

ビジネスレジスターの観光統計への応用¹

宮川幸三（慶應義塾大学産業研究所）

Ossi Nurmi（フィンランド統計局）

1. 本研究の目的と概要

2012年度より日本でもビジネスレジスター（事業所母集団データベース）の運用が開始された。ビジネスレジスターには、全ての事業所をカバーできるという大きな利点がある反面、収集される各企業や事業所の項目は基礎的なものに限られている。そのため日本におけるビジネスレジスターは、主に、標本調査を実施する際の母集団名簿や標本データから母集団推計を行う際の情報として使用され、『経済統計を正確に作成するための名簿情報の提供及び管理のための重要なインフラ』²とされている。

このような、標本調査を実施する際の補助的な役割に加えて、諸外国では、ビジネスレジスターの情報それ自体に基づいた統計の作成や分析が行われているケースも多い。そのうちの1つは、事業所や企業の動態に関する統計の作成である。ここで動態と呼んでいるのは、企業や事業所の参入（誕生）（Birth）や退出（死亡）（Death）、生き残り（Survival）や成長（Growth）等、個別企業や事業所の変化を意味しており、これらの企業・事業所の動態を表した統計は、**Business Demography Statistics**と呼ばれている³。ビジネスレジスターは、常に個別企業や事業所の最新の情報を保持していることから、ビジネスレジスターのみを用いて精度の高い**Business Demography Statistics**を作成することが可能となる。

もう1つの方向性として、事業所の立地に関する情報を用いた分析をあげることができる。前述のように、一般的なビジネスレジスターには、各企業や事業所の基礎的な情報のみが収録されているケースが多い。これは例えば、名称、所在地、資本金、従業員規模、売上高、開業年次、産業分類、といった項目である。このうち所在地情報に関しては、地理情報システム（**Geographic Information Systems : GIS**）を合わせて使用することによって、企業や事業所が立地する地域の詳細な情報を入手することが可能となり、それら立地にまつわる様々な変数を分析に利用することができる。例えば **Jarmin and Miranda (2006)**では、アメリカのビジネスレジスターより得られる事業所の所在地情報を用いて、ハリケーンが地域経済に及ぼした影響について分析を行っている。また **Miyagawa et al.**

¹ 本稿における研究の一部は、平成26年度文部科学省科学研究費補助金（基盤研究（C）、研究課題：地理情報システムを利用した地域の観光GDP推計手法の開発、研究代表者：宮川幸三、研究課題番号：25380270）の助成を受けている。

² 総務省統計局サイト（<http://www.stat.go.jp/data/jsdb/gaiyou.htm>）より引用。同サイトには、ビジネスレジスターの主な機能として、・各種統計調査へ母集団情報を提供、・母集団情報提供の際、調査対象者の負担を平準化（重複是正）、・各種統計調査実施の支援等、の3点があげられている。

³ 例えば、各国で比較可能な **Business Demography Statistics** の作成に向けたマニュアルとして、**Eurostat-OECD (2007) "Eurostat-OECD Manual on Business Demography Statistics"**があり、そこでは **Birth** や **Death** といった各動態の定義や計算方法の詳細が示されている。

(2014)では、ビジネスレジスターの利用方法の1つとして、事業所の所在地情報を用いて地域の観光客向けの売上額を推定する方法を示し、事業所の所在地情報およびGISの利用によって推定精度が改善されることを明らかにしている。これらの事例は、ビジネスレジスターが、存在する企業や事業所のすべてをカバーしているからこそ可能になる分析であり、企業や事業所の所在地情報を利用した分析は、ビジネスレジスターの重要な役割の1つであると言える。

この他に、ビジネスレジスターを利用して小規模な企業や事業所のデータを推定する試みも行われている。一般的に、小規模な企業や事業所の数は多く、従って統計調査にかかるコストは大きい。生産額や従業者数等のボリュームはそれほど大きくないため、費用対効果の観点から言えば小規模な企業や事業所の統計調査を行うことは大規模のケースと比較して非効率である。また調査に対する回答率に関しても、多くの統計調査では小規模な企業や事業所の回答率が大規模に比較して低い。以上の観点から言えば、小規模な企業や事業所のデータを統計調査を実施することなく収集できれば、費用面や統計の精度面からそのメリットは大きいものになるであろう。Nurmi (2014)では、フィンランドのビジネスレジスターを用いて、小規模な宿泊事業所の延べ宿泊者数を推定する方法が示されており、その推定結果はフィンランドの実際の宿泊統計として公表されている。

このようにビジネスレジスターは、単に標本調査を実施する際の母集団名簿や標本データから母集団推計を行う際の情報としての利用のみならず、多岐に渡って活用されている。ここであげたいいくつかの活用事例のうち、Miyagawa et al. (2014)およびNurmi (2014)は、いずれも観光に関連してビジネスレジスターを利用した推定を行うものである。一般的に、観光客は常に流動的であるため、観光客を対象として調査を行う際には多くの制約がある。一方で観光に関連する事業所や企業は小規模なものも多く、事業所や企業を対象として観光に関連する調査を実施することも容易ではない。特に一国よりも小さな地域を対象として観光の規模を把握することは多くの困難を伴い、従って地域の観光統計調査では、通常の家計や企業・事業所を対象とした調査とは大きく異なる方法によって調査が行われている。しかし、これまでに行われている調査手法が、精度の面で問題を抱えていることも否定できない⁴。このような状況の中で、観光統計に関連してビジネスレジスターを活用した推定手法を導入することは、より精度の高い地域観光規模の把握を達成するための1つの有用な方法になりうるものである。そこで本稿では、上述のNurmi (2014)およびMiyagawa et al. (2014)の詳細を紹介し、ビジネスレジスターの観光統計への活用に関する展望を明らかにする。

以下では、次節および第3節において、それぞれNurmi (2014)およびMiyagawa et al. (2014)の研究内容を紹介した上で、第4節では、ビジネスレジスターの活用に関する今後の展望について簡単に述べている。

⁴ 具体的には、①季節変動を把握することが困難、②詳細な費目について調査を行うことができない、③母集団名簿の欠如、といった問題がある。日本の観光統計の状況と課題や問題点については、宮川 (2009) に詳しい説明がある。

2. ビジネスレジスターの観光統計への活用事例 I

2.1. フィンランドの観光統計と EU レギュレーション

Nurmi (2014)は、フィンランドにおける小規模な宿泊事業所の延べ宿泊者数の推定にビジネスレジスターを活用したものである。内容を紹介するのに先立って、まずはフィンランドにおける観光統計の現状とそれを取り巻く EU の統計レギュレーションを明らかにする。

表 1 は、フィンランドの観光統計の現状を表している。

	アウトバウンド	国内観光	インバウンド
日帰り	フィンランド 旅行調査	フィンランド 旅行調査 (2018 年～)	国境インタビュー (2012 年)
宿泊 (支払いを伴う)	フィンランド 旅行調査	宿泊統計	宿泊統計
宿泊 (支払いを伴わない)	フィンランド 旅行調査	フィンランド 旅行調査	国境インタビュー (2012 年)

