

消費者物価地域差指数用ウエイトの作成に用いる支出金額の 推計手法及び推計結果の評価

櫻井 智章[†]

An evaluation of the estimation methods and the estimation results of expenditures for the weights used to calculate the regional difference index

SAKURAI Tomoaki

平成25年の小売物価統計調査（構造編）創設に伴い、消費者物価地域差指数用ウエイト作成に使用する、品目別・市町村別の1ヶ月・一世帯当たり支出金額の新たな推計手法が導入された。本稿の目的は、新たに導入された手法の定量的な評価を行うことである。また、本稿では推計結果の精度を安定的に確保するため、推計結果を定量的に評価するための手法を提案した。導入した評価手法の解説も行う。

キーワード：消費者物価地域差指数、小売物価統計調査（構造編）、ウエイト、家計調査、小地域推計

A new estimation method of expenditures by item for municipalities per a month of one household for the weights used to calculate the regional difference index was established at the beginning of the Retail Price Survey (Structural survey). The purpose of this paper is to evaluate this estimation method in quantitative manner. Moreover, the paper will introduce an evaluation method of the estimation results of expenditures in order to maintain quality. This paper will also explain this estimation method.

Key words: Regional difference index, Retail Price Survey (Structural Survey), Weights, Family Income and Expenditure Survey, Small area estimation

[†] 総務省統計局統計調査部消費統計課物価統計室

1. はじめに

物価水準の地域間比較のために用いられる消費者物価地域差指数（以下「地域差指数」という。）は、平成 19 年までは 5 年ごとに全国物価統計調査を実施して公表していた。しかし、5 年という周期では、加速する近年の消費・物価構造分析のニーズに十分に 대응することができないため、全国物価統計調査を発展的に見直すことで小売物価統計調査（構造編）を創設し、平成 25 年分の結果より、地域差指数を毎年公表している¹。

地域差指数は、小売物価統計調査（動向編及び構造編）で収集した価格に加え、消費構造を指数に反映させるためのウェイトを使用する。ウェイトの作成には、総務省統計局で毎月実施している家計調査結果から作成した品目別・市町村別の一ヶ月・一世帯当たり支出金額（以下「品目別・市町村別支出金額」という。）の推計結果を使用する。家計調査は全国の 168 市町村において調査を行っているが、県庁所在市及び政令指定都市を除き、市町村別の公表を前提とした標本設計となっていないため（総務省統計局(1)）、別途何らかの手法により品目別・市町村別支出金額の推計結果を安定化させる必要がある。これは一般的に小地域推計問題と呼ばれ、多くの分野において非常に関心の高い問題として扱われており、今日までに多くの推計手法が提案されている（久保川(2006)、坂田(2011)、Ghosh and Rao(1994)）。

全国物価統計調査における地域差指数の作成においては、間接推計と呼ばれる手法により、品目別・市町村別支出金額の推計を行っていた（総務省統計局(2)）。これは、当該市町村の値を推計するために、当該市町村に属さない標本（周辺の市町村）を利用して推計を行うものである。しかしこの手法は、当該市町村と周辺の市町村の値が同水準という前提の下では非常に有効であるが、この前提が成り立たない場合、推計結果と真の値とのバイアスが拡大することになり、品目別・市町村別支出金額の推計結果の精度、ひいては地域差指数の精度を著しく損なう恐れがある。

我々は、構造編の創設に伴い、品目別・市町村別支出金額の推計結果の精度を安定的に確保するため、全国物価統計調査で採用していた推計手法を見直し、平成 25、26 及び 27 年結果の作成に新たな推計手法を適用した。本研究の目的は、導入した推計手法の定量的な評価を行うことである²。推計手法の評価は、他の推計手法との比較検証により行う。比較対象は、単年分の家計調査結果の単純集計による手法及び全国物価統計調査において採用していた手法である。

検証の結果、導入した手法は、多くの品目において、平均二乗誤差及び標準誤差の市町村平均が、上記 2 手法よりも縮小するという意味で改善していることを確認した。一方、平均二乗誤差及び標準誤差の市町村平均が拡大した品目の存在も確認出来た。このことは、推計結果の精度を確保することが困難なことがあり得ることを示唆する。

推計結果の品質を保つため、我々は、「誤差の縮小」、「支出金額の水準の維持」に加え、市町村別の推計値を得るといった目的に鑑みて、「地域性³の維持」を評価基準に加えた、簡易的な評価手法を提案した。これについても解説を行う。

本稿の構成は以下のとおりである。「2. 推計手法」では、構造編の創設に伴い導入した、品目別・市町村別支出金額の推計手法について解説する。「3. 推計手法の評価」では、平均二乗誤差及び標準誤差を使用した評価の方法及び導入した推計手法の評価結果について解説する。「4. 推計結果の評価」では、推計結果の品質を保つための評価手法を提案し、

¹ 詳細は、総務省統計局ホームページ「小売物価統計調査（構造編）」(<http://www.stat.go.jp/data/kouri/kouzou/index.htm>)を参照。

² 検証を実施した時点では、平成 26 年分結果まで公表済であった。従って、本稿における検証は、その時点における最新のデータ（平成 26 年）までを使用して行った結果である。

³ 本研究における「地域性」の定義は、付録 1. (1)に記載する。

推計結果の評価結果について考察を行う。「5. おわりに」では、まとめと今後の展望について述べる。

2. 推計手法

2.1. 導入

一般的に小地域推計の手法は、「直接推計」、「間接推計」及び「モデル推計」の3つのカテゴリーに分けられる。直接推計は、当該地域の個票データのみを使用して推計する手法であり、個票データの単純集計が該当する。また、個票データ以外に補助情報（当該地域の母集団の大きさなど）が既知であれば、比推定などにより標本誤差を縮小することが可能である。家計調査の公表値作成に使用する集計用乗率は、国勢調査結果を用いて作成した地域ごとの世帯数を補助情報として、比推定により作成していることから（総務省統計局(1)）、家計調査の集計用乗率を市町村ごとに単純集計すること（以下、集計した値を「単年集計結果」という。）は直接推計であると言える。間接推計は、1.でも述べたとおり、当該市町村に属さない標本（周辺の市町村）を利用して推計を行うものであり、全国物価統計調査で採用していた手法が挙げられる。モデル推計は、推計したい値に適切なモデルを仮定し、モデルに基づき推計結果を得る手法で、時系列回帰モデルや線形混合モデルなどが知られている。近年はモデル推計が主流となっており、官庁統計を含む多くの分野において、モデル推計手法の導入が行われている（久保川(2006)、小泉(2004)、坂田(2011)、高部(2004)、元山・山口(2007)）。また、精度面を考慮するとモデル推計の導入を検討することが望ましいことである。しかし、我々が取り扱う品目数は500近くあるため、導入のためには長期的な検討が必要である。加えて、運用の容易さといった実務的な面を考慮すると、簡易的な手法により、高い効果が得られる手法が望ましい。

このようなことから、我々は、品目別・市町村別支出金額を推計するために、以下に述べる手法を考えた。

従来は単年分の家計調査結果を使用していたところを、複数年分（具体的には直近5年分）の家計調査結果を使用することで、各市町村の標本数を増やすこととした⁴。また、各品目の支出金額を、それが属する上位類の支出金額を適切に按分することにより推計し、標本誤差の縮小効果を高めることを考えた。本手法は、各市町村の値の推計に周辺の市町村の個票データを使用しておらず、各市町村の個票データから直接推計していることから、直接推計の手法に分類される⁵。

なお、当初、按分処理は、耐久財やサービス品目など、購入頻度（利用頻度）が低く、使用する年数を増やすことによる標本誤差の縮小効果がそれ程得られない品目への適用を想定していたが、却って誤差が拡大してしまうことや、過剰な推計により、真の値とのバイアスが拡大することによる精度の低下を避けるため、別途基準を設けて、対象とする品目を選定することとした。

以下に、それぞれの考え方について考察を行い、具体的な推計手法について述べる。

2.2. 使用するデータ

複数年分のデータを使用することは、標本誤差を縮小させる上で有効な方法の一つであり、

⁴ 精度を保ちつつ標本誤差を縮小する（安定化させる）上で最も確実な方法は標本数を増やすことである。しかし、一回の抽出における調査世帯数を増やすことは、家計調査の方法を変更することが必要となるため、困難である。そこで、集計に使用するデータの年数を増やすことで、これを回避した。

⁵ 按分処理について、当該品目の推計に、当該品目以外（上位類）の支出金額を使用することから、本手法を間接推計と見ることが可能であるが、間接推計は、当該地域の推計に当該地域に属さない個票データを利用する推計手法であり、また、上位類の値は補助情報とみなすことが可能であるため、本手法は直接推計であると言える。

使用する年数が多い方が、得られる効果は大きくなる。しかし、年数を増やすことで、消費支出の経年変化の影響を受け、推計したい年次における真の値とのバイアスが拡大することが考えられる。このようなことから、標本誤差を縮小させる点では、使用する年数は多い方がよいが、経年変化の影響を考慮すると、直近の数年分のデータを使用することが望ましいと考えられる。そこで問題となるのは、何年分を使用するのが適切かということである。

また、家計調査は5年ごとに調査品目の一部入れ替えを行っているため、複数年分のデータを使用する場合、新旧品目の接続という問題が発生する。新品目が複数の旧品目を合わせたものである場合、新旧の接続は、該当する旧品目の支出金額を単純に合算させることで行える。しかし、新品目が旧品目からの分割である場合、新旧の接続において、旧品目の支出金額をどう扱うのかが問題となる。ここで、後者の場合の、旧品目から新品目への接続における基本的な考え方は、以下の2通りある。

- ①旧品目を新品目に合わせるために按分する
- ②旧品目の年次データは使用しない

①については、接続結果が按分比の作成方法に依存するため、場合によっては真の値とのバイアスが拡大し、地域性が不整合となる恐れがある⁶。②については、品目ごとに使用する年数を変えるため、継続品目との扱いが整合的ではない。また、使用する年次が減るため、標本誤差の縮小効果が低下することも考えられる。

5年ごとに入れ替えを行う品目数は必ずしも多いわけではないが、標本誤差の縮小効果を最大限にしつつ、新旧品目の接続による推計結果への影響を最小限に抑えるという観点から（複数年分を使用することから、必ず1回の接続は必要）、1回分の見直しの対応で済ませることが可能な範囲で最大の5年分のデータを使用することとした。また、品目の入れ替えが行われた際の接続方法については、使用する年次が減ることによる標本誤差の縮小効果の低下よりも地域性を温存することを優先させるため、新旧品目の接続は②を採用した。

2.3. 上位費目・類から按分する品目の選定⁷

当初の想定では、上位費目・類からの按分を導入する品目は耐久財やサービス品目など、購入頻度（利用頻度）が低い品目としていた。しかし、上位費目・類からの按分による推計が有効となるのは、上位費目・類の購入頻度がある程度確保出来ており、標本誤差の縮小が見込める場合である。また、上位類には費目、中分類及び小分類と異なるレベルがあり、選定するレベルによって推計結果に与える影響がそれぞれ異なるものと考えられ、特に、上位費目・類と当該品目の地域性が極端に不整合であると、推計結果と真の値の地域性が不整合となる恐れがある。加えて、按分処理は標本誤差を縮小させる上で有効であると期待されるが、一方で、加工度が高いため、やみくもに使用することは避ける必要がある。

このような考察から、対象とする品目の選定にはある一定の基準を設け、出来るだけ誤差の拡大や、真の値との地域性の不整合が起らないように、上位費目・類と品目の組を選定することとした。また、効率的な品目選定を行うため、按分により推計を行う品目を個別に選定するのではなく、ある程度まとめて選定を行うため上位類を選定対象とし、選定された上位類に属する全ての品目を按分対象とする方針とした。

