

Discussion Paper No.425

国勢調査と社会生活基本調査を用いたデータリンケージの  
取り組み状況について

中央大学経済学部

伊藤 伸介

(公財)統計情報研究開発センター

村田 磨理子

(公財)統計情報研究開発センター

滝澤 有美

熊本学園大学経済学部

林田 実

April 2026



INSTITUTE OF ECONOMIC RESEARCH  
Chuo University  
Tokyo, Japan

# 国勢調査と社会生活基本調査を用いたデータリンケージの

## 取り組み状況について

中央大学経済学部 伊藤 伸介

(公財)統計情報研究開発センター 村田 磨理子

(公財)統計情報研究開発センター 滝澤 有美

熊本学園大学経済学部 林田 実

### 1. はじめに

海外では、学術研究目的のために、公的統計マイクロデータだけでなく、行政記録情報の個票データを含む公的大規模データの利活用が進められてきた。さらには、デンマーク等の北欧諸国やスコットランドのように、各種の社会人口的属性や経済的属性を包含する公的大規模データと医療・健康・福祉等の他分野のマイクロデータの連結(リンケージ、linkage)がなされることによって、リンケージデータの広範な利活用が展開されている事例が存在する(伊藤(2025a))。

世帯・人口系の公的大規模データを例にすると、統計作成部局やデータ提供機関によるデータリンケージは、①個人に備わっている共通 ID (個体識別番号) に基づいて、公的大規模データ間のリンケージを行うことが可能な場合、及び②リンケージ用の ID を生成することによって、公的大規模データのリンケージを行うことができる場合に類別される(伊藤(2020), 伊藤(2024), 伊藤(2025a), 伊藤(2025b))。わが国では、各種の公的大規模データを連結するための ID は存在しないが、海外におけるデータリンケージの展開状況を踏まえると、各種の公的大規模データに共通に付与される ID の整備と ID に基づくデータリンケージ技法の追究は、わが国の公的大規模データにおける利活用のさらなる可能性を図る上で検討すべき課題と言える<sup>1</sup>。そのための基礎となる実証研究として、わが国でも、利用

---

<sup>1</sup> 2025年6月に、内閣官房デジタル行財政改革会議より出された『データ利活用制度の在り方に関する基本方針』によれば、「4. 行政保有データの利活用」の「(2) 分野間におけるデータ連携の推進、識別子」の中で、「プライバシーを保護しつつ分野間のデータ連携・解析を進めるための枠組みの構築を目指し、関係府省庁の協力も得ながら、仮名化の手法、連携用の識別子の在り方、利活用環境の整備、利活用主体の資格に関する枠組みなど、諸外国の先進的な取組等について調査研究を行い、その結果を踏まえた所要の措置を講ずる」(p.13)ことが記されており、識別子を用いて各種のデータを連携することの重要性が明示されている(伊藤(2025b, p.31))。

可能な公的統計の調査票情報を用いて、データリンケージの有効性に関する検証を行うことが求められる。

本稿の目的は、わが国の公的大規模データを主たる研究対象とした上で、リンケージされた公的大規模データの利用可能性を方法的に模索することである。そのため本研究では、公的統計マイクロデータを中心としたデータリンケージの方法論を追究することを指向している。また、本研究においては、複数のデータセットに共通の直接識別子や ID を用いたレコード同士の完全照合(exact matching)とは異なる、複数の変数の組み合わせによる統計的照合(statistical matching)の技法を適用することによって、データリンケージの可能性を探究する。

## 2. わが国の公的統計マイクロデータを用いたリンケージの可能性

わが国における事業所・企業系の統計調査間のデータリンケージについては、企業の名称や住所、事業所母集団データベース等で把握可能な事業所番号、さらには法人番号といった識別子を用いた完全照合を行うことが可能である<sup>2</sup>。その一方で、前節で述べたとおり、国勢調査のような世帯・人口系の統計調査を用いてデータリンケージを行う場合に、わが国では共通 ID 等の個体識別子を利用したリンケージは困難である。そのため、世帯・人口系のデータ間に共通する属性群を用いて、統計的照合を行う必要がある。具体的な属性指標としては、都道府県番号、市区町村番号、性別、出生の年月等が考えられる。

本研究は、異種の世帯・人口系の統計調査の調査票情報(個票データ)同士を用いたデータリンケージの実証実験を行うことを指向している。対象となる統計調査は、全数調査である国勢調査と標本調査である社会生活基本調査である。本研究では、照合のための属性指標として、都道府県番号や市区町村番号だけでなく、調査区ごとに照合を行うためのキーを生成した上で、そのようなキーを使用した場合と使用しなかった場合のそれぞれの照合結果を比較することによって、データリンケージの精度検証を行う。

なお、わが国では、2025年3月末よりリモートアクセス<sup>3</sup>の運用が開始されたことから、

---

<sup>2</sup> 例えば、村田・伊藤(2016)は、経済センサス及び事業所・企業統計調査で把握可能な事業所番号を介して、賃金構造基本統計調査と経済産業省企業活動基本調査の個票データによる完全照合と統計的照合を行った上で、それぞれの照合技法で作成されたリンケージデータの精度に関する比較・検証を行っている。

<sup>3</sup> リモートアクセス (remote access) とは、大学の研究室のような研究を行う場所で使用する端末から個票データが保管されているサーバにリモートでアクセスすることを可能にする安全な分析環境を表す(伊藤(2020))。近年海外では、統計作成部局、大学やリサーチデータセンター内に設置される、物理的に安全な分析環境(オンサイト施設)での個票データの利活用を停止して、個票データをリモートアクセスでのみ提供する動きが出てきている(オランダ、オーストラリア等)(伊藤(2024))。一方、わが国においては、オンサイト施設での調査票情報の利用だけでなく、研究室や自宅等の利用者自身の拠点から、インターネ

本研究でのデータリンケージに関する実証実験は、リモートアクセスのサーバ上で実施されている。

本研究では、世帯・人口系の統計調査のマイクロデータに関するリンケージについては、以下の手順で行う(図 1)。

- ① 国勢調査と他の標本調査(ex.就業構造基本調査)を対象に、調査区番号(及び調査区符号)と調査区に関する対応表(調査区対応表)から、図 1 で示されるリンケージ用の「共通 ID」(以下「調査区マッチングキー」と呼称)をオンサイト施設の安全な分析環境において生成する。なお、オンサイト環境で格納されている調査区対応表は、総務省統計局から本研究のために、特別に提供されたものである。
- ② 図 1 に示されているように、「持ち出し表(1) 就業構造基本調査」と「持ち出し表(2) 国勢調査」のおおのの、元の調査区対応表を 2 つに分割して編成された調査区マッチングキー、及び両調査の集計区分と調査年に関わる情報(調査 ID)が連結された「接続情報」をそれぞれ設定した上で、リモートアクセスのサーバに移管する。
- ③ リモートアクセスのサーバ上で、調査区マッチングキーも用いた上で、国勢調査と他の標本調査(図 1 では就業構造基本調査)における個票データ間のリンケージを行う。この場合、市区町村レベルの地域情報を用いたが、調査区マッチングキーを使用しなかった場合の照合と、調査区マッチングキーも用いて行った照合では、精度が大きく異なることが推察される。
- ④ リモートアクセスのサーバ上で、③の作業によって作成されたリンケージデータを用いて実証分析を行う。なお、リモートアクセスで行った分析による成果物については、必要に応じてサーバの外に持ち出すことが可能になっている。

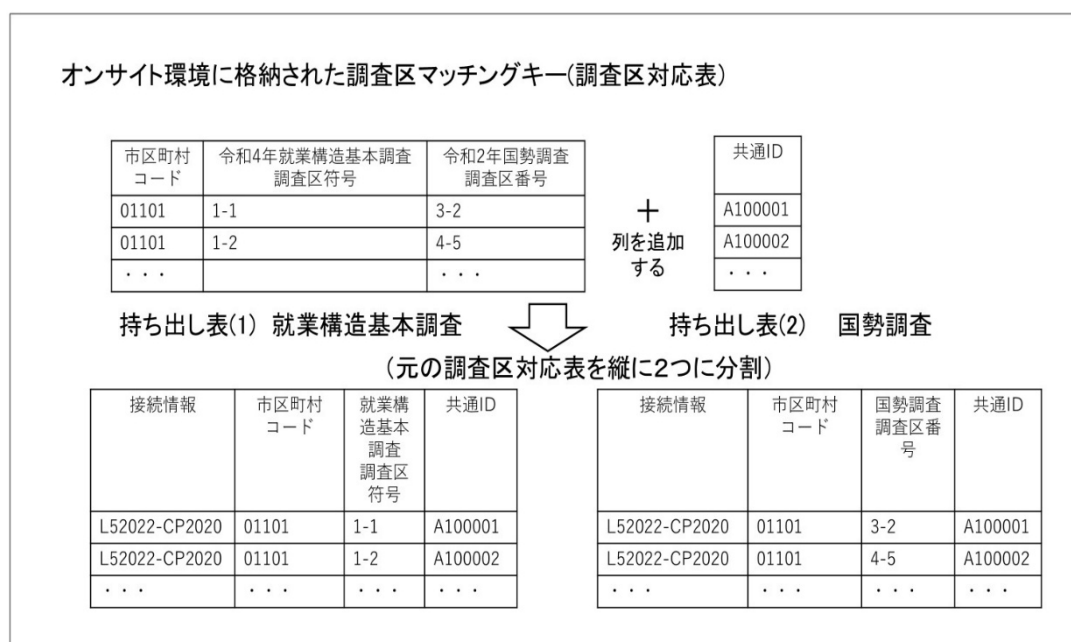
---

ット回線を経由して調査票情報を格納したサーバにアクセスし、サーバ上の仮想 PC を遠隔操作して調査票情報を利用することができる。

わが国でリモートアクセスを推進するにあたっては、提供者側と利用者側の信頼関係についてどのような形で制度設計を行うかが重要だと考える。具体的には、利用者の調査票情報や匿名データの利用経験の有無が、リモートアクセスにおける個票データの適正管理義務と安全管理措置の観点から重要であって、その意味では、オンサイト施設での調査票情報の利用と比較して、わが国のリモートアクセスにおける調査票情報の利用においては、安全な分析環境を担保する上で「自己管理・自己責任が原則」となっていることに留意する必要がある。その理由は、リモートアクセスの調査票情報の提供が、統計法 33 条における磁気媒体の提供に準じているからであり、海外とは異なり、リモートアクセスのサーバ上から分析結果を持ち出す場合の統計作成部局によるチェックが求められないからである。