出典：Nurmi (2014)より筆者翻訳。

フィンランドでは、観光関連の統計調査としてフィンランド旅行調査 (Finnish travel survey) が行われている。この調査はサンプル調査であり、15 歳から 84 歳のフィンランド国内居住者を対象として毎月電話調査によって実施されている。サンプルサイズは 2,350 人/月であり、主にアウトバウンドと国内観光に関する日帰り旅行、宿泊旅行の目的地や移動手段、宿泊の形態などを明らかにしている。ただし国内の日帰り旅行者については、2018 年から調査を行う予定となっている。

一方でインバウンド観光に関しては、国境インタビュー (Border Interview Survey) が行われている。この調査は、空港や港、国境検問所などにおけるインタビュー調査として実施されており、インバウンド旅行者の国籍や居住国、訪問理由、目的地、訪問期間、宿泊手段などが調査されている。またいくつかの地域では、観光消費等も尋ねている。

以上は、観光旅行者に対する需要サイドの統計調査であるが、国内観光およびインバウンド観光の支払いを伴う宿泊客に関しては、宿泊統計 (Accommodation statistics) においてホテル等の施設を対象に月次の調査を行っている。調査項目は、部屋数やベッド数、客室稼働率、宿泊の目的、宿泊の価格等であり、国内宿泊客とインバウンド宿泊旅行を分けて調査している。調査対象となる宿泊施設数は、およそ 1,500 施設である。

これら 3 種類の調査に基づいて、フィンランドにおけるアウトバウンド観光、国内観光、

インバウンド観光の実態が網羅的に把握されており、これらの結果を用いて TSA (Tourism Satellite Account) 等の加工統計が推計されることになる。

一方でEU諸国については、観光統計に関するレギュレーション (EU 692/2011) が定められており⁵、その中では以下のように観光統計における宿泊施設の取り扱い方法が定められている。

1. 10 ベッド以上の宿泊施設を統計の対象とする。
2. 当該国の 1 年間の延べ宿泊者数 (人×泊) が EU 全体の 1% よりも小さい場合は、20 ベッド以上の宿泊施設を対象としても良い。
3. 2 番目の基準を適用した国は、調査対象規模以下の毎年の延べ宿泊者数を推定し、提供しなければならない。

フィンランドは、このうち 2 番目の基準が適用される国であるため、前述の宿泊統計では、20 ベッド以上の施設のみを対象として調査が行われている。従って、上述の 3 番目の基準にあるように、20 ベッド未満の宿泊施設に関しては、毎年の延べ宿泊者数を何らかの方法で推定しなければならない。そこでフィンランドでは、この小規模宿泊施設分の推定にビジネスレジスターを利用した方法を用いている。次項以降では、その方法の詳細を明らかにする。

2.2. 小規模宿泊施設の選定

小規模宿泊施設の推定を行うためには、まず当該事業所の規模を把握し、宿泊調査の調査対象であるのか、あるいは小規模施設推定の対象となるのかを選定する必要がある。宿泊施設の規模把握に使用されるソースは以下の 3 つである。

1. 宿泊施設名簿
2. ビジネスレジスター
3. 農業企業の納税データ

1 は、宿泊調査の母集団名簿としても使用されるものであり、常に情報のアップデートを行いながら 20 ベッド以上の宿泊施設約 1,500 施設が含まれている。施設の単位は、地域活動単位⁶ (Local Kind of Activity Unit) であり、各地域活動単位のキャパシティに関する情報とともに連絡先情報が保持されている。また地域活動単位には、共通のビジネス ID

⁵ 観光統計に関するレギュレーション (EU 692/2011) については、Eurostat(2014)において詳細な説明がなされている。

⁶ 地域活動単位は、日本の産業統計における「事業所」の概念に類似したものである。ただし地域活動単位の定義が、①単一の主体に所有されており、②単一の場所に立地しており、③主に単一の生産活動を行っていることであるのに対し、日本では複数の生産活動を行っていたとしても、①および②の条件に従っているものは事業所とみなされることになる。

および地域活動単位IDが付されているため、宿泊施設名簿に掲載される宿泊施設については、ビジネスレジスター等の他の行政データに容易にリンクすることが可能である。

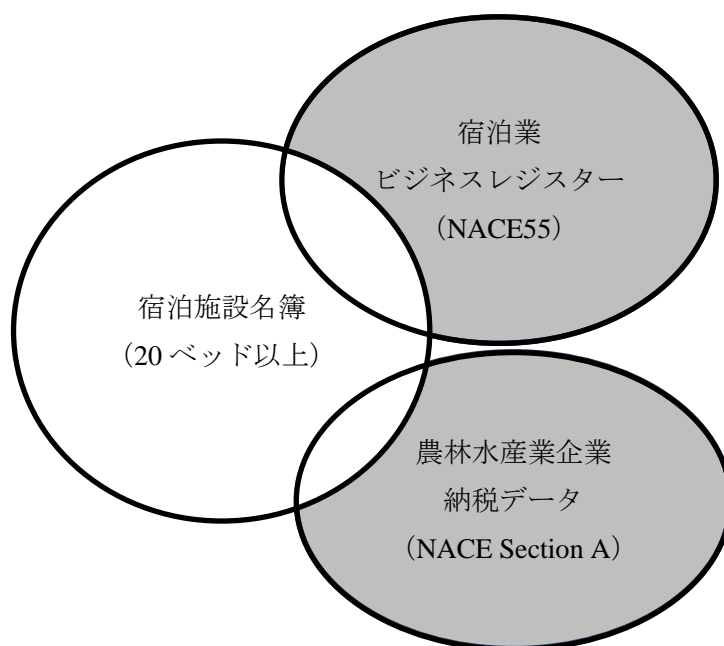
2 のビジネスレジスターは、フィンランドで活動する全ての企業、地域活動単位および企業グループの毎年のデータを収録している。フィンランドのビジネスレジスターでは、NACE⁷ Rev.2 と互換性のあるTOL2008 と呼ばれる標準産業分類に従って地域活動単位が格付けられており、また全ての地域活動単位について売上額の情報も収録されている。情報は毎年 11 月に更新される。

宿泊業は、NACE で言えば、「55. 宿泊業 (Accommodation)」に該当するものであり、1 の宿泊名簿より把握される 20 ベッド以上の宿泊施設と、2 のビジネスレジスターにおける NACE55 に格付けられる地域活動単位を照合することによって、20 ベッド未満の宿泊施設を特定化することができる。このようにして把握される 20 ベッド未満の宿泊施設は、2012 年時点で約 3,500 施設である。