以下に(1)按分対象とする上位費目・類のレベルの設定、(2)按分対象とする上位類の選定、

⁶ 真の値は知ることができないことから、真の値との地域性の整合性を見ることはできない。仮に真の値を知り得た場合、それとの地域性が整合的ではなくなる可能性がある、という意味で「地域性が不整合となる恐れがある」と述べたものである。

⁷ 本稿において使用している収支項目分類は、平成22年基準の家計調査における収支項目分類に基づく。詳細は、総務省統計局ホームページ「家計調査 収支項目分類」(<http://www.stat.go.jp/data/kakei/9.htm>)を参照。

(3)按分による推計手法について解説する。

2.3.1. 按分対象のレベル

按分処理を導入した目的は、購入頻度が低く、使用する年数を増やすことによる効果がそれ程期待出来ない品目について、購入頻度の不足を補い、誤差の縮小効果を得るためである。そのため、按分対象とする上位類は、十分な購入頻度が確保出来ている必要がある。類のレベルを上げれば購入頻度は高くなることから、この目的からは「可能な限り上位レベルのものを選定」することが望ましい。一方で、類のレベルを上げすぎると、当該類は地域性が大きく異なる品目を含む可能性が高くなり、当該品目との地域性が整合的ではなくなることから、推計結果と真の値との不整合を拡大させてしまう恐れがある。従って「必要以上に類のレベルを上げすぎない」ことも重要と考えられる。

このような観点から、按分対象とする上位類のレベルを、中間的な位置にある「中分類」と定めることが妥当であると考えられる。ただし、10大費目のうち「光熱・水道」については、構成品目が7品目と少ないため、費目全体を按分処理の選定対象としている。

また、支出項目のうち次の項目については、一般に市場が存在しないこと、支払う費用と受ける対価の関係が明確ではないこと、あるいは他の世帯などへの所得移転であることなどの理由から、消費支出から除外する⁸。

信仰・祭祀費、寄付金、贈与金、他の負担費、仕送り金

2.3.2. 按分対象の選定基準

按分対象とする中分類は、購入頻度が低く、使用する年数を増やすことによる標本誤差の縮小効果が、それ程期待出来ない品目を相当程度含むものを選定する。ここで、相当程度含むとしているのは、按分処理のやみくもな使用を避ける目的と、地域差指数用のウェイトを作成するという目的から考えて、消費支出に占めるウェイトが高い中分類を優先的に選定するためである。

ところで、購入頻度が低い品目は、標本誤差が相対的に大きいいため、単年集計結果の市町村間のばらつきが大きくなる傾向にある。また、極端に高額な商品の影響を受けやすいため、場合によっては単年集計結果に外れ値を含むことがある。外れ値は推計結果の精度に影響を与えるため、適切に抑える必要があるが、購入頻度が低い品目は複数年分のデータを使用することによっても外れ値を抑える効果が望めない。そのため、選定する中分類は、このような品目を多く含むものを按分対象として優先的に選定する必要がある。このようなことから、当初、按分対象の選定には購入頻度や標本誤差を選定指標とすることを検討していたが、外れ値を見逃す恐れがあるため、按分対象の選定には、集計結果の市町村間のばらつきの大きさを指標とする基準を設定した。ただし、ばらつきの大きさのみを指標として品目を選定すると、本来であれば地域性と見るべき外れ値を除外してしまう恐れがあるため⁹、ばらつきが大きく、外れ値とみなされる市町村を含む場合であっても、それが定常的なものであれば、選定対象から除外すべきである。そのため、地域性が安定しているかを評価するための指標も設定した。選定指標の作成方法及び選定基準の詳細は、以下のとおりである。

(1) 選定指標

各指標は、支出金額の推計に使用する直近5年分の市町村別支出金額から作成した支出構成比（消費支出に対する万分比）を使用し、中分類または品目ごとに作成する。ここで、指標の作成に「支出金額」ではなく「支出構成比」を使用している理由は、経年変化による支

⁸ 消費者物価指数におけるウェイト作成においても、同理由により消費支出から除外をしている。

⁹ 北海道・東北地方の「灯油」、高知における「かつお」、四国における「うどん」など。

出金額の変動の影響を除くためである。

①指標 1 (市町村別支出構成比の全国平均)

地域差指数に対する影響が大きいものを選定するために設定した指標。単年の市町村別支出構成比を5年平均した値を、家計調査における最新年の市町村別調整集計世帯数で加重平均して算出する。

②指標 2 (市町村別支出構成比の変動係数の対数化)

支出構成比の市町村間のばらつきの大きさを測るために設定した指標。変動係数ではなくその対数化を指標とするのは、変動係数の品目分布が対数正規分布に近く、当指標に対する閾値の設定に標準正規分布を使用するためである。(厳密には正規分布ではない。) 変動係数は品目ごとに、下記により算出する。

$$\text{変動係数} = \frac{\text{市町村別支出構成比の標準偏差}}{\text{市町村別支出構成比の単純平均}}$$

③指標 3 (単年分の市町村別支出構成比と市町村別支出構成比の5年平均との相関係数の5年分平均)

支出構成比の地域性の安定度を測るために設定した指標。

(2) 選定基準

基本的な考え方は、市町村間のばらつきが大きく、かつそのばらつきが安定的でない品目を多く含む中分類で、中分類そのものの消費支出に占めるウェイトが高いものを選定対象とするというものである。具体的には、以下の基準にすべて合致する中分類を按分対象に選定する。なお、各指標における閾値は、経験則により定めた。

①支出全体に占める割合が高い中分類

指標 1 の累積上位 80%の中分類を選定

②支出構成比の市町村間変動が大きい品目の、当該類における累積支出構成比が高い中分類

指標 2 が閾値以上である品目について、指標 1 の当該類内の累積が 10%以上となる中分類を選定。指標 2 に対する閾値は、以下により算出する。

$$\text{閾値} = \text{変動係数の対数値の全品目平均} + z(\alpha) \times \text{変動係数の対数値の標準偏差}$$

ただし、 $z(\alpha)$ は標準正規分布における上側 $\alpha\%$ 点を表す。

③市町村別支出構成比の地域性が安定していない品目の、当該類における累積支出構成比が高い中分類

指標 3 が 0.7 未満である品目について、指標 1 の当該類内の累積が 10%以上の中分類を選定。

2.3.3. 按分比の作成

按分対象となった類Iについて、直近5年分の品目別・市町村別支出金額 e_{iky} 及び、家計調査における最新年の市町村別調整集計世帯数 α_k (k は調査市町村)を使用し、類内の各品目*i*に対し、下記算式により、地方区分C別按分比を算出する。

$$R_{ic} = \frac{1}{5} \sum_{y \in Y} \frac{C \text{地域における} y \text{年の品目} i \text{に対する平均支出構成比}}{C \text{地域における} y \text{年の類} I \text{に対する平均支出構成比}}$$

$$= \frac{1}{5} \sum_{y \in Y} \left\{ \frac{\sum_{k \in C} \alpha_k \frac{e_{iky}}{e_{ky}}}{\sum_{k \in C} \alpha_k} \bigg/ \frac{\sum_{k \in C} \alpha_k \frac{e_{Iky}}{e_{ky}}}{\sum_{k \in C} \alpha_k} \right\} = \frac{1}{5} \sum_{y \in Y} \left\{ \frac{\sum_{k \in C} \alpha_k \frac{e_{iky}}{e_{ky}}}{\sum_{k \in C} \alpha_k \frac{e_{Iky}}{e_{ky}}} \right\}$$

R : 按分比

i : 類内の各品目

I : 費目・類

C : 地方区分

k : 調査市町村

e_{iky} : 調査市町村 k における y 年の品目 i に対する支出金額

e_{ky} : 調査市町村 k における y 年の消費支出金額

e_{Iky} : 調査市町村 k における y 年の費目・類 I に対する支出金額

α_k : 家計調査における最新年の市町村別調整集計世帯数

2.3.4. 按分による品目別支出金額の推計

按分対象となった中分類の支出金額を、算出した按分比を使用して下位の内訳品目に按分し、市町村別支出金額を推計する。ただし、県庁所在市及び政令指定都市は、家計調査において公表を前提とした標本設計となっており、按分のような特殊処理を行わなくても十分に安定化出来ているものと考えられることから、按分による推計値ではなく、単年集計結果の5年分の平均値を使用する。

平成25、26年分の結果作成において、第2.3.2項の基準を適用し、按分対象として選定した中分類は、「家賃地代」、「設備修繕・維持」、「交通」、「自動車等関係費」、「授業料等」、「教養娯楽サービス」及び「他の諸雑費」であり、品目数は全83品目である。

なお、中分類と品目の間に小分類のあるものがあるが、小分類を介して複数段階により按分した結果「中分類→小分類→品目」と、中分類を直接品目に按分した結果「中分類→品目」は一致する。(付録1.(2)を参照。)

3. 推計手法の評価

本検証の目的は、構造編創設に伴い導入した品目別・市町村別支出金額の推計手法を、定量的に評価することである。検証は、家計調査個票データを母集団とみなし、そこからのリサンプリングシミュレーションにより推計結果の特性値（真の値からの平均二乗誤差、標準誤差）を推計し評価する。なお、検証方法については、坂田(2011)を参考にした。

3.1. 検証方法

3.1.1. 使用するデータ

家計調査結果「2010年1月分～2014年12月分」を使用する。また、評価対象とする品目は、平成22年基準の家計調査における収支項目分類のうち、「信仰・祭祀費」、「寄付金」、「贈与金」、「他の負担費」及び「仕送り金」を除いた491品目である。

3.1.2. 比較対象及び評価対象

(1) 比較対象

① 単年分の家計調査結果の単純集計

既存の集計用乗率を単純に集計する。以下、当手法を「単年集計」という。

② 全国物価統計調査で採用していた手法

①により得られた市町村別支出金額を、家計調査における当該年の市町村別調整集計世帯数により、家計調査における地方・都市階級区分ごとに加重平均を算出し、県庁所在市及び政令指定都市を除く市町村に、対応する地方・都市階級区分の平均値を代入する。県庁所在市及び政令指定都市は、当該市における単年集計の結果をそのまま使用する。以下、当手法を「地方・都市階級集計」という。

(2) 評価対象

評価対象は、構造編創設時に導入した推計手法であるが、按分処理の効果についても検証を行うため、下記2手法を評価対象とする。

①実際に導入した推計手法（単年集計の結果5年分の単純平均と按分処理）

導入した推計手法の評価を行う。以下、当手法を「5年平均・按分」という。

②単年集計の結果5年分の単純平均のみ

5年分の家計調査結果を使用することの効果の評価する。また、按分処理の効果についても確認する。以下、当手法を「5年平均のみ」という。

3.1.3. 評価方法

評価指標には、推計値の標準誤差（以下「SE」という。）及び、真の値からの平均二乗誤差（以下「MSE」という。）を使用する。評価手順は以下のとおり。

推計手法ごとに作成した品目別・市町村別 MSE 及び SE を基に、「平均 MSE（市町村別 MSE の市町村平均）」及び「平均 SE（市町村別 SE の市町村平均）」を品目ごとに算出。以下に基づき評価を行う。

①単年集計との比較

単年集計の平均 MSE 及び平均 SE に対する、推計結果の平均 MSE 比及び平均 SE 比が共に 1 未満であるとき、当該品目の支出金額の推計値は、単年集計よりも改善したと評価。この評価を、単年集計5年分に対して行う。

