側の円グラフですが、833、母集団ございますが、このうち 527 から回答をいただいております、回収率といたしましては 63%という結果になってございました。

ただ、833 の中に 527 とは別に、実質廃業しています 100 事業所がございます。

それと、この中の 206、ピンクのところですが、この 206 について回収状況をさらに確認していますのが右のグラフにございます。

このうち、96 事業所については、追加回答という形で月報の公表後に提出いただいております。

そのほか、製造設備が調査対象とならなかった事業所というものが 5 事業所、それと、いわゆる実質廃業しておりました事業所が 11。あと、廃業していた事業所が 7 ということで、残りが 53 と 34 を足して 87 事業所になりますが、これについて経済センサスとのマッチング調査を実施しておりまして、そのうち、マッチングしたものが 53 事業所、マッチングしなかったものが 34 事業所という形になってございます。

ということで、もともとでいきますと、833 分の 527 の 63%ということで御説明差し上げておったわけですが、今回、もともと回収があった 527 と、追って回答のあった 96、それと、経済センサスとマッチングした 53、これを足した 676 を分母といたしまして、回収のあった 527 と 96 を足した 623 を分子といたしますと、回収率が 92%という結果になってございます。

1 枚おめくりください。

次が、造機調査の結果です。

造機調査についても同様の調査方法を用いておりまして、こちらは結果だけ御説明しますと、92%ということになってございます。

ということで、回収率は造船・造機ともに 9 割を達成しているということで、欠測値補完については特に行わず、更なる回収率の向上を図っていくということをもって品質の維

持を図りたいと考えています。

続きまして、5ページを御覧ください。

先ほど速報という形で公表しておるとい御説明差し上げたかと思いますが、造船調査については主要44事業所、造機については43事業所を対象に速報という形で公表しています。

あと、確定値ということで、造船統計月報と造機統計四半期報という形で公表していますが、その後、追送で来たものについては、当該調査年の最終月に遡及改訂をするということをしています。

その遡及改訂をした最後の結果を用いて、それぞれの補完率を出していますのが下の表になります。

右側の船価のところで見ますと、改訂後の数字、1兆6,847億円になっていますが、当初、出てきた月報での公表については99.2%、速報値での補完では92%という形になってございます。

それと、6ページを御覧ください。

未回収事業所に係る検証ということで、先ほど経済センサスとのマッチングをしたという御説明をしましたが、このマッチングで一致した、造船については53事業所、造機については34事業所についての経済センサスー活動調査の調査票情報に出てくる品目別製造品出荷額と造船・造機それぞれの調査結果を比較したものです。

造船調査につきましては、合計が1兆6,000億円ほどで、経済センサスー活動調査の出荷額の該当部分にある事業所、53事業所ですが、こちらの金額をトータルすると115億円ということで、捕捉率としては0.7%ほど落ちているという状態になっていますので、99.3%捕捉できていると考えています。

造機調査も同様で、0.3%が経済センサスー活動調査での出荷額のシェアということになっていますので、全体で99.7%カバーできていると考えています。

駆け足ですが、造船造機統計調査についての整理結果です。

以上です。

○樫分科会長 どうもありがとうございました。

ただ今の国土交通省様の御説明につきまして、御質問、御意見があれば、よろしく願います。いかがでしょうか。

○美添臨時委員 説明ありがとうございました。結論はこれでいいような気はしますが、教えていただきたいことが幾つかあります。まず、3ページ目、4ページ目あたりにある整理結果なのですけれども、これまで、平成28年で指摘された段階での名簿の事業者数が、造船だと833ということですが、指摘を受けて検討したところ、そのうちの100は所在不明であるということですよ。

所在不明というのは、調査の方法によりますが、例えば郵送なら返ってこない、電話なら使われていないということだとすると、母集団名簿の整備をどのくらいの周期で実施していたのかが気になります。

今回の検討の結果、この100という所在不明は対象外であるという整理であれば、母集

団サイズが少なくなったという理解でいいと思うのですが、そういう作業は、例えば経済センサスであれば5年間に2回は実施している。負担にはなると言いながら、これが一番大事な話なので、そこの実態をお聞かせいただきたいと思います。

もう1つは、例えば造船調査だと、3ページですが、206事業所に対して、右側に行つて、再度、訪問や電話で96事業所を回収できた。これは丁寧な作業で、立派だと思います。

そこで、更にマッチしたものが53ある。ここも経済センサスから情報を持ってきて比較したというところはよく分かるのですが、話を聞いていてよく分からなかったのは、左下にある34、ピンク色の事業所です。

これは、206の中で調査対象ではない。でも、センサスとマッチングしなかった事業所で存在はしているわけですね。業種として正しいのであれば、経済センサスでの漏れという可能性があるのですが、この辺はどうお考えなのか、教えていただけますか。

○樁分科会長 お願いします。

○土屋国土交通省総合政策局情報政策課交通経済統計調査室課長補佐 まず、1つ目の御質問です。

母集団の廃業の部分ですが、造船法の第6条というところで、届出義務をすることになっていますが、廃業するときも同様に6条のところで、廃業届を出すことになってございます。

ただ、廃業届をきちんと出していない事業所が結構いるようで、実質的には廃業しているような状態になっているものが100ほどあったということです。

実際にどれほど母集団の整理をしているかということについては、恥ずかしながら、造船法のデータをいただいてきて、それをベースに毎年調査を実施するという形で、その母集団名簿を信じてしまっていたところが若干ございまして、それについての事実確認というものができていなかったということがございます。

ですので、今後については、その辺について、実際にその届出するのは運輸局ですので、運輸局だとかにきちんと把握しながら、母集団名簿の整理をきちんと実施していきたいというふうに考えています。

それと、もう1つ、34の関係です。

34については、調査の方法といたしまして、まず、運輸局だとかにきちんと確認をしておりまして、基本的には、まず、ここでいう5、11と7の部分については運輸局だとかと調整をしてきれいに抜けているところがございますが、34と54のところについては、運輸局の方では実は存在がなくなっているという確認がとれなかったというものですので、経済センサスー活動調査の結果でなかったということをもって、現在、こちらの34を抜いているという形をとってございます。

ですので、美添先生のおっしゃるとおり、その確認をしていく必要が今後はあるのかとは思っていますが、現在のところ、こちらを仮に抜いた処理して、その34を分母の方に加えたとしても、回収率が88%ほどになるということで、大きな違いは出てこないのかと思っています。この34については引き続き、我々の方でもきちんと確認はしていこうとは思いますが、結論については変わらないかと思っています。

