



RIETI Technical Paper Series 19-T-001

市町村合併を考慮した市区町村パネルデータの作成 (改訂版)

近藤 恵介
経済産業研究所



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所

<https://www.rieti.go.jp/jp/>

市町村合併を考慮した市区町村パネルデータの作成*

近藤恵介[†]
(経済産業研究所)

要旨

本稿では、1980年から2020年までの市区町村単位のパネルデータの作成方法について説明する。平成の大合併として知られているように、2000年代中頃には全国的に市町村合併が行われ、市区町村数は約半数にまで減少している。それに伴い、長期的な市区町村パネルデータを作成するには非常に煩雑な作業が必要となっているのが現状である。このような問題を解決するため、1980年から2020年までの市町村合併を考慮した市区町村コードのコンバータ作成方法について解説するとともに、作成したコンバータを一般に利用できるように公開する。

JEL classification: Y10

Keywords: 市区町村パネルデータ, 市町村合併

論文に述べられている見解は執筆者個人の責任で発表するものであり、所属する組織及び(独)経済産業研究所としての見解を示すものではありません。

*本稿は、独立行政法人経済産業研究所で実施した「RIETI データ整備・活用」プロジェクトの研究成果である。Online Appendix として全都道府県に関する1980年の行政区域と2015年の行政区域の比較を地図上で示している。本稿で使用している地図データは、地理情報分析支援システム MANDARA における「日本市町村」である。本稿で作成した市区町村コンバータは以下の URL より入手可能である。

(URL: <https://keisukekondokk.github.io/data/>)

[†]独立行政法人経済産業研究所, 上席研究員, 東京都千代田区霞が関 1-3-1 経済産業省別館 11 階.
(E-mail: kondo-keisuke@rieti.go.jp)

1 はじめに

平成の大合併として知られているように 2000 年代中頃に市町村合併が全国的に行われ、1980 年から 2015 年にかけて市区町村数は半数近くにまで減少している。表 1 で示すように、1980 年には市区町村数は 3278 あったが、2020 年には 1741 となっている。さらに表 1 では市区町村数の変化の内訳も示しているが、合併を通じて町村の数が大きく減少する一方で、政令指定都市や市の数が上昇していることが分かる。

近年、エビデンスに基づく政策形成の推進や内閣官房まち・ひと・しごと創生本部 (2019) によって提供されている地域経済分析システム RESAS の普及とともに、市区町村パネルデータを用いた統計分析への需要が高まる一方で、長期的な市区町村パネルデータの作成には市区町村合併をすべて考慮しなければならず、非常に煩雑な作業が必要とされる。その結果、データセットの作成から統計分析を開始するまでに非常に多くの時間を費やさなければならない。このような問題を解決するため、本稿では、1980 年から 2020 年までの市町村合併を考慮した市区町村コードのコンバータ作成方法について解説する。

本稿で使用する「コンバータ」とは、「2 つ以上のデータを接続するための対応表」のことを意味する。図 1 では、コンバータのイメージを示している。ここでは 1980 年時点の市区町村データと 2015 年時点の市区町村データがある状況を考える。この間に市町村合併が行われているため、調査時点の市区町村コードを用いて単純に 2 つのデータを接続することはできない。コンバータの役割とは、1980 年時点の市区町村データと 2015 年時点の市区町村データを接続できるように両年の市区町村コードを統一することである。

市町村合併ということは、2015 年時点の市区町村が 1980 年時点の複数の市区町村によって構成されるということである。したがって、1980 年時点の市区町村に 2015 年時点の市区町村コードを付与することができれば、2015 年時点の市区町村単位で統一して 1980 年と 2015 年の比較ができるようになる。つまり、市区町村コンバータを利用すれば、2015 年時点の市区町村コードを軸にした市区町村パネルデータが作成できることを意味する。

市区町村コンバータの利用は、市区町村パネルデータの作成のみに限らず、様々な状況で利用できる。例えば、学術研究では家計や企業に関する個票データを用いた実証分析を行うが、地理情報として市区町村コードが利用できることがある。しかしながら、合併がある場合は市区町村コードが時系列で変化してしまうが、市区町村コンバータを利用すれば時間を通じて統一した地理単位のコードを新たに付与できるようになり、都市・地域経済の実証分析での利用が期待できる。また、市区町村合併による市区町村区分の変化を自然実験として利用した実証研究でも市区町村コンバータの利用が期待できる。