②地方・都市階級集計との比較

地方・都市階級集計の平均 MSE 及び平均 SE に対する、推計結果の平均 MSE 比及び平均 SE 比が共に 1 未満であるとき、当該品目の支出金額の推計値は、地方・都市階級集計よりも改善したと評価。この評価を、地方・都市階級集計5年分に対して行う。

3.1.4. SE 及びMSE の推計

リサンプリングごとの推計結果を \hat{e}_b 、全リサンプリング推計結果の平均を \hat{e}^* （推計結果の期待値）、真の値を e_0 としたとき、品目別・市町村別の MSE 及び SE を、次式により推計する。ただし、 B はリサンプリング回数を表す。また、真の値 e_0 は全個票データの単年集計結果であり、リサンプリングの平均値 \hat{e}^* は、リサンプリングごとの推計結果 \hat{e}_b の単純平均により算出する。本検証におけるリサンプリング回数は $B = 1000$ とし、リサンプリングは、抽出率を 1/4 とする非復元抽出により実施した¹⁰。なお、 \hat{e}_b 算出の際の乗率再作成は行わない。

¹⁰ 標本理論によると、実際の調査では、抽出率が高いほど精度のよい推計が可能である。しかし、個票データを母集団と見立て、そこからの抽出によるシミュレーションを行う場合、真の値を個票データの単年集計結果としていことから、抽出率が高いほど、平均二乗誤差が縮小するという意味で、単年集計が最もよいという結果が得られることになる。実際の調査における抽出率は 1 よりもかなり低い値であるため、シミュレーションにおいても、現実に近い状況で抽出を行うことが望ましい。そのため、抽出率は大きすぎない値を設定する必要がある。ただし、個票データ数の下限を割るような値を設定することは適切ではないため、個票データ数に合わせて抽出率を設定した。

$$MSE = \frac{1}{B} \sum_{b=1}^B (\hat{e}_b - e_0)^2$$

$$SE = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (\hat{e}_b - \hat{e}^*)^2}$$

3.2. 検証結果

3.2.1. 単年集計との比較

表3-1に、単年集計と、5年平均のみ及び5年平均・按分の比較結果を掲載する。5年平均のみ、5年平均・按分共に、ほぼ全ての品目において平均SEが縮小しており、全体の約8割の品目について、平均MSE及び平均SEが縮小していることが確認できる。

表3-1 平均MSEと平均SE比が1未満の品目数
(対単年集計)

	5年平均のみ					5年平均・按分				
	2010	2011	2012	2013	2014	2010	2011	2012	2013	2014
平均MSE比が1未満	386	389	417	410	381	384	391	411	406	380
全体に対する割合 (%)	78.6	79.2	84.9	83.5	77.6	78.2	79.6	83.7	82.7	77.4
うち平均SE比が1未満	386	389	417	410	381	384	391	411	406	380
全体に対する割合 (%)	78.6	79.2	84.9	83.5	77.6	78.2	79.6	83.7	82.7	77.4
(参考) 平均SE比が1未満	487	488	486	482	490	489	489	487	483	488
全体に対する割合 (%)	99.2	99.4	99.0	98.2	99.8	99.6	99.6	99.2	98.4	99.4

※割合は、全491品目に対する値を算出

次に、5年分の家計調査結果を使用する効果について確認する。付録2に、5年平均のみについての、購入頻度¹¹と平均SE比の関係及び購入頻度と平均MSE比の関係を掲載する。付録2(左)より、購入頻度が低い程、平均SE比が大きくなる傾向にあることが確認できる。これは、購入頻度が低いと、使用する年数を増やすことによる標本誤差の縮小効果が全体的に低下するというを表しているといえる。一方、付録2(右)より、平均MSEは購入頻度によらず、拡大している品目が一様に分布していることが確認できる。また、平均MSEの拡大は、主に非耐久財及びサービス品において発生していることも見て取れる。(表3-2参照。)平均二乗誤差は標準誤差の二乗(標本分散)及び真の値からのバイアスの二乗の和に分解できるため、ほぼ全ての品目において平均SEが縮小していることから、平均MSEの拡大は、真の値からのバイアスの拡大によるものであることがわかる。これは、5年分のデータを使用していることから、消費支出の経年変化の影響を受けたものと考えられる。

続いて、按分処理の効果について確認する。付録3に、5年平均・按分についての、購入頻度と平均SE比の関係及び購入頻度と平均MSE比の関係を掲載する。按分による推計を行った品目は主に、購入頻度が比較的低いサービス品からなるが、付録3(左)をみると、サービス品の平均SEについて、5年平均のみよりも縮小していることが確認できる。一方で、按分処理により平均SEが拡大している品目もあり、必ずしも按分処理により誤差が縮小するわけではないことも確認できる。付録3(右)の平均MSEについてみると、5年平均のみと比較して、按分処理により一部の品目において拡大していることが確認できる。こちらは、

¹¹ (引用) 総務省統計局「家計調査結果(家計収支編) 詳細結果表(二人以上世帯、全国、年平均)」の2010年~2014年結果

5年平均のみの場合と異なり、経年変化によるバイアスの拡大ではなく、按分処理により地域性が崩れたことによるバイアスが発生していることが考えられる。

表3-2 平均MSEが拡大した品目数（財・サービス区分別）

	5年平均のみ					5年平均・按分				
	2010	2011	2012	2013	2014	2010	2011	2012	2013	2014
非耐久財	47	43	17	22	42	47	43	17	22	42
割合 (%)	19.0	17.4	6.9	8.9	17.0	19.0	17.4	6.9	8.9	17.0
半耐久財	5	8	8	7	9	4	7	8	7	9
割合 (%)	5.7	9.1	9.1	8.0	10.2	4.5	8.0	9.1	8.0	10.2
耐久財	10	10	14	15	14	9	8	13	15	14
割合 (%)	25.0	25.0	35.0	37.5	35.0	22.5	20.0	32.5	37.5	35.0
サービス品	43	41	35	37	45	47	42	42	41	46
割合 (%)	37.1	35.3	30.2	31.9	38.8	40.5	36.2	36.2	35.3	39.7

※割合は、財・サービス区分ごとの全品目数に対する値を算出

表3-3に、地方・都市階級集計について、同様の分析を行った結果を掲載する。同表より、ほぼ全ての品目の平均SEが縮小している一方で、全品目の約4割程度で平均MSEが拡大していることが確認できる。これは、市町村別支出金額を地方・都市階級にまとめ上げたことで、地域性が崩れたことにより、真の値とのバイアスが拡大したことが要因と考えられる。付録4に、地方・都市階級集計についての、購入頻度と平均SE比の関係及び購入頻度と平均MSE比の関係を掲載する。付録4（左）より、5年平均のみ及び5年平均・按分と同様に、地方・都市階級集計についても、購入頻度が低い程、平均SE比が大きくなる傾向にあることが確認できる。一方で、付録4（右）より、平均MSEについては、主に購入頻度が高い非耐久財において拡大していることが確認できる。

表3-3 平均MSEと平均SE比が1未満の品目数
(対単年集計)

	地方・都市階級				
	2010	2011	2012	2013	2014
平均MSE比が1未満	316	288	309	319	308
割合 (%)	64.4	58.7	62.9	65.0	62.7
うち平均SE比が1未満	316	288	309	319	308
割合 (%)	64.4	58.7	62.9	65.0	62.7
(参考) 平均SE比が1未満	491	491	491	489	489
割合 (%)	100.0	100.0	100.0	99.6	99.6

※割合は、全491品目に対する値を算出

3.2.2. 地方・都市階級との比較

表3-4に、地方・都市階級集計と、5年平均のみ及び5年平均・按分の比較結果を掲載する。5年平均のみ、5年平均・按分共に、9割以上の品目において平均SEが縮小しており、全体の約8割の品目について、平均SE及び平均MSEが縮小していることが確認できる。付録5及び付録6それぞれに、5年平均のみ及び5年平均・按分についての、購入頻度と平均SE比の関係及び購入頻度と平均MSE比の関係を掲載する。同図より、5年平均のみ及び

5年平均・按分共に、単年集計結果との比較結果と同様の傾向であることが確認できる。

表3-4 平均MSEと平均SE比が1未満の品目数
(対地方・都市階級集計)

	5年平均のみ					5年平均・按分				
	2010	2011	2012	2013	2014	2010	2011	2012	2013	2014
平均MSE比が1未満	385	416	433	426	412	377	413	417	417	402
割合 (%)	78.4	84.7	88.2	86.8	83.9	76.8	84.1	84.9	84.9	81.9
うち平均SE比が1未満	383	415	433	425	412	375	412	417	416	402
割合 (%)	78.0	84.5	88.2	86.6	83.9	76.4	83.9	84.9	84.7	81.9
(参考) 平均SE比が1未満	463	466	468	463	470	466	467	467	462	468
割合 (%)	94.3	94.9	95.3	94.3	95.7	94.9	95.1	95.1	94.1	95.3

※割合は、全491品目に対する値を算出

3.3. まとめ

本節のシミュレーションの結果、構造編創設時に導入した推計手法は、単年分の集計用乗率の単純集計及び全国物価統計調査で採用していた手法と比較し、全491品目の約8割の品目について、標準誤差及び平均二乗誤差の市町村平均が縮小するという意味において、改善していることが確認できた。しかし一方で、平均二乗誤差の市町村平均が拡大している品目が少なからずあることも確認した。また、平均二乗誤差の拡大は、誤差の拡大ではなくバイアスの拡大によるものであることもわかった。

家計調査における各層の抽出率は168/83326以下であり、今回のシミュレーションにおける抽出率(1/4)よりも極めて小さい値であることから、実際には全体の8割よりも高い割合で改善していることが期待される。しかし、正確なことはわからないため、導入した推計手法によっても、真の値とのバイアスが拡大するという意味で改善しない品目があり得ることを考慮しなければならない。そのためにも、実際の推計結果について、誤差の評価に加えて、真の値とのバイアスを適切に評価し、品質の管理に努めることが必要である。第4節では、本研究で提案した、推計結果を評価するための手法について述べる。

4. 推計結果の評価

4.1. 基本的な考え方

推計結果に最低限期待することは、少なくとも単年集計結果よりも誤差が拡大しないことである。加えて、公表値である単年集計結果と極端に矛盾しない(不整合とならない)ことも必要である。県庁所在市及び政令指定都市以外は公表を前提とした標本設計となっていないため、安定的であるとは言えないが、単年集計結果は真の値の不偏推定量であることから、理論的にもこれをもある程度の目安として推計結果の評価を行うことは、合理的であると考えられる。このようなことから、我々は、単年集計結果を一つの目安と考え、推計結果を評価する手法を提案した。

さて、誤差を拡大させないということの意味は明らかであるが、単年集計結果と不整合とならないということの意味が曖昧である。そこで、「単年集計結果と不整合とならない」ことについて考察する。明らかに、推計結果と単年集計結果の間に次のいずれかのパターンが見られる場合、両者は整合的ではないと言って差し支えないと考えられる。

(a) 多くの市町村において、推計結果が単年集計結果から見て一定方向(上方または下方)にシフトしている。

(b) 地域のあるブロック単位で、推計値の高低関係が、単年集計結果のそれと大きく入れ替

わっている。

1つ目は「水準」が維持されていない場合であり、2つ目は「地域性」が整合的でない（維持されていない）場合である。逆に、推計結果が上記いずれにも当てはまらない場合、推計結果は単年集計結果と整合的であると言って差し支えないと考えられる。本研究では、単年集計結果と推計結果が整合的でないということを、上記いずれか（両方含む）のパターンが表れる場合と定義し、これらのパターンを識別するための指標（評価指標）を導入した。

1つ目は、市町村ごとに算出した単年集計結果と推計結果の差（バイアス）を1つの分布とみなし、当分布の代表値を品目ごとの単年集計結果からの乖離の大きさを測る指標と定め、この値が0または大きくないとみなせる場合に水準が維持されていると評価する。