また、利用者側の研究グループのメンバー間における信頼関係の担保についても、慎重な議論が求められる。共同研究では、もしメンバーの 1 人でも不適切利用があった場合、同一グループ内のメンバーにも共同責任が問われる可能性があることから、メンバーのおおのの個票データの利用については、「自己責任」による対応が求められつつも、不適切利用があった場合には、メンバー全員の「連帯責任」が問われる。このことは、実際のリモートアクセスの運用においては、自己責任と連帯責任の両方が求められることを意味している。

図1 調査区対応表作成の方法(イメージ)



注 調査区対応表に合併・分割情報が記載されている場合、必要に応じ、持ち出し表に適宜追加する。

### 3. 使用するデータの特性とリンケージデータの作成手順

本研究で使用するデータは、令和2年国勢調査(以下「国勢調査」と呼称)<sup>4</sup>と令和3年社会生活基本調査(調査票A)(生活行動編)(以下「社会生活基本調査」と呼称)<sup>5</sup>の調査票情

<sup>4</sup> 令和2年国勢調査は、令和2年10月1日現在でわが国の地域において実施され、本邦内に常住している者を調査対象としている。

<sup>5</sup> 令和3年社会生活基本調査は、令和3年10月20日現在で実施され、10歳以上の世帯員を集計対象としている。標本の抽出方法は、第1次抽出単位を平成27年国勢調査調査区とし、第2次抽出単位を世帯とする層化2段抽出法によって行い、次のものを除く国勢調査の調査区から抽出している。

- a 山岳・森林・原野地帯等の調査区(後置番号:2)
- b 相当規模の工場・学校等のある調査区(後置番号:3)
- c 社会施設・相当規模の病院のある調査区(後置番号:4)
- d 刑務所・拘留所等のある調査区(後置番号:5)
- e 自衛隊地域の調査区(後置番号:6)
- f 駐留軍地域の調査区(後置番号:7)
- g 水面調査区(後置番号:9)

報(個票データ)、及びこれらに含まれる調査区番号(調査区符号)と調査区対応表である。

本研究で用いる世帯・人口系のデータについては、リモートアクセスの環境の下で検証作業を行っている。なお、総務省統計局から提供される調査区対応表は、オンサイト施設に格納されており、調査区対応表を加工して調査区マッチングキーを生成した後に、リモートアクセス環境に転送される。なお、個票データの処理には SAS が使用されている。

つぎに、データリンケージの作業については、前節で示した手順を踏まえて、以下のとおり実施した(図2)。

- ① オンサイト施設に格納されている国勢調査と社会生活基本調査の調査区対応表に基づいて編成された調査区マッチングキーを付与した上で、国勢調査と社会生活基本調査の持ち出し表(以下「加工済み調査対応表」と呼称)をそれぞれ作成する。
- ② 2種類の加工済み調査区対応表に両調査の調査IDを連結した「接続情報」を設定した上で、リモートアクセス環境に移管する。
- ③ リモートアクセス環境で、(1)市区町村レベルまでの地域情報を用いた統計的照合と(2)加工済み調査区対応表に含まれる調査区マッチングキーも用いた統計的照合を行った上で、これらにおける照合の精度について比較・検証する。

なお、データリンケージに用いた個票データの基本統計量は、表1と表2で示されている。表1は、国勢調査の公表された統計表と社会生活基本調査の個票データのそれぞれを用いて算出した都道府県別の世帯員数を示している。なお、社会生活基本調査を用いて算出された世帯員数は個票データに含まれるレコード数に該当しており、集計用乗率は考慮して算定されていないことに留意されたい。

つぎに表2は、年齢5歳階級別の世帯員数を示したものである。表2を見ると、社会生活基本調査の個票データで算出された世帯員数は、国勢調査における世帯員数の結果数値と比較して、20~34歳の年齢階層における比率が低く、60~74歳の年齢階層での比率が相対的に高くなっていることが確認される。

図2 調査区対応表の加工方法

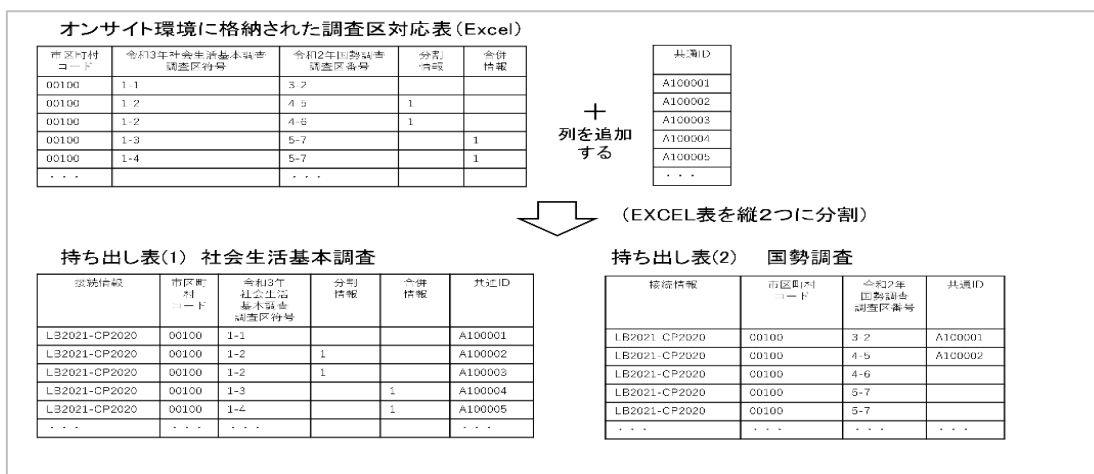


表1 令和2年国勢調査及び令和3年社会生活基本調査における世帯員数の比較  
—全国、都道府県

		令和2年国勢調査 世帯員数		令和3年社会生活基本調査 調査票A：生活行動編 世帯員数 (※調査対象：10歳以上)		
			うち、10歳以上			
			世帯員数	世帯員割合		
00	全国	126,146,099	113,609,086	90.1%	171,360	
01	北海道	5,224,614	4,813,698	92.1%	5,175	
02	青森県	1,237,984	1,137,368	91.9%	3,243	
03	岩手県	1,210,534	1,113,187	92.0%	3,418	
04	宮城県	2,301,996	2,083,037	90.5%	3,399	
05	秋田県	959,502	894,353	93.2%	3,515	
06	山形県	1,068,027	983,458	92.1%	3,665	
07	福島県	1,833,152	1,668,919	91.0%	3,601	
08	茨城県	2,867,009	2,601,915	90.8%	3,451	
09	栃木県	1,933,146	1,754,220	90.7%	3,523	
10	群馬県	1,939,110	1,757,296	90.6%	3,587	
11	埼玉県	7,344,765	6,577,758	89.6%	6,153	
12	千葉県	6,284,480	5,680,327	90.4%	4,670	
13	東京都	14,047,594	12,565,969	89.5%	3,496	
14	神奈川県	9,237,337	8,323,177	90.1%	5,273	
15	新潟県	2,201,272	2,018,581	91.7%	3,606	
16	富山県	1,034,814	943,875	91.2%	3,584	
17	石川県	1,132,526	1,024,769	90.5%	3,365	
18	福井県	766,863	695,971	90.8%	3,538	
19	山梨県	809,974	733,436	90.6%	3,216	
20	長野県	2,048,011	1,856,050	90.6%	3,569	
21	岐阜県	1,978,742	1,788,623	90.4%	3,764	
22	静岡県	3,633,202	3,317,633	91.3%	3,822	
23	愛知県	7,542,415	6,710,275	89.0%	6,028	
24	三重県	1,770,254	1,597,253	90.2%	3,524	
25	滋賀県	1,413,610	1,258,876	89.1%	3,445	
26	京都府	2,578,087	2,308,079	89.5%	2,845	
27	大阪府	8,837,685	7,927,372	89.7%	4,520	
28	兵庫県	5,465,002	4,861,102	88.9%	4,512	
29	奈良県	1,324,473	1,209,704	91.3%	3,305	
30	和歌山県	922,584	848,036	91.9%	3,214	
31	鳥取県	553,407	501,378	90.6%	3,444	
32	島根県	671,126	609,678	90.8%	3,420	
33	岡山県	1,888,432	1,672,463	88.6%	3,284	
34	広島県	2,799,702	2,517,921	89.9%	3,406	
35	山口県	1,342,059	1,226,530	91.4%	3,146	
36	徳島県	719,559	642,050	89.2%	3,088	
37	香川県	950,244	855,945	90.1%	3,143	
38	愛媛県	1,334,841	1,202,959	90.1%	3,347	
39	高知県	691,527	631,591	91.3%	2,837	
40	福岡県	5,135,214	4,536,417	88.3%	4,383	
41	佐賀県	811,442	728,888	89.8%	3,427	
42	長崎県	1,312,317	1,195,323	91.1%	3,096	
43	熊本県	1,738,301	1,565,518	90.1%	3,420	
44	大分県	1,123,852	1,014,903	90.3%	3,100	
45	宮崎県	1,069,576	963,333	90.1%	3,084	
46	鹿児島県	1,588,256	1,411,515	88.9%	3,002	
47	沖縄県	1,467,480	1,278,357	87.1%	2,707	