本稿で作成する市区町村コンバータは、もちろん学術研究だけでなく、政策立案や教育における利用でも有用だと考えられる。長期的な市区町村パネルデータは、内閣府 (2015) のウェブサイトでも一部の変数について一般向けに公開されているが、他の変数を利用したい場合は自分自身でデータセット作成しなければならない。その場合は、総務省統計局 (2019) の e-Stat から市区町村データをダウンロードし、市区町村コードを両データで揃えて再集計

表 1 1980年から2015年までの市区町村数の推移

市区町村の区分	1980年	1985年	1990年	1995年	2000年	2005年	2010年	2015年	2020年
東京都特別区	23	23	23	23	23	23	23	23	23
政令指定都市	10	10	11	12	12	14	19	20	20
市	636	641	644	652	659	736	767	770	772
町	1993	2001	2003	1992	1991	1178	757	745	743
村	616	601	587	576	570	288	184	183	183
市区町村総数	3278	3276	3268	3255	3252	2239	1750	1741	1741

注) 総務省統計局(2019)のe-Statにおける「統計分類・調査項目>市区町村・コード>市区町村数を調べる」の情報より著者作成。年次の基準日は国勢調査の10月1日時点に合わせている。なお北方四島の6村(色丹村, 泊村, 留夜別村, 留別村, 紗那村, 薬取村)では統計調査が行われていないため, 上記の数字には含まれていない。

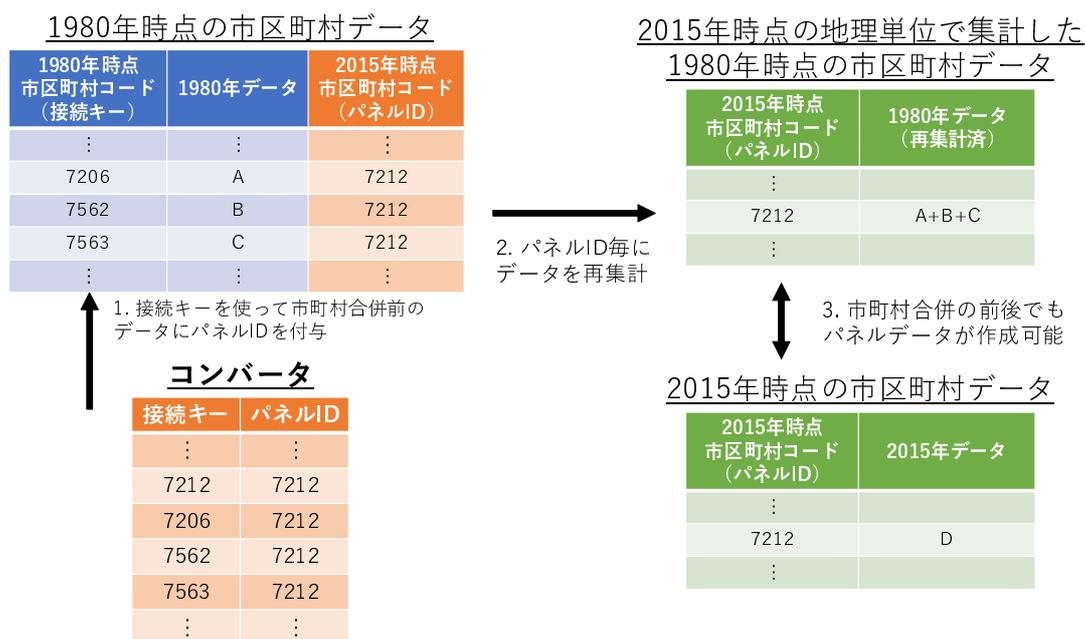


図 1: コンバータとは

注) 著者作成．市区町村コンバータを過去の市区町村データに適用することで，市町村合併の前後であっても時間を通じて同一の地理単位でデータを再集計できるようになる．

する必要がある．市区町村コンバータを用いれば，そのような煩雑な作業を簡略化することができ，スムーズにデータ分析へ移行できるようになる．市区町村パネルデータの利用および分析が進むことで，地域経済に関する議論がより活発になることが期待される．

本稿の構成は，以下の通りである．第 2 節では，市区町村の識別に使用される市区町村コードの概要とコンバータの概念について詳細に解説する．第 3 節では，コンバータの作成方法とファイルのレイアウトについて説明する．第 4 節では，市区町村コンバータを利用してどのように市区町村パネルデータを作成するのか簡単な手順を紹介する．第 5 節では，作成したコンバータが本当に正しく作成されているのか検証した結果を述べる．最後に，第 6 節を結語とする．

