

平成 30 年度 (2018 年度) 統計法施行状況報告書
(抜粋 : V-4 統計技術の評価に関する事項)

令和元年 6 月 27 日

総 務 省

政 策 統 括 官
(統 計 基 準 担 当)

V その他

4 統計技術の評価に資する事項

(1) 統計技術評価の取組の概要

① 統計技術評価の取組の根拠

法第55条の規定では、総務大臣は国の行政機関の長等に対し法の施行の状況について報告を求めることができ、毎年度、当該報告を取りまとめ、その概要を公表するとともに、統計委員会に報告しなければならないとされている。

また、法第55条第3項により、統計委員会は、法第55条第2項の規定による報告があったときは、この法律の施行に関し、総務大臣又は関係行政機関の長に対し、意見を述べることができるとされている。

さらに、統計委員会令では、第1条第1項において「統計委員会（以下「委員会」という。）に、評価分科会（以下「分科会」という。）を置く。」とされ、同条第2項において「分科会は、委員会の所掌事務のうち、統計法第55条第3項の規定により委員会の権限に属させられた事項（同法の施行に関し、主として統計技術の観点から評価を行い、その結果に基づき意見を述べることに限る。）を処理することをつかさどる。」とされていることから、これらに基づき統計技術評価の取組を実施した。

② 統計技術評価に資する報告

ア) 建築着工統計調査の標本設計移行時の推計方法に関する報告

「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30日総務省統計委員会。以下「統計精度検査報告書」という。）においては、建築着工統計調査補正調査について、

- i) 平成33年（2021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。
- ii) この切り替えの際、a) 見直し後の標本設計に基づき抽出された調査対象と、b) 現行の抽出方法に基づき抽出され都道府県から国土交通省に移管された調査対象、の両方を混在させて調査を行うことが必要となり、その移行期の結果推定には統計技術的な対応が必要である。そのため、総務省は、移行時期の結果推定の方法の開発について国土交通省に協力し、円滑な調査の移行を支援することが必要である。

との趣旨の記載がある。

このため、総務省において国土交通省から建築着工統調査の調査票情報の提供を受け、リサンプリングによる標本抽出法の移行期におけ

る推計方法のシミュレーションを行い、その結果を（２）①のとおり取りまとめた。

イ) 経済産業省生産動態統計調査における欠測値補完に関する報告

統計精度検査報告書においては、

- i) 商業動態統計調査（丙調査（百貨店・スーパー）及び丁3調査（ドラッグストア））、経済産業省生産動態統計調査及び経済産業省特定業種石油等消費統計調査では、回収率90%以上を保持していることから、致命的な対応とは判断されないが、平成30年度から総務省の支援を得て、計画的にシミュレーションを行いデータの使用期限も含めた補完方法の検証を行うことが必要である。
- ii) 経済産業省が所管する
 - ・ 商業動態統計調査
 - ・ 経済産業省生産動態統計調査
 - ・ 経済産業省特定業種石油等消費統計調査

のうちいずれかの調査を総務省及び経済産業省が選択して、過去の回答結果を使用した単一補完を実施する場合の使用データの期限の検討に資するシミュレーションを実施し、その結果について統計委員会に報告するとともに、経済産業省に対し欠測値への対応のより適切な対応を提案する。

とされている。このため、検討の端緒として、経済産業省生産動態統計調査の109種類の調査票のうち、我が国の代表産業である自動車等を調査する調査票を選択し、品目「軽自動車」、「小型自動車」、「普通自動車」、「バスシャシー（完成車を含む）－小型バス」、「バスシャシー（完成車を含む）－大型バス」、「トラックシャシー（完成車を含む）－軽自動車」、「トラックシャシー（完成車を含む）－小型自動車－ガソリン車」、「トラックシャシー（完成車を含む）－小型自動車－ディーゼル車」、「トラックシャシー（完成車を含む）－普通自動車－ガソリン車」、「トラックシャシー（完成車を含む）－普通自動車－ディーゼル車」、「トラックシャシー（完成車を含む）－けん引車」、「特殊自動車」及び「トレーラ」の13品目について、総務省において経済産業省から同調査の調査票情報の提供を受け、ジャックナイフ法を応用した欠測値補完法のシミュレーションを行い、その結果を（２）②のとおり取りまとめた。