造機の方についても、更に9割を超えている。34に類する9のところ、ピンクの9のところを仮に加えたとしても、9割を超えていますので、結果としては変わらないのかと考えています。

以上です。

○樫分科会長 はい、美添臨時委員。

○美添臨時委員 ありがとうございます。もう一言言うと、大規模のところはきちんと捉えられているから、実はそれほど大きな影響を与えてないということですね。

回収率が低いことをそのまま公表すると、統計の信頼性が疑問視されますが、実は大規模な事業所は回答しているので信頼性は高い。是非このようなメッセージを工夫して出していただけるといいと思います。ありがとうございます。

○樫分科会長 ほかに何か御指摘の点等ありますでしょうか。

恐らくこの御説明あったとおり、当面、欠測値補完云々という問題よりは、母集団名簿の整備といいますか、ここにもあったように、捕捉された本来の母集団の中での回答率を上げていくということが、むしろ品質上一番重点であるという、この取りまとめ自体は全くそのとおりではないかというふうに思ったのですけれども、よろしいでしょうか。

是非今後、今回実施していただいたような整備事業で、事実上、本当は非常に加工段階では回収率も高いということも見える化していただければと思いますので、よろしく願います。どうもありがとうございます。

それでは、この御説明についても委員会としては了承したという形にさせていただきます。

それでは、次に、自動車輸送統計調査の取組について、これも事務局並びに国土交通省様からよろしくお願いします。

○栗原総務省統計委員会担当室次長 自動車輸送統計調査、まず、上段がバスの関連となつてございます。こちらの方は、過去の回答結果を使用して単一補完を実施する場合の使用データの期限のないものということですが、自動車輸送統計調査（バス関連調査）では、下の項目でも示していますとおり、第Ⅲ期基本計画において、自動車輸送統計調査そのものの見直しが指摘されている。当該見直しと整合するよう、並行してデータの使用期限も含めて検証を行うことが必要であるということで、補完は実施されているのだけでも、その使用期限も含めて検証を行うことが必要とされています。

併せて下段の方ですけれども、こちらはその調査そのものの見直しが必要であるものという指摘でございまして、同調査では、調査票に応じまして、単一補完、ウェイト調整などの欠測値への対応の措置がとられているのですが、一部の調査票様式につきまして、回収率が50%を下回るなど、欠測値に対する統計技術的な対応では補正し切れない状態に置かれているものと判断される。このため、第Ⅲ期基本計画において、調査そのものの見直しが指摘されていることに併せて、推計方法や欠測値への対応についても検討を行い、結論を得ることが必要であるとされているものです。

これについての説明の方をよろしくお願いします。

○樫分科会長 よろしく願います。

○土屋国土交通省総合政策局情報政策課交通経済統計調査室課長補佐 それでは、続きまして、私の方から自動車輸送統計調査における欠測値補完に関する取組について、資料7に沿って御説明します。1ページを御覧ください。

自動車輸送統計調査の概要です。

調査の目的ですが、自動車による貨物及び旅客の輸送の実態を明らかにし、我が国の経済政策及び交通政策を策定するための基礎資料を得るということです。

例えばI O表の作成やI T F国際交通フォーラム、資源エネルギー庁の総合エネルギー統計などで活用いただいています。

調査の概要について、先ほど、総務省の御説明の中でもありましたとおり、第Ⅲ期基本計画に基づく改訂を行っていますので、次の2ページに沿って内容を御説明します。

この概要の中で赤字が付いている部分については、今回の改正をしている箇所ということです。

調査の概要です。一般乗合・一般貸切・特定旅客自動車、いわゆるバス事業者です。これを営む事業所、それと、あと普通自動車等であって、主として貨物・旅客の輸送の用に供する自動車で、そのうち国土交通大臣が選定する事業所及び自動車というものを対象としています。

ここでいう普通自動車というのは、道路運送車両法に基づく普通車ですので、例えばトラックで言うと1ナンバーの自動者であったり、バスの2ナンバーの自動者が該当することになります。

調査の対象範囲、対象数、調査事項についてはこの表のとおりですが、自動車の母集団でいいますと、約1,021万両が母集団になりますが、そのうち毎月2万両程度を調査対象として抽出して調査を実施しています。

調査の調査組織、調査方法ですが、郵送、オンライン調査で実施しておりまして、ここでは第1号様式及び第3号様式と記載させていただいていますが、現在、第2号様式と4号様式についてもオンライン調査できないかということで、総務省と調整をしております、できれば4月から実施したいと考えています。

あと、結果の公表です。それぞれ速報、月報、年報という形で公表しています。

それでは、1枚おめくりください。

同じく平成30年3月の審議結果報告書です。まず、1点目、先程総務省の方から御説明ございましたが、⑥単一補完を実施する場合の使用データの期限のないものということで、バス関連調査について、データ使用期限を含めて検証しなさいということをおっしゃっております。

それと、あと⑦調査そのものの見直しが必要であるということで、一部の調査票様式、具体的に言いますと、営業用トラックの第1号様式になりますが、こちらの回収率が50%を下回っているということで、欠測値に対する統計技術的な対応では補正し切れない状態になっているということで、推計方法や欠測値への対応についても検討しなさいという御指摘を頂いています。

それでは、資料の4ページを御覧ください。

まず、バスの関係についての御説明です。

分析・検証の経緯ですが、全数バス調査につきましては、欠測となった事業所について、過去のデータを用いて欠測値の補完を実施しています。

ただし、更に欠測が連続して発生している場合については同様の方法で補完を行っているということで、いわゆる補完に補完を重ねているということをもって、使用するデータの期限がないのではないのかという御指摘を頂いているものです。

具体的な内容については、5ページを御覧ください。

現行手法ですが、当該月前月の値に総走行キロの前年前月と前年同月の伸び率を乗じることによって推計するというので、総走行キロをベースとした伸び率を乗じて、それぞれの項目に値を掛けていくというやり方をしています。

全体的な統計精度、推計精度自体は比較的保たれているのかと考えていますが、デメリットとして3点挙げています。

1つ目が、異なる調査項目の伸び率を適用する。先程、御説明した総走行キロを適用するということになりますので、相対的な関係性が薄い、一定でないようなものについては、推計値に乖離が生じるのではないかということ。

2つ目が、過去データを用いて推計しますので、13カ月分のデータの蓄積が必要になってくるということ。

それと、3つ目が、同じく前年のデータを用いますので、例えばオリンピック・パラリンピックのような急激な需要変動が発生するようなどきには対応がし切れないということが問題になるのかと考えています。