2 市区町村コードの変更とコンバータの概念

市区町村を一意に識別するために市区町村コードが割り振られている¹⁾．統計データの調査票にも市区町村コードの記入欄があり，総務省統計局 (2019) の e-Stat 等，市区町村毎に調査の集計結果が公表される際は一般に市区町村コードも同時に付与されている．本コンバータの作成においても市区町村コードを利用している．

¹⁾ 総務省の以下のウェブページによると，「情報処理の効率化と円滑化に資するため，コード標準化の一環として，総務省 (当時：自治省) が昭和 43 年に全国の都道府県及び市区町村のコードを設定した」とある．

総務省トップ > 政策 > 地方行財政 > 電子自治体 > 全国地方公共団体コード

<http://www.soumu.go.jp/denshijiti/code.html> (2019 年 2 月 20 日確認)

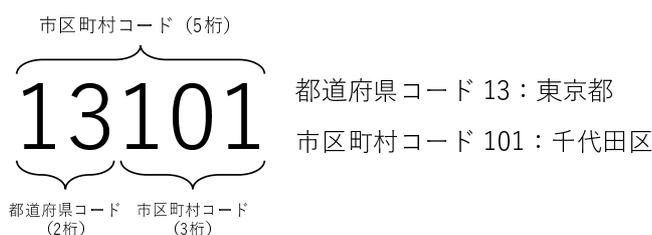


図 2: 市区町村コード (5桁) の構成

注) 著者作成。都道府県コードは北海道の 01 番から沖縄県の 47 番までになる。市区町村コード (3 桁) は、100 番台が政令指定都市、201 から 299 番までが市、301 から 799 番までが町村になる。詳細は総務省 (2007) を参照のこと。市区町村コードと市区町村名称の対応表は総務省統計局 (2019) の e-Stat の「統計分類・調査項目 > 市区町村・コード > 市区町村を探す」より入手可能である。

図 2 において示すように、市区町村コードは 5 桁の数字から構成される²⁾。最初の 2 桁は都道府県コードを表し、北海道の 01 番から沖縄県の 47 番までが割り振られている。残り 3 桁が都道府県内の市区町村コードを表している。なお本稿において 5 桁の市区町村コードと 3 桁の市区町村コードを区別する必要がある場合は、括弧書きで補足する。市区町村コード (3 桁) は、100 番台が政令指定都市、201 から 299 番までが市、301 から 799 番までが町村として割り振られている。

市区町村合併が行われると、市区町村コードの新設、変更、欠番処理が行われる。また合併がなくても市区町村コードの変更が起こる状況はある。このような市区町村コードの変更を整理してコンバータを作成する。市区町村コードの変更基準の詳細は総務省 (2007) において記述されているが、本稿ではコンバータ作成を理解する上で必要な点のみを説明する。

表 2 では、市区町村コードの変更が行われる状況を一覧にしている。表 2 の「市区町村コードの変更」における記号は、● が必ず市区町村コードが変更される場合、▲ は市区町村コードの変更がない市区町村もある場合、✖ は市区町村コードの変更が行われない場合に対応する。

合併には新設と編入の 2 つの方法があり、どちらの場合も少なくとも 1 つ以上の市町村において市区町村コードの変更が行われる。編入合併では、編入を受け入れる市町村では市区町村コードの変更が行われず、編入を行う市町村では市区町村コードの変更が行われる。新設合併では、新たな市町村名とともに新設される場合のみ、合併を行うすべての市区町村コードが変更される。しかしながら、新設合併の場合であっても、既存の市区町村名を引き継ぐ場合は該当する市町村のコードが継続して利用される³⁾。

図 3 では、宮城県の市町村合併に伴う行政区域の変更を表している。図 3(a)、(b) とともに同じ地図であるが、表示されている市区町村名の時点が異なっている。ここでは仙台市の合

²⁾ 市区町村コードの詳細は、総務省 (2007) の「全国地方公共団体コード仕様」の文書を参照のこと。データ提供時に市区町村コード 5 桁ではなく 6 桁で提供される場合もある。6 桁目は検査コードとされており、データ分析上で使用する必要性はない。

³⁾ 既存の市町村名を引き継ぐ新設合併の場合、市区町村コードの変更という意味では編入合併の場合と全く同じになる。しかしながら、行政手続き上は新設合併と編入合併は全く異なるため、両者の違いを区別するように注意する必要がある。