NACE55 に格付けられる宿泊業を主業とする地域活動単位に加えて、フィンランドでは、農林水産業を主業とする企業が宿泊サービスを提供している、いわゆるアグリツーリズムあるいはグリーンツーリズムと呼ばれるようなケースが数多く存在する。このような農林水産業企業の提供する宿泊サービスを把握するために、3 の納税データが使用される。このデータは、NACE で言えば「Section A. 農林水産業 (Agriculture, Forestry and Fishing)」に格付けられる企業の納税データを収集したものであり、その中でこれらの企業が提供した宿泊サービスより得られる収入を分離して申告している。この農林水産業に関する納税データに関しても、1 のベッド数 20 以上の宿泊施設名簿と照合することによって、20 ベッド未満の施設を特定化することが可能となる。言うまでもなく、ここで述べたようなデータソース間の照合作業は、全てのデータに共通のビジネス ID および地域活動単位 ID が存在するために容易に行うことができるものである。

図 1 は、ここで述べた 3 種類のデータの関係を表している。図 1 に描かれる 3 つのデータを照合することによって、灰色部分に該当する 20 ベッド未満の小規模宿泊施設を把握することが、推定のための第 1 段階となる。2011 年の結果としては、ビジネスレジスターの宿泊業 (NACE55) に含まれる小規模施設数が 1,952 (図 1 右上円内の灰色部分に該当)、農業企業納税データに含まれる小規模施設数が 2,475 (図 1 右下円内の灰色部分に該当)であった。

⁷ NACE は、ヨーロッパの標準産業分類 (Statistical Classification of Economic Activities in the European Community) である。詳細については、Eurostat (2008) を参照のこと。



出典：Nurmi (2014)より筆者翻訳。

図 1. 宿泊規模識別のためのデータソース

次の段階としては、把握された 4,427 (=1,952+2,475) の地域活動単位のうち、推計の対象にならない地域活動単位を取り除く作業が行われる。例えば、宿泊業を主業とする企業の地域活動単位であったとしても、本社や営業所等のように、地域活動単位自体が宿泊サービスを提供していないケースについては、これを推計対象から取り除くことになる。また、ビジネスレジスターでは把握されているが宿泊施設名簿では把握されておらず、かつ 20 ベッド以上の規模である施設が存在することもあり、このような地域活動単位も小規模施設推定の対象から除く必要がある。これらの選別作業は手作業で行われ、結果として 3,522 の地域活動単位が小規模推計の対象となる。

2.3. 延べ宿泊者数の推定

延べ宿泊者数の推定に際しては、「同一産業に格付けられる施設である場合、施設自体の規模(ベッド数)に関わらず 1 泊当たりの売上額は一定である」という仮定を置いている⁸。20 ベッド以上の規模を持つ宿泊施設に関しては、3.1 項で述べた宿泊統計より 1 泊当たりの売上額を求めることができる。また全ての宿泊施設は、宿泊業に関する 3 つの NACE クラス(「55.1 ホテルおよび類似の宿泊施設 (Hotels and similar accommodation)」、「55.2 休暇施設およびその他の短期滞在宿泊施設 (Holiday and other short-stay

⁸ 通常の売上額には、レストラン等の宿泊サービス以外の売上が含まれているため、本来は「1 泊あたり売上額」ではなく「1 泊あたり宿泊料金」を一定として推定を行うべきである。しかし後に示すように、ビジネスレジスターでは各地域活動単位の売上額のみが収録されているため、「1 泊あたり売上額」を一定とするような仮定が用いられている。

accommodation)」、「55.3 キャンプ場, RVパークおよびトレーラーパーク (Camping grounds, recreational vehicle parks and trailer parks)」のいずれかに格付けられており、このそれぞれの部門について1泊当たり売上額が一定であるという仮定を置くことになる。

表 2 は、推定の詳細を表したものである。

表 2. 規模以下宿泊施設 延べ宿泊者数推定結果

NACE クラス		(1) 地域活動 単位数	(2) 売上額 (百万€)	(3) 1泊 あたり 売上額 (€)	(4) 延べ 宿泊者数 (千人泊)	(5) 非居住者 シェア	(6) 延べ 宿泊者数 居住者 (千人泊)	(7) 延べ 宿泊者数 非居住者 (千人泊)
55.1	ホテルおよび類似 の宿泊施設	237	11.7	84.56	139	23%	107	32
55.2	休暇施設およびそ の他の短期滞在宿 泊施設	3,234	46.7	25.49	1,834	21%	1,451	384
55.3	キャンプ場, RV パ ークおよびトレイ ラーパーク	51	1.1	45.58	25	16%	21	4
合計 (2011年推定値)		3,522	59.6	29.84	1,998	21%	1,579	420

出典：Nurmi (2014)より筆者翻訳。

表 2 の(1)列は、前項において選定された、推定の対象となる 20 ベッド以下の宿泊施設 (地域活動単位) の数を表している。3,522 施設のうち大半が、NACE クラスで言えば 55.2 に格付けられていることがわかる。(2)列は、ビジネスレジスターおよび農林水産業企業納税データより得られる売上額を、部門ごとに合計したものである。

一方(3)列は、宿泊統計より計算される 20 ベッド以上の施設に関する「1泊当たり売上額」(施設の売上額/延べ宿泊者数)である。ただし NACE クラス 55.2 については、そもそも小規模施設が多く、20 ベッド以上の施設を対象とした宿泊統計ではほとんどその実態を把握できていないため、別途、休暇施設およびその他の短期滞在宿泊施設のための集中予約サービス (centralized reservation services) より得られる情報に基づいて、(3)列の 1泊当たり売上額を計算している。結果として、NACE クラス 55.2 の 1泊当たり売上額がもっとも小さく、55.1 がもっとも大きくなっている。

(4)列は、(2)列の売上額を(3)列の 1泊当たり売上額で割ることによって計算した延べ宿泊者数である。更に延べ宿泊者数は、居住者と非居住者に分割して求められなければならない。そこで(5)列には、宿泊統計より入手される「延べ宿泊者数に占める非居住者のシェア」を掲載している。このシェアを用いて(4)列の値を案分することによって、(6)列および

(7)列において居住者、非居住者別の延べ宿泊者数が推定されることになる。

2.4. 推定手法の評価とまとめ

前項までに説明した推定手法を適用することによって、20ベッド未満の小規模宿泊施設に関する延べ宿泊者数を、新たに統計調査を実施することなく把握することができる。20ベッド以上で宿泊統計の対象となる施設の数約 1,500 であるのに対し、推定の対象となる 20ベッド未満の小規模施設数は約 3,500 に上っており、ここで紹介した推定手法を適用することによって、調査のコストや記入者負担が大きく削減されることは明らかであろう。

また一般的には、小規模主体による調査への回答率は低いものであるため、ビジネスレジスターを用いた今回の推定手法を適用することによって全ての小規模施設をカバーできることを考えれば、精度面からもメリットがあるだろう。しかし一方で、脚注 8 でも述べたように、推定時に使用される施設ごとの「売上額」には宿泊サービス以外の売上が含まれているため、そのことが推定結果にバイアスをもたらす可能性を否定することはできない。