これらを踏まえて、次のページ、6ページ目と7ページ目で、国土交通省で実施した3点の手法について御説明していますが、時間の都合もありますので、最終的に選んだ手法について、7ページ目で御説明させていただきます。7ページ目を御覧ください。

統計手法につきましては、層別前月比法というものを採用しています。

用途別、こちらで言うと、分類のところにある表ですが、用途別3区分、地方運輸局別9区分、実在車両数別3区分、この表で言うと、下側にある事業所カバー率というところになります。3区分、これを設定し、欠測している事業所と同一層の調査項目ごとの前月比の平均値を乗じて推計するという形をとっております。

結果としては、メリットとして、書かさせてもらっているところですが、現行の方法と同等程度の推計精度は出ているという結果になっています。

それと、1年以内に仮に開業した、過去のデータ、実績がないようなところであっても適用が可能になるということ。

先程御説明したオリンピック・パラリンピックのような急激な需要変動についても対応が可能になるということで、この手法を採用したいと考えています。

それと、もう1つ、使用期限を定めて検証することという御指摘がございました。

この手法を仮にとったとしても、連続補完をしていくと、結局、同じ結果になりますので、連続補完を行う場合、乗合バスと貸切バス、特定バスでそれぞれ区別して、12か月、6か月、12か月ということで、それぞれ補完期限を決めています。

次の8ページ、9ページについては、今の検証の結果を散布図だったり、ヒストグラムであったりとかで分析したものですので、こちらの説明は割愛します。

また、10ページ、11ページにつきましては、連続補完をする時に、どこまで対応が可能かということを検証した結果ですが、こちらについても、説明は割愛します。

それでは、次に、13ページを御覧ください。

先程の審議結果の中で、⑦の方で御指摘いただいている欠測値に対する統計技術的な対応では補正し切れない状態があるという部分の検討結果になります。

先程御説明したように、営業用トラック調査が該当してくるわけですが、こちらについての想定回収率は59%で設定をしています。

ただ、調査票の回収率については、実際40%程度ということで、達成精度が5%以上になってしまっているということが問題であると考えています。

そこで要因分析を行ったところですが、現行調査につきましては、1事業所に対してより多くの情報を得たいと考えておりましたので、1事業所に対して、事業所の情報を教えてもらう調査票と、自動車票ということで、普通車、小型車、特種車、軽自動車、それぞれの車種について調査票を2枚ずつの計9枚の調査票を配布し、こちらについての記入を求めているということになっています。

ですので、事業所からすると、かなりの負担が大きいということで、例えば調査拒否を招き、回収率の低下に至っていたのではないかと考えています。

そうしたことから、14ページを御覧ください。新調査での対応です。

新調査につきましては、調査単位を事業所から自動車のみに変更することで、事業所票を廃止するということ。

それと、あと調査期間については、これまで事業所票1か月と自動車票7日間ということで調査をしておりましたが、事業所票の廃止に伴って、7日間分の自動車票のみの調査ということで、調査の簡素化を図っています。

また、行政記録情報を活用しまして、プレプリントなどを実施することで報告者の負担を軽減するということを行っています。

その結果、平成29年度に予備的調査というものを実施しておりますが、そのときの回収率が約60%に改善したということで、標本設計時の目標精度についてもそれぞれ達成したということになってございます。

それでは、15ページを御覧ください。

推計方法についての検証が必要であるという御指摘を頂いていますので、その対応ということになります。それがこの15ページです。

行政記録情報、いわゆる車検データを用いた比推定を導入することとしています。比推定にあたりましては、集計項目と補助変量に一定の相関がある場合には精度が向上するということになっているかと思いますが、集計項目の相関が想定される指標について、この表で言うと①、②、③の部分にあたりますが、指標を設けて、それぞれの相関係数Rを試算いたしまして、適切な補助変量の特定を行っています。

その結果、一番下の表で言いますと、集計項目は輸送トン数から輸送トンキロ、実車キ

ロ、能力トンキロ、走行キロの5つですが、そのうち、輸送トンキロと能力トンキロについては、ここで言う、真ん中にある③にある最大積載量と1日当たりの平均走行キロの乗じたものとの相関係数R、それと、実車キロと走行キロについては、1日当たり平均走行キロ、走行距離について相関関係が高いということで、おおむね0.5から0.6という結果になってございます。

ただ、輸送トン数については相関関係が非常に低かったということ、具体的には0.2以下ということになってしまいましたので、輸送トン数以外については比推定を導入するという形をとりたいと考えています。

続きまして、資料16、17、18ページにつきましては、資料として添付させていただいていますが、内容についての説明は割愛しますが、簡単に言いますと、資料17、18ページにつきましては、基本計画の中で御指摘頂いている内容についての答えという形で資料を付けさせてもらっています。

以上です。

○樫分科会長 どうもありがとうございました。

ただ今の御説明につきましても、御質問、コメントなどあれば、よろしく願います。

久我臨時委員、よろしく願います。

○久我臨時委員 御説明ありがとうございました。

とても単純な疑問なのですが、資料の13ページで想定回収率を59.0%に設定しているという記載があって、調査票を工夫されておおむね6割達成しているという御判断なのですが、ほかの調査の回収率なども鑑みますと、そもそもこの59.0%というのはどんな根拠に基づいて設定しておられるのかということをお伺いしたいと思います。

○樫分科会長 よろしく願います。

○小林国土交通省総合政策局情報政策課交通経済統計調査室専門官 この13ページにある59%についてなのですが、こちらは現在の現行調査の想定回収率でございまして、当初これを設定したのは、平成21年度にある部分的な地域の予備調査をさせて頂いています。

これは中部運輸局管内の試験調査を実施していますが、その際に回収率が、この当方で6割程度あったということも聞いています。

その際に、その結果の実績としまして、営業貨物につきましては59%台で、全国の輸送トン数、全国の輸送トンキロで精度5%以内に設定しているところです。

○樫分科会長 平成23年当時の実際の回収率が基本になっていたという御説明でよろしいですね。

○久我臨時委員 説明はよく分かりました。ありがとうございます。

多分、この分科会で議論する話ではないと思うので、コメントなのですが、そもそも調査によって回収率が恐らく大きく違って、それをどういう方向に持っていくのかというのは、恐らく統計委員会様の方でやられていると思うので、かなり目立つ値だったので、質問させていただきました。ありがとうございます。