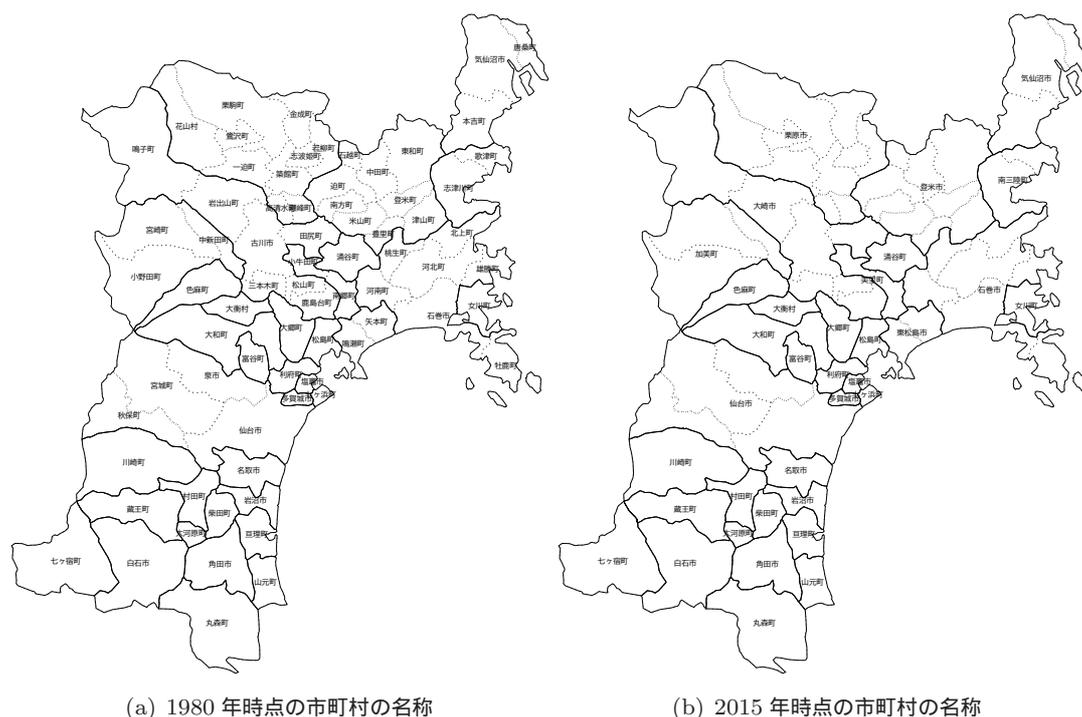


図 3: 宮城県の行政区域の変遷

注) 著者作成。実線は 2015 年時点の行政区域，点線は 1980 年時点の行政区域を表す。全都道府県の地図については Online Appendix として掲載している。

併に着目する。図 3 で示されているように，2015 年時点の仙台市 (04100) は，1980 年時点の仙台市 (04201)，泉市 (04210)，秋保町 (04382)，宮城町 (04405) の 4 区域から構成される。仙台市 (04201) は，1987 年と 1988 年に編入合併を経験しているが，編入合併のため仙台市の市区町村コードは 04201 のまま変更はなく，宮城町，泉市，秋保町の市区町村コードが 04201 へ変更される。

表 1 で示したように，1980 年から 2015 年にかけて政令指定都市の数は 10 から 20 に上昇している。この場合は必ず市区町村コードが変更され，下 3 桁が市の 201-299 番から 100 番台へ変更される。先ほどの仙台市 (04201) の例では，1988 年の編入合併の後，1989 年 4 月 1 日に政令指定都市へ移行するため，仙台市は 04201 から 04100 へ変更される。

町から市へ移行する市制施行の場合も同様に必ず市区町村コードの変更が行われ，下 3 桁が 301-799 番から 201-299 番への変更となる。ただし，村から町へ移行する町制施行の場合，市区町村コードの変更は行われず同じ値が使われる。

町村では郡の区域の新設もしくは変更があった場合は市区町村コードの変更が行われる。この場合は町村名の変更はないので一見すると市区町村コードの変更に気づかないが，郡の変更が理由で市区町村コードの変更が行われる。市区町村の名称変更のみの場合は市区町村コードの変更は行われない。表 2 の一つのみではなく，複数が同時に起こることもあるので注意する必要がある。

表 2 市区町村コードの変更

分類	市区町村コードの変更	説明
新設合併	▲	新設される市区町村コードに変更される。ただし、新設であっても既存の市区町村名を引き継ぐ市区町村は市区町村コードの変更はなし
編入合併	▲	編入を行う市区町村は編入先の市区町村コードに変更される。編入を受け入れる市区町村では市区町村コードの変更はなし。
政令指定都市施行	●	市区町村コード(3桁)が100番台へ変更
市制施行	●	町村から市の変更に伴い、市区町村コード(3桁)が301-799番から201-299番へ変更
町制施行	✕	市区町村コードの変更なし
名称変更	✕	市区町村コードの変更なし
区域変更	●	郡の区域の新設や変更の場合、市区町村コード(3桁)の変更あり