表2 年齢5歳階級別世帯員数—令和2年国勢調査（10歳以上）と令和3年社会生活基本調査（調査票A）<sup>6</sup>

	令和2年国勢調査 10歳以上		令和3年社会生活基本調査 調査票A：生活行動編 (※調査対象：10歳以上)		(b)-(a) ポイント差
	世帯員数	割合(a)	世帯員数	割合(b)	
10-14歳	5,350,517	4.7%	8,574	5.0%	0.3
15-19歳	5,617,440	4.9%	8,342	4.9%	-0.1
20-24歳	5,931,306	5.2%	6,316	3.7%	-1.5
25-29歳	6,031,964	5.3%	6,039	3.5%	-1.8
30-34歳	6,484,594	5.7%	7,420	4.3%	-1.4
35-39歳	7,311,567	6.4%	9,769	5.7%	-0.7
40-44歳	8,291,077	7.3%	11,569	6.8%	-0.5
45-49歳	9,650,293	8.5%	13,984	8.2%	-0.3
50-54歳	8,539,851	7.5%	13,497	7.9%	0.4
55-59歳	7,767,482	6.8%	12,244	7.1%	0.3
60-64歳	7,297,190	6.4%	13,194	7.7%	1.3
65-69歳	8,075,268	7.1%	14,893	8.7%	1.6
70-74歳	9,011,795	7.9%	17,475	10.2%	2.3
75-79歳	6,930,928	6.1%	11,217	6.5%	0.4
80-84歳	5,296,728	4.7%	8,731	5.1%	0.4
85-89歳	3,669,823	3.2%	5,375	3.1%	-0.1
90歳以上	2,351,263	2.1%	2,721	1.6%	-0.5
合計	113,609,086	100.0%	171,360	100.0%	

#### 4. 国勢調査と社会生活基本調査を用いたデータリンケージの実証実験の概要

##### 験の概要

本節では、国勢調査と社会生活基本調査の個票データを用いて世帯員レベルでのデータリンケージを行った上で、照合のために使用したキーの違いによるマッチング率の変化を検証する。

本研究では、データリンケージの検証に用いる地域として、四国ブロック（36 徳島県、37 香川県、38 愛媛県、39 高知県）を選定した。その理由としては、第1に人口規模が比較的小さいこと、第2に令和2年国勢調査が実施された後、令和3年社会生活基本調査が実

<sup>6</sup> 令和2年国勢調査の割合は、10歳以上の世帯員合計を100.0%として算出した。

施されるまでの間に市区町村の廃置分合がなく、これに伴う調査区の大幅な組み換えが発生しないことが指摘される。以下では、国勢調査と社会生活基本調査を用いたデータリンクの実証実験の概要について述べる。

実証実験においては、マッチングキーの選定が求められる。本実験では、調査区マッチングキーだけでなく、国勢調査と社会生活基本調査で共通する調査項目である①市区町村コード、②男女の別、③出生の年、④出生の月、⑤配偶者の有無と⑥教育をマッチングキーとして用いる。なお、各マッチングキーにおける分類区分の内容、及びその調整方法については後述する。

本実証実験の詳細は、以下のとおりである。

本研究では、2種類の統計的照合を行う。第1は、調査区マッチングキーを除く、市区町村コード等の6つのマッチングキーによる照合である。照合の結果に関しては、社会生活基本調査の特定のレコードに対して、①令和2年国勢調査の個票データの中に照合するレコードが1つも存在しない場合、②令和2年国勢調査の個票データの中で照合するレコードと1対1で対応付けられる場合、③令和2年国勢調査の個票データの中で照合するレコードが複数存在する場合の3つのパターンが該当する。これらのパターンについて、全レコードに対する比率が比較・検証される。

なお、本実証実験では、完全照合で使用される共通ID等の個体識別子是用いられないことから、1対1で対応付けられたレコードが、「正しく」照合されたかどうかは、本実験における検証の対象外となっていることに留意されたい。

第2は、市区町村コードを含む6つのマッチングキーに加えて、調査区マッチングキーも使用した照合である。これについても同様に、上記の3種類の照合のパターンに関して国勢調査の個票データに含まれるレコード数に対する比率を比較した。さらに、マッチングキーである教育がマッチング率に与える影響と、調査区単位での照合の状況を確認した上で、各マッチングキーにおける分布特性の確認を行った。

つぎに、市区町村コードを用いるが、調査区マッチングキーを使用しない照合と、調査区マッチングキーも用いて行った照合について、それぞれの手順を詳述する。

前者の照合に関しては、マッチングキーとして、先述の通り、市区町村コード、男女の別、出生の年、出生の月、配偶者の有無と教育が使用される。マッチングキーの作成にあたっては、2つの調査において分類区分が異なる出生の年、配偶者の有無と教育に関して、以下のとおり区分の変更を行った。

出生の年については、和暦表記と西暦表記が併存しているだけでなく、令和2年国勢調査と令和3年社会生活基本調査では和暦に関する分類区分が異なっていることを踏まえ、本研究では、マッチングキーの調整と年齢5歳階級への対応を勘案し、西暦の表記に統一した(表3)。

つぎに、配偶者の有無に関しては、表4で示されるとおり、国勢調査の「死別」及び「離別」を社会生活基本調査における分類区分に合わせて「死別・離別」に統一した。

表3 出生の年における分類区分の調整方法

(1)令和2年国勢調査「出生の年月 元号・西暦」の分類区分

分類区分
明治
大正
昭和
平成
令和
西暦
不詳（基本項目記入不備世帯のみ）

(2)令和3年社会生活基本調査「出生元号」の分類区分

分類区分
明治
大正
昭和
平成
西暦

(3)マッチングキーの調整方法

分類区分	マッチングキーの調整方法
明治	出生の年に1867を加算
大正	出生の年に1911を加算
昭和	出生の年に1925を加算
平成	出生の年に1988を加算
西暦	調整なし

表4 「配偶者の有無」のマッチングキーの調整方法

(1)令和2年国勢調査「配偶者の有無」の分類区分

分類区分
未婚（幼児などを含む）
配偶者あり
死別
離別
不詳

(2)令和3年社会生活基本調査「配偶者の有無」の分類区分

分類区分
未婚
配偶者あり
死別・離別
不詳

### (3) マッチングキーの調整方法

令和2年国勢調査 配偶者の有無		調整後の分類区分		令和3年社会生活基本調査 配偶者の有無
未婚（幼児などを含む）	→	未婚	←	未婚
配偶者あり	→	配偶者あり	←	配偶者あり
死別	→	死別・離別	←	死別・離別
離別				
不詳	→	不詳	←	不詳

教育については、社会生活基本調査における「専門学校」の3分類及び「短大・高専」を国勢調査に合わせて「短大・高専」に変更した(表5)。さらに、国勢調査の調査年である令和2年(2020年)と社会生活基本調査の調査年である令和3年(2021年)で進学により学校の種類が異なると推定される世帯員(2021年度末で19歳、16歳、13歳)を対象に、令和3年社会生活基本調査の「教育(学校の種類)」における分類区分を令和2年国勢調査の「教育状態」のそれに合わせる調整を行った。

表5 「教育」のマッチングキーの調整方法

(1)令和2年国勢調査「教育状態\_19区分」における分類区分<sup>7</sup>

分類区分
卒業者 小学校
卒業者 中学校
卒業者 高校・旧中
卒業者 短大・高専
卒業者 大学
卒業者 大学院
卒業者 不詳
在学者 小学校
在学者 中学校
在学者 高校・旧中
在学者 短大・高専
在学者 大学
在学者 大学院
在学者 不詳
未就学者 幼稚園
未就学者 保育園・保育所
未就学者 認定こども園
未就学者 その他
未就学者 不詳
就学の有無不詳(世帯の種類が不詳の世帯を含む)
上記以外

<sup>7</sup> 国勢調査における分類区分の「未就学者」については、6歳以下である子どもが対象となっているため、それは10歳以上を調査対象とする社会生活基本調査のデータとはリンクしない。また、分類区分の「上記以外」には、「在学したことがない」者が該当する。

(2)令和3年社会生活基本調査「教育（学校の種類）」の分類区分<sup>8</sup>

分類区分
小学
中学
高校・旧制中
専門学校 1年以上2年未満
専門学校 2年以上4年未満
専門学校 4年以上
短大・高専
大学
大学院
不詳
対象外

(3)マッチングキーの調整方法<sup>9</sup>

令和2年国勢調査 教育状態_19区分		調整後の 分類区分		令和3年社会生活基本調査 5 教育（学校の種類）
小学校	→	小学	←	小学
中学校	→	中学	←	中学
高校・旧中	→	高校・旧制中	←	高校・旧制中
短大・高専	→	短大・高専	←	専門学校 1年以上2年未満
				専門学校 2年以上4年未満
				専門学校 4年以上
				短大・高専
大学	→	大学	←	大学
大学院	→	大学院	←	大学院
不詳	→	不詳	←	不詳

(4)学校の種類が変わる年齢の世帯員の「教育」の調整方法

令和3年（2021年）社会生活基本調査の出生年月	学校の種類の調整 （令和2年（2020年）国勢調査時点）
2002年4月～2003年3月（2021年度末時点で19歳）	高校・旧中
2005年4月～2006年3月（2021年度末時点で16歳）	中学校
2008年4月～2009年3月（2021年度末時点で13歳）	小学校

つぎに、後者の調査区マッチングキーも使用した場合の照合の方法について述べる。前述の通り、本研究では、マッチングキーとして、市区町村コード、男女の別、出生の年、出生の月、配偶者の有無と教育に加えて、調査区(調査区マッチングキー)を用いた照合を行う<sup>10</sup>。なお、調査区マッチングキーについては、加工済み調査区対応表に含まれる調査項目に基づ