産業のカバレッジという観点から言えば、今回の手法では、宿泊業を主業とする地域活動単位だけでなく、農林水産業でありながら宿泊サービスを提供するケースも含んで推定を行っている。この点はこの推定手法のメリットの 1 つであるが、一方で、他の産業部門において宿泊サービスを提供しているケースが存在する可能性もあり、これを考慮していない点は本手法の欠点でもある。産業のカバレッジを拡張することは、今後の課題の 1 つとなるであろう。

特に観光統計の場合には、データ周期の問題も重要な要素となる。これは、観光関連の支出に関しては季節変動が激しいためであり、理想的には月次周期でのデータ収集が望まれる。フィンランドにおいても 20ベッド以上の施設を対象とした宿泊統計では、月次データが収集されている。しかしビジネスレジスターにおいて収録されている売上額データは、月次で更新されるものではない。従って小規模施設の推定結果は、基本的には年に 1 回のペースで改定されることになる。小規模施設に対する統計調査を行わずにこの問題を解決することは困難であり、ここであげたような周期の問題は、ビジネスレジスターを用いる手法の限界であろう。

3. ビジネスレジスターの観光統計への活用事例Ⅱ

3.1. 分析の目的と概要

本節では、ビジネスレジスターの観光統計への活用事例の 1 つとして、Miyagawa et al. (2014)の内容を明らかにする。この研究の目的は、ビジネスレジスターを利用して一国より小さな地域の観光規模を推定する手法を開発することである。

一般的に、地域の観光規模を把握する際には、観光地において観光客を対象として消費額に関する調査を実施するケースが多い。これは言うてみれば、需要サイドからのアプロ

一チである。日本でも、地域観光規模把握を目的とした観光統計として、各都道府県の手によって「観光入込客統計」が実施されている。この統計は、観光庁において定められた「観光入込客統計に関する共通基準」⁹に従って共通の方法によって作成されるものである。共通基準では、観光地点における観光客を対象としたパラメータ調査を通じて観光客の訪問地点数や消費額を調査し、最終的に地域の入込客数や観光消費額を推定する手法が細かく決められている。しかし、費用面の問題からパラメータ調査は四半期に一度行われるのみであり、従って観光消費の季節変動を厳密に把握することはできず、また観光客の回答者負担の観点から詳細な調査を行うことができない、といった問題があることも事実である。

そこで観光庁では、上述のような需要サイドアプローチの問題点を考慮し、供給サイドから観光規模を把握する試みとして 2012 年に「観光地域経済調査」を実施した。「観光地域経済調査」は、宿泊業や飲食業、小売業等、観光に関連する産業部門に格付けられる事業所を対象とした調査である。調査対象事業所は約 9 万事業所にのぼり、国内の 904 の地域を対象として調査を実施している。調査事項は、事業所の従業者数や経営組織、売上額などの一般的な項目に加えて、売上額に占める観光の割合や、観光に関連する事業所の特性（観光協会に加盟しているかどうか、観光ガイドブックに掲載されているかどうか、大型バス用の駐車場があるかどうか、など）等多岐にわたっている。

この調査はサンプル調査であることから、調査の結果を用いて地域内の全事業所の観光向け売上総額を求める、といったような場合には、事業所の母集団名簿、つまりビジネスレジスターを用いて母集団推定を行わなければならない。地域の観光向け売上額は、当該地域の政府や観光関連事業者にとって最低限把握すべき重要な情報であることから、できる限り精度の高い手法によって母集団推定を行うことが望まれる。しかしながら、一般的なビジネスレジスターでは、事業所の産業分類や従業者数、場合によっては生産額といった基礎的な情報が収録されているのみで、観光に関連する事業所の情報を保有しているケースはほとんどない。そのため、母集団推定を行うとしても、産業ごとに分割して推定を行う、従業者数や生産額のデータを用いて規模別の推定や比推定を行う、といった程度の手法を採らざるを得ない。観光向け売上額は、当該事業所の立地によって大きく異なるものであることから、ここで述べたような一般的な手法によって精度の高い観光向け売上額を推定することは困難であろう。

そこで Miyagawa et al. (2014)では、事業所の立地に着目し、ビジネスレジスターより得られる住所情報から割り出される緯度・経度に基づき、GIS を用いて事業所の周辺情報を利用した新たな母集団推定手法を提案している。次項以下では、その詳細を明らかにする。

3.2. ビジネスレジスターを用いた観光売上額推定

⁹ <http://www.mlit.go.jp/kankocho/siryoutoukei/irikomi.html>

図 2 は、一般的に想定される、観光地域経済調査とビジネスレジスターを用いた地域観光売上総額推定のフローを表したものである。

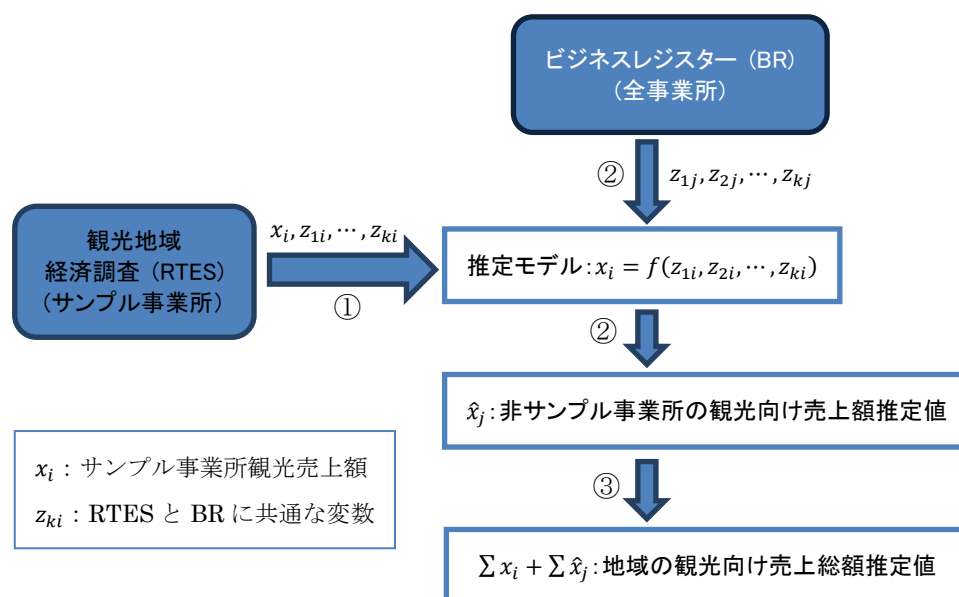


図 2. ビジネスレジスターを用いた地域観光売上推定のフロー

推定は以下のようなプロセスで行われる。

- ① 最初に、観光地域経済調査より得られる個別事業所の観光売上額 x_i と各事業所にまつわる変数 z_{1i}, \dots, z_{ki} を用いて、事業所の観光売上額を推定するモデル $x_i = f(z_{1i}, z_{2i}, \dots, z_{ki})$ のパラメータを決定する。ここで z_{1i}, \dots, z_{ki} は、観光地域経済調査とビジネスレジスターに共通な変数であり、一般的には産業格付けや従業員数、売上額等を想定している。
- ② z_{1i}, \dots, z_{ki} はビジネスレジスターにも収録されている変数であるため、観光地域調査の対象となっていない非サンプル事業所の z_{1i}, \dots, z_{ki} についても、ビジネスレジスターより入手することができる。これを①で構築した推定モデルに代入することによって、非サンプル事業所の観光向け売上額 \hat{x}_j を推定することができる。
- ③ 観光地域経済調査より得られる全てのサンプル事業所の観光売上総額 $\sum x_i$ と、②で推定した非サンプル事業所の観光向け売上総額 $\sum \hat{x}_j$ の合計が、当該地域の観光向け売上総額推定値となる。