2つ目は、単年集計結果と推計結果の地域性が整合的か評価するもので、本研究では、単年集計結果と推計結果の間の相関係数により評価を行うこととしている。（付録1．(1)参照。）続く第4.2節では、評価指標の算出方法と評価基準について解説する。

4.2. 評価指標と評価基準

ここでは、推計結果の評価に使用する評価指標と評価基準について解説する。なお、評価指標の作成には、品目別・市町村別に作成した標準誤差及び推計結果の期待値を使用するが、評価指標は、これらの値を品目ごとに集約したものである。評価指標を品目ごとに作成する理由は、本研究で扱う評価対象は非常に多いため（491品目×168市町村）、個別に推計結果の評価を行うことは困難だからである。本研究では、実用的な評価手法の確立を目的としており、品目ごとに評価を行えることが望ましいと考えたため、品目ごとに市町村の情報を集約して評価指標を作成した。導入した指標は、「誤差の指標」、「水準の指標」及び「地域性の指標」に加え、補助的に使用するための「補助指標」の4種である。なお、各指標に対する基準（閾値）は経験的に定めた。

4.2.1. 標準誤差及び期待値の算出

ここでは、評価指標の算出に使用する、単年集計結果、推計結果の品目別・市町村別の「標準誤差」及び「期待値」の推計方法について解説する。本研究では、これらの値をノンパラメトリックブートストラップ法¹²により推計した。（ブートストラップ法については、小西・越智・大森(2008)、Efron and Tibishirani(1993)を参照。）リサンプリング回数は $B = 1000$ である。詳細は以下のとおり。

(1) 期待値及び標準誤差の推計式

ノンパラメトリックブートストラップ法による期待値及び標準誤差の算出方法に基づき、以下により推計する。リサンプリングごとの単年集計結果または推計結果を \hat{e}_b としたとき、期待値及び標準誤差は下記算式により算出する。

①期待値

$$\hat{e}^* = \frac{1}{B} \sum_{b=1}^B \hat{e}_b$$

②標準誤差

¹² ブートストラップ法は、データ解析の確からしさを評価するための統計手法の一つであり、従来の複雑な数式に基づく理論を莫大な数値計算による単純なシミュレーションで置き換える。標準誤差や信頼区間の推計など、通常は理論式により推計する統計量を、反復計算により近似的に推計することが可能。ブートストラップ法は、ノンパラメトリック法とパラメトリック法に分類される。

$$SE = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (\hat{e}_b - \hat{e}^*)^2}$$

(2) リサンプリングごとの \hat{e}_b の算出

家計調査の本集計結果作成は、地方10区分別・世帯人員階級別の世帯数（労働力調査結果より作成）をベンチマークとした比推定により集計用乗率を作成し、集計している（総務省統計局(1)）。本研究におけるブートストラップ集計においても、リサンプリングごとに同様の方法により集計用乗率を作成し、 \hat{e}_b を算出する。なお、本研究で使用したベンチマークは、公表結果の作成時とは異なり、平成22年国勢調査結果から作成した地方10区分別・世帯人員階級別の一般世帯数（2人以上）をそのまま使用している¹³。

4.2.2. 誤差の評価

評価指標：平均誤差の比

単年集計結果からの誤差の縮小度合いを評価する指標であり、「誤差の指標」と呼ぶ。品目ごとに、推計結果の標準誤差の市町村平均（以下「平均誤差」と呼ぶ。）の比（単年集計結果の平均誤差に対する比）を算出する。

評価基準：平均誤差の比が1未満のとき、誤差は縮小したと評価する。

4.2.3. 水準の維持の評価

評価指標：市町村別標準化バイアスの中央値

単年集計結果からの乖離の大きさを測定するための指標であり、「水準の指標」と呼ぶ。標準化バイアスは、下記算式により算出する¹⁴。（便宜上100倍した値を使用する。）

$$\text{標準化バイアス} = \frac{\text{推計結果の期待値} - \text{単年集計結果の期待値}}{\text{推計結果の期待値}} \times 100$$

なお、指標の作成にバイアスではなく、標準化バイアスを使用した理由は、品目間で乖離の大きさを比較するためである。また、指標に中央値を採用した理由は、市町村によっては極端に大きなバイアスがあり、そのような外れ値の影響を受けない（ロバストな）指標を作成するためである。代表値として最頻値（最頻階級の中央値）を採用することも考えられるが、乖離の大きさが階級区分の取り方に依存し、恣意的な評価につながる恐れがあること、品目ごとに階級区分を決めることは現実的に非常に困難であることから、採用しなかった。

評価基準：絶対値が20未満のとき、乖離はないもしくは小さいと評価する。

4.2.4. 地域性の維持の評価

評価指標：単年集計結果と推計結果の相関係数

単年集計結果との地域性の整合性を評価するための指標であり、「地域性の指標」と呼ぶ。

評価基準：相関係数が0.4以上のとき、整合的であると評価する¹⁵。

¹³ 参考のために、集計用乗率の再作成を行わずにブートストラップ集計を行い、期待値及び標準誤差の算出を行ったが、どちらを用いても同様の結果が得られた。

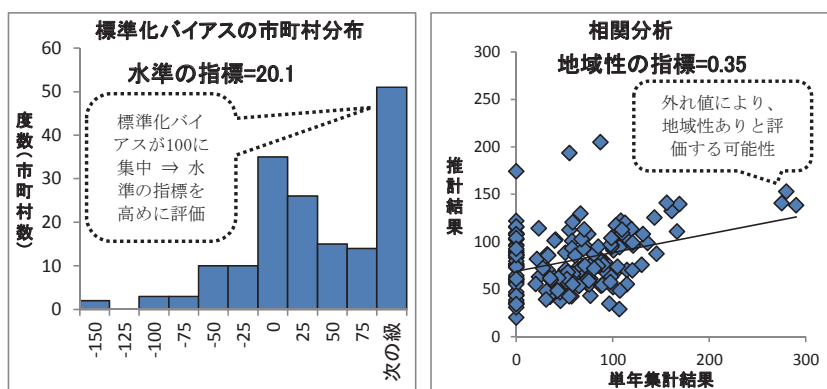
¹⁴ ブートストラップ法では、推計結果の真の値とのバイアスは、「推計結果の期待値－真の値の不偏推定量」により推計する（Efron and Tibishirani (1993)の第10章を参照。）。単年集計結果は真の値の不偏推定量であるため、ここでのバイアスは、推計結果の真の値とのバイアスの推計値である。

¹⁵ 相関の強さについては、一般的に「ギルフォードの規則」と呼ばれる以下の基準があるが、実際には相関の有無を検定するための検定統計量を算出し、同検定に基づき、相関の有無を判定する必要がある。なお、相関の有無を検定するための検定統計量は、データ数にのみ依存し、本研究で取り扱うデータについては、家計調査の調査市町村である168がそのままデータ数となる。このとき、有意水準5%で考えた場合、帰無仮説「相関はない」が棄却

4.2.5. 補助指標

本研究では、単年集計結果を目安としているが、単年集計結果が0の市町村数が多い品目については注意が必要である。該当品目は購入頻度が低く、元々誤差が大きいことに加え、値が0となる市町村数が多いことから、単年集計結果が目安としての機能を十分に果たせず、水準及び地域性について、誤った評価をしてしまう可能性がある。(図4-1参照。)このことから、補助指標として、「単年集計結果が0の市町村数」を設定し、補助指標が一定値以上の品目については、誤差を縮小させるという観点からの評価のみを行う¹⁶。補助指標に対する基準は、経験的に20(全168市町村の12%未満)と定めている。

図4-1 補助指標が大きいことによる水準及び地域性の指標への影響



4.3. 評価結果

表4-1に、推計結果の評価結果を掲載する。同表より、ほぼ全ての品目の平均誤差が縮小していることが確認できる。

表4-1 推計結果の評価結果

	5年平均					5年平均・按分				
	2010	2011	2012	2013	2014	2010	2011	2012	2013	2014
誤差の指標が1未満(①)	488	488	487	482	490	489	490	488	484	488
割合(%) (※1)	99.4	99.4	99.2	98.2	99.8	99.6	99.8	99.4	98.6	99.4
水準の指標の絶対値が20未満	441	432	442	441	437	417	413	423	425	418
うち補助指標が20未満(②)	403	402	409	412	404	395	397	404	405	401
割合(%) (※2)	96.4	96.4	97.8	98.1	97.3	94.5	95.2	96.7	96.4	96.6
地域性の指標が0.4以上	464	467	464	460	462	437	440	438	437	437
うち補助指標が20未満(③)	409	410	409	405	401	394	394	396	391	392
割合(%) (※2)	97.8	98.3	97.8	96.4	96.6	94.3	94.5	94.7	93.1	94.5
①かつ②かつ③	396	395	402	396	395	380	379	387	380	381
割合(%) (※2)	94.7	94.7	96.2	94.3	95.2	90.9	90.9	92.6	90.5	91.8

される(すなわち、相関があると判定される)ときの相関係数は、約0.15である。従って、本研究においては、相関係数が0.15以上であれば相関があると判断してよいことになる。しかし、一般的に相関係数が0.15というのはかなり低い値であるため、ギルフォードの規則における中程度以上の相関を、総合的であると評価することとした。

【ギルフォードの規則】

0.2未満 ⇒ 相関なし、0.2~0.4 ⇒ 弱い相関、0.4~0.7 ⇒ 中程度の相関、0.7以上 ⇒ 強い相関

¹⁶ ただしこの場合でも、推計結果と単年集計結果を視覚的に比較し、極端な不整合が生じていないことを品目ごとに確認を行う。

	5年平均					5年平均・按分				
	2010	2011	2012	2013	2014	2010	2011	2012	2013	2014
(参考) 補助指標が20未満	418	417	418	420	415	418	417	418	420	415

※1) 全491品目に対する割合を算出

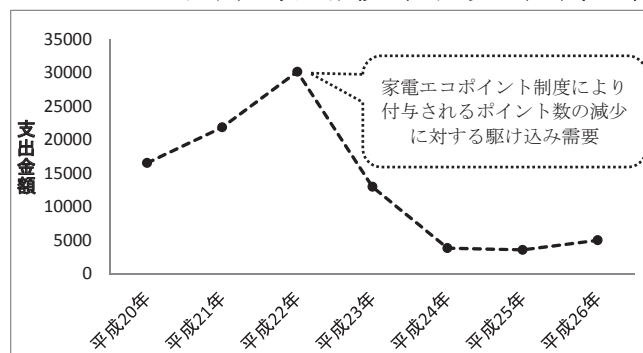
※2) 各年次における補助指標が20未満の品目数に対する割合を算出

付録7に、購入頻度と誤差の指標の関係を掲載する。同図左より、5年平均結果では、購入頻度が低くなると、誤差の指標が大きくなる傾向にあり、購入頻度が低い品目の一部で、誤差の指標が1を超えていることが確認できる。一方、同図右より、按分の効果をみると、購入頻度が高い品目の一部で、誤差の指標が1を超えていることが確認できる。2014年分の単年集計結果との比較では、「ガソリン」及び「年極・月極有料駐車場」が該当する。これらは購入頻度が高く、元々誤差が大きい品目である。このようなことが起こったのは、按分処理により、中分類を介して他の品目の誤差が伝播したためと考えられる。(上記2品目について、5年平均結果では誤差の指標が1未満である。)