（２）統計技術評価に資する報告

① 建築着工統計調査の補正調査の標本設計移行時の推計方法に関するシミュレーション結果

ア) 建築着工統計調査の補正調査の概要

建築着工統計調査は、

- i) 全国の建築物（着工建築物）の着工状況（建築物の数、床面積合計、工事費予定額等）を延べ床面積10㎡超の建築物に届出が義務付けられている「建築工事届」から集計する本体調査（全数調査）
- ii) 本体調査の建築物のうち、住宅の着工状況（戸数及び床面積合計）を構造、建て方、利用関係、資金等に分類して集計する住宅着工統計（全数調査）
- iii) 本体調査の建築物のうち、建築物の竣工時に実際にかかった費用（工事実施額）を実地に調査し、着工時における工事費予定額とのかい離を明らかにする補正調査（標本調査）から構成されている。

このうち補正調査の標本抽出は、①調査実施市区を抽出し、②更に調査実施市区内の建築物に関する「建築工事届」を調査実施市区ごとに定められた抽出率で抽出する方法が採用されている。

イ) 建築着工統計調査の補正調査の見直しの概要と課題

建築着工統計調査の補正調査については、標本抽出法を次の i) から v) までのとおり見直すことで、同じ標本数で標本誤差を $1/5$ に縮小することが見込まれている。

- i) 木造・非木造別に工事予定額階級3区分（20億円以上、1～20億円、1億円未満）の6層に標本を区分する。
- ii) 20億円以上の2層は全数を抽出する。
- iii) 残りの4層はネイマン配分により標本を配分する。
- iv) 合計の標本数を5,000件とする。
- v) 令和3年（2021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。

このような対応を行うことにより、令和3年1月から12月までに完成する調査対象は、旧抽出方法で抽出された標本と見直し後の抽出方法で抽出された標本が混在した状態で調査が行われる。一方、抽出切り替え後、一定期間が経過すれば、全て見直し後の抽出方法で抽出された標本で調査が行われることから、集計方法は単純化される。

このため、移行期間のために特別な集計を行う必要があることから、効率的かつ偏りのない方法での集計が必要と考えられる。

ウ) 新旧の方法により抽出された標本の混在期間における集計方法

建築着工統計調査の補正調査については、本体調査において「工事予定額」及び「完成予定時期」が全数で調査されており、工期は1か月から36か月で完成したものが主な調査対象となっている。このため、例えば、令和元年12月まで旧標本設計で抽出し、令和2年1月以降は新標本設計で抽出を行い、令和3年1月より抽出された標本のうち完成した建

建築物に対し新たな調査が開始されることが考えられる。このような場合、不偏推定量を構成するためには、

- i) 届出年（抽出した年）及び（完成年を近似する）完成予定年（＝令和3年）の3区分ごとに、構造（2区分）及び工事予定金額階級（3区分）を場合分けして全数集計を行い、それぞれの部分母集団の総数を計算する。
- ii) 令和3年に完成し調査対象となった標本の数についても届出年（抽出した年）の3区分ごとに構造（2区分）及び工事予定金額階級（3区分）を場合分けして数える。
- iii) 上記ア）及び上記イ）で計算された部分母集団数及び標本数に基づき、以下の推計式で総額推計を行う。

【推計式 1】

$$\sum_{i,j,k} \sum_{l=\text{令和3年}} \sum_m \overset{\text{計測できない値}}{\downarrow} N(i,j,k,l) \frac{X(i,j,k,l,m)}{n(i,j,k,l)}$$

$$\equiv \sum_{i,j,k} \sum_{\substack{l=\text{令和3年} \\ l'=\text{令和3年}}} \sum_m N'(i,j,k,l') \frac{X(i,j,k,l,m)}{n(i,j,k,l)}$$

- i : 構造（木造、非木造）
- j : 完成予定金額階級（1億円未満、1～20億円、20億円以上）
- k : 届出年（令和元年、令和2年、令和3年）
- l : 完成年（令和3年）
- l' : 完成予定年（令和3年）
- $N(i,j,k,l) \equiv N'(i,j,k,l')$: 構造 i、完成予定金額階級 j、届出年 k、完成年 i の建築物の総数
- $n(i,j,k,l)$: 構造 i、完成予定金額階級 j、届出年 k、完成年 i の建築物の標本数
- $X(i,j,k,l,m)$: 構造 i、完成予定金額階級 j、届出年 k、完成年 l の m 番目の標本の工事実施額等の観測値