ここで重要な問題は、非サンプル事業所の観光向け売上額 \hat{x}_j の推定精度であり、それは z_{1i}, \dots, z_{ki} としてどのような変数を用いるかによって大きく左右されることになる。Miyagawa et al. (2014)では、観光地域経済調査を利用して行う推定シミュレーションを行

い、 z_{1i}, \dots, z_{ki} として事業所周辺の立地特性変数を用いることによって、他の方法に比較して推定精度が飛躍的に高まることを示している。ここで立地特性変数と呼んでいるのは、各事業所の経度・緯度情報に基づき、GISを用いて得られる事業所周辺の状況を表す変数であり、具体的には、事業所から観光地点や鉄道の駅までの距離等である。小売業や飲食店における観光向け売上を考えれば、観光地点や鉄道までの距離が大きく影響することは明らかであろう。このような立地特性変数は、事業所の住所情報のみから作成することができるため、住所情報を保持するビジネスレジスターより、全ての事業所の観光売上額を推定することが可能となる。次項では、推定シミュレーションの詳細を明らかにする。

3.3. 観光売上額推定シミュレーション

図3は、Miyagawa et al. (2014)における推定シミュレーションの手法を示したものである。

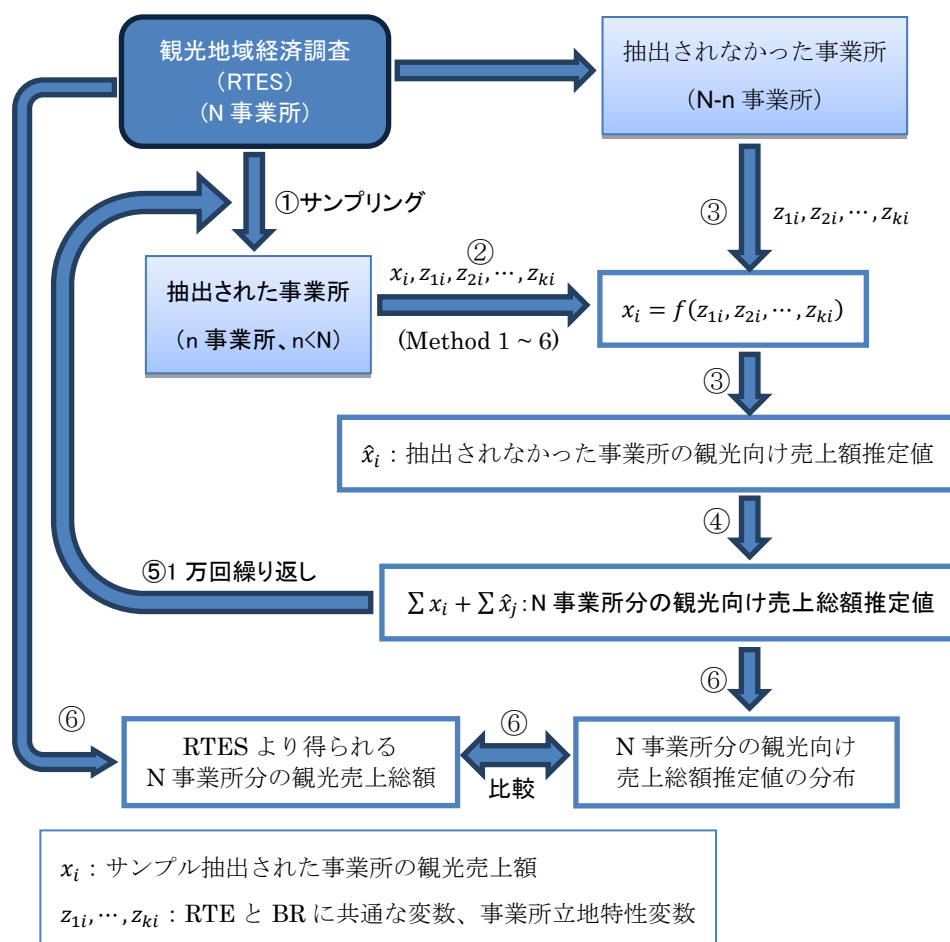


図 3. 推定シミュレーションのフロー

シミュレーションは以下のような方法で行われる。

- ① まず、観光地域経済調査の結果データ（N 事業所分）から無作為抽出を行い、n 事業所分を抽出する。実際の計算では、抽出率は 50%となっている。（ $N=2n$ である。）
- ② 抽出された事業所の観光売上額 x_i と各種変数 $z_{1i}, z_{2i}, \dots, z_{ki}$ を用いて、観光売上額推定モデル $x_i = f(z_{1i}, z_{2i}, \dots, z_{ki})$ のパラメータを決定する。この際には、Method 1~6の異なる 6つのモデルを用いる。詳細は後に説明するが、Method 1~4は立地特性変数を用いない手法であり、5・6は立地特性変数を用いた手法である。論文の目的は、5・6の手法による推定精度が、1~4の手法に比較して高いことを示すことである。
- ③ ①で行ったサンプリングにおいて抽出されなかった事業所（ $N-n$ 事業所分）について、観光地域経済調査より得られる各種変数 $z_{1i}, z_{2i}, \dots, z_{ki}$ を②で構築した観光売上額推定モデルに代入することによって、抽出されなかった事業所の観光向け売上額推定値 \hat{x}_j を計算する。
- ④ ①で抽出された事業所の実際の観光売上総額 $\sum x_i$ （これは観光地域経済調査のデータそのものである）と、③で推定した抽出されなかった事業所の観光向け売上総額推定値 $\sum \hat{x}_j$ の合計を、観光地域経済調査の対象となった N 事業所分の観光向け売上総額推定値とする。
- ⑤ ①~④のプロセスを、6つの手法それぞれについて 1 万回ずつ繰り返し行う。これは、推定結果の分布を描くためである。
- ⑥ ⑤の繰り返し計算の結果に基づき、各手法ごとに観光向け売上総額推定値（1 万回分）の分布を描き、観光地域経済調査から得られる N 事業所分の実際の（調査結果としての）観光向け売上総額と比較する。もしも推定結果が観光地域経済調査から得られる売上総額の近くに集中していれば、その手法の推定精度は高いものであると言える。