付録8に、水準及び地域性の指標の関係を掲載する。同図より、5年平均、5年平均・按分共に、一部の品目において、水準の指標もしくは地域性の指標が基準外となっていることが確認でき、特に耐久財や半耐久財が基準外となる傾向が強いことがみてとれる。また、同図右より、按分処理により、一部の品目において、水準または地域性の不整合が拡大している様子が見える。一方、按分により推計を行う品目ではないが、2010年分の単年集計結果との比較において、テレビの水準が極端に外れていることが確認できる。これは、2010年12月に実施された、家電エコポイント制度により付与されるポイント数の減少に対する駆け込み需要により、2010年のテレビに対する支出が他の年次よりも極端に高くなったことが要因である。(図4-2参照。)

導入した推計手法では、5年分のデータを使用しているため、経年変化の影響を無視することは出来ない。そのため、テレビの例のように、社会の制度変更など、さまざまな要因により消費構造が大きく変化する場合、それらは適切に検出されるべきであるが、この例では、それが適切に検出されたといえる。

図4-2 テレビへの支出金額の推移(二人以上世帯、全国平均)¹⁷



5. おわりに

本研究では、構造編創設時に導入した品目別・市町村別支出金額の推計手法について、単年分の集計用乗率の単純集計及び全国物価統計調査で採用していた手法と比較し、全491品

¹⁷ (引用) 総務省統計局「家計調査結果(家計収支編) 詳細結果表(二人以上世帯、全国、年平均)」の2008年～2014年結果

目の約8割の品目について、標準誤差及び平均二乗誤差の市町村平均が縮小していることを確認した。一方で、平均二乗誤差の市町村平均が拡大している品目が少なからずあることも確認した。また、平均二乗誤差の拡大は、誤差の拡大ではなく、経年変化や按分処理の影響を受けたバイアスの拡大によるものであることもわかった。

また、推計結果の品質を保つための定量的な評価手法を提案し、推計結果について評価を行った。その結果、按分処理による誤差の拡大や、経年変化及び按分処理の影響を受けた水準または地域性の不整合が起こることを確認した。一方で、提案した評価手法は、制度変更などの影響を受けた単年集計結果との極端な不整合を、適切に検出できることも確認できた。このことから、提案した評価手法は、購入頻度が低いため、単年集計結果が0となる市町村数が多くなる品目についての評価が適切に行えていないという課題はあるものの、簡易的な手法としては実用に耐えうるものであると考えられる。

今後の展望として、提案した推計結果の評価手法の、実務への組み込みが挙げられる。しかし、本研究では、標準誤差及び期待値の推計にブートストラップ法を利用したが、これによると、標準誤差及び期待値の推計に莫大な時間を要するという難点がある。実用化に向けて、推計結果に対する標準誤差及び期待値を解析的に得られることが望ましい。また、評価結果に基づく推計手法の改善や、モデル推計手法導入時の評価などへの適用も考えられる。

付録1

(1) 本研究における地域性の基本的な考え方

本研究では、地域性を「市町村ごとの値の高低関係についてのパターン」と定義する。ここでいうパターンとは、市町村間の順序関係のみではなく、市町村ごとの値の相対的な位置関係を含めたものをいう¹⁸。またこの定義から、異なる2つの品目間で、地域性が整合的かどうかということ、両者間の正の線形関係の有無によりみることができる。線形関係の有無は相関係数によりみることができることから、地域性の整合度は、相関係数により把握される。また、地域性の整合性は、異なる2つの品目間だけでなく、異なる推計手法間や、2時点間でも考えることが可能である。

ただし、注意として、相関係数は外れ値の影響を受けやすいため、地域性が整合的ではないにもかかわらず、地域性が整合的であると評価してしまう可能性があるため、実際の分布をみて判断する必要がある。

(2) 按分のステップに関する不変性について

品目*i*の調査市町村*k*における支出金額を e_{ik} と表す。費目・類を*I*と表すとき、調査市町村*k*における費目・類*I*の支出金額は以下により算出される。

$$e_{Ik} = \sum_{i \in I} e_{ik}$$

また、上位地域を*K*で表し、市町村別支出金額を上位地域へ畳み上げる際の、調査市町村*k*における調整集計世帯数を α_k と表すと、地域*K*における品目*i*の支出金額は以下により算出される。

$$e_{iK} = \sum_{k \in K} \alpha_k e_{ik}$$

次に、中分類を*H*、小分類*I*をと表したとき、*H*を*I*に按分する際の、地域*K*における按分比 r_{IK} 及び、小分類*I*を品目*i*に按分する際の、地域*K*における按分比 r_{iK} は、定義より以下で与えられる。

$$r_{IK} = \frac{e_{iK}}{\sum_{I \in H} e_{iK}}$$

$$r_{iK} = \frac{e_{iK}}{\sum_{i \in I} e_{iK}}$$

このとき、品目*i*の調査市町村*k*における支出金額を、二段階按分「中分類→小分類→基本品目」により推計した値と、直接按分「中分類→基本品目」により推計した値は、互いに一致することが示される。

$$\hat{e}_{ik} = r_{iK} r_{IK} e_{HK} = \frac{e_{iK}}{\sum_{i \in I} e_{iK}} \times \frac{e_{iK}}{\sum_{I \in H} e_{iK}} \times e_{HK} = \frac{e_{iK}}{\sum_{I \in H} e_{iK}} \times e_{HK} = \tilde{r}_{iK} e_{HK}$$

ここで、右辺の係数 \tilde{r}_{iK} は、中分類*H*を品目*i*に直接按分する際の按分比である（左辺＝二段階按分による推計値、右辺＝直接按分による推計値）。また、3番目の等式は、以下が成り立つことによる。

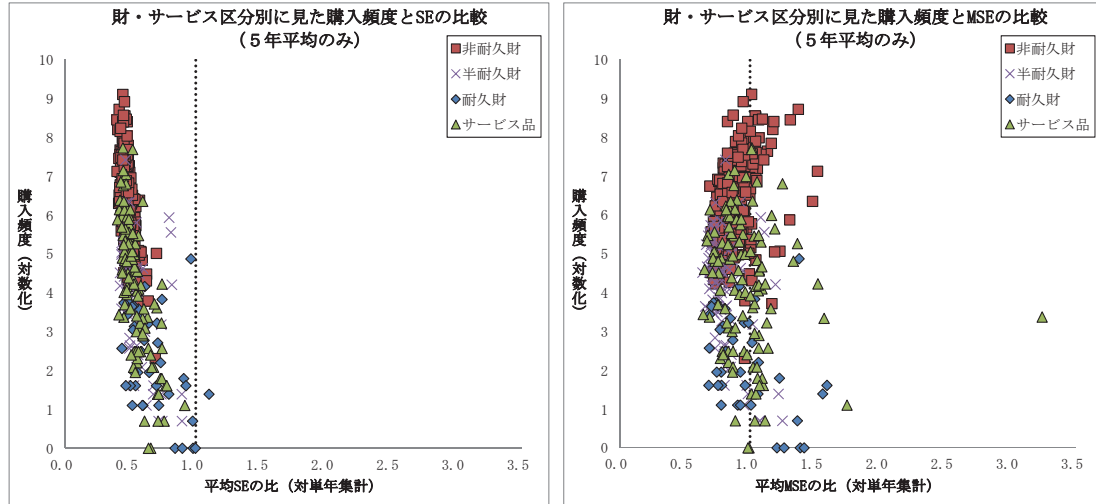
$$e_{iK} = \sum_{k \in K} \alpha_k e_{ik} = \sum_{k \in K} \alpha_k \sum_{i \in I} e_{ik} = \sum_{i \in I} \sum_{k \in K} \alpha_k e_{ik} = \sum_{i \in I} e_{iK}$$

なお、三段階以上の複数段階の按分に対しても、同様のことが示せる。

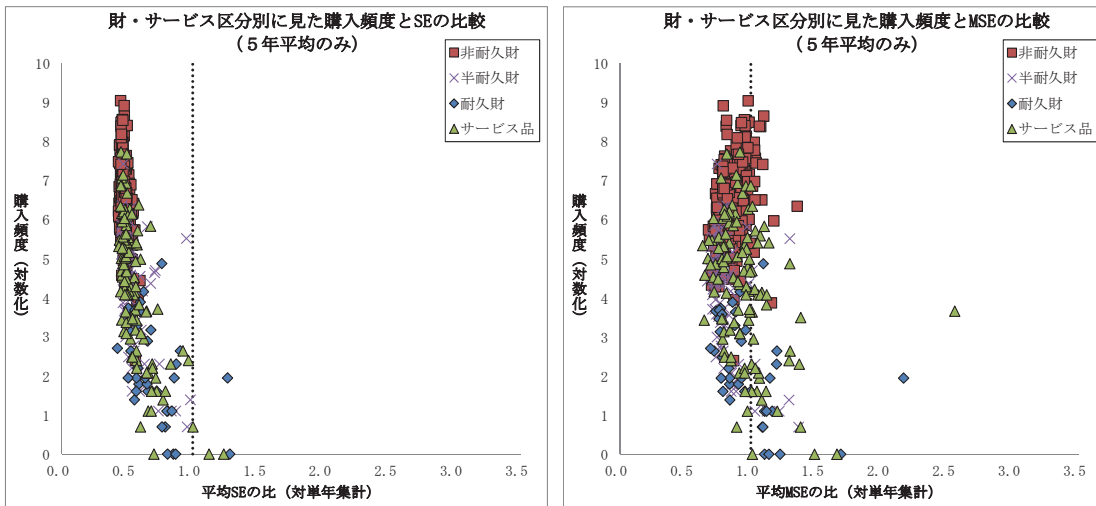
¹⁸ 北海道・東北地方の「灯油」に対する消費支出が全国的に極めて高いことや、高知における「かつお」の消費支出が全国的に極めて高いことなどは、地域性といえる。一方で、全国で消費支出が同程度であることも、1つの地域性である。（NHK放送受信料など）

付録2 図 (推計手法の評価結果) 財・サービス区別に見た購入頻度と SE 及び MSE の比較 (単年集計との比較) (5年平均のみ)

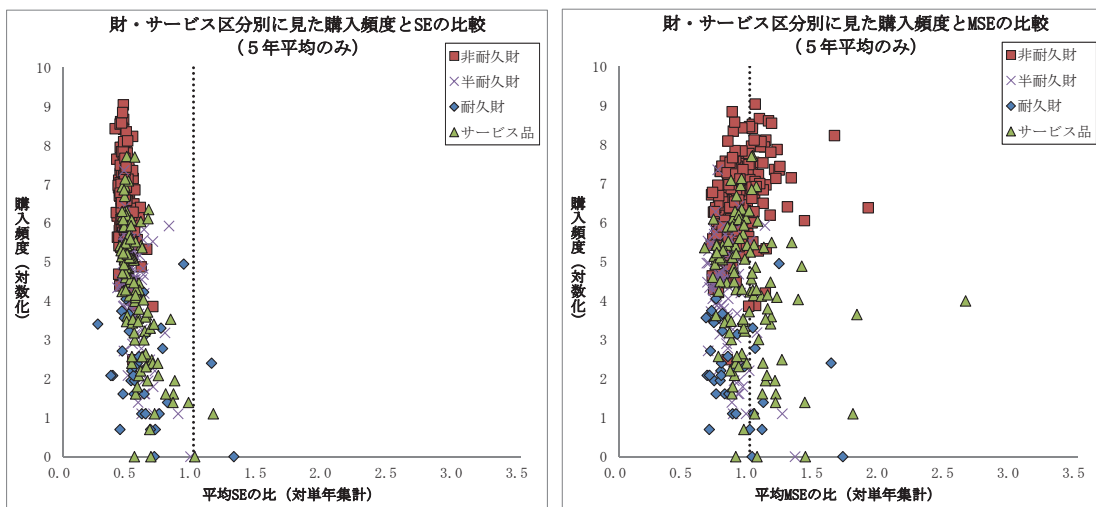
(2014年)