しかしながら、過渡的な期間のために場合分けを多用する推計式を用いて集計することは非効率的と考えられる。そのため、本来の理論上の算式は完成時期の条件付確率の逆数をウェイトとして加重平均する算式とすべきところ、完成時期の条件を無視し、単に抽出時の抽出確率の逆数をウェイトとする算式が考えられる【推計式2】。この簡易な算式は、これまでの建築着工統計調査の補正調査のデータ特性上では漸近的に不偏であることが後述するシミュレーションにより確認されている。

【推計式 2】

$$\sum_{i,j,k} \sum_{l=\text{令和3年}} \sum_m N(i,j,k) \frac{X(i,j,k,l,m)}{n(i,j,k)}$$

i : 構造 (木造、非木造)

j : 完成予定金額階級 (1 億円未満、1 ~20 億円、20 億円以上)

k : 届出年 (令和元年、令和 2 年、令和 3 年)

l : 完成年 (令和 3 年)

l' : 完成予定年 (令和 3 年)

N(i, j, k) : 構造 i、完成予定金額階級 j、届出年 k の建築物の総数

n(l, j, k) : 構造 i、完成予定金額階級 j、届出年 k、完成年 i の建築物の標本数

X(i, j, k, l, m) : 構造 i、完成予定金額階級 j、届出年 k、完成年 i の m 番目の標本の工事実施額等の観測値

エ) 簡易な集計方法の不偏性の確認

建築着工統計調査の補正調査は、届出時に抽出し、完成時期に調査をかけるが、工期が長いものほど工事実施額が高額となる相関関係が見られる。そのため、推計式 2 のように抽出率の逆数を乗じるだけでは、推定結果に偏りが生じる可能性も否定できない。

このため、総務省では、単純ランダムサンプリングで抽出されたと想定される 2010 年から 2016 年までの間の建築着工統計調査の補正調査の調査票情報を国土交通省から提供され、これを基に次のようなシミュレーションを実施した。なお、提供を受けた調査票情報は、比例配分法により無作為抽出していると見なすことができる方法で抽出されたものである。

i) 2010~2012年に届出された調査票情報でのシミュレーション

a) 2010年に届出された調査票情報から、重複を許して旧標本抽出法で5,000件をリサンプリングする。

b) 2011年及び2012年に届出された調査票情報から、重複を許して新抽出法であるネイマン配分で配分された標本数をそれぞれの層からリサンプリングする。

c) リサンプリングされた調査票情報から、2012年に完成するもののみを選定し、上記【推計式 2】の推計式に基づき計算される値から改定率等の統計量を計算する。これを1,000回繰り返す。

d) 2010~2012年に届出された調査票情報のうち2012年に完成するものの全体で計算された改定率等の統計量を計算する (標本空間上の平均)。

e) 上記 c) の統計量は上記 d) を推定していることとなるため、c) の結果が d) の値を偏りなく推定しているか確認する。

ii) 上記 i) を2011年~2013年、2012~2014年、2013~2015年、2014~2016年、の調査票情報の組で繰り返し計算し、いずれの年でも偏り

がないことを確認する。

オ) 簡易な集計方法のシミュレーションの結果

総務省によるシミュレーションの結果は表23～25のとおりであり、前述ウ)の【推計式2】はいずれの年においても、標本空間の平均とシミュレーションの結果に大きな差はなく、概ね偏りなく推計できていると考えられ、新旧の方法で抽出された標本が混在した状態で処理を行う期間が短期であることも踏まえれば、簡易な方法による集計においても、公的統計として問題なく単価及び補正率が推計されるものと考えられる。

表 23 シミュレーション結果 1 (単位面積当たり予定額単価)

木造			非木造		
リサンプリングデータ	標本空間の平均	試算結果 1000件の平均 (標準偏差)	リサンプリングデータ	標本空間の平均	試算結果 1000件の平均 (標準偏差)
2010-2012年	15.8	15.80(0.143)	2010-2012年	16.6	16.71(0.504)
2011-2013年	15.8	15.77(0.113)	2011-2013年	17.3	17.25(0.280)
2012-2014年	16.7	16.76(0.127)	2012-2014年	17.8	17.76(0.176)
2013-2015年	16.8	16.76(0.140)	2013-2015年	18.7	18.49(0.240)
2014-2016年	16.8	16.79(0.173)	2014-2016年	20.5	20.35(0.385)