注)「市区町村コードの変更」の記号は、●が必ず市区町村コードが変更される場合、▲は市区町村コードの変更がない市区町村もある場合、✕は市区町村コードの変更が行われない場合を表す。ここでは分割、分離、廃止等については省略している。詳細は総務省(2007)を参照のこと。既存の市町村名を引き継ぐ新設合併の場合、市区町村コードの変更という観点からは編入合併の場合と区別がつかなくなる。しかしながら、行政手続き上は新設合併と編入合併は全く異なるため、両者の違いを区別するように注意する必要がある。

市区町村コンバータに含まれる情報とは、これまでの市区町村合併とともに市区町村コードがどのように変更されてきたのかという履歴である⁴⁾。コンバータを「2つ以上のデータを接続するための対応表」と表現した意味は、対象期間中であればどの時点の市区町村データであっても、市区町村番号の変更履歴を持つコンバータを仲介させることで異時点間でも共通の市区町村コードを追加できるということである。その共通コードを用いれば2つ以上のデータが接続できるようになる。

3 市区町村コンバータの作成方法

市区町村コンバータの作成方法について解説する。コンバータの最小限の構成要素として、「接続キー」と「パネルID」を含む必要がある。市区町村コンバータにおいて、接続キーは、(1)コンバータ内で一意のコードであること、(2)接続キーによって対象期間中のどの時点のデータであっても接続できることの2つの条件を満たすように作成している。本稿では対象とする期間は1980年から2020年までなので、上記の2つの条件を満たす接続キーとは、廃止された市区町村を含む1980年から2020年までに存在した全市区町村コードの集合となる。ここで重要な点として、廃止された市区町村コードは欠番処理されているため、過

⁴⁾ 市区町村コードの変更履歴の一覧は、総務省統計局(2019)のe-Statの「統計分類・調査項目>市区町村・コード>廃置分合等情報を探す」より調べることができる。本稿で作成した市区町村コンバータは、ここでの変更履歴の情報に基づいている。

去に使用された市区町村コードが再利用されることはない。本コンバータには 3,627 件の接続キーがある。

次の作業として、接続キーの市区町村について、対象期間内の市区町村コードの変遷を記述する。本稿では国勢調査を基準にして市区町村コンバータを作成しているため、1980 年から 2020 年までの 5 年毎に市区町村コードの変遷を記述している。

コンバータの接続キーの 1 レコードは、1980 年から 2020 年までに存在していた 1 つの市区町村と対応する。コンバータには、その市区町村が期間中にどのような市区町村コードの履歴を持っているのかがわかるよう情報を追加する。例えば、ある市が 1980 年には存在しており、2020 年までの間に一回も市区町村コードの変更の経験していなければ期間中は同じ市区町村コードが続いていることを意味する。次に、ある市が 2006 年に合併等の理由によって市区町村コードの変更が生じたという状況を考える。この場合、1980 年から 2005 年までの従来の市区町村コードが続き、2010 年から 2020 年までは新たな市区町村コードが付与される。さらに、もし 2006 年に新設合併により新たな市区町村コードとともに新設された市を考える。この場合、市区町村コードの変遷は 2 レコードによって構成される。まず 1 レコード目には 1980 年から 2005 年まで従来の市区町村コードが記録され、2010 年から 2020 年までは空白が続く。2 レコード目には 1980 年から 2005 年までが空白で、2010 年から 2020 年までは新設された市区町村コードが付与される。

コンバータ内の市区町村コードのみでは、具体的にどの市区町村かすぐに判別できないため、同時に市区町村名の変数を含めている。また市区町村コードの変更が行われたときに、どのような事由によるのかも知る必要がある。表 2 の区分と対応するように、コンバータ内に情報を含めている。

表 3 は、コンバータにおける変数の一覧を示している。接続キーとなる変数は `merge_id_muni` とし、1980 年から 2020 年までの市区町村コードに対して接続できるようになっている。パネル ID は、`id_muni[yyyy]` (`[yyyy]` は調査年を表す) になる。例えば、2015 年時点の市区町村単位に集計したい場合は、接続キーを用いて `id_muni2015` を原データに付与することになる。必要に応じて市区町村名の変数 `name_muni[yyyy]` (`[yyyy]` は調査年を表す) も付与することができる。

コンバータの具体的な例として、岐阜県における市町村合併について紹介する。図 4 は、岐阜県における市町村合併を通じた行政区域の変遷が地図上で表されている。実線は 2015 年時点の行政区域、点線は 1980 年時点の行政区域を表している。