(2012年)

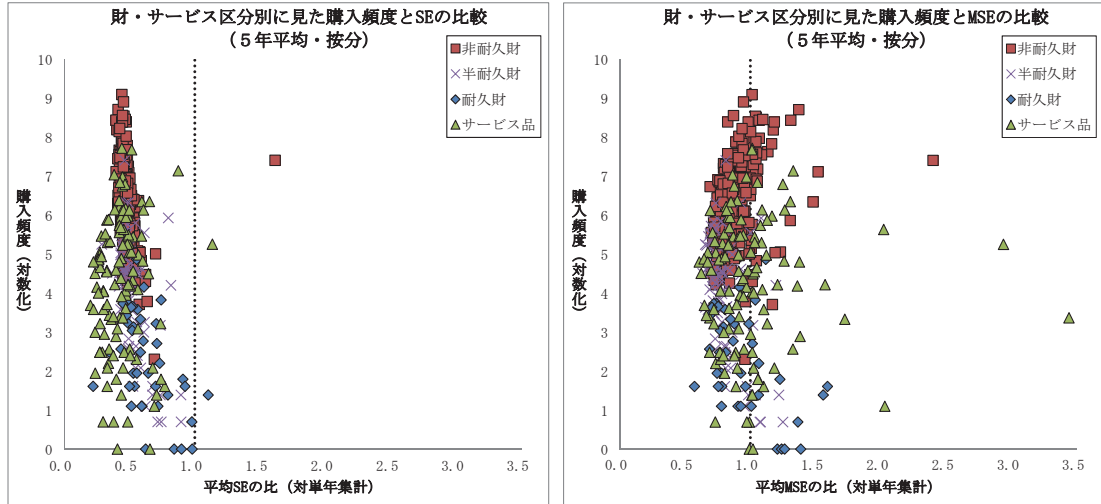


(2010年)

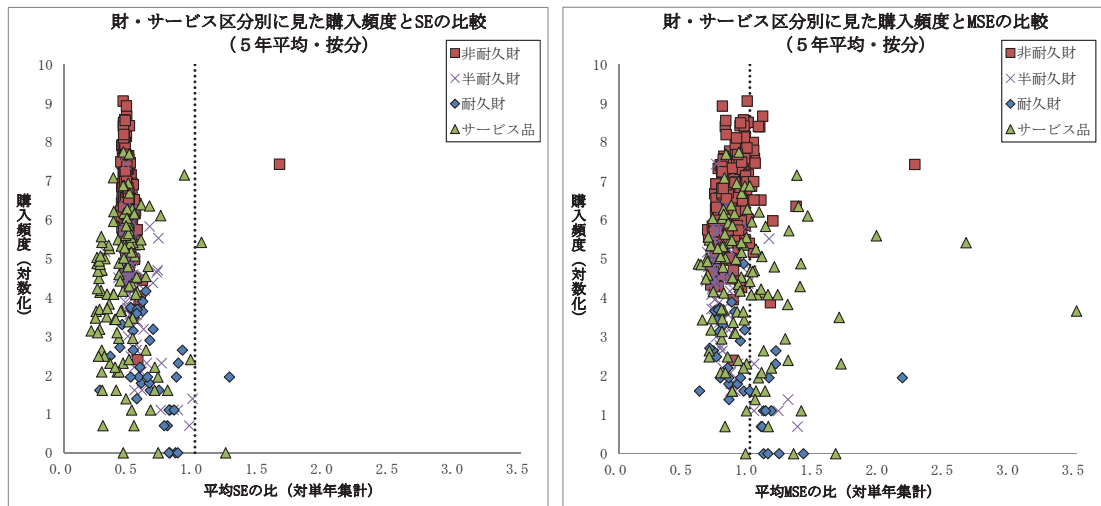


付録3 図 (推計手法の評価結果) 財・サービス区分別に見た購入頻度と SE 及び MSE の比較 (単年集計との比較) (5年平均・按分)

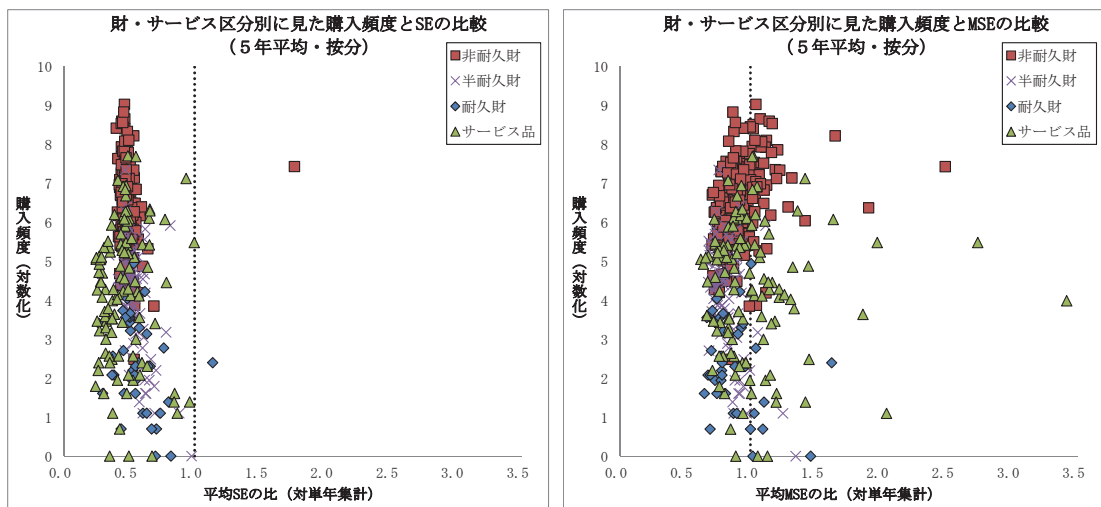
(2014年)



(2012年)

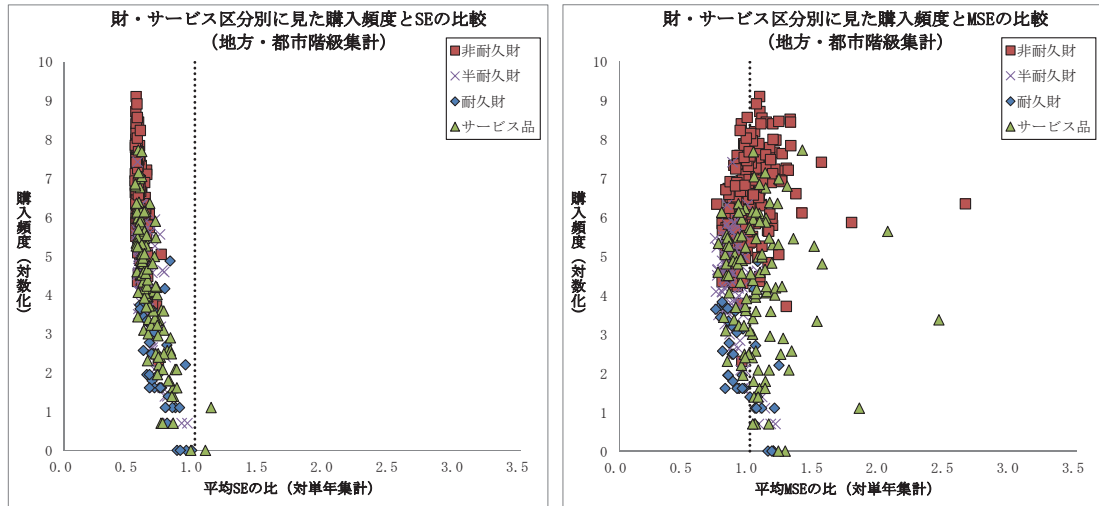


(2010年)

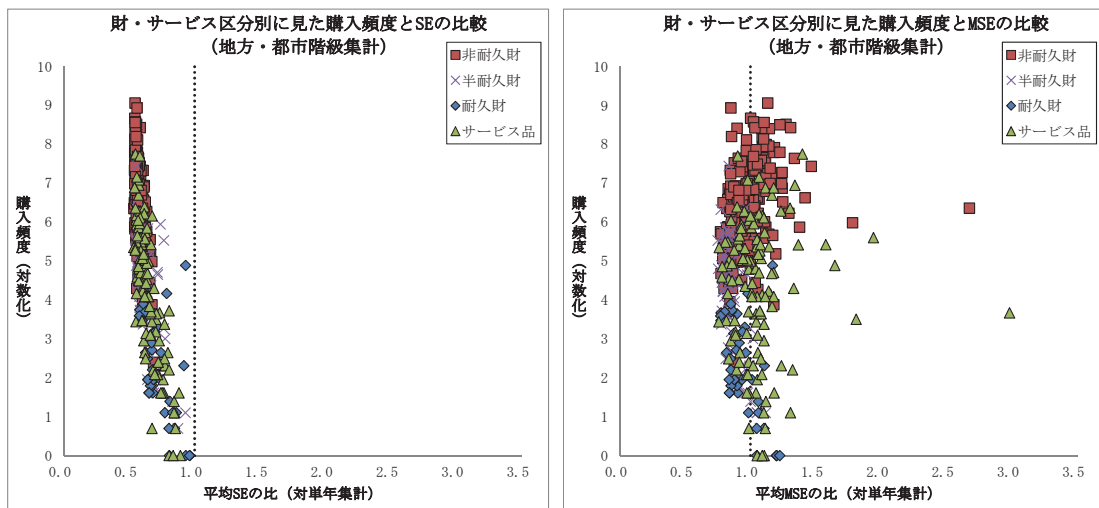


付録4 図 (推計手法の評価結果) 財・サービス区別に見た購入頻度と SE 及び MSE の比較 (単年集計との比較) (地方・都市階級集計)

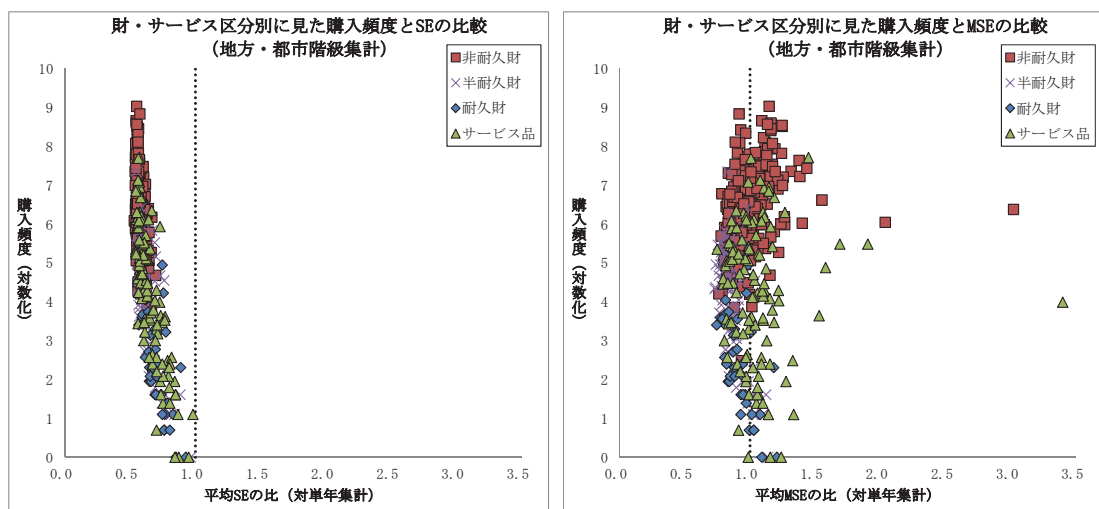
(2014年)