表 24 シミュレーション結果 2 (単位面積当たり実施額単価)

木造			非木造		
リサンプリングデータ	標本空間の平均	試算結果 1000件の平均 (標準偏差)	リサンプリングデータ	標本空間の平均	試算結果 1000件の平均 (標準偏差)
2010-2012年	16.1	16.02(0.154)	2010-2012年	17.7	17.86(0.565)
2011-2013年	16.2	16.11(0.133)	2011-2013年	17.5	17.50(0.271)
2012-2014年	17.2	17.25(0.144)	2012-2014年	18.5	18.63(0.245)
2013-2015年	17.3	17.27(0.150)	2013-2015年	19.5	19.29(0.257)
2014-2016年	17.2	17.19(0.184)	2014-2016年	21.5	21.12(0.369)

表 25 シミュレーション結果 3 (改定率)

木造			非木造		
リサンプリングデータ	標本空間の平均	試算結果 1000件の平均 (標準偏差)	リサンプリングデータ	標本空間の平均	試算結果 1000件の平均 (標準偏差)
2010-2012年	1.01	1.014(0.004)	2010-2012年	1.07	1.069(0.009)
2011-2013年	1.02	1.022(0.002)	2011-2013年	1.02	1.014(0.007)
2012-2014年	1.03	1.029(0.003)	2012-2014年	1.04	1.049(0.015)
2013-2015年	1.03	1.031(0.004)	2013-2015年	1.04	1.043(0.005)
2014-2016年	1.03	1.023(0.003)	2014-2016年	1.05	1.037(0.003)

このため、移行期の集計は、【推計式1】で対応するか、簡易な集計式である【推計式2】で推計することが考えられる。

② 経済産業省生産動態統計調査における欠測値補完に関するシミュレーション結果

ア) 経済産業省生産動態統計調査の概要

経済産業省生産動態統計調査は、次の i) 及び ii) までの内容の基幹統計調査である。

- i) 特定の製品を生産する事業所に対し、一定の要件を満たす事業所（全体で約14,000事業所）を全数で毎月調査
- ii) 109種類の調査票を用い、次の調査事項を調査
 - 製品に関する事項
 - ・生産(数量・重量・金額等)
 - ・受入(数量・重量等)
 - ・消費(数量・重量等)
 - ・出荷(数量・重量・金額等)
 - ・在庫(数量・重量等)
 - 原材料に関する事項
 - ・受入(数量・重量等)
 - ・消費(数量・重量等)
 - ・在庫(数量・重量等)
 - 労務に関する事項
 - ・月末従事者数(事業所、当該製造部門及び当該品目群)
 - 生産能力・設備に関する事項
 - ・生産能力
 - ・月末保有台数 等

イ) 経済産業省生産動態統計調査における欠測値補完の課題

経済産業省生産動態統計調査は回収率が95%程度であり、欠測値が精度に多大な影響を与えている状況にはない。しかし、欠測値が発生した場合、過去の回答結果を使用した単一補完（Last Observation Carried Forward: LOCF）が明確な使用期限を定めことなく行われており、長期にわたって同一調査対象に対しLOCFを続けた場合、統計精度の悪化につながる懸念がある。

このため、総務省統計研究研修所の協力を得て、検討の端緒として、調査票「機械器具月報(その40)自動車」（以下「自動車等の調査票」という。品目は、「軽自動車」、「小型自動車」、「普通自動車」、「バスシャシー（完成車を含む）－小型バス」、「バスシャシー（完成車を含む）－大型バス」、「トラックシャシー（完成車を含む）－軽自動車」、「トラックシャシー（完成車を含む）－小型自動車－ガソリン車」、「トラックシャシー（完成車を含む）－小型自動車－ディーゼル車」、「トラックシャシー（完成車を含む）－普通自動車－ガソリン車」、「トラックシャシー（完成車を含む）－普通自動車－

ディーゼル車」、「トラックシャーシ（完成車を含む）－けん引車」、「特殊自動車」及び「トレーラ」）について試行的に検証を行い、LOCFのデータの使用期限について検討を行った。