以上のような方法によって、立地特性変数を用いた推定手法と用いない推定手法の精度を比較することができる。実際の分析では、静岡県の小売業と飲食サービス業のそれぞれについてこのシミュレーションを行い、結果を示している。ここで述べた事業所別観光向け売上額推定の手法（Method 1~6）の詳細は、以下のようなものである。

Method 1

Method 1は、図 3の①で抽出された事業所についての 1 事業所あたり平均観光向け売上額を、抽出されなかった事業所の観光向け売上額として使用するものである。

Method 2

Method 2は、図 3の①で抽出された事業所についての 1 事業所あたり平均観光向け売

上割合を、抽出されなかった事業所の観光向け売上割合として使用し、観光向け売上額を推定するものである。

Method 3

Method 3 では、被説明変数を観光向け売上割合、説明変数を各事業所の売上額（観光向け以外も含む）としたロジスティック回帰によって観光向け売上額を推定している。¹⁰

Method 4

Method 4 も、ロジスティック回帰の手法を用いたものである。Method 3 と同様に被説明変数は観光向け売上割合であり、説明変数としては事業所の従業者数を用いている。

最終的にビジネスレジスターを用いた母集団推定を行うことを考えた場合、ここで説明変数として使用できる変数はビジネスレジスターに収録されている変数でなければならない。ビジネスレジスターには、通常、売上額や従業者数等の基本的な情報のみが収録されているため、Method 3 や 4 のようなモデルによって観光売上額の推定を試みている。

Method 5

Method 1～4 が立地特性変数を使用しない手法であったのに対し、Method 5 は、立地特性変数を説明変数としてロジスティック回帰を行うものである。立地特性変数とは、具体的には以下のようなものである。

表 3. 事業所立地特性変数リスト

変数名	内容
観光地点	事業所から観光地点までの距離を表す変数。1 つの事業所が複数の異なる規模の観光地点の近隣に立地しているケースもあるため、観光地点リストで把握されている全ての観光地点について、各観光地点の入込客数をその観光地点から当該事業所までの距離で割った値を求め、その合計を変数として使用した。これによって、事業所の距離が遠い観光地点は、この変数の中でほぼ0に近い評価をされ、距離が近く入込客数の大きな観光地点ほど大きく評価されることになる。近傍に大規模な観光地点があれば、当該事業所の観光向け売上割合も大きくなると考えられるため、期待されるパラメータの符号は(+)である。
新幹線駅	事業所から最も近い新幹線の駅までの距離を表す変数。新幹線を利用する観光客が駅の近くで消費する可能性があるため、期待されるパラメータの符号は(-)である。
鉄道駅	事業所から最も近い新幹線以外の JR の駅からの距離を表す変数。JR 線を利用する観光客が駅の近くで消費する可能性があるため、期待されるパラメータの符号は(-)である。
高速道路出口	事業所から最も近い高速道路の出口からの距離を表す変数。高速道路を利用する観光客が出口付

¹⁰ ロジスティック回帰については、Dobson and Barnett (2008)などに詳しい説明がある。

	近で消費する可能性があるため、期待されるパラメータの符号は(-)である。
宿泊施設	事業所から宿泊施設までの距離を表す変数。観光地点のケースと同様、1つの事業所が複数の異なる規模の宿泊施設の近隣に立地しているケースもあるため、経済センサスで把握されている全ての宿泊施設について、各宿泊施設の売上額をその宿泊施設から当該事業所までの距離で割った値を求め、その合計を変数として使用した。これによって、事業所の距離が遠い宿泊施設は、この変数の中でほぼ0に近い評価をされ、距離が近く売上額の大きな宿泊施設ほど大きく評価されることになる。近傍に大きな宿泊施設があれば、当該事業所の観光向け売上割合も大きくなると考えられるため、期待されるパラメータの符号は(+)である。

これらの事業所立地特性変数は、事業所の住所から割り出される緯度・経度情報と、一般的に使用されている地図や経済センサス等より得られる事業所周辺の状況のみによって求められるものであるため、ビジネスレジスターを含め、住所情報を保有する全てのデータにおいて上述のような変数を用いた分析を行うことができる。

Method 6

Method 5 は、立地特性変数に加えて事業所の観光に関連する特性変数（以下では観光特性変数と呼ぶ）を説明変数としてロジスティック回帰を行うものである。観光特性変数は、具体的には以下のようなものである。

表 4. 事業所観光特性変数リスト

変数名	内容
観光協会	当該事業所が観光協会に加盟している場合に1をとるダミー変数。
地域の観光パンフレット	当該事業所が地域の観光パンフレットに掲載されている場合に1をとるダミー変数。
観光ガイドブック	当該事業所が観光ガイドブックに掲載されている場合に1をとるダミー変数。
クレジットカード	当該事業所においてクレジットカードを使用できる場合に1をとるダミー変数。
大型バス用駐車場	当該事業所に大型バス用の駐車場がある場合に1をとるダミー変数。
自社ホームページ	当該事業所が自社ホームページを持っている場合に1をとるダミー変数。
観光協会ホームページ	当該事業所が地域の観光協会のホームページに掲載されている場合に1をとるダミー変数。
旅行会社ホームページ	当該事業所が旅行会社のホームページに掲載されている場合に1をとるダミー変数。

これらの観光特性変数に関して、期待されるパラメータの符号はすべて(+)である。

3.4. 分析結果

推定シミュレーションの結果を明らかにする前に、観光地域経済調査で収集された全事業所のデータを用いて、Method 5 および 6 で用いたロジスティック回帰を行った際の推定結果を示す。分析対象は、静岡県内の事業所であり、小売業および飲食サービス業のそれぞれについて被説明変数を観光向け売上割合とした推定を行っている。表 5 が推定結果である。