○樫分科会長 どうもありがとうございます。むしろ今回は、予備調査というのを実施している、予備的調査で調査事項をかなりとりやすいものにされたという対策を打ったとい

うことは、先ほど言ったとおりですけれども、これがあるべき水準かどうかというのはもう少し議論を続けなければならないことだと思います。

美添臨時委員、よろしくお願いします。

○美添臨時委員 バスとトラックと両方精査していただいて、そもそも指摘の後に調査そのものの見直しがあるということなので、指摘は解決されているように見えます。細かいことですが、バスの方の資料で7ページあたりに連続補完という言葉が出てくるのですが、これは無回答が続いたときに何か月か連続して補完するという意味ですね。では、それで結構です。

これは改善の余地が非常に大きく、引き続き検討が必要と思うのですが、これはこの評価分科会ではなくて、統計委員会マターですか。

先ほどの久我臨時委員の指摘のように、基幹統計で回収率60%が目標というのは、そもそも調査負担が大きいということだと思います。

そこで、今回されたような工夫をやることは求められることだと思いますので、直接この分科会の検討課題ではないような気もしますが、調査の改善ということを考えると、記入者負担を軽減するような設計をする一方で、使える限り行政資料を活用すること、これが今後の統計改革の方向だと思います。

○土屋国土交通省総合政策局情報政策課交通経済統計調査室課長補佐 ありがとうございます。

一応バスにつきましては、回収率が88%ほど達成しておりますが、営業用トラックが特に悪いという状態になっていますので、営業用トラックについては引き続き努力をしていきたいと思っています。

○樫分科会長 神林専門委員、よろしくお願いします。

○神林専門委員 済みません、欠測値の補完とは少々異なるのですが、先ほどの造船機統計調査のやり方と同様に、その母集団の改訂というのはこの場合はどういうふうに行われているのでしょうか。

○小林国土交通省総合政策局情報政策課交通経済統計調査室専門官 全数バスの調査につきましては、自動車局が保有します監査台帳等がありまして、その台帳に基づいて、年1回、母集団名簿の修正をしています。

また、そのほか、貨物のトラックや、タクシーにつきましては、車検データを母集団としておりまして、こちらは3か月に一度、母集団の更新をして、それを用いて調査対象を抽出している状況です。

○樫分科会長 よろしいでしょうか。

まだいろいろな御質問があると思いますが、少し私の方の進行がまずくて、大分時間が押してしまっていますので、一応この統計に関しましては、先ほどありましたように、統計委員会自体で少し考えなければいけないことがあることは前提とした上で、一方で、先ほどありましたように、バスの方では、補完に関して連続補完という言葉を使いましたけれども、それに対して一定の歯止めをかけるということ。

これもほかの省庁の調査でも非常に必要なことではないかと思えますし、そういうこと

をきちんと明確にさせていただいた上で、回収率を上げていただくための方策をとっていくということ。

あと、一応、欠測値的なものについていろいろ検討してもらったことに関しては妥当であるというふうに判断していくということですのでよろしいでしょうか。

どうもありがとうございます。

国土交通省、ありがとうございました。

それから、次は、建築着工統計調査補正調査の見直しについてということで、これも事務局と国土交通省です。よろしくお願ひします。

○栗原総務省統計委員会担当室次長 続きまして、最後、5個目になりますけれども、建築着工統計調査の補正調査に関してということですよ。

こちらは、標本設計の見直しに関係するものでございまして、既にこちらの評価分科会でも昨年来議論して、こちらのほうは1つ結論は出ておるのですが、その状況を踏まえまして、国土交通省の方では、30年度に試験調査を実施いたしまして、その後必要な準備を進め、平成33年1月に完成する建築物から見直し後の標本設計に基づき抽出された対象に切り替えて調査を開始することが必要であるということございまして、この30年度に行われました試験調査の実施の状況について、今回、報告いただいて確認するというものです。

○樫分科会長 よろしくお願ひします。

○相部国土交通省総合政策局情報政策課建設経済統計調査室課長補佐 よろしくお願ひします。国土交通省の建設経済統計調査室ですよ。

お手元の資料8に基づき説明させていただきます。

1ページですが、建築着工統計の体系という表がございまして。

建築着工統計につきましては、毎月、悉皆調査ですよ、地方公共団体に出てくる建築物の着工届というのを基に、今月、何平米着工されているとか、住宅が何戸着工されているというのを出しているというのがメジャーなところというか、大きな話ですよ、それを③の補正調査ということよ、1年に1回、本当にその平米が建っているのか、本当にその工事予定額はそうなのかということよ補完している調査がありまして、この補完している補正調査というのを今回見直しということが課題になっておりまして、試験調査等を実施し、見直しを行っております。

2ページですが、そのような概況が書いてありますが、2ページの真ん中のところですよ。

現在、補正調査につきましては、建築物の着工統計の中からサンプル抽出しているということよ、全数につきましては約60万件程度のデータがあるのですが、その中から5,000ほどを抽出して実施しているところですよ。

これは何が問題かといいますと、ある程度有意抽出しているといいますか、都道府県を指定しまして、更に市まで指定しまして、その指定したところで何割抜いてこいと決めておりまして、それがなかなか問題だということですよ。

済みません、この補正調査の無作為抽出はちょっとあれですよ。無作為抽出というよりも、市と対象物件を何割抜いてくださいということよ、その市が持っているデータから帳

票の順番ごとに抜いていくということを実施しているということです。

その精度がやはり偏りがあるとかという問題がございまして、標本設計について、統計研究研修所の御協力をいただきまして見直してきたという次第です。

3 ページに移らせていただきます。

今回の新しい補正調査の枠組みといたしましては、これが試験調査の概要ですが、都道府県が実施するというよりも、我々、国土交通省本省が一括して民間事業者に発注して、それぞれの建築主に、事後どうでしたかという価格を聞きにいくという調査にしています。

実際、2 カ月ほどかかって実施したところとございまして、調査対象物件につきましては、サンプル数につきましては、おおむね、地方公共団体ごと、47 都道府県 10 戸ぐらいを目指しまして、450 ぐらい抽出したところです。

結局、完成した物件がないとかというので、そのぐらい減っておるのですが、その程度の、サンプル数で試験調査を実施させていただきました。

4 ページを御覧いただくと、個人が持つ物件、会社・団体が持つ物件とあるのですが、左の上の方です、回収状況。

いかんせん全国の建築物に郵送しまして、どのぐらい建築物にお金がかかりましたかとか、平米どのぐらいでしたかと聞くので、なかなか内容がテクニカルで難しいという状況とございまして、1 回督促を郵送して送った程度では回答率が低いという状況とございまして、個人の方とかかなり難航したという状況とございました。

さはさりながら、集めて集計結果として出してみればそれなりの数字が出たところであるのですが、やはり個人の方に建築物の価格とか、平米どうでしたかとか、変更はありましたかとかというのを聞くのは難しいかという状況とございまして、実際、個人の方なども施工業者に聞かないと分からないとかという話も多々ございまして、この点課題であったかというところが最大のポイントです。