ウ) 試行的な検証の方法

試行的に検証を行った自動車等の調査票は、事業所数が8～17と限られていることから、検討期間である2003年から2017年までに回答を得られている事業所の回答データに対して、1事業所の回答のみ期首から人工的に欠測させ補完を行い販売金額の合計を算出し、これをすべての事業所について順番に計算することにより、①現行のLOCFのパフォーマンスについて他の方法と比較評価を行うとともに、②どの程度の期間LOCFを継続すると精度が悪化するかについての評価を行った。

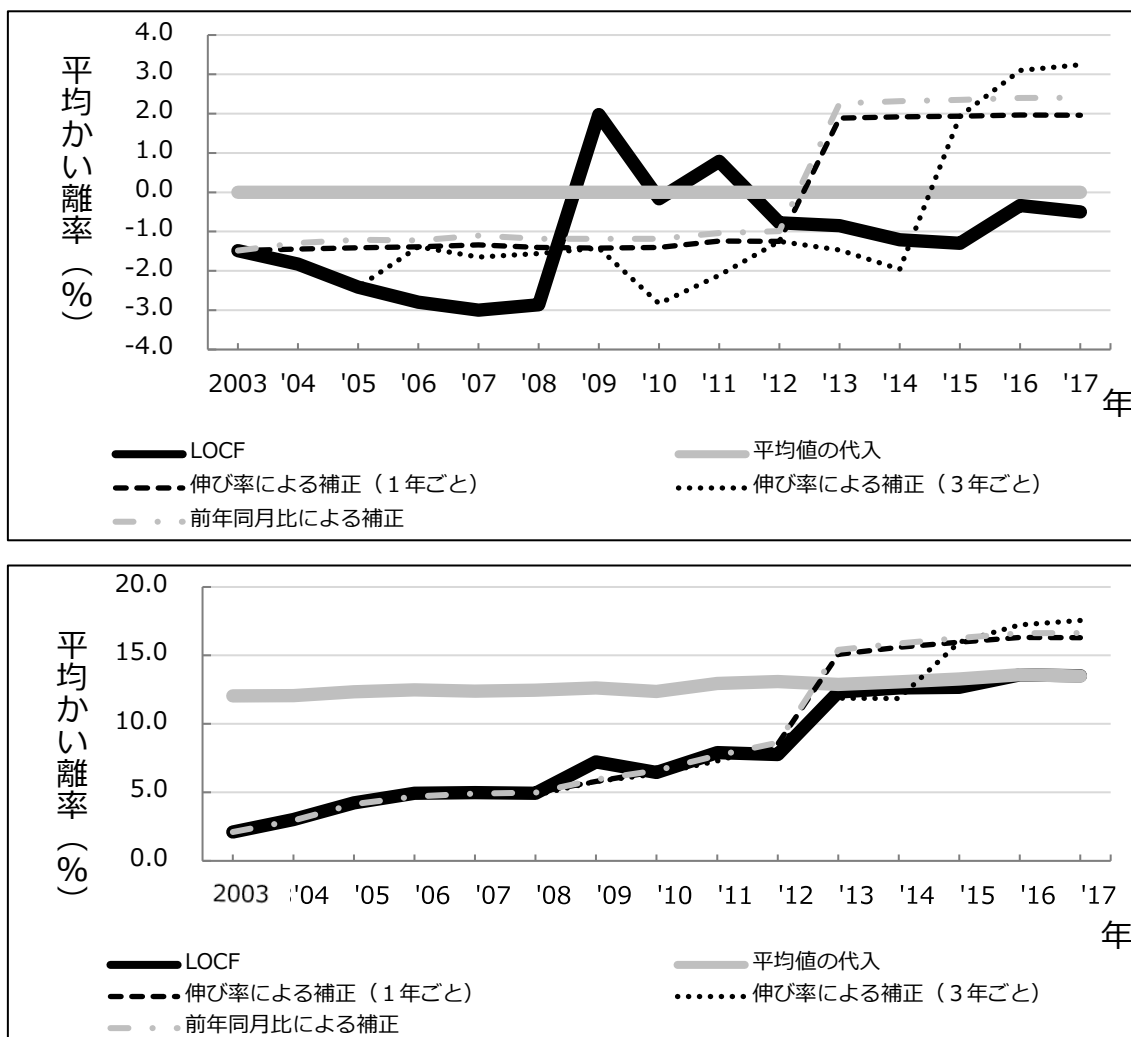
エ) 試行的な検証の結果

い) 現行の補完方法のパフォーマンス評価

試行的な検証として、自動車等の調査票の調査票情報について、現在採用されているLOCFに加え、平均値代入法、年間の伸び率によるLOCFの補正及び前年同月比によるLOCFの補正を行い、欠測及び補完を行わなかった結果とのかい離により評価した。