<sup>8</sup> 分類区分の「対象外」については、「在学したことがない」者が該当する。

<sup>9</sup> 在学者と卒業者の区別は行っていない。

<sup>10</sup> 市区町村コード、男女の別、出生の年、出生の月、配偶者の有無と教育については、市区町村コードのみを用いたマッチングと同様のマッチングキーの調整を行った。

き、以下の変数群が採用されている。

令和2年国勢調査	「市区町村コード」及び「調査区番号のうち、主番号及び後置番号」
令和3年社会生活基本調査	「市区町村コード」及び「調査区符号」

総務省統計局から提供された調査区対応表には、令和3年社会生活基本調査の調査区符号と平成27年国勢調査の調査区番号との対応表が記されている。そのため、本実験では、平成27年国勢調査における調査区番号を令和2年国勢調査のそれに変換する作業を行った。

調査区番号の変換作業にあたっては、平成27年国勢調査及び令和2年国勢調査の「小地域集計第1表 男女別人口及び世帯数—基本単位区」を用いた。この統計表には調査区番号と基本単位区番号との対応表が表示されている。調査区番号は2つの国勢調査間での変更があり得るが、基本単位区番号は原則として固定されている。こうした特徴を踏まえて、平成27年国勢調査から令和2年の国勢調査への調査区番号の変換を行った。

図3の①で示されるとおり、1つの調査区は、通常1つまたは複数の基本単位区を組み合わせて設定される。なお、1調査区内には、原則として40世帯から70世帯、平均すると約50世帯が、調査対象世帯として含まれている。

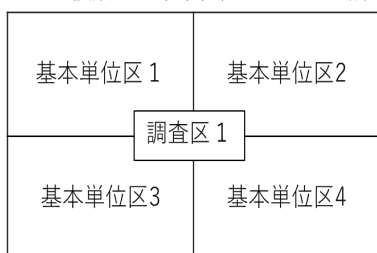
一方で、世帯数の多い基本単位区においては、図3の②のように、基本単位区を分割して調査区が設定される場合がある。例えば、中高層共同住宅がある基本単位区等に発生し得るケースがそれに該当する。

以上の点を踏まえて、データリンケージの実証実験における予備的作業として、調査区番号の変換を行った。図4で示されるように、変換された結果を見ると、一部の調査区で調査区番号が変換できないケースが存在している。図4における変換の可否の①は、平成27年と令和2年の国勢調査が、基本単位区番号、調査区番号のいずれについても1対1で対応していることから、正常に変換することが可能なケースである。また、変換の可否の②は、基本単位区番号が1対1、調査区番号（主番号-後置番号）は1対2で対応するため、これに関しても正常に変換することができる。それに対して、変換の可否の③は、基本単位区番号が不一致であるため変換不可となっている。さらに、変換の可否の④は、基本単位区番号は一致しているが、基本単位区内の調査区の数で平成27年国勢調査と令和2年国勢調査では異なることから、調査区番号を昇順に並べて1対1で対応付けを行っても、どのレコードにも連結されない調査区が発生するケースが存在する場合である。

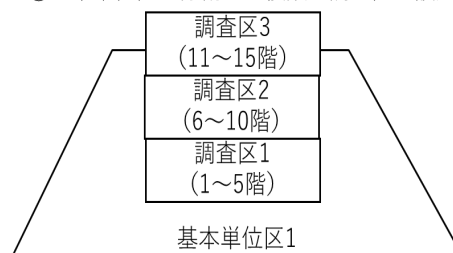
これらを踏まえた調査区マッチングキーを用いたデータリンケージのための確認作業の結果、平成27年国勢調査と令和2年国勢調査の間で調査区番号を変換することができず、その結果として令和3年社会生活基本調査とのリンケージができない調査区は、令和3年社会生活基本調査の全国調査区7,152件の中で163件（約2.3%）であることがわかった。

図3 調査区と基本単位区のイメージ

① 1つまたは複数の基本単位区から1つの調査区を設定



② 基本単位区を分割して複数の調査区を設定



出所 『令和2年国勢調査調査区 調査区関係資料 利用の手引』(総務省統計局)を参考に、筆者が作成。

図4 平成27年国勢調査から令和2年国勢調査への変換結果のパターン

平成27年国勢調査 小地域集計 第1表			変換の可否	令和2年国勢調査 小地域集計 第1表		
市区町村 コード	基本単位区 番号	調査区番号 主番号-後置番号(-単位番号)		市区町村 コード	基本単位区 番号	調査区番号 主番号-後置番号(-単位番号)
00100	010100000	225.1.	① 平成27年の1調査区と、令和2年の1調査区が対応	00100	010100000	351.1.
00100	010200000	226.1.1.		② 平成27年の1調査区が、令和2年の2調査区に対応 (226-1が、352-1と353-1に対応)	00100	010200000
00100	010300000	226.1.2.	00100		010300000	353.1.
00100	010400000	452.1.	③ 変換不可 (基本単位区番号が不一致)	00100	010400001	573.1.1.
				00100	010400002	573.1.2.
00100	010500000	771.1.	④ 部分的に変換不可 (基本単位区番号は同一だが、調査区の数異なる。 平成27年の1-3番目までは1対1対応できるが、4番目の調査区の対応先がない)	00100	010500000	891.1.
00100	010500000	772.1.		00100	010500000	892.1.
00100	010500000	773.1.		00100	010500000	893.1.
00100	010500000	774.1.				

## 5. データリンケージに関する実証実験の結果

本節では、国勢調査と社会生活基本調査の個票データを対象に、市区町村コードを用いたが、調査区マッチングキーを使用しなかった場合の照合と、調査区マッチングキーも使用して行った照合に関する実証実験の結果を比較・検証する。なお、社会生活基本調査の1レ

コードは1世帯員に関する情報を表しており、四国ブロックの総レコード数は12,415レコードである。

#### 5.1 市区町村コードのみを用いた照合に関する結果

国勢調査と社会生活基本調査を対象に、市区町村コードを用いたが、調査区マッチングキーを使用しなかった場合のレコード間の照合結果(表6)を見ると、「1対1でマッチングした」世帯員は、対象となるレコード群の中の8.4%に留まることがわかった。

また、「マッチングできなかった」比率が4.6%、「1対多でマッチングした」比率が87.0%であった。このことは、同一市区町村内で、マッチングキーである男女の別、出生の年月、配偶者の有無と教育を用いてデータリンケージを行うと、1レコードも照合しないことは稀であるのに対して、照合相手が複数のレコードになる割合が大幅に高くなることを示唆している。

以上の結果から、調査区マッチングキーを使用しない照合の場合、照合の対象となるレコードと1対1で対応付けすることは困難だと考えられる。

#### 5.2 調査区マッチングキーを用いた照合結果

調査区マッチングキーも用いて照合を行った場合の結果は、表7で示されている。表7を見ると、「1対1でマッチングした」比率が71.5%となり、調査区マッチングキーを使用しなかった場合の1対1のマッチング(8.4%)と比較して、約63%も増加するという興味深い結果が得られている。

「マッチングできなかった」比率は26.5%であり、市区町村コードを用いたが、調査区マッチングキーを使用しなかった場合の比率(4.6%)と比較して、約22%増加している。この中には、第4節で述べた調査区が変換不可であるために、対応付けが困難なレコードが含まれることに留意する必要がある。

「1対多でマッチングした」比率は2.0%であり、市区町村コードを用いたが、調査区マッチングキーを使用しなかった場合で照合する比率(87.0%)と比較して、約85%減少した。このことは、同一調査区内では、男女の別、出生の年月、配偶者の有無と教育のマッチングキーが重複するケースが非常に少ないことを意味している。

以上の結果から、調査区マッチングキーが、「1対1でマッチングした」レコード数の増大に大きく寄与したことがわかる。

なお、表8を見ると、市区町村コード、男女の別、出生の年、出生の月、配偶者の有無と調査区マッチングキーで照合した場合の「1対1でマッチングした」比率は、81.2%であって、これらの変数に教育も含めて照合を行った場合と比較して、「1対1でマッチングした」比率が、約10%上昇していることが確認された。この点についてはさらなる検証が求められることから、次節で検証の内容と結果について詳述する。

表6 市区町村コードのみを用いたマッチング結果—四国ブロック

四国ブロック マッチングキー 男女の別、出生の年、出生の月、 配偶者の有無、教育	度数			割合		
	マッチングで できなかった	1対1でマッチ ングした	1対多でマッチ ングした	マッチングで できなかった	1対1でマッチ ングした	1対多でマッチ ングした
市区町村コードのみ	568	1,048	10,799	4.6%	8.4%	87.0%

表7 加工済み調査区対応表を使用したマッチングの結果—四国ブロック

四国ブロック マッチングキー 男女の別、出生の年、出生の月、 配偶者の有無、教育	度数			割合		
	マッチングで できなかった	1対1でマッチ ングした	1対多でマッチ ングした	マッチングで できなかった	1対1でマッチ ングした	1対多でマッチ ングした
加工済み調査区対応表を使用	3,291	8,879	245	26.5%	71.5%	2.0%
市区町村コードのみ（再掲）	568	1,048	10,799	4.6%	8.4%	87.0%

表8 マッチングキー「教育」の有無によるマッチング率の変化—四国ブロック

四国ブロック マッチングキー	度数			割合		
	マッチングで できなかった	1対1でマッチ ングした	1対多でマッチ ングした	マッチングで できなかった	1対1でマッチ ングした	1対多でマッチ ングした
市区町村コード、調査区、男女の別、出生の年、出生の月、配偶者の有無	1,732	10,084	599	14.0%	81.2%	4.8%
市区町村コード、調査区、男女の別、出生の年、出生の月、配偶者の有無、教育	3,291	8,879	245	26.5%	71.5%	2.0%