試験調査はこんな感じで終わったということですが、そういうものをどう反映したかというのが 3 ページ以降に書かれているところです。

済みません、ページが間違っていますが、次のページ、3 ページと下に書いてありますが、5 ページ、調査変更概要です。

全体的に現行につきましては、調査対象数 5,000 ということで実施しています。建築工事費調査につきましては、変更案として 1 万を想定していますが、現行の調査につきましては、必ず都道府県が回収してきますので、回収率は 100% みたいなものですので、そうなのですが、最初の建築工事費調査になると、この試験調査の結果を踏まえると、50% 行くかという不安がございまして、一応倍の目標値で設定しているということです。

抽出方法につきましては、補正調査につきましては層化二段の無作為抽出ということで、市区を固定して建築物を層化抽出しているということですが、今度の新しい調査にしては、層化の無作為抽出ということで、しかも 20 億円以上は全数調査にしたいということです。

推計方法にいたしましても、単純集計であったのを、抽出率及び回収状況を加味した線形推定及び比推定です。

工事実施額が幾らかということですが、これでおおむね日本の建築物の 1 年間の工事実

施額が分かるということを目指している統計になるということです。

層化基準につきましては、今まで都道府県や木造・非木造という大きな枠だったのですが、これからは木造・非木造と、あと工事費予定額の金額ベースで分ける3区分。

ですから、6個の升ができるということです。日本の建築物の工事費別に木造・非木造と分けて、6個の升で検討していくということになっています。

標本配分法につきましては、今までは層別に抽出率を設定していたのが、きちんとネイマン配分で、20億以上は悉皆なのですが、残り4升につきましてネイマン配分で調整することになっています。

以上の標本設計を行いまして、次のページです、補正調査の建築工事費調査への変更概要。

名称を、今までみたいに親の建築着工統計の金額の補正をするという意味ではなくて、きちんと日本の全国の建築工事費を推計する調査というふさわしい名称に変更したことですし、調査方法につきましても、都道府県がやる調査から民間事業者を活用した郵送・オンライン調査に変更して、全国统一していきたいと考えています。

続きまして、次のページです。補正調査の建築工事費調査への変更概要（その3）。

調査事項につきまして、細かい内訳を聞いていたところですが、この建物は一体建築本体工事と設備工事は幾らなのかという内訳を聞いていたのですが、これがかなり負担がかかるということでごさいます、あと、利活用の面からも、なかなか使っている利用者のニーズというのがなかったものですから、廃止しています。

あと、本当の実着工日というのは建築着工統計は分からないものですので、建築予定日で集計して発表しているものです。

本当にトンカチやった日というのを聞きたい。そして、工事の完了日を聞きたいということで、きちんと工事予定と完成時期の乖離が把握できるようにしたいということです。

あと、⑤番、集計事項につきましては、調査計画の全面的な見直しに伴い、構造別、今みたいな木造・非木造別、あと、工事費予定額階級別、今の予定額は20億円以上だったら幾らぐらい、1億円から20億円ぐらいだったら幾らぐらいという6升の、先ほど説明した6升に入るような集計にしたいということです。

なお、都道府県別の標本設計を取りやめていますので、都道府県別集計は今のところ、見せないというか、積極的にはやらない予定です。

そして、最後になりました、公表期日につきましても、試験調査をやったときに、一個一個丁寧に実施していくのでなかなか時間がかかるということですので、年単位の調査ではございますが、昨年12月に行ったものを4月でやるには間に合わないかと思ひまして、9月末まで公表期日を延期して対応したいと考えているところです。

以上、試験調査を踏まえた変更案概要ということで説明させていただきました。

○椿分科会長 どうもありがとうございました。試験調査を踏まえて、標本設計をかなり大幅に変更したという御説明だったと思います。

何か御質問、コメント等あれば、よろしくお願ひします。

神林専門委員。

○神林専門委員 済みません、1点、2点でしょうか。1つは、試験調査と本調査の間の乖離がどれぐらいあるのかということです。それは報告書には書いた方がよいのではないかということが1件で、もう1点が、都道府県別集計を廃止することについても、多分これは既に統計委員会で決まっていることだと思うのですが、補足しておいた方がよいのかと思います。

○相部国土交通省総合政策局情報政策課建設経済統計調査室課長補佐 はい、ありがとうございます。

○樫分科会長 ありがとうございます。この点は報告書に記載いただくということでもよろしいでしょうか。

○相部国土交通省総合政策局情報政策課建設経済統計調査室課長補佐 そうですね。報告書といたしますか。

○樫分科会長 この分科会の。

○相部国土交通省総合政策局情報政策課建設経済統計調査室課長補佐 そうですね。その乖離がどのくらい。

あと、都道府県の集計の廃止につきましては、本親の統計委員会で指摘がございまして、もう少し細かい、ブロック別とか、少し大きなレベルでもいいのですが、出せるのはいかとか、いろいろ今後の検討課題とされていますので、どう見せていくか、ちょっとこれから検討していきたいと思っています。

○樫分科会長 ほかいかがでしょうか。山本臨時委員。

○山本臨時委員 今の御議論に関係するのですが、地域は全国1層にしてしまうのですか。

○相部国土交通省総合政策局情報政策課建設経済統計調査室課長補佐 1層で抽出する。20億円のものは全部悉皆で出して、1億円から20億円のものと1億円以下のものを木造・非木造の4層に分けて、全部ネイマン配分にして、割合を出して無作為抽出するというようにしていますが。

そうしないと、なかなか精度が保てないということです、都道府県の層を作ると。

○山本臨時委員 いや、ブロック別の層も設けないで、全国一律というののもかなりドラステックな変更だと思ったもので、その中間として、表章はしないまでも、もちろん郵送調査なので、調査の実施には都道府県の枠のあるなしは全然影響ないのですが、調査結果に対して、全国で1層にしてしまうことでどんな影響があるのかがちょっと分からなかったもので、単にちょっと驚いての質問なのですから。

○相部国土交通省総合政策局情報政策課建設経済統計調査室課長補佐 それは統計研究研修所のシミュレーションで、一番精度が出るのが全国1層にしてやる。

○山本臨時委員 全国の工事費の総額の精度がそこで一番高まるということ。

○相部国土交通省総合政策局情報政策課建設経済統計調査室課長補佐 そこで一番高まるということでも伺ってございまして、一応都道府県別を断念しているのですが、しかしながら、公表の仕方としてそういうものも入れられるのではないかという意見が多々ありましたので、そういうものも検討して、どこまで出せるのか。