表 5. 静岡県内全事業所推定結果

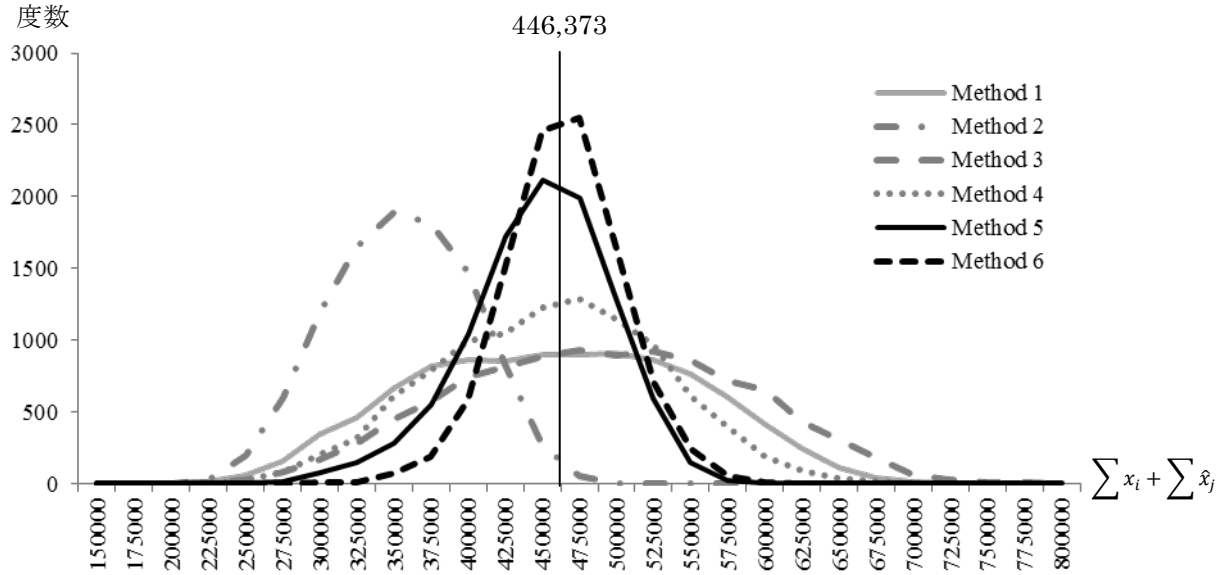
部門	飲食サービス業		小売業	
	Method 5	Method 6	Method 5	Method 6
観光地点	0.145 (0.000) **	0.063 (0.000) **	0.019 (0.000) **	0.022 (0.000) **
新幹線駅	0.036 (0.000) **	0.054 (0.000) **	0.063 (0.000) **	0.061 (0.000) **
鉄道駅	0.112 (0.001) **	0.068 (0.001) **	0.078 (0.000) **	0.045 (0.000) **
高速道路出口	-0.069 (0.001) **	-0.017 (0.001) **	-0.208 (0.001) **	-0.141 (0.001) **
宿泊施設	0.130 (0.002) **	0.167 (0.002) **	0.696 (0.002) **	0.719 (0.002) **
観光協会		0.144 (0.007) **		0.976 (0.004) **
観光パンフレット		0.290 (0.007) **		0.411 (0.006) **
観光ガイドブック		0.145 (0.006) **		2.045 (0.005) **
クレジットカード		0.248 (0.005) **		0.681 (0.004) **
大型バス用駐車場		1.704 (0.006) **		
自社 HP		0.182 (0.006) **		
観光協会 HP		0.440 (0.007) **		
旅行会社 HP		0.648 (0.007) **		
自由度	668	660	743	739
尤度比インデックス	0.385	0.605	0.436	0.577

1. カッコ内は標準誤差を表す。
2. * および** は、それぞれ有意水準 5% および 1% で有意であることを示している。
3. 定数項を含めて推定を行っているが、表中には結果を表示していない。

この結果を見れば、Method 5 では、飲食サービス部門と小売業の両部門について全ての立地特性変数について有意な結果が得られている。ただし「新幹線駅」と「鉄道駅」に関する予想されるパラメータの符号はマイナスであったのに対し、推定結果は両部門についてプラスとなった。これは、駅から近い事業所ほど観光向け売上比率が低いことを示唆するものである。駅の利用者は観光客だけでなく、むしろ地域の居住者の多くが駅およびその付近の飲食店や小売店を利用しているため、このような結果が得られたと解釈することもできる。一方で Method 6 については、全ての立地特性変数について有意な結果が得られたものの、小売業では大型バス用駐車場やホームページに関する観光特性変数が有意ではなかったため、それらを除いた推定結果を掲載している。しかし、有意であった観光特性変数のパラメータの符号については、全て予想と同じ結果となっている。以上の結果から、ここで用いている立地特性変数や観光特性変数が、事業所の観光向け売上げに影響することは明らかとなった。実際に行ったシミュレーションでは、観光地域調査から無作為抽出

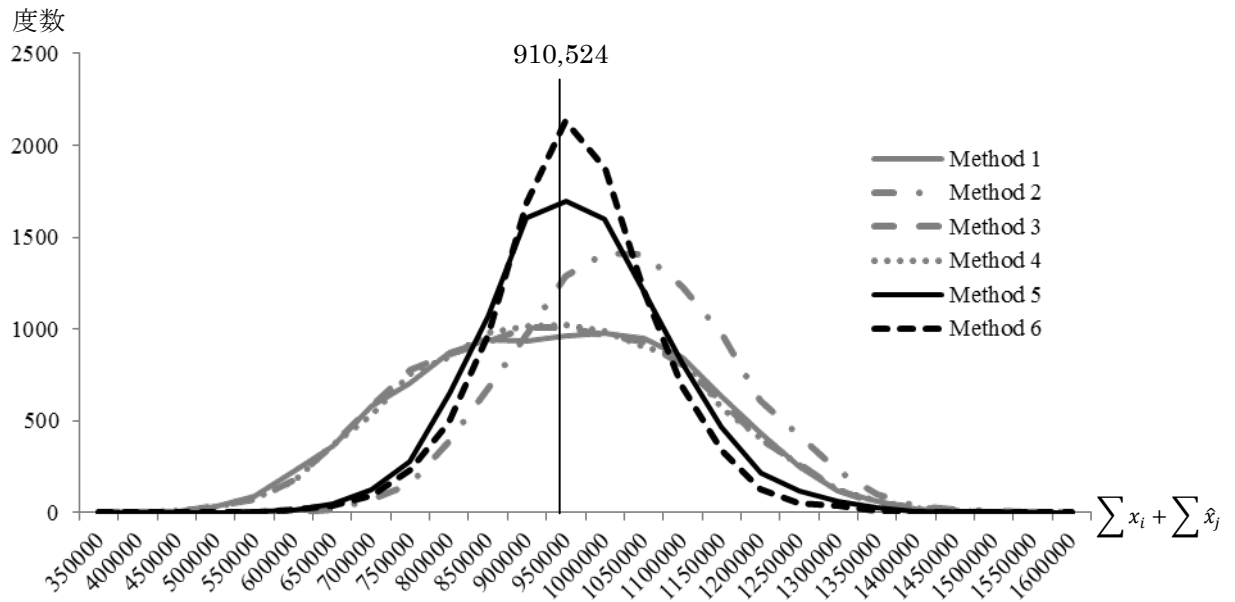
した事業所のデータを用いて、表 5 のモデルを推定している。

図 4 および図 5 は、図 3. 推定シミュレーションのフロー図 3 で示した推定シミュレーションを各手法について 1 万回行ったうえで、その分布を描いたものである。



(単位：10,000 円)

図 4. 飲食サービス業シミュレーション結果



(単位：10,000 円)

図 5. 小売業シミュレーション結果

図 4 および図 5 では、横軸が推定された観光売上総額を、縦軸が度数を表している。各手法について 1 万回の繰り返しシミュレーションを行っているため、手法ごとの度数の合計はいずれも 1 万である。図 4 の 446,373 および図 5 の 910,524 は、それぞれ飲食サービス業と小売業の観光地域経済調査における全事業所の観光向け売上総額であり、これが推定のターゲットとなる金額である。推定シミュレーションの結果がこの金額に近ければ近いほど、その推定結果は望ましいものであるということになる。