ここでは岐阜県東濃地方にある中津川市の 2005 年 2 月 13 日における編入合併と合わせて、コンバータのレイアウトの具体例を紹介する。図 4 より、2015 年時点の中津川市 (21206) は、1980 年時点の中津川市 (21206)、坂下町 (21561)、川上村 (21562)、加子母村 (21563)、付知町 (21564)、福岡町 (21565)、蛭川村 (21566)、長野県山口村 (20431) との編入合併により構成されている。なお岐阜県中津川市は長野県との県境を越えた編入合併を経験している。この市区町村コードの変遷がどのようにコンバータ上で表されているのかが表 4 において示されている。コンバータの 2005 年とは 2005 年 10 月 1 日時点であり、市区町村コードの変更は 2005 年 2 月 13 日に行われているので、2005 年から市区町村コードおよ

表 3 コンバータに含まれる変数

変数名	説明
merge_id_muni	各データとの接続キー．コンバータ内で一意の市区町村コード
id_muni1980	1980年10月1日時点における市区町村コード(5桁)
id_muni1985	1985年10月1日時点における市区町村コード(5桁)
id_muni1990	1990年10月1日時点における市区町村コード(5桁)
id_muni1995	1995年10月1日時点における市区町村コード(5桁)
id_muni2000	2000年10月1日時点における市区町村コード(5桁)
id_muni2005	2005年10月1日時点における市区町村コード(5桁)
id_muni2010	2010年10月1日時点における市区町村コード(5桁)
id_muni2015	2015年10月1日時点における市区町村コード(5桁)
id_muni2020	2020年10月1日時点における市区町村コード(5桁)
name_muni1980	1980年10月1日時点における市区町村の名称
name_muni1985	1985年10月1日時点における市区町村の名称
name_muni1990	1990年10月1日時点における市区町村の名称
name_muni1995	1995年10月1日時点における市区町村の名称
name_muni2000	2000年10月1日時点における市区町村の名称
name_muni2005	2005年10月1日時点における市区町村の名称
name_muni2010	2010年10月1日時点における市区町村の名称
name_muni2015	2015年10月1日時点における市区町村の名称
name_muni2020	2020年10月1日時点における市区町村の名称
note_merge_consol	新設合併 (consolidation-type merger) に関する事由
note_merge_incorp	編入合併 (incorporation-type merger) に関する事由
note_seirei	政令指定都市施行に関する事由
note_city	市制施行に関する事由
note_town	町制施行に関する事由
note_change_zone	区域変更に関する事由
note_change_name	名称変更に関する事由

注) コンバータでは 2020 年国勢調査 (2020 年 10 月 1 日時点) までのパネル化に対応している．コンバータは日本語と英語で提供しているが、英語版では事由に関して省略されており、note が flag となっており、市区町村コードの変更が区分に該当するなら 1、しなければ 0 という変数へ変更している．

び市区町村名の変更が反映されるようになっている．

コンバータの接続キーを用いれば対象期間中であればどの時点の市区町村データとも接続できるようになっている．ただし市区町村パネルデータを作成する際は注意点もある．文字通り合併ということは、現在の 1 つの市区町村が過去の複数の市区町村から構成されるということである．したがって、市区町村パネルデータを作成する際には、対象期間の最終年以降の市区町村コードを用いてパネル化しなければならない．1980 年から 2020 年までの市区町村パネルデータの作成であれば、2020 年時点の市区町村コードを用いる必要がある．1980 年から 2005 年までの市区町村パネルデータの作成であれば、2005 年、2010 年、2015