この結果、LOCFは、平均値代入法よりよいパフォーマンスを示し、年間の伸び率によるLOCFの補正や前年同月比によるLOCFの補正と比べてもパフォーマンスは劣らなかった（図2）。品目ごとに見ても、おおむね同様の結果となっている。このため、現状の補完方法を伸び率によるLOCFの補正などに変更しても、それほど大きな精度改善は見込まれないと考えられる。ただし、「バスシャーシ（完成車を含む）－小型バス」、「トラックシャーシ（完成車を含む）－けん引車」及び「特殊自動車」のように、急激に金額の水準が変化している品目では、LOCFのパフォーマンスが低下する傾向及び補正を行うことでLOCF開始時の水準までパフォーマンスが改善される傾向がみられた。このため、急激な金額の水準変化が発生した場合は、何らかの補正を検討することが有効であると考えられる。

図2 各種補完法による平均かい離率の全品目平均（上図：符号付、下図：絶対値）



ii) 継続的に補完に用いるデータの期間

試行的な検証の結果、継続的に補完を行った場合であっても、一部の特殊な品目を除き、多くの品目で長期的にかい離が極端に拡大する傾向は観測されなかった（図3）。

また、LOCFの開始年を2003年だけでなく1年ずつ移動させてLOCF開始後経過年数ごとの平均かい離率を求めた（図4）。この結果からも一部の特殊な品目を除き長期的にかい離が極端に拡大する傾向は観測されなかった。

個別の品目を見ると、リーマン・ブラザーズが破綻し数値の変動が大きかった2008年前後でかい離が変動する傾向が見られ、また、「特殊自動車」のように長期的に金額の水準が大きく変化する場合は時間経過と共にかい離が増大する傾向が見られた。

図3 LOCFの継続による販売金額及び平均かい離率の全品目平均の推移
 (上図：符号付、下図：絶対値)