図 4 についていえば、Method 1、3、4 の結果は、類似した分布の形状を描いている。理論的な観点からいえば、Method 1 による推定値は不偏性を持ち、Method 3、4 の推定値は漸近的不偏性を持つ。図 4 においても、Method 1、3、4 の分布の中心は、ほぼターゲットとなる金額 (446,373) の近くにあることがわかる。Method 2 は、分布の中心が左方向に大きく偏った形状となっている。Method 2 では、1 事業所あたり平均観光向け売上割合を用いた推定を行っているが、各事業所の観光向け売上割合と売上総額の間に関連がある場合には、このような偏りが発生することになる¹¹。Method 1~4 に比較して、Method 5 および 6 の分布の分散は明らかに小さい。理論的には Method 5 および 6 の推定値も漸近的不偏性を持ち、分布の中心もほぼターゲット金額近傍にあることがわかる。この結果は、事業所の立地特性変数を用いて推定を行った Method 5 および 6 の手法が、それらを用いない Method 1~4 に比較して精度の高い推定法であることを示唆している。

図 5 の小売業の結果も図 4 とほぼ同様であり、Method 1~4 に比較して、Method 5 および 6 が明らかに望ましい推定手法であることがわかる。

表 6 は、各手法の推定値とターゲットとなる観光売上総額 (飲食サービス業 446,373、小売業 910,524) の差の絶対値を計算し、その 1 万回分の平均値を計算したものである。

表 6. 推定値と現実値の差の絶対値平均

	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6
飲食サービス業	76,816	104,255	81,852	61,160	38,266	30,569
小売業	145,964	131,396	143,380	141,946	96,116	82,174

値が小さいほど、平均的な推定の誤差が小さいため、より精度の高い推定手法であると判断することができる。この結果からも、Method 1~4 に比較して Method 5 および 6 がより望ましい手法であることがわかる。また、飲食サービス業と小売業の両部門ともに、Method 5 よりも Method 6 の値が小さいことから、立地特性変数だけでなく、観光特性変数も合わせて使用することによって、推定の精度がより一層高まることも明らかにされて

¹¹ Method 2 では、全ての非サンプル事業所について単一の観光向け売上割合を仮定することになるため、例えば、実際には観光以外の売上も含む売上額の大きな事業所ほど観光売上げ割合も高いようなケースでは、図 4 のように左方向に偏った過小推定となる。実際に飲食サービス業について売上額と観光売上げ割合の相関係数を求めたところ、その値は 0.181 であった。

いる。この点は、観光地域経済調査のような観光統計において、観光特性を調査することの重要性を示唆しているといえる。

以上の結果より、住所情報のみから求めることができる立地特性変数を使用することによって、地域の観光向け売上額の推定精度は大きく高まることが明らかとなった。これは、ビジネスレジスターの1つの有効な活用方法を示す事例であるといえよう。

4. 本稿のまとめとビジネスレジスターの活用に関する今後の展望

本稿では、ビジネスレジスターの観光統計への活用事例として、Nurmi (2014)およびMiyagawa et al. (2014)を紹介した。

Nurmi (2014)では、フィンランドのビジネスレジスターを利用して、宿泊統計における小規模な宿泊施設の延べ宿泊者数を推定する手法が示された。このような小規模事業所の推定は、観光統計に限るものではなく、経済センサス等の事業所を対象とした様々な統計調査にも応用できる可能性がある。小規模事業所については、一般的に事業所数が多く、統計調査への回答率も大規模事業所に比較して低いことから、調査コストや調査の精度という面からいえば、ビジネスレジスターの活用によって小規模事業所を対象とした統計調査を省略することができれば、そのメリットは大きいものであろう。

ただしフィンランドの事例において、小規模宿泊施設の推定を行うことができたこと背景には、2つのポイントがある。1つは、フィンランドのビジネスレジスターにおいて、各事業所の売上総額に関する情報が収集されていたことであり、もう1つは事業所の詳細な産業分類が把握されていたことである。Nurmi (2014)においては、ビジネスレジスターより得られる事業所の売上額を、NACEのクラス別の「1泊あたり売上額」で割ることによって延べ宿泊者数を求めていた。これは、「1泊あたり売上額」のレベルが、NACEで言えば同じ「55. 宿泊業」であっても、更に詳細なNACEクラスごとに大きく異なっていたためである。これらの結果は、ビジネスレジスターの活用には、事業所の売上額と産業格付けの情報が重要や役割を果たすことを物語るものであった。

日本においても、「平成26年経済センサス-基礎調査」では事業所の年間総売上（収入）金額に関する調査が行われ、その結果は「事業所母集団データベース」に収録されることになる。今後は、「事業所母集団データベース」を利用した小規模事業所の推定に関しても、その可能性を探ってゆく必要がある。

一方でMiyagawa et al. (2014)では、ビジネスレジスターに収録される事業所の住所情報に基づいて、事業所の立地特性に関する変数を導出し、それを利用することによって観光売上額の推定精度が高まることが示された。近年のGISの発展によって、地図から得られる情報は膨大なものであり、それを有効に活用することができれば、観光に限らず様々な分野の統計調査に関してより精度の高い推定を行うことが可能になるであろう。

言うまでもなく、このような事業所立地特性を利用した推定を行うためには、各事業所

の緯度・経度情報がビジネスレジスター上で正確に把握されていることが必要不可欠である。今後は、日本の「事業所母集団データベース」においても、全ての事業所について正確な緯度・経度情報を収集することが望まれる。

参考文献

Dobson, Annette J. and Adrian Barnett (2008) *An Introduction to Generalized Linear Models*, Third Edition, Chapman&Hall.

Eurostat (2008) *NACE Rev. 2 Statistical classification of economic activities in the European Community*, Eurostat Methodologies and Working papers.

Eurostat (2014) *Methodological manual for tourism statistics version 3.1, 2014 edition*, Eurostat manuals and guidelines.

Eurostat-OECD (2007) *Eurostat-OECD Manual on Business Demography Statistics*, European Communities/OECD.

Jarmin, Ron and Javier Miranda (2006) “The Impact of Hurricanes Katrina, Rita, and Wilma on Business Establishments: A GIS Approach”, Center for Economic Studies Discussion Paper CES 06-23.

Miyagawa, Kozo, Hiroyuki Kamiyama, Ryuta Shimamura, Fumikado Yamamoto (2014) “Estimating establishment-level tourism sales using the Regional Tourism Economic Survey and Geographical Information System”, presented in the 12th edition of the Global Forum on Tourism Statistics, <http://tsf2014prague.cz/programme/15-may>.

Nurmi, Ossi (2014) “Estimating Nights Spent in Finnish Accommodation Establishments Based on Business Registers” presented in the 12th edition of the Global Forum on Tourism Statistics, <http://tsf2014prague.cz/programme/15-may>.

宮川幸三 (2009) 「我が国の観光統計をめぐる現状と課題－地域観光統計体系の整備に向けて－」『産業連関』 Vol.17, No.1,2.