(2012年)

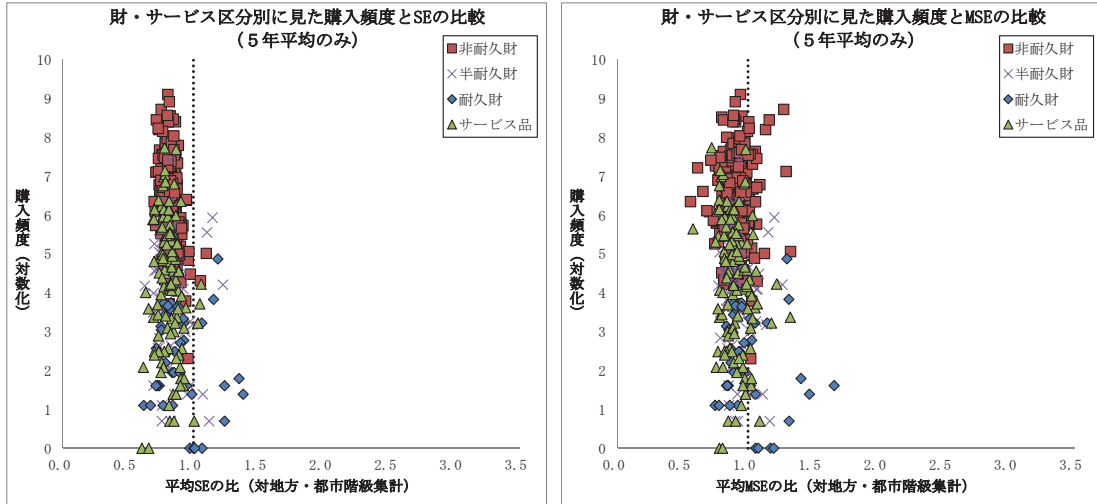


(2010年)

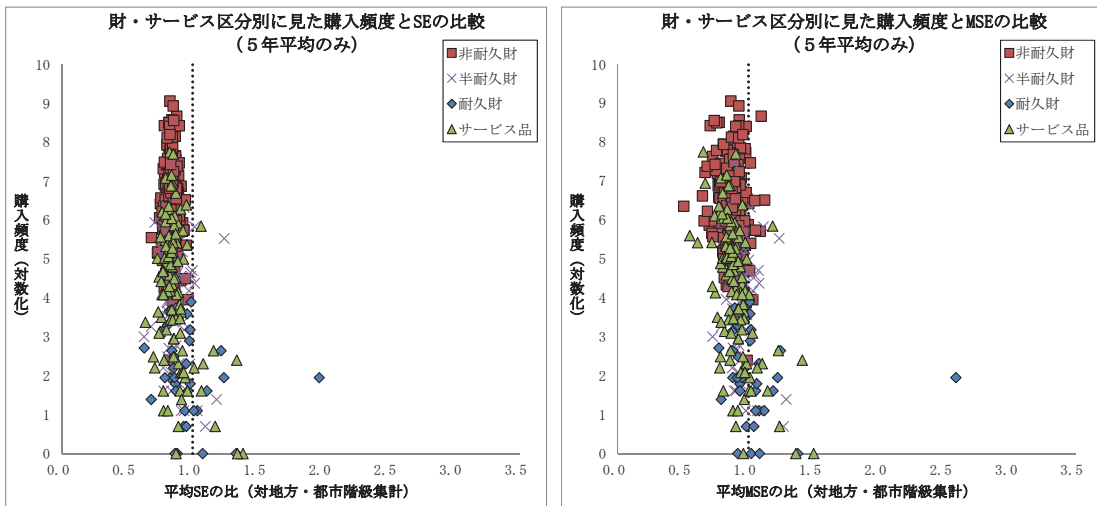


付録5 図 (推計手法の評価結果) 財・サービス区別に見た購入頻度とSE及びMSEの比較 (地方・都市階級集計との比較) (5年平均)

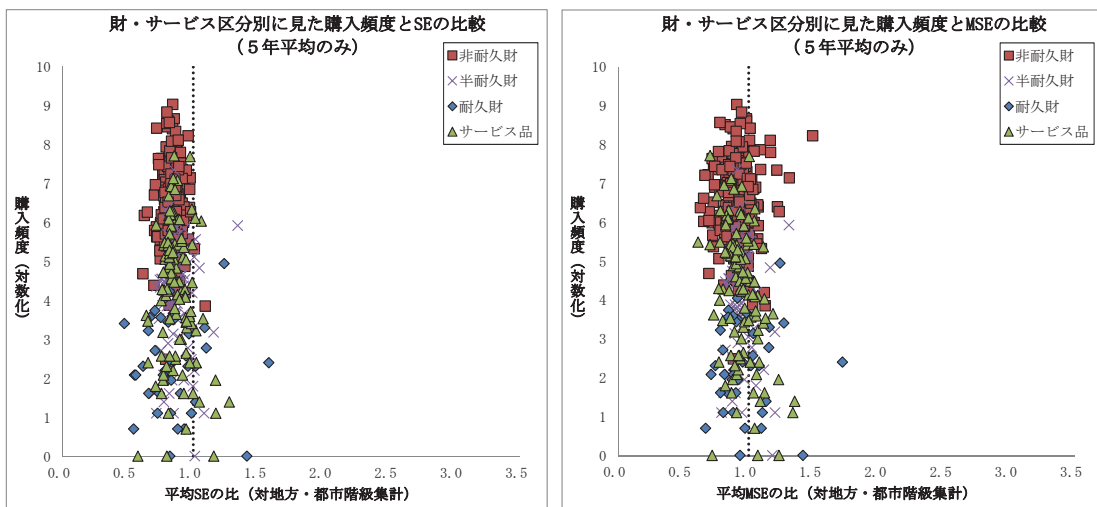
(2014年)



(2012年)

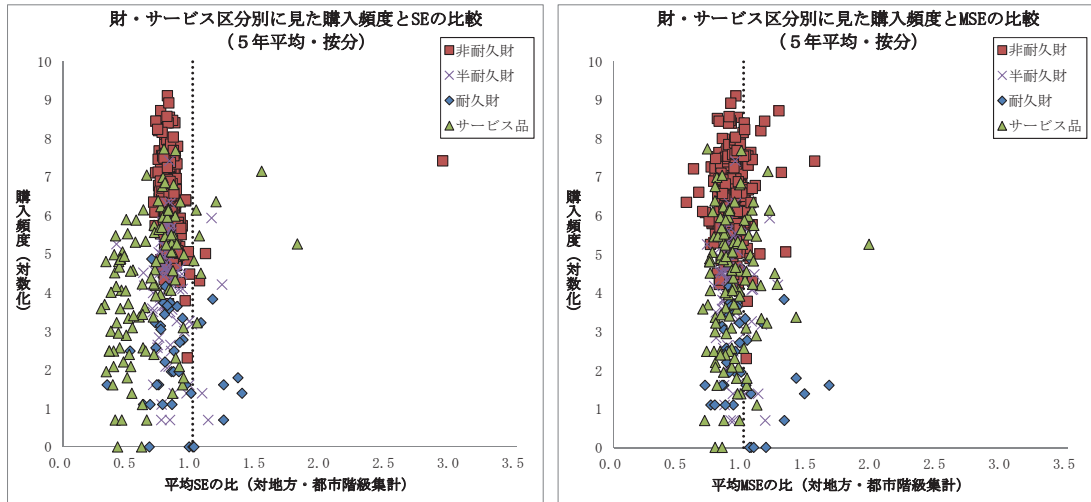


(2010年)

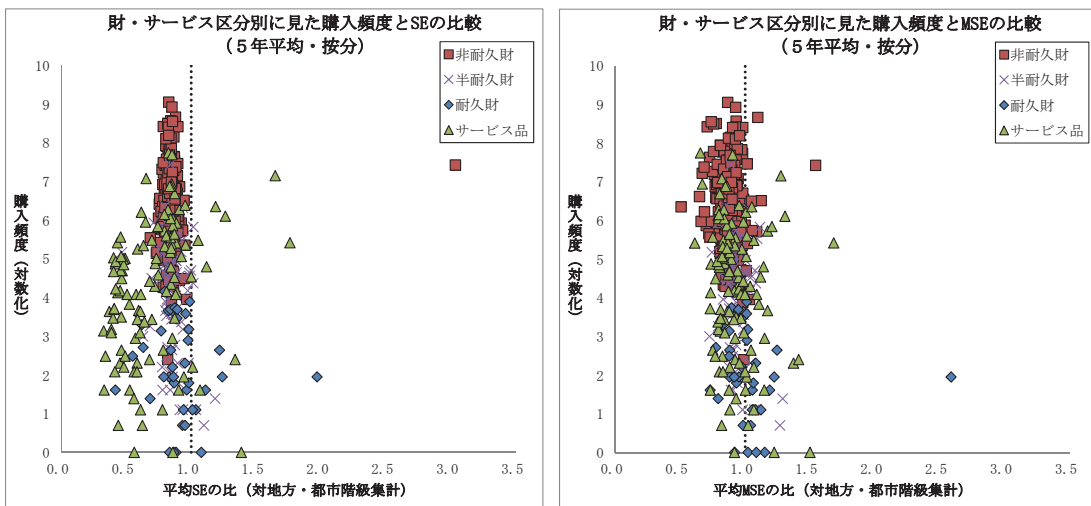


付録6 図 (推計手法の評価結果) 財・サービス区分別に見た購入頻度と SE 及び MSE の比較 (地方・都市階級集計との比較) (5年平均・按分)

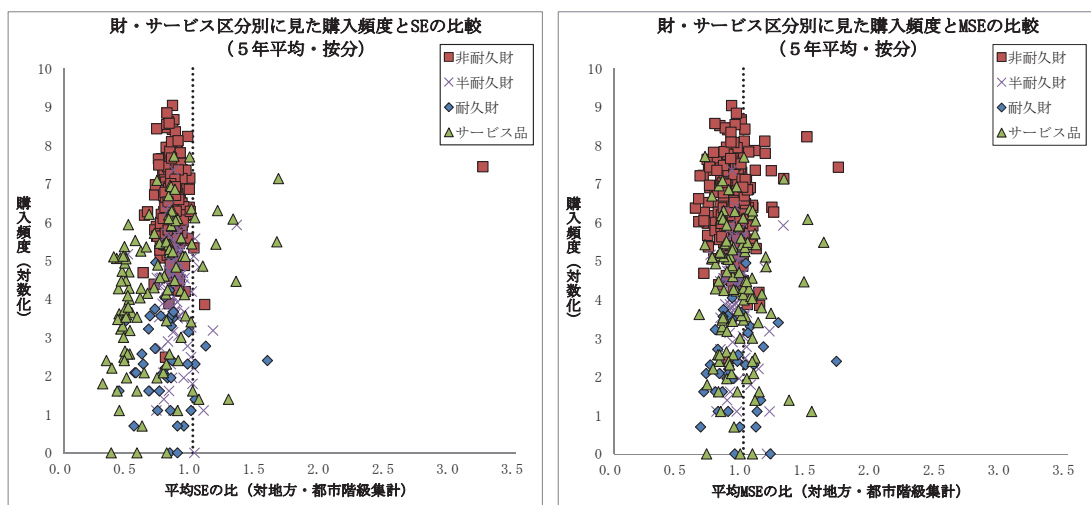
(2014年)



(2012年)