やはりサンプル数が少ないので、それだけで出してしまうと、そうなのかというあれが

ありますので。

○上田総務省政策統括官付参事官 済みません、我々の方で審査させていただきましたので、補足だけさせていただきますと、もともと標本設計を見直す際に、結果をお見せすれば一目瞭然なのですが、非木造に関しては相当暴れている数字になっているというのが現状だということで、まずは、とにかく全国が暴れてしまっている状態だったものですから、まずは全国の数値をきちんと誤差の少ないものにしようということで、統計研究研修所の方で設計をしていただいたということが1つで、そうしますと、当然、結果として。

ただし、その標本、負担は増やせないという前提を置いたものですから、正直申し上げて、大都市に票を集めるという構図に改めまして、大きなものを当てていくという構図は大都市に票を集めるという構図になりましたので、そういう意味で、都道府県別の小さなところはまず無理でしょう。

現状でも相当暴れているのは間違いないのですが、まず無理でしょうということで、一旦は諦めましょう。

まずは、今現状、全国が本当に、怪しいというと語弊がありますがけれども、誤差が大きい状態なので、まずは全国を直しましょうということでチャレンジしていただいているのが現状ということです。

ただし、先般、実は答申も出ていますのであれですが、一応地域別の集計に関しては、きちんと得られた情報を基に検証してくださいという宿題を背負っていただいていますので、そういった検証を踏まえながら、また必要であれば、そもそも調査を見直していくというステップになる。

まずは、1つは、全国をきちんとした結果で出しましょうということを背負っていただいているということで御理解いただければと思います。

○山本臨時委員 ありがとうございます。

○椿分科会長 どうもありがとうございました。

○山本臨時委員 あと1点だけ。調査対象数が出ているのですが、想定回収数幾らぐらいを想定されているのですか。

○相部国土交通省総合政策局情報政策課建設経済統計調査室課長補佐 今、50%なのですが、法人や会社に聞くと高いのです。8割ぐらいの回収を見込めるということで、これは調査対象者も建築主から、やはり工事情報とか積算情報を持っている施工業者に移すことを考えています。

そうすると、かなり高い回収率が見込める、8割ぐらいを見込めると踏んでいます。

○山本臨時委員 ありがとうございます。

○椿分科会長 どうもありがとうございます。これは、統計委員会の方からも宿題が出ているということで、先ほど神林専門委員からもありましたように、報告書には若干その辺のことを書いていただくということで、またよろしくお願いします。

基本的には今の形のことはかなりドラスティックな変化であることは間違いないのですが、御報告自体は了承という形にさせていただいた上で、今のようなことを書いていただくことにさせていただきます。

よろしいでしょうか。申し訳ありません。まだ議題がかなり残っておりますのですけれども、最初の方の審議で大分時間を使ってしまいました。

済みません、引き続きまして、諸外国における欠測値補完及び総務省による各府省の統計作成支援にということで、欠測値の対応につきまして、いわゆる各府省研究成果の共有化の一環として、統計研究研修所で蓄積されました諸外国における欠測値補完の状況についての知見ということを御紹介いただくとともに、統計研究研修所で実施されている各府省の統計作成支援について御説明いただきたいと思います。資料は9と10を使っています。

○坂下総務省統計研究研修所統計研修研究官 欠測値補完ということですが、これについて、日本ではあまり公開の場での議論というのが行われてこなかったのですけれども、海外に目を向けますと、特にこれからの国際会議や統計機関などで議論が活発に行われるようになっていきます。

そういったデータもいろいろ出ていますので、情報収集を行いまして、その要点を今回解説したいと思います。

お手元の資料の概要の次の2のところ、公的統計における主な欠測値補完方法ですが、経済統計学会の学会誌に載せられた論文なのですけれども、UNEC Eの国連欧州委員会の会合に参加している23機関に対してアンケートを行ったところ、回帰代入法、比率代入法、平均値代入法、ホット・デック法を例に挙げて使用状況を尋ねたところ、ホット・デック法が100%、ほかの手法も95%の機関で使用されている。

一般的に使用されている手法ではホット・デック法が多くて65%、次いで比率代入法60%というのが挙げられています。統計の種別によると、事業所・企業の統計で比率代入法が80%、世帯統計でホット・デック法が、これも80%で多くなっているとしています。

それが先行調査なのですけれども、今回、分かったものは参考資料として、6ページ以下に挙げています。

そちらの方を見ていただきたいのですが、これは非常に多岐にわたりますので、要点だけ説明しますと、大きいものはアメリカの人口センサス。これは以前、内閣府の方で行った調査で判明したことなのですけれども、Assignmentというのをまず行う。

これは、ある項目欠測をほかの記入から論理的に推測する方法で、ほかのところでは、演繹的インピュテーションと呼ばれることもあります。

これができないものについてはAllocationというのがあって、これをほかの世帯員あるいは近隣世帯員から代入を行って、具体的には何種類かのホット・デック法を行っているということです。

それもできないような、世帯員全員が欠落した場合、Substitutionといいまして、ほかの世帯から代替する。日本語では代替という言い方をしています。

その下にありますACSというのは地域社会調査ですが、これも人口センサスのlong formを独立させた調査なのですけれども、同じ手法をとっています。

ただ、一方で、多変量モデルを使った手法も検討中ということでございました。

飛ばしまして、7ページ、経済センサスにいきますと、これは従来分野ごとに、産業分

野によって異なる欠測値補完を行っていたのですが、次回と書いてありますが、今回なっているところですが、分類が大幅に変わって、産業ごとの手法をとることができなくなったということで、幾つかの手法を検討していました。

その結果、結局、比率によるインピュテーションあるいは多変量の代入、回帰などを検討していたのですが、結果的にはホット・デック法になったということが報告されています。

それと、8ページに参ります。労働統計局というところで職業雇用統計というものを実施していますが、これもホット・デック法の一つを行っているということです。使っている用語が違いますけれども、内容的には同じことです。

農務省もその下にありますが、これも用語が異なっているのですけれども、演繹的な手法と、コールド・デック、ホット・デックというのを行っています。

9ページに参りまして、イギリスの人口センサスですが、これは、CANCEISという、これはカナダで開発した人口統計用のエディットとインピュテーションのシステムを使っています。

この中身は、具体的に言うと、ホット・デック法の一つによる最近隣法という方法をとっています。

現在、次回センサスのために、行政情報の利用を検討しているということです。

10ページに参りまして、ドイツのセンサスですが、ドイツの人口というのは、御存じのように、レジスタ・ベースで行っているのですけれども、調査部分についてはCANCEISを現在、検討しているということです。