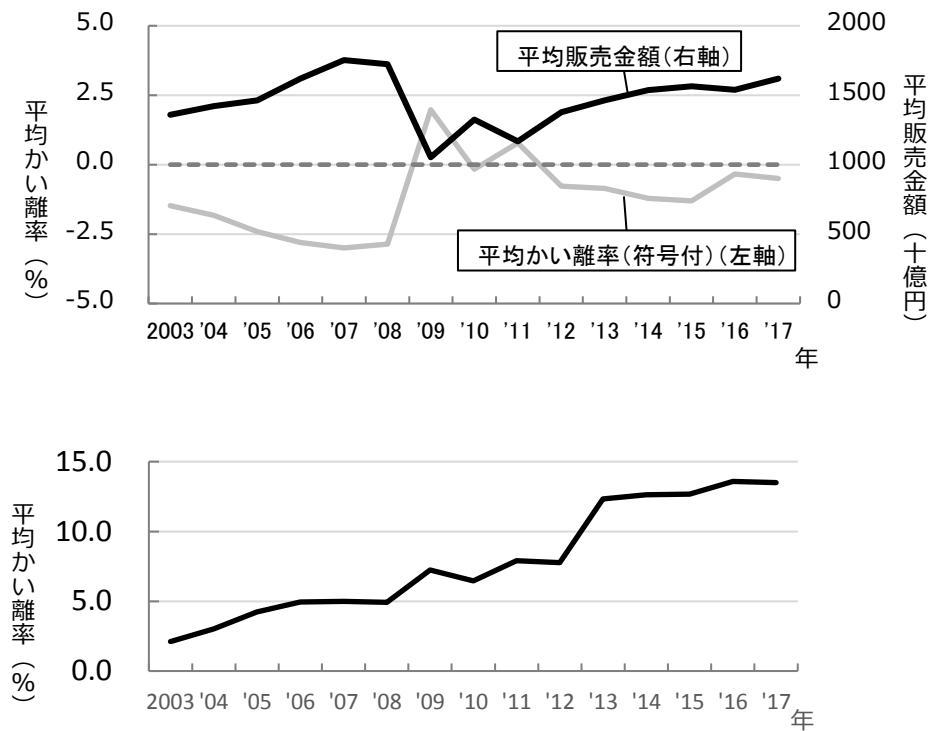
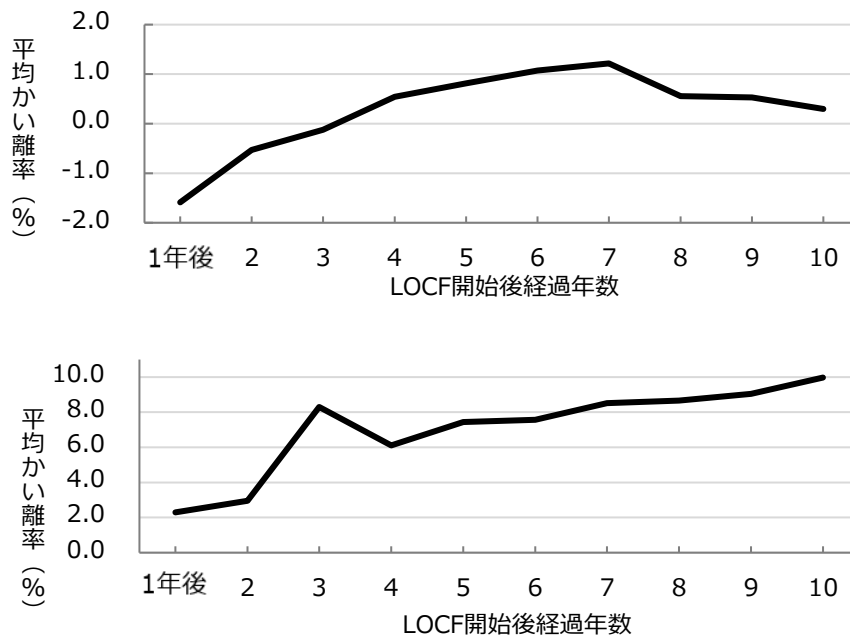


図4 開始年移動分析によるLOCF開始後経過年数による平均かい離率の全品目平均の推移
 (上図：符号付、下図：絶対値)



これらの結果から、自動車等の調査票に含まれる品目については、極端な金額の水準変化がない場合において、LOCFは数年間それを継続したとしても有効に機能していると整理される。

一方、極端に長期のLOCFについては信頼性に疑義が投げかけられる可能性もあることから、次のように対応することが考えられる。

- a) 自動車等の調査票に含まれる品目のうち、金額の水準変化が大きくない品目であって補完期間が5年に至るものについて、5年ごとに実施される経済センサス-活動調査を活用し、経済センサス-活動調査の当該事業所の伸び率等でLOCF値を補正するなどの対応を行う。
- b) 自動車等の調査票に含まれる品目のうち、複数年にわたり金額の大きな水準変化が続いている品目については、2年以内に同じ品目内の伸び率等を乗じてLOCF値を補正し、あわせて、5年ごとの経済センサス-活動調査の当該事業所の結果を用いて補正するなどの対応を行う。

オ) 他の品目への転用

今回の検証は、検討の端緒として自動車等の調査票に限ってジャックナイフ法を応用した補完精度の検証を行ったものであり、厳密には他品目にこの成果は応用できない。しかしながら、経済産業省生産動態統計調査の全調査票に対し同じような検証を行うには相当な時間がかかる。

一方、経済産業省生産動態統計調査の結果作成は絶えず行われており、自動車等の調査票以外の調査票に対しても、期限を定めないLOCFが継続されている状況にある。

自動車等の調査票の検証結果は、

- a) 安定した推移をしている場合、経済センサスを活用し5年ごとにチェックを行う
- b) 安定した推移をしていない場合（複数年にわたって増加が続く場合、又は減少が続く場合）には5年を待たず品目ごとの伸び率で補完値の補正を行い、補正結果について、経済センサスを活用し5年ごとにチェックする。

といった、常識的かつ包括的な考え方で対応することとしており、検証を行っていない品目についても、当該ルールを暫定的に導入して対応することが考えられる。