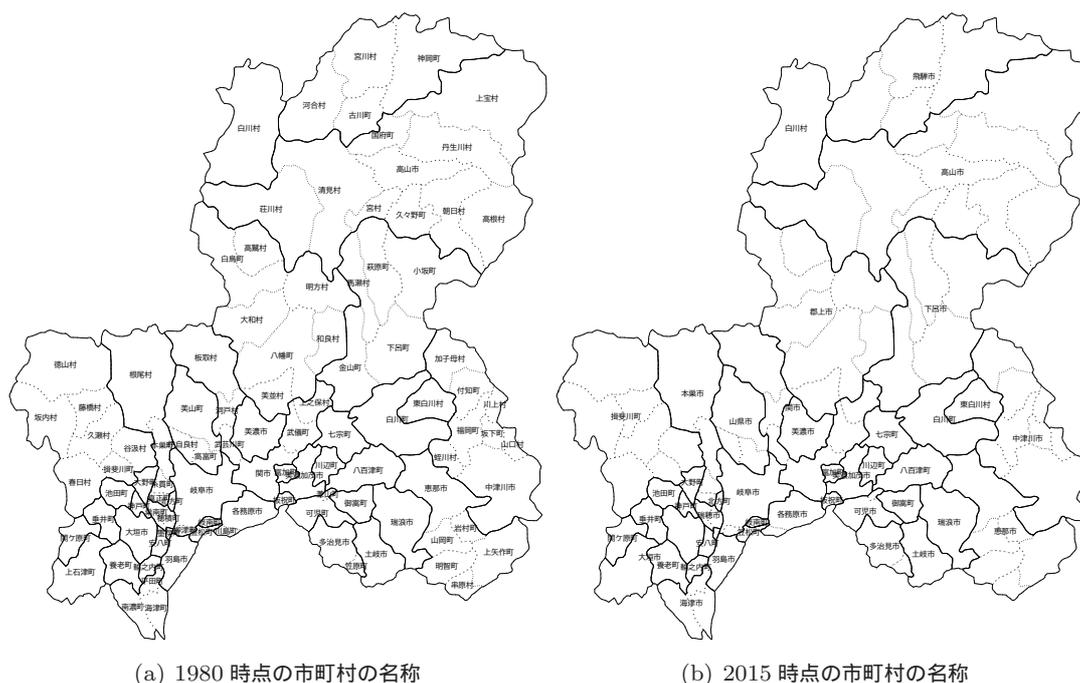


図 4: 岐阜県における合併に伴う市町村の区域の変更

注) 著者作成．実線は 2015 年時点の行政区域，点線は 1980 年時点の行政区域を表す．なお山口村 (20431) は 1980 年時点では長野県に属するが，2005 年 2 月 13 日に岐阜県中津川市 (21206) に県境を越えて編入合併している．

年，2020 年の市区町村コードから 1 つを選択することになる⁵⁾．

4 市区町村コンバータによる市区町村パネルデータの作成方法

図 1 でも示したように，ここでは市区町村コンバータを用いた市区町村パネルデータの作成手順を整理しておく．1980 年から 2020 年までの各年の市区町村データを持っているが，市区町村コードが調査時点のコードになっているとする．そして，2020 年時点の市区町村の地理単位に基づいて過去の市区町村データを再集計することで 1980 年から 2020 年まで

⁵⁾ 問題は，合併と分割が同時に起こった場合に市区町村コンバータでは単純に対応できないことである．この事例は 1 件のみ存在し，2006 年 3 月 1 日に山梨県上九一色村 (19341) が分割され山梨県甲府市 (19201) と山梨県富士河口湖町 (19430) に編入合併している事例である．これを市区町村単位のままコンバータ内で記述すると山梨県上九一色村の接続キーが重複してしまい，もし単純に 2015 年時点の市区町村コードで過去のデータを再集計すると山梨県上九一色村のデータが甲府市と富士河口湖町の双方で含まれるため二重集計されてしまう．本コンバータでは甲府市側への編入合併という扱いで処理しているが，甲府市と富士河口湖町において 2005 年から 2010 年にかけて部分的に断絶が生じてしまうことに注意する．もしこの処理上の問題が大きい場合の対応方法としては，2010 年以降について甲府市と富士河口湖町を 1 つの地理単位として再定義すれば 1980 年から 2015 年までで同一の地理単位として市区町村データを集計することもできる．より正確に分割をパネルデータで反映させるために町丁単位で集計されたデータを利用する．総務省統計局 (2019) の e-Stat の「統計 GIS > 統計データダウンロード」における小地域集計から山梨県上九一色村の分割地域毎にデータを集計し，それぞれ甲府市と富士河口湖町にそれぞれ足し上げる作業が必要となる．他にもメッシュデータからおおよその分割比率を計算し，過去の上九一色村のデータに適用することも考えられる．

表 4 市区町村コンバータのレイアウト（岐阜県中津川市の合併を抜粋）

merge_id_muni	id_muni1980	id_muni1985	id_muni1990	id_muni1995	id_muni2000	id_muni2005	id_muni2010	id_muni2015
21206	21206	21206	21206	21206	21206	21206	21206	21206
21561	21561	21561	21561	21561	21561	21206	21206	21206
21562	21562	21562	21562	21562	21562	21206	21206	21206
21563	21563	21563	21563	21563	21563	21206	21206	21206
21564	21564	21564	21564	21564	21564	21206	21206	21206
21565	21565	21565	21565	21565	21565	21206	21206	21206
21566	21566	21566	21566	21566	21566	21206	21206	21206
20431	20431	20431	20431	20431	20431	21206	21206	21206