今までCANCEISというのが何回か出てきましたけれども、13ページを御覧ください。

カナダの頭のところで、人口向きのシステムとしてCANCEISというものを作ってまして、更に経済統計についてはBanffというのを開発してまして、これを統計局での採用を検討しているということです。

最初のページに戻りまして、今の結果、非常に大ざっぱに見てきたわけですが、まとめますと、研究マターではいろいろ検討されているのですが、現時点で何しているかという意味では、ホット・デック法あるいは比率代入法の伝統的な手法が多いということです。

一方で、多変量の回帰といったような分布モデルを用いた方法は検討課題とされていることが多い。ペーパーベースでは、アメリカやイタリアなどで出てきています。

一方で、最近では行政情報の統計作成への利用というのが進展していて、そのために、概念が合わないために欠測値を補完する必要があるという例、あるいは、欠測値補完に行政情報を反映させるといった検討が見られます。

ただ、自動的に行政情報で埋めるというよりは、その行政情報をすぐに持ってくるというよりは、行政情報をどういうふうに使えるかといったような検討が多い。あるいはホット・デック法の中に行政情報を使うなどといった手法が検討されていることが多い。

国を超えた手法の共有化としてCANCEIS、今出てきたものが、英国、ドイツ、ニュージーランド、イスラエルといったところで検討されています。

アメリカでは、採用はしてないのですけれども、研究用にCANCEISと比べるといったような研究が多いです。

2ページに入ります。

Banffは今、出てきましたけれども、ニュージーランドあるいはアメリカの経済統計局で採用が検討されています。

現状はそのようになっていますが、では、海外ではどういうふうに分測値補完を受けとめているのかということで、定義とか用語の分類というものについて言いますと、3のところに書いてありますように、現在、一覧表でも見たように、用語は一定しているわけでもないのですけれども、もともと実務的な話ということで体系が定まっていないということが背景にあります。

ただ、調査した限りでも、1980年代以降は豊富な書籍あるいは研究論文が蓄積されてきて、それを踏まえて議論されているということで、混乱は起きていないと思われま

す。論文は多数出てきますけれども、ここに掲げている2つがよく引用されています。

最初のKalton&Kasprzykはかなり古典的な本で80年代の本ですが、分測値の補完方法としては、先ほど出てきました演繹的インピュテーション、先ほどAssignmentと言ったものです。それと、平均によるインピュテーションは全体の平均あるいは分類内層化して、その各層の中の平均という方法があります。

そのdとeに全体でのランダムなインピュテーション、分類内でのランダムなインピュテーションとありますが、これはホット・デック法といわれるものの別称でして、類似したデータをランダムに持ってくるというものです。

同じ調査内のほかのデータで埋めるもので、dとeがランダムなのですが、次のシークンシャルなホット・デックというのは、ホット・デックに用いるデータを順番に並べておいて、順番に使っていくという手法です。

あと、階層的なホット・デックインピュテーションというのは、分類内でのホット・デックなのですが、その分類内で使えるデータがなくなった場合は、分類を統合して上の方の、もう少し大きい分類で埋めていくという手法です。

次に、回帰によるインピュテーションあるいは距離関数によるマッチングという方法を挙げています。

最後の距離関数によるマッチングというのは、最近隣法と呼ばれるものと同じで、データ同士の距離をはかる関数を定義して、最も近いものの値を持ってくるということで、通常、これはホット・デック法の一つとされます。

このあたり、分類とか用語の定義が若干揺れのあるところ

です。基本的にこういうふうに分けたうち、それに残差を加えて分布の揺れを再現する。その場合はどういう残差を加えるかといった視点で分類しているということ。

それと、回帰によるインピュテーションというのが出てきているわけですが、その推定値の代わりに実際のデータがとった値で、それに最も近いものを入れる。それによってあり得ないデータが入ることを防ぐという予測平均値マッチングという手法も、この段階で挙げられています。

これは最近よく話題になっていますが、これは実際にあるデータを持ってくるという意味ではホット・デックの一種とされています。

次に挙げました Andridge and Little、2010年という、これはわりと新しい論文ですが、これはホット・デック法について論じた方法で、このホット・デック法というのは、海外では昔からわりと用いられていたのですが、実務的な手法でありこれに関する理論的な文献というのがなかったということで、その裏づけを行ったものです。

ただ、このホット・デック法というのは日本であまり行われていなくて、統計用語集でも定義がわりとはっきりしていなくて、コールド・デック法との違いも明確ではなかったのですが、単純に言いますと、ホット・デック法というのは、同じ調査の中の別のデータの値を持ってきて代入する方法、コールド・デック法というのは、代入する値をあらかじめ決めておく手法というふうないうことができると思います。

その下にある4番は、海外で商業出版されている書籍で、よく引用されているものをレビューしたものです。

4つ挙げていますが、最初の Rubin の 1987 年の論文というのは、「Multiple Imputation for Nonresponse in Surveys」という題名のとおり、多重代入法を論じた本で、一番最後に出てくる Van Buuren の本によりますと、これ以降、多重代入法に関する文献というのが非常に、指数関数的に増え続けているということです。

その間にコンピューターが各段に進歩して、最初は理論的な話だった反復シミュレーションというのがやりやすくなったということと、あと、多重代入法というのは統計だけではなくて、統計を使った分析の精度を評価する手法であるために、マイクロデータの提供が統計の精度評価以外に非常に要請が増えてきたことが背景にあると考えられます。

その次にある Little and Rubin の 2002 年の本と、最後の Van Buuren というのも理論的な書物で、多重代入法と、あとはそれを使った分析に焦点を当てています。

3つ目の De Waal というのが、公的統計の実務における欠測値の補完に焦点を絞った本で、これは 2011 年という比較的最近のものなのですが、商業出版されています。

こういうものが海外でも商業出版されているのが興味を引くところであります。

最後、5番目の統計機関による補完方法の解説。

特に見た感じでは、オランダ統計局というのは統計の方法論でかなり進歩的な情報公開をしているところですが、90年代まで遡って情報がとれます。

時間がないのと、非常に多いので、内容は省略しますが、御覧のようなものが出ています。

4ページに参りまして、アメリカのセンサス局でも 2000 年の人口センサス、それ以前は確認できなかったのですが、それについて、項目非回答についての報告書を作成しています。それで、発生状況とか補完手法の解説などを公開している。