### 5.3 マッチングキーとして教育を追加した場合の比率の変化に関する検証及びマッチングキーの分布特性

本節では、市区町村コード、調査区マッチングキー、男女の別、出生の年、出生の月及び配偶者の有無で「1対1でマッチングした」及び「1対多でマッチングした」世帯員のレコードを対象に、マッチングキーとして教育を追加した場合の比率の変化を検証した。その結果、表9のとおり、マッチングキーである教育の追加で、「1対1でマッチングした」世帯員（10,084人）の15.0%（1,514人）が、「マッチングできなかった」世帯員に変化した一方で、「1対多でマッチングした」の世帯員（599人）の51.6%（309人）が、「1対1でマッチングした」世帯員に変わったことがわかった。

教育を含まない場合、「1対1でマッチングした」比率が81.2%で、「1対多でマッチングした」比率が4.8%になっていることを考慮すると、マッチングキーとして教育を用いることは、照合に関する精度に関してマイナスの影響をもたらしていることが考えられる。

そこで、市区町村コード、調査区マッチングキー、男女の別、出生の年、出生の月、配偶者の有無による照合の結果、「1対1でマッチングした」が、教育をマッチングキーに追加

表9 マッチングキー「教育」の有無によるマッチング状況の変化—四国ブロック

四国ブロック	市区町村コード、調査区、男女の別、出生の年、出生の月、配偶者の有無、教育					
	度数			割合		
市区町村コード、調査区、男女の別、出生の年、出生の月、配偶者の有無	マッチングできなかった	1対1でマッチングした	1対多でマッチングした	マッチングできなかった	1対1でマッチングした	1対多でマッチングした
1対1でマッチングした	1,514	8,570		15.0%	85.0%	
1対多でマッチングした	45	309	245	7.5%	51.6%	40.9%

すると「マッチングなし」に変化したレコード群（以下「教育不一致データ」と呼称）を対象に、国勢調査と社会生活基本調査における教育に関する分類区分のデータ特性を確認した。

その結果、表10のとおり、教育不一致データには「小学」と「中学」、「中学」と「高校・旧制中」、「短大・高専」と「高校」の混同が数多く見られることがわかった。なお、他の属性とのクロス集計を行うことによって、教育不一致データの傾向が明らかになる可能性がある。留意点としては、「高校・旧制中」の教育不一致データにおいて、令和2年国勢調査の教育が不詳であるデータが54.9%と半数を超えていることが指摘される。照合対象となるレコードの一方が不詳である場合の処理については、さらなる検討が必要である。

つぎに、令和3年社会生活基本調査の調査区レベルで、調査区内の全データにおける「マッチングできなかった」レコード及び「1対1でマッチングした」レコードの割合を検証し、その分布特性を確認した。図5と図6は、これらの検証結果を示したものである。「マッチングできなかった」レコードの割合の分布は10%から35%に、「1対1でマッチングした」データの割合は65%から85%にそれぞれ集中していることがわかった。「マッチングできなかった」レコードの割合が高い調査区、及び「1対1でマッチングした」レコードの割合が低い調査区は、第4節で述べた平成27年国勢調査から令和2年国勢調査への調査区番号の変換の不具合が生じている可能性がある。

つぎに、マッチングに用いた全レコード（以下「全データ」と呼称）と「マッチングできなかった」、「1対1でマッチングした」、及び「1対多でマッチングした」の3種類の照合のパターンについて、各マッチングキー及び年齢5歳階級別の分類区分ごとにその分布状況を比較・検証した。また、全データと「1対1でマッチングした」レコードについては、各分類区分の全体に対する割合のポイント差も比較した。

男女の別(表11)については、全データと「マッチングできなかった」、「1対1でマッチングした」と「1対多でマッチングした」それぞれのパターンにおいて、概ね男性の比率が46~48%、女性の比率が52~54%であるという結果が得られており、照合のパターンによる大きな分布差は見られなかった。

年齢5歳階級については、表12で示されたとおり、20~24歳、25~29歳及び80~84歳が「1対1でマッチングした」データの割合が-0.5ポイントとやや低く、10~14歳が+0.6ポイント、65~69歳が+0.7ポイントとやや高いが、大きな分布差はないと考える。なお、「1対多でマッチングした」レコードの比率は、2.0%にすぎないが(表7)、10~14歳、15~19歳における割合はそれぞれ、12.7%と17.6%であり、全データにおいて対応する年齢

表 10 「教育不一致データ」におけるマッチングキー「教育」の分類区分の分布特性、  
四国ブロック

教育不一致データ-四国ブロック				
令和3年社会生活基本調査 調査票A:生活行動編 教育	度数	総計に占める割合	内数	
			度数	小計に占める割合
在学したことがない	4	0.3%		
小学	29	1.9%		
うち令和2年国勢調査で「中学」			23	79.3%
中学	242	16.0%		
うち令和2年国勢調査で「高校」			124	51.2%
うち令和2年国勢調査で「不詳」			86	35.5%
高校・旧制中	397	26.2%		
うち令和2年国勢調査で「中学」			109	27.5%
うち令和2年国勢調査で「不詳」			218	54.9%
短大・高専	626	41.3%		
うち令和2年国勢調査で「高校」			408	65.2%
うち令和2年国勢調査で「不詳」			103	16.5%
大学	125	8.3%		
うち令和2年国勢調査で「高校」			42	33.6%
うち令和2年国勢調査で「短大・高専」			27	21.6%
うち令和2年国勢調査で「不詳」			46	36.8%
大学院	11	0.7%		
不詳	80	5.3%		
うち令和2年国勢調査で「中学」			21	26.3%
うち令和2年国勢調査で「高校」			40	50.0%
総計	1,514	100.0%		

注 教育不一致データの「うち令和2年国勢調査」の「学校の種類」については、度数の小さい区分があるため、主な区分のみの「内数」表示とした。

図5 全データに対する「マッチングできなかった」データの割合に関する調査区レベルの分布状況、四国ブロック

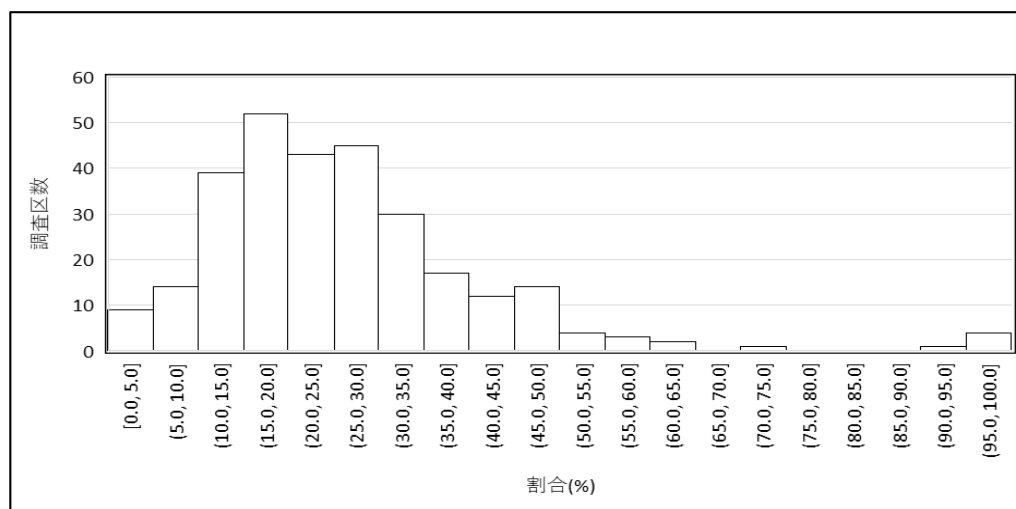


図6 全データに対する「1対1でマッチングした」データの割合に関する調査区レベルの分布状況、四国ブロック

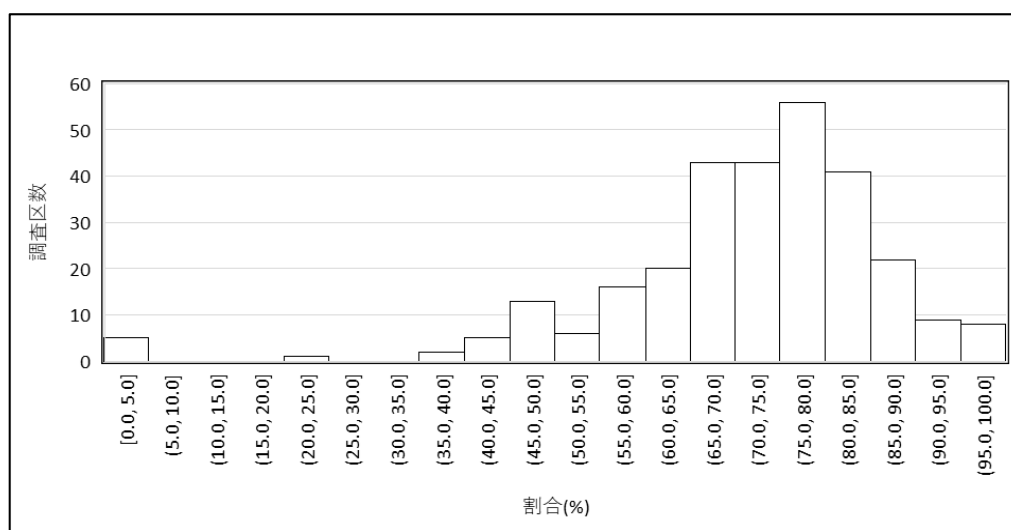


表11 照合パターン別で見た男女の別に関する分類区分の分布状況—四国ブロック

四国ブロック	度数				分類計に占める割合				ポイント差 (b)-(a)	
	マッチングキー 男女の別	全データ	マッチングでき なかった	1対1でマッチ ングした	1対多でマッチ ングした	全データ (a)	マッチングでき なかった	1対1でマッチ ングした(b)		1対多でマッチ ングした
男		5,873	1,513	4,246	114	47.3%	46.0%	47.8%	46.5%	0.5
女		6,542	1,778	4,633	131	52.7%	54.0%	52.2%	53.5%	-0.5
合計		12,415	3,291	8,879	245	100.0%	100.0%	100.0%	100.0%	