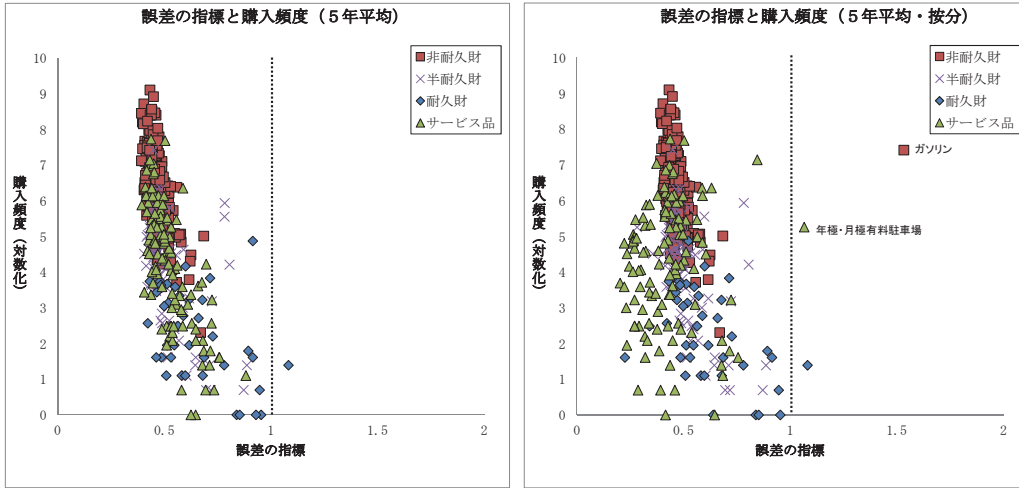


(2010年)

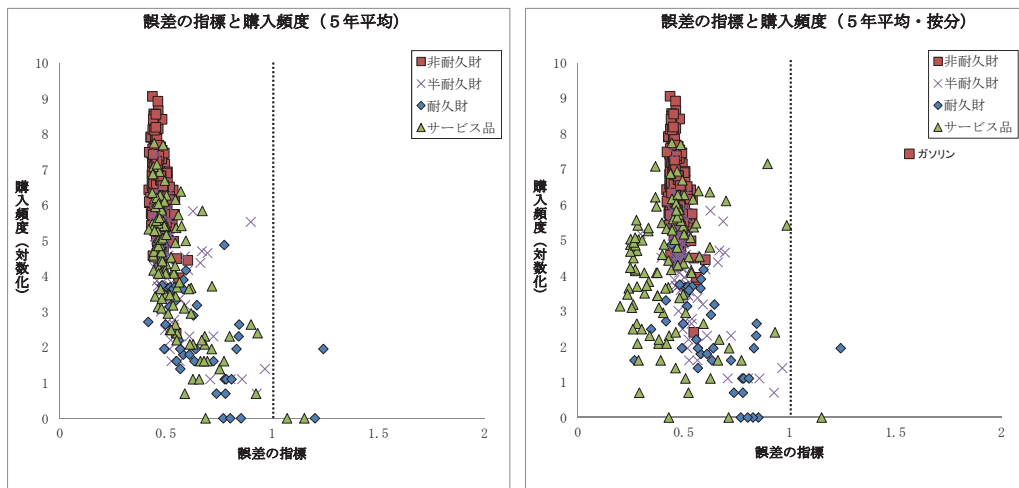


付録7 図 (推計結果の評価結果) 財・サービス区分別に見た分布1 (購入頻度と誤差の指標) (左: 5年平均のみ、右: 5年平均・按分)

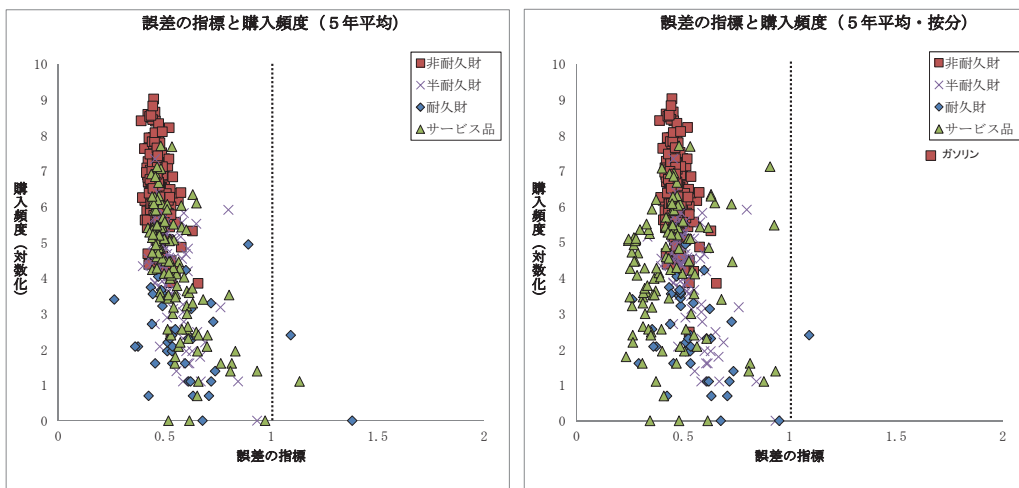
(2014年)



(2012年)

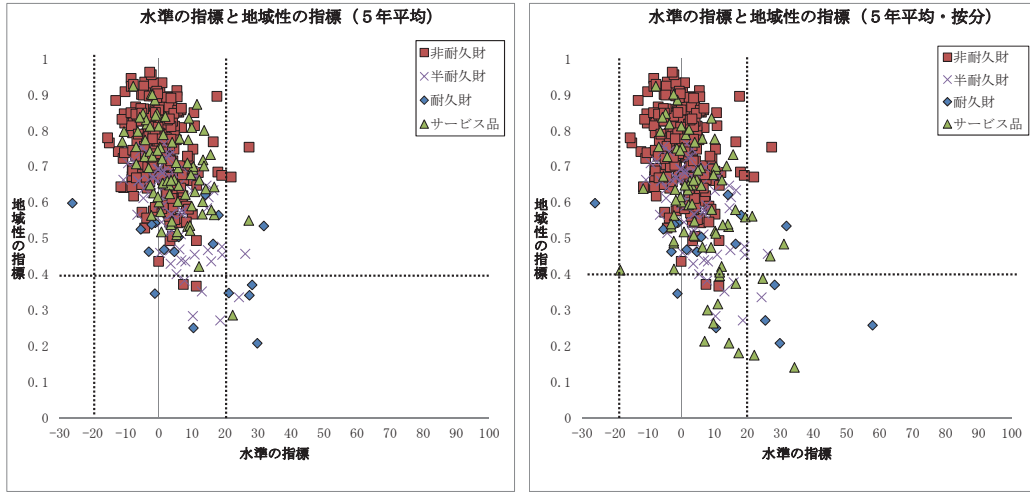


(2010年)

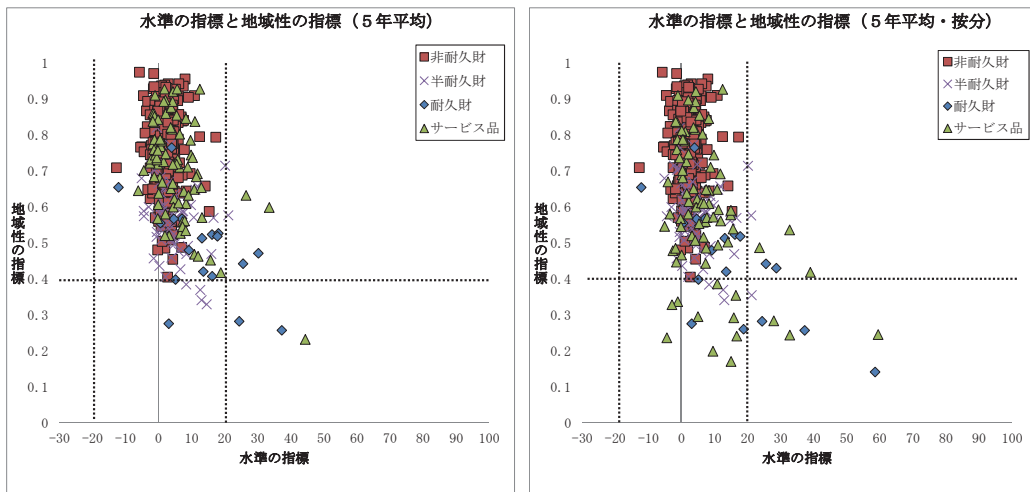


付録8 図 (推計結果の評価結果) 財・サービス区別に見た分布1 (水準の指標と地域性の指標) (左：5年平均のみ、右：5年平均・按分)

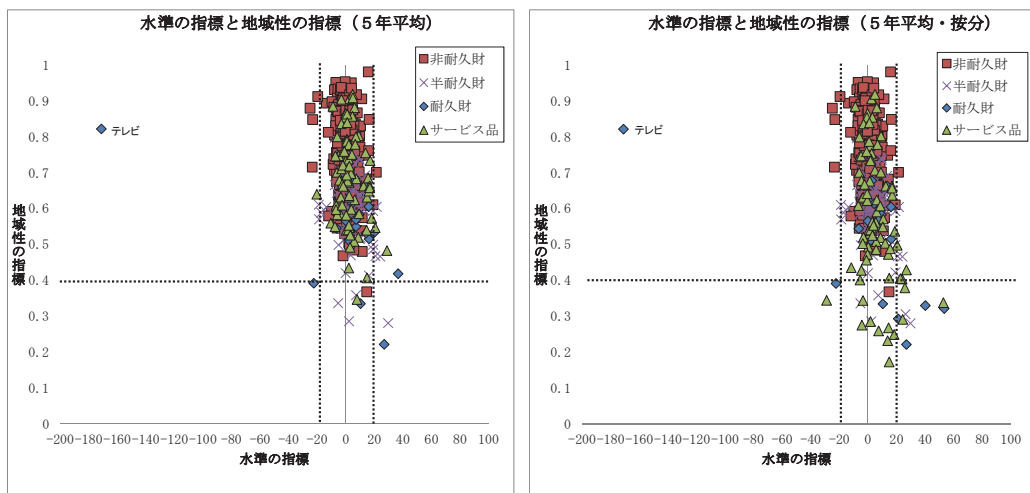
(2014年)



(2012年)



(2010年)



参考文献

- [1] 久保川達也(2006), 線形混合モデルと小地域の推定, 応用統計学, Vol.35, No.3, 139-161.
- [2] 小泉英希(2004), いくつかのSteinタイプの推定量の導入と評価方法, 統計研究彙報, 第61号, 139-179.
- [3] 小西貞則, 越智義道, 大森裕造 (2008), 計算統計学の方法—ブートストラップ・EM アルゴリズム・MCMC—, 朝倉書店.
- [4] 坂田幸繁 (2011), 標本調査データからの小地域情報の抽出可能性—都道府県別業況DIの推定をめぐる検証—, 「統計学」(2011年3月), 第100号, 41-56.
- [5] 総務省統計局(1), 「家計調査の推計方法」
<http://www.stat.go.jp/data/kakei/pdf/suikei.pdf>
- [6] 総務省統計局(2), 「全国物価地域差指数の作成方法」
<http://www.stat.go.jp/data/zenbutu/2007/pdf/sisu.pdf>
- [7] 高部勲 (2004), 小地域推計各手法の労働力調査への適用—都道府県別完全失業率の推定—, 統計研究彙報, 第61号, 1-179.
- [8] 元山齊, 山口幸三(2007), 小地域推計と労働力調査への適用, 「統計」(2007年3月), 66-73.
- [9] B. Efron and R. J. Tibishirani (1993), *An Introduction to the Bootstrap*, Chapman & Hall, New York.
- [10] M. Ghosh and J. N. K. Rao (1994), Small area estimation: An appraisal, *Statistical Science*, Vol. 9, No. 1, 55-93.