あるいは人口動態調査、あるいはセンサス局以外の部署でもそういうものが公開されているというのが海外の現状ということです。

非常に駆け足でございましたが、報告させていただきました。

○樫分科会長 恐縮です。多分理論的ないろいろな関連する御質問等あったら、各委員の

方、実際に研究研修所の方にまたメール等を出していただいたらと思いますけれども、よろしいでしょうか。

どうもありがとうございます。

資料 10 の業務相談窓口ということについては。

○石原総務省統計研究研修所研究開発課長 それでは、業務相談窓口について説明いたします。

資料 10 を御覧ください。

業務についてですが、統計の作成や利用に対する各府省からの相談を受け付け、支援を行っていくということであり、背景・必要性は、資料 10 に記載されているとおりです。

具体的な相談業務の流れですが、各府省において統計に関しての疑問等がありましたら、統計研究研修所で受付をして、内容整理をした上で、相談を各統計の専門機関に振り分けて、回答の集約をまた研修所で行い、それを基に回答・対応するというものでございます。昨年の 8 月に設置以降、今のところ十数件の相談をいただいております。内容につきましては、調査を行う場合の適正なサンプルの大きさ、省内研修を行う場合や研修講師派遣に関するアドバイスといったものがありました。

以上です。

○樫分科会長 どうもありがとうございます。

是非各府省におかれましては、この種の各国の動向や、研究研修所からいただいた成果というものについて、欠測値への対応について、今後、生かしていただければと思いますし、この委員の皆様方に関しましても、こういうところが本来必要なのではないかということに関しては、是非研究研修所の方に対して御質問やコメントを頂戴して、これがまた各府省への支援につながるかと思っておりますので、よろしくお願いします。

申し訳ありません。既に定刻を過ぎてはいますが、最後に、欠測値への対応に関する内閣府の研究成果について御説明いただければと思います。

○石井内閣府経済社会総合研究所研究官室上席主任研究官 時間もございませんので、ごくかいつまんで御紹介させていただきたいと思っております。お手元の資料 11 を御覧ください。

1 枚めくっていただきまして、内閣府で平成 28 年度に欠測値補完の手法について、有識者の先生方にも入っていただきまして、調査研究を行いました。

内容としては、基本的な欠測データ処理の考え方ですとか、どんな処理方法があるのかというのを一覧にできるようにということをやったものです。

内容につきましては、恐らく今日、御出席の方々にはあまり目新しいものはないとは思いますが、1 つ 2 つ PR させていただきますと、もう 1 枚めくっていただきまして、スライドの右下に 2 という番号が振ってあるページですけれども、欠測処理をする実務の方にどんな手順を踏めばいいのかというのが分かるように、全体像を示したというのがこの調査研究の報告の 1 つのポイントになっています。

その下のスライドの 3 というところですが、最初のステップの段階では、例えばですけれども、欠測が起きているときにどういう状況で起きているのかというのを見てみましょう。

客体企業によっては、例えば調査票に記入する担当者が異動すると戻りが悪いですとか、あるいは、業績がちょっと悪いときはしばらく回答回数が落ちる傾向があるとか、そういったところを確認することも大事ではないかということを行っています。

それから、5番というふうに打ってあるスライドですけれども、これは欠測データのメカニズムを3種類に分類しています。

①、②、③で、①の完全にランダムというのは、これはバイアスが生じないので、そもそも問題になりませんが、②のランダムな欠測と③のランダムでない欠測というのでは対処方法が異なるということで注意を促しています。

②の方は、例えば調査対象が学生ですとか無職者の方ですと回答率が高いけれども、就業者の場合には落ちてしまうといったようなケースですけれども、この場合には補助変数を利用して、バイアスを緩和できるでしょう。

それに対して、下の③の方は、例えば金融資産保有額の調査で資産が高いほど、恐らく隠したいというようなこともあって回答率が落ちる。こういう場合には、変数の値自体に依存して欠測が生じますので、この場合には何らかのモデル化が必要ではないかということに注意喚起しています。

次のページのスライドの6番ですけれども、こちらは代表的な処理方法を一覧として載せています。今日、各府省から御紹介のありました取組でも既にこのようなものが使われていますし、先ほどの総務省の外国のサーベイでもいろいろ専門的なものがありましたけれども、実務に携わる方はこのようなものがあるのだということを入れて検討を始めていただければと思っています。

ごく簡単ではございますけれども、以上とさせていただきます。

○樫分科会長 どうもありがとうございます。

非常に具体的な手順で、本来だったら大変理論的には難しい部分もあるのですが、これは非常に具体的な手順に落とし込んでくださって、大変高く評価したいと思います。

何か簡単に御質問、コメントがあれば、よろしく願います。山本臨時委員。

○山本臨時委員 本当に簡単なのですけれども、この研究の報告書は、普通にどこかインターネット上で公開されているのでしょうか。

○石井内閣府経済社会総合研究所研究官室上席主任研究官 内閣府のホームページに報告書本体を載せていますので、もしよろしければご覧いただけますとありがたいです。

○山本臨時委員 分かりました、ありがとうございます。

○樫分科会長 先ほど申し上げましたように、非常に分かりやすくまとめていただいているのではないかと思いますので、是非各府省の方々がこの内容、研究成果なども実務に対する検討に生かしていただければと思いますので、これは私の方からも是非よろしく願いたいと思います。

ほかに先生方、よろしいでしょうか。

それでは、まさにそういうふうにしていただいて、研究研修所や何かの研修にもこういうところを含めていただくと非常にいいのではないかと考えながらコメントをさせていただきます。

どうもありがとうございました。大変失礼いたしました。非常に時間が長くなりましたけれども、予定された議事が終了しましたので、本日の分科会はここまでとさせていただきます。

今回は、精度検査報告書のフォローアップとして賃金構造基本統計調査及び薬事工業生産動態統計調査における取組の状況について、報告を聴取して審議を行うとともに、総務省と経済産業省による欠測値への対応に関する研究成果を御紹介いただく予定です。

最後に、事務局から次回の日程について連絡をお願いします。

○福田総務省統計委員会担当室室長補佐 次回は2月19日水曜日の午後4時に開催する予定です。場所は、今回と同じ若松庁舎6階のこの会議室で開催しますので、よろしくお願いします。詳細につきましては、別途、御連絡します。

また、第7回につきましては、3月25日火曜日の午後4時から開催する予定です。場所は現在、調整中ですので、別途、お知らせします。よろしくお願いします。

○椿分科会長 どうもありがとうございました。

それでは、本日はこれまでとします。

皆様方には、大変お忙しいところ、審議に御協力いただいたことを心から感謝申し上げます。若干の遅延になったこともお詫び申し上げます。

どうもありがとうございました。