表 12 照合パターン別で見た年齢 5 歳階級に関する分類区分の分布状況—四国ブロック

四国ブロック 年齢5歳階級	度数				分類計に占める割合				ポイント 差 (b)-(a)
	全データ	マッチン グできな かった	1対1で マッチン グした	1対多で マッチン グした	全データ (a)	マッチン グできな かった	1対1で マッチン グした(b)	1対多で マッチン グした	
10-14歳	594	82	481	31	4.8%	2.5%	5.4%	12.7%	0.6
15-19歳	593	101	449	43	4.8%	3.1%	5.1%	17.6%	0.3
20-24歳	398	143	244	11	3.2%	4.3%	2.7%	4.5%	-0.5
25-29歳	403	154	244	5	3.2%	4.7%	2.7%	2.0%	-0.5
30-34歳	471	145	323	3	3.8%	4.4%	3.6%	1.2%	-0.2
35-39歳	676	210	462	4	5.4%	6.4%	5.2%	1.6%	-0.2
40-44歳	823	238	571	14	6.6%	7.2%	6.4%	5.7%	-0.2
45-49歳	1,059	276	767	16	8.5%	8.4%	8.6%	6.5%	0.1
50-54歳	926	250	660	16	7.5%	7.6%	7.4%	6.5%	0.0
55-59歳	874	204	656	14	7.0%	6.2%	7.4%	5.7%	0.3
60-64歳	986	222	747	17	7.9%	6.7%	8.4%	6.9%	0.5
65-69歳	1,112	238	853	21	9.0%	7.2%	9.6%	8.6%	0.7
70-74歳	1,392	345	1,015	32	11.2%	10.5%	11.4%	13.1%	0.2
75-79歳	834	235	587	12	6.7%	7.1%	6.6%	4.9%	-0.1
80-84歳	638	223	414	1	5.1%	6.8%	4.7%	0.4%	-0.5
85-89歳	430	139	286	5	3.5%	4.2%	3.2%	2.0%	-0.2
90歳以上	206	86	120	0	1.7%	2.6%	1.4%	0.0%	-0.3
合計	12,415	3,291	8,879	245	100.0%	100.0%	100.0%	100.0%	

階層についてのそれらと比較して高くなっている。これに関しては、10～19歳の年齢階層の大部分は在学者かつ未婚であり、同じ年齢であれば、教育について同一の属性値となる可能性が高いことから、マッチング対象を特定できないケースが多くなると推察される。

配偶者の有無については、表 13 のとおり、「1対1でマッチングした」レコードにおいて有配偶者が多く、死別・離別に該当する者が少なくなっていることが確認される。死別・離別の場合、「マッチングできなかった」レコードにおける割合が 20.6% であって、全データでの割合である 14.6% よりも 6 ポイント多くなっている。このことは、両調査の実施間における変化（有配偶からの変化、死別・離別に伴う転居等）に伴い、照合されなかったレコード数が増加したことを意味する。

さらに、教育に関しては、表 14 に含まれる「ポイント差」が示すとおり、「1対1でマッチングした」データにおいて「高校・旧制中」や「大学」が多く、「短大・高専」が少なくなっている。なお、「短大・高専」については、「マッチングできなかった」データにおける割合が 29.5% で、全データの 20.1% との差異が見られるため、マッチング率の低さについて検証が必要である。

なお、四国ブロックにおける県別で見た各マッチングキーの分類区分の構成比については、付図 1～付図 4 のグラフを参照いただきたい。

表 13 照合パターン別で見た配偶者の有無に関する分類区分の分布状況—四国ブロック

四国ブロック	度数				分類計に占める割合				ポイント差 (b)-(a)
	全データ	マッチングで きなかった	1対1でマッチ ングした	1対多でマッチ ングした	全データ (a)	マッチングで きなかった	1対1でマッチ ングした(b)	1対多でマッチ ングした	
マッチングキー 配偶者の有無									
未婚	3,105	780	2,227	98	25.0%	23.7%	25.1%	40.0%	0.1
有配偶	7,449	1,783	5,532	134	60.0%	54.2%	62.3%	54.7%	2.3
死別・離別	1,811	678	1,120	13	14.6%	20.6%	12.6%	5.3%	-2.0
不詳	50	50	0	0	0.4%	1.5%	0.0%	0.0%	-0.4
合計	12,415	3,291	8,879	245	100.0%	100.0%	100.0%	100.0%	

表 14 照合パターン別で見た教育の分類区分に関する分布状況—四国ブロック

四国ブロック	度数				分類計に占める割合				ポイント 差 (b)-(a)
	全データ	マッチ ングでき なかつた	1対1で マッチ ングした	1対多で マッチ ングした	全データ (a)	マッチ ングでき なかつた	1対1で マッチ ングした(b)	1対多で マッチ ングした	
マッチングキー 教育									
在学したことがない	12	12	0	0	0.1%	0.4%	0.0%	0.0%	-0.1
小学	367	82	272	13	3.0%	2.5%	3.1%	5.3%	0.1
中学	1,887	528	1,325	34	15.2%	16.0%	14.9%	13.9%	-0.3
高校・旧制中	5,006	1,063	3,814	129	40.3%	32.3%	43.0%	52.7%	2.6
短大・高専	2,496	971	1,497	28	20.1%	29.5%	16.9%	11.4%	-3.2
大学	2,305	464	1,800	41	18.6%	14.1%	20.3%	16.7%	1.7
大学院	186	46	140	0	1.5%	1.4%	1.6%	0.0%	0.1
不詳	156	125	31	0	1.3%	3.8%	0.3%	0.0%	-0.9
合計	12,415	3,291	8,879	245	100.0%	100.0%	100.0%	100.0%	

## 6. むすびにかえて

本稿では、公的統計におけるデータリンケージの精度検証を行うために、四国地方を対象にして、令和2年国勢調査と令和3年社会生活基本調査の世帯員レベルのデータリンケージに関する実証実験を行い、調査区マッチングキーの有無による統計的照合の程度を比較・検討した。その結果、地域情報としての調査区マッチングキーを統計的照合における追加的な変数と設定した上で、「1対1でマッチングした」割合は、市区町村コードを用いるが、調査区マッチングキーを使用しなかった場合のそれと比較して、約63%上昇することが明らかになった。このことから、調査区対応表が、データリンケージに大きく寄与したことが確認された。

また、マッチングキーとして教育を追加することにより、「1対1でマッチングできた」レコードの中の15.0%が、「マッチングできなかった」レコードに変化したことがわかった。これについては、リンケージデータの利用目的によっては、教育をマッチングキーとして含めずに照合を行うことも考えられる。さらに、マッチングキーの分類区分の分布特性に関する比較・検証の結果から、「1対1でマッチングした」レコードのうち、配偶者の有無及び

教育に関する分類区分の分布特性が、全データにおけるそれとは異なることがわかった。こうした結果は、リンケージデータを様々な分析に用いる上での留意すべき点であると考えられる。

その一方で、調査区マッチングキーの編成にあたって、平成 27 年国勢調査調査区を令和 2 年調査区に変換することができないために、特定の国勢調査調査区がそれと対応する令和 3 年社会生活基本調査の調査区と全く照合できない事例が、調査区全体の約 2.3%で発生することが確認された。これに関しても、その影響が大きい都道府県や市区町村を含む地域を対象にした実証分析を行う場合には、留意する必要があるだろう。

なお、本稿では、四国ブロックを対象に実証実験を行ったが、今後は、他地域を含む全国の個票データを用いて同様の検証を行うことが求められる。さらに、本研究で実施した検証の方法を他の統計調査とのデータリンケージに応用することによって、世帯単位での分析において作成されたリンケージデータをより広範に活用することが考えられる。これについては、今後の研究課題としたい。

## 謝辞

本稿は、JST 経済安全保障重要技術育成プログラム【JPMJKP24U5】の支援を受けている。本研究において使用した「国勢調査」及び「社会生活基本調査」の調査票情報は、統計法第 33 条に基づき提供を受けたものであり、本稿で作成した集計表等は提供を受けた調査票情報を独自集計したものである。記して関係各位に御礼申し上げたい。また、本稿で実施した実証実験において、国勢調査と社会生活基本調査のデータ特性の視点から貴重な助言をいただいた小林良行先生（(公財)統計情報研究開発センター研究員、中央大学経済研究所客員研究員）に謝意を表したい。

## 参考文献

伊藤伸介(2020)「デンマークとオランダにおける医療健康データの二次利用について」『日本統計学会誌』, 第 50 巻第 1 号, pp. 109-138.

伊藤伸介(2024)「海外におけるマイクロデータの利活用推進に向けた取り組み状況—リモートアクセスを中心に—」『経済学論纂(中央大学)』第 65 巻 1 号, pp.199-218.

伊藤伸介(2025a)「海外における公的大規模データの利活用の現状」, 『日本労働研究雑誌』, No.779, pp.53-64.

伊藤伸介(2025b)「海外における行政記録情報の利活用に関する取り組み状況」『統計』2025

年 9 月号, pp. 27-31.

村田磨理子・伊藤伸介(2016)「事業所・企業系のマイクロデータを用いたデータリンケージの可能性―賃金構造基本統計調査を例に―」, 『統計学』 110 号, pp.1-17.

以下については、2026 年 3 月 3 日にアクセス

<総務省統計局ホームページ>

令和 2 年国勢調査関係

「令和 2 年国勢調査の概要」 <https://www.stat.go.jp/data/kokusei/2020/gaiyou.html>

「統計表で用いられる地域区分の解説」 <https://www.stat.go.jp/data/kokusei/1995/04-02.html>

令和 3 年社会生活基本調査関係

「令和 3 年社会生活基本調査の概要」 <https://www.stat.go.jp/data/shakai/2021/gaiyou.html>

「標本設計及び結果の推定方法」 <https://www.stat.go.jp/data/shakai/2021/pdf/suikai-2.pdf>

<国勢調査調査区 調査区関係資料>

総務省統計局(2017) 平成 27 年国勢調査調査区 調査区関係資料 利用の手引 pp.1-6

総務省統計局(2022) 令和 2 年国勢調査調査区 調査区関係資料 利用の手引 pp1-6

<マイクロデータ利用ポータルサイト>

調査票情報の利用 (リモートアクセス利用)

<https://www.e-stat.go.jp/microdata/data-use/remote>

調査票情報の利用 (オンサイト利用)

<https://www.e-stat.go.jp/microdata/data-use/on-site>

<政府統計の総合窓口>

総務省統計局 (2017)

平成 27 年国勢調査 人口等基本集計に関する集計 小地域集計

第 1 表 男女別人口及び世帯数－基本単位区

徳島県

[https://www.e-stat.go.jp/stat-](https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001080615&cycle=0&tclass1=000001094495&tclass2=000001094535&cycle_facet=cycle&tclass3val=0&metadata=1&data=1)

[search/files?page=1&layout=datalist&toukei=00200521&tstat=000001080615&cycle=0&tclass1=000001094495&tclass2=000001094535&cycle\\_facet=cycle&tclass3val=0&metadata=1&data=1](https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001080615&cycle=0&tclass1=000001094495&tclass2=000001094535&cycle_facet=cycle&tclass3val=0&metadata=1&data=1)

香川県

<https://www.e-stat.go.jp/stat->

[https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001080615&cycle=0&tclass1=000001094495&tclass2=000001094536&cycle\\_facet=cycle&tclass3val=0&metadata=1&data=1](https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001080615&cycle=0&tclass1=000001094495&tclass2=000001094536&cycle_facet=cycle&tclass3val=0&metadata=1&data=1)

愛媛県

[https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001080615&cycle=0&tclass1=000001094495&tclass2=000001094537&cycle\\_facet=cycle&tclass3val=0&metadata=1&data=1](https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001080615&cycle=0&tclass1=000001094495&tclass2=000001094537&cycle_facet=cycle&tclass3val=0&metadata=1&data=1)

高知県

[https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001080615&cycle=0&tclass1=000001094495&tclass2=000001094538&cycle\\_facet=cycle&tclass3val=0&metadata=1&data=1](https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001080615&cycle=0&tclass1=000001094495&tclass2=000001094538&cycle_facet=cycle&tclass3val=0&metadata=1&data=1)

総務省統計局（2022）

令和2年国勢調査 人口等基本集計に関する集計 小地域集計

第1表 男女別人口及び世帯数－基本単位区

徳島県

<https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001136464&cycle=0&tclass1=000001136472&tclass2=000001159909&tclass3val=0>

香川県

<https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001136464&cycle=0&tclass1=000001136472&tclass2=000001159910&tclass3val=0>

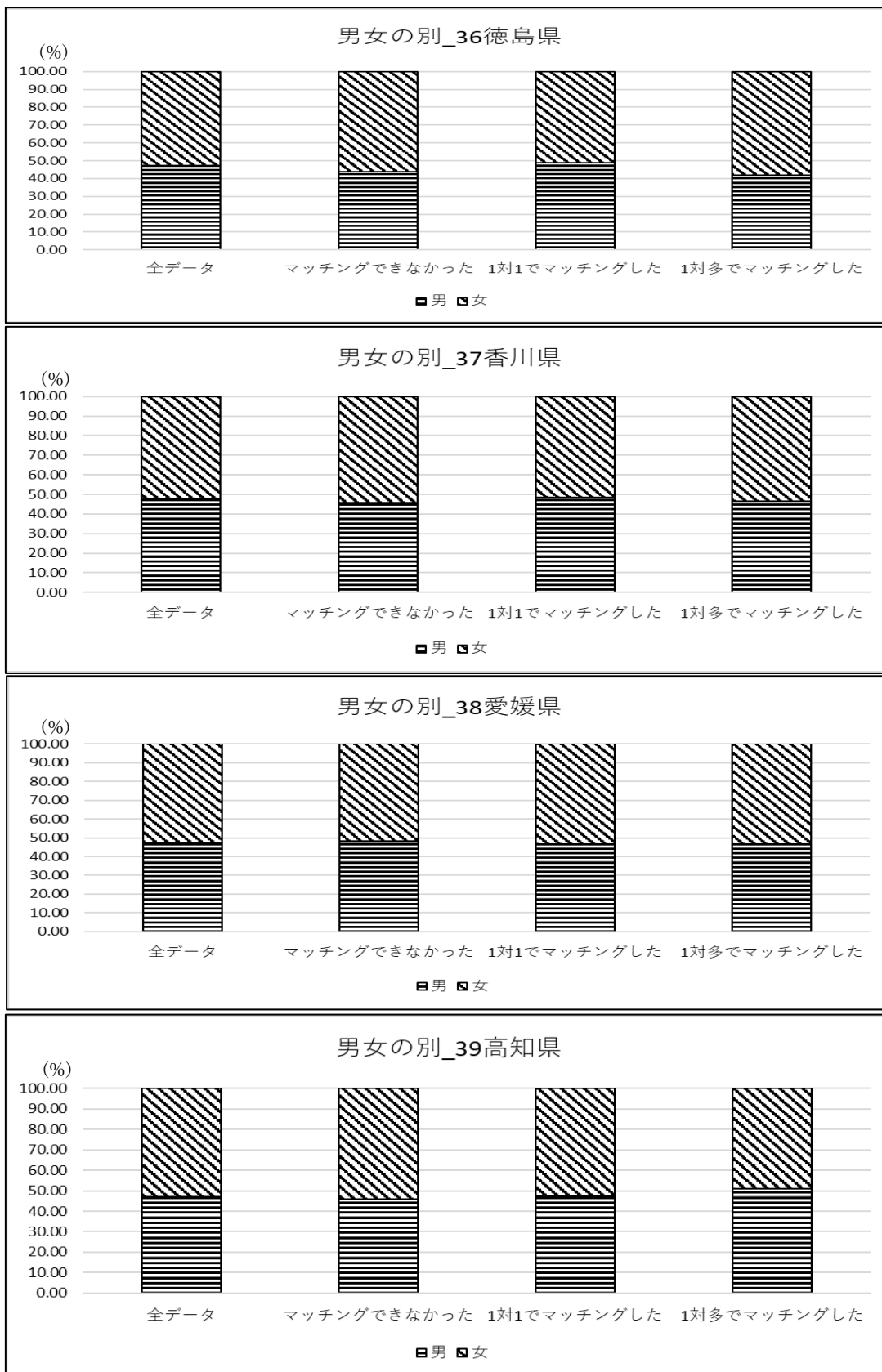
愛媛県

<https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001136464&cycle=0&tclass1=000001136472&tclass2=000001159911&tclass3val=0>

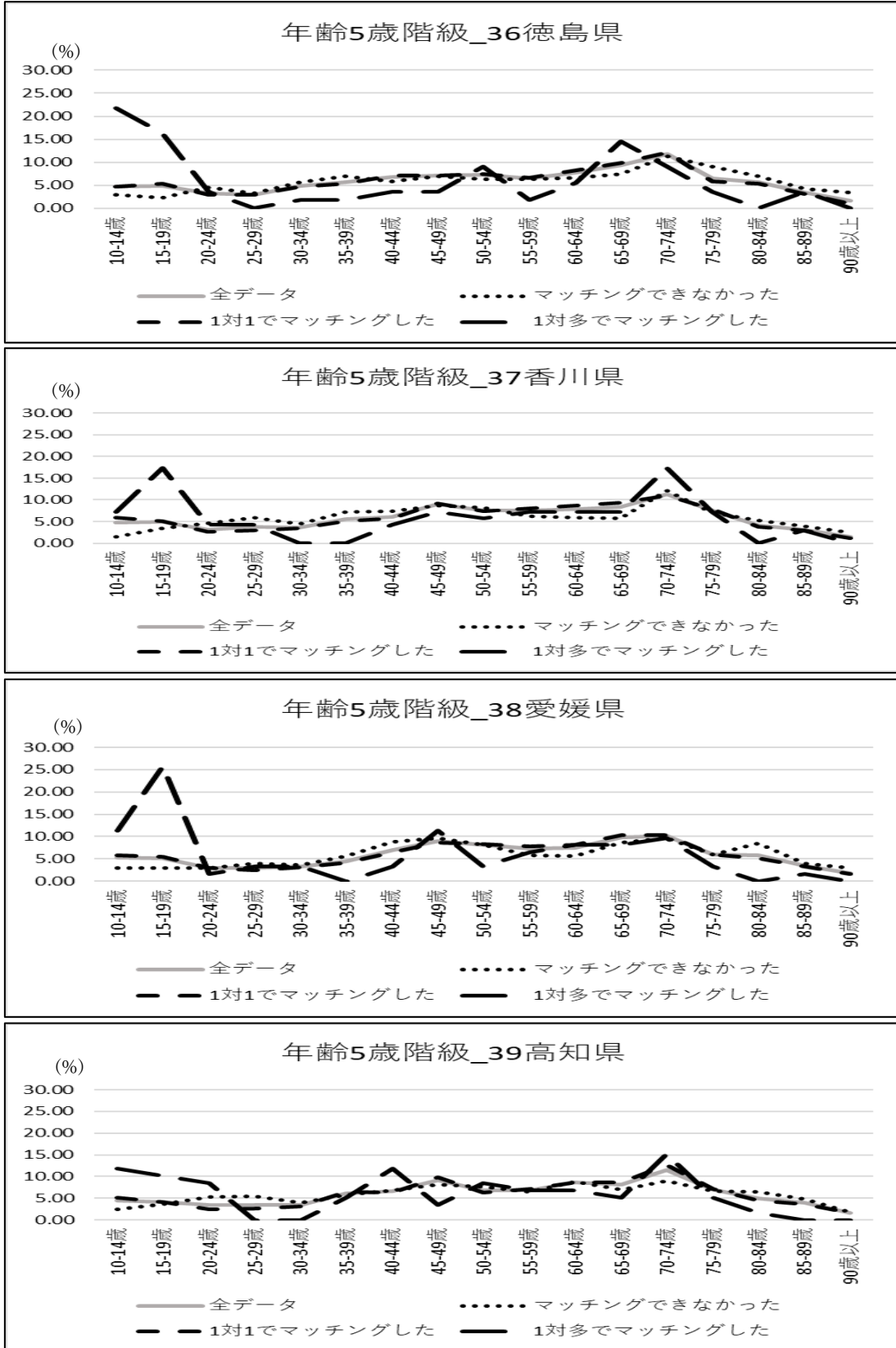
高知県

<https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001136464&cycle=0&tclass1=000001136472&tclass2=000001159912&tclass3val=0>

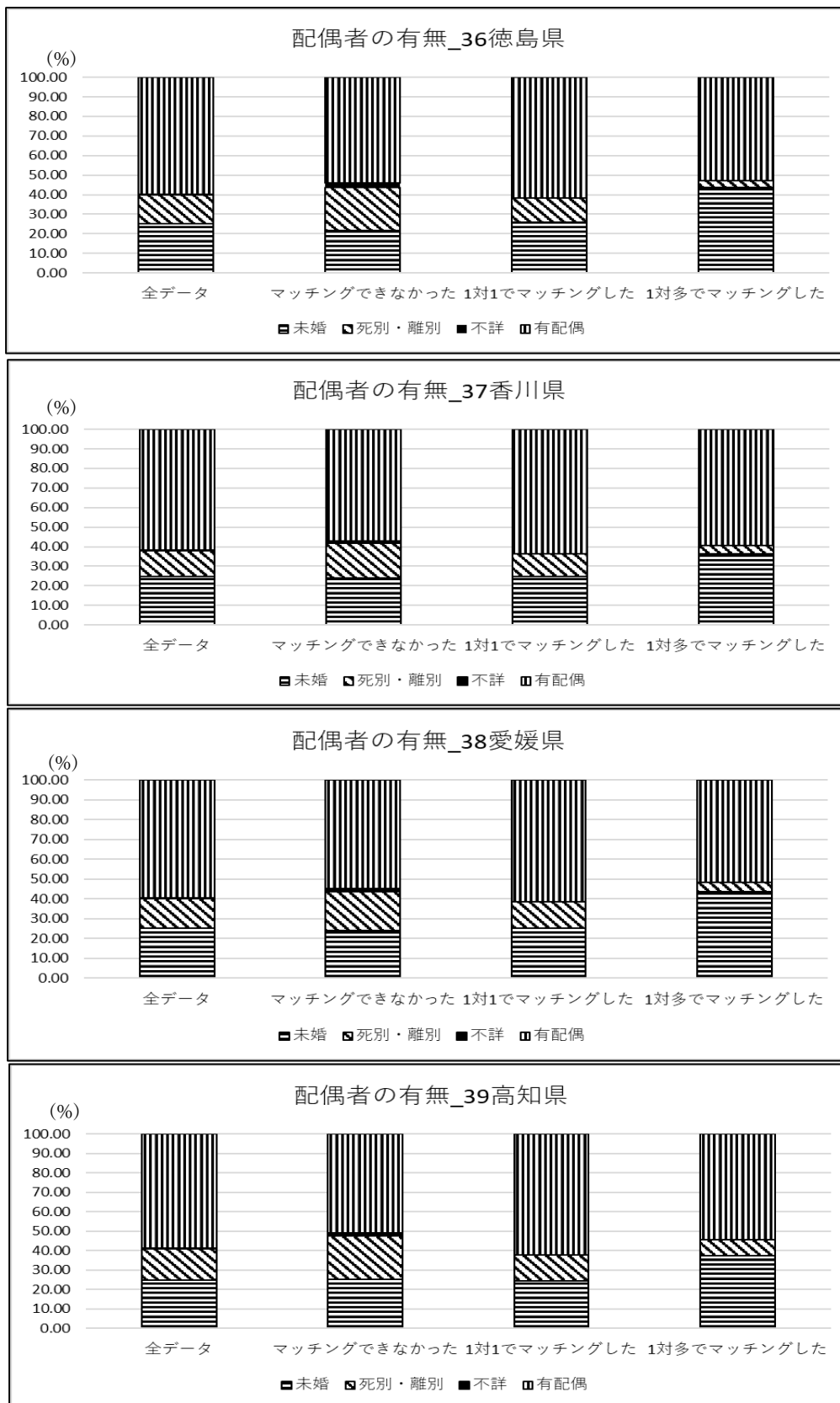
付図1 マッチングの状況別に見た「男女の別」の構成比一県別



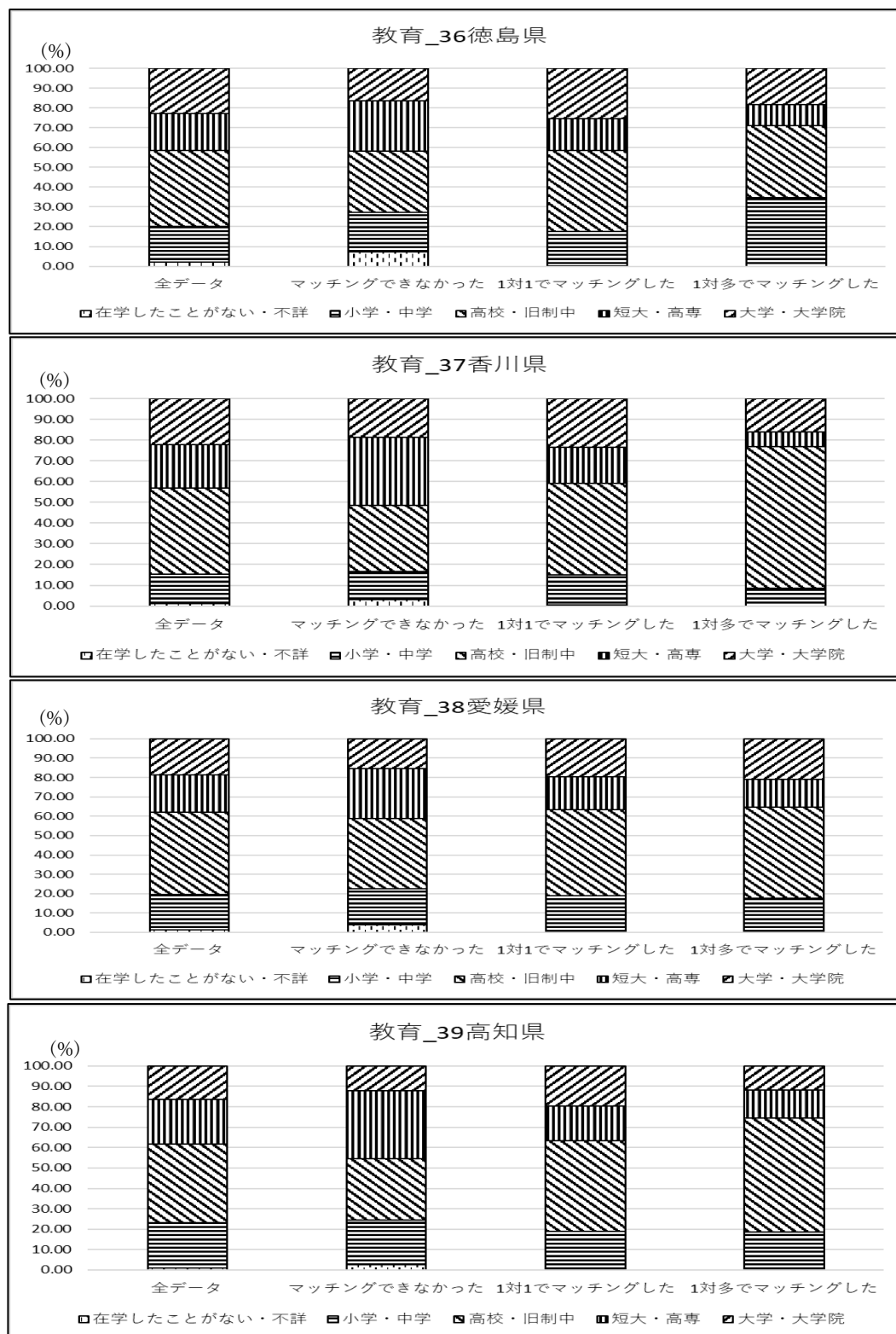
付図2 マッチングの状況別に見た「年齢5歳階級」の分布一県別



付図3 マッチング状況別に見た「配偶者の有無」の構成比一県別



付図4 マッチング状況別に見た「教育」の構成比―県別



注 県によっては度数の小さい区分があるため、「在学したことがない」と「不詳」、「小学」と「中学」、「大学」と「大学院」の構成比を合算している。

中央大学経済研究所  
( INSTITUTE OF ECONOMIC RESEARCH, CHUO UNIVERSITY)  
代表者 阿部 顕三 (Director: Kenzo Abe)  
〒192-0393 東京都八王子市東中野 742-1  
(742-1 Higashi-nakano, Hachioji, Tokyo 192-0393 JAPAN)  
TEL: 042-674-3271 +81 42 674 3271  
FAX: 042-674-3278 +81 42 674 3278  
E-mail: keizaiken-grp@g.chuo-u.ac.jp  
URL: <https://www.chuo-u.ac.jp/research/institutes/economic/>