merge_id_muni	name_muni1980	name_muni1985	name_muni1990	name_muni1995	name_muni2000	name_muni2005	name_muni2010	name_muni2015
21206	岐阜県 中津川市							
21561	岐阜県 恵那郡 坂下町	岐阜県 中津川市	岐阜県 中津川市	岐阜県 中津川市				
21562	岐阜県 恵那郡 川上村	岐阜県 中津川市	岐阜県 中津川市	岐阜県 中津川市				
21563	岐阜県 恵那郡 加子母村	岐阜県 中津川市	岐阜県 中津川市	岐阜県 中津川市				
21564	岐阜県 恵那郡 付知町	岐阜県 中津川市	岐阜県 中津川市	岐阜県 中津川市				
21565	岐阜県 恵那郡 福岡町	岐阜県 中津川市	岐阜県 中津川市	岐阜県 中津川市				
21566	岐阜県 恵那郡 蛭川村	岐阜県 中津川市	岐阜県 中津川市	岐阜県 中津川市				
20431	長野県 木曾郡 山口村	岐阜県 中津川市	岐阜県 中津川市	岐阜県 中津川市				

注) merge_id_muni は、原データとの接続に用いる接続キーである。id_muni の後ろの数字は年数を表しており、各年 10 月 1 日時点における市区町村コード(5桁)と対応する(2020年の市区町村コードは掲載を省略)。国勢調査の市区町村データを前提しているため、国勢調査の調査日である10月1日に対応させている。岐阜県中津川市は、県境を越えた編入合併を2005年2月13日に行っている。編入合併した町村は2005年国勢調査では岐阜県中津川市として調査対象となっているため、id_muni2005では岐阜県中津川市(21206)へ市区町村コード(5桁)の変更がコンバータ上で行われている。都道府県と市区町村名の間は半角空白で分割されている。なお町村の場合は郡名を追加している。

の市区町村パネルデータを作成するという状況を想定する。市区町村コンバータを用いた操作方法の手順は、以下の通りである。

1. 過去の市区町村データの市区町村コードとコンバータの接続キー (`merge_id_muni`) を用いて、過去の市区町村データに 2020 年時点の市区町村コード (`id_muni2020`) を付与する。
2. 過去の市区町村データにおいて、2020 年時点の市区町村コード (`id_muni2020`) を基準にデータの集計を行う。
3. 過去の市区町村データにおいて、2020 年時点の市区町村コード (`id_muni2020`) に関して重複削除を行う。2020 年時点の市区町村コードと再集計されたデータのみを残したデータを保存する。なお 1980 年時点の市区町村コードおよび原データは不要なので削除しておく。
4. 過去の必要な年次だけ上記の手順を繰り返す。
5. 作成された過去データを統合すれば、2020 年時点の市区町村コードと調査年に基づいて市区町村パネルデータが作成される。

作成方法の直感的なイメージが持てるように、図 5 では Stata のコマンドの一例を紹介している。実際に作業する際は、各自の状況に応じてデータ名や変数名を適宜変更する必要がある。

5 市区町村コンバータの正確性の検証

本節では、作成した市区町村コンバータの正確性について検証する。コンバータが正しく作成されていれば、コンバータを用いて再集計された結果を信頼して分析に用いることができる。どのようにコンバータの正確性の検証するのかというと、信頼性の高い第三者が独立に作成したデータと同じ結果になるのかどうかを比較するという方法で正確性の判断を行う。第三者が作成したデータが仮に正しいとして、ここでは解答データと呼ぶことにする。この解答データとして、内閣府 (2015) の「選択する未来」委員会において作成された市区町村パネルデータが一般利用可能になっており、このデータを検証用に使用する。コンバータを用いて独自に再集計した結果が解答データと同じであればコンバータが正しく作成されていると判断する。なお内閣府 (2015) のデータにおいて検証で利用する項目は、国勢調査による人口総数と工業統計調査の製造品出荷額の 2 つである。表 5 において、検証内容の一覧をまとめている。

まず国勢調査における人口総数の検証結果は、2010 年の山梨県甲府市の人口が異なっていたこと以外は、すべての市区町村で 1980 年から 2010 年の間で同一数値が得られる。脚注 5) で議論したように、2006 年 3 月 1 日に山梨県上九一色村 (19341) は編入合併しているが、「大字梯および古関」が甲府市 (19201) に、「大字精進、本栖および富士ヶ嶺」が富士河口湖町 (19430) へ分割されて編入している。市区町村コンバータでは市区町村が分割される場合への対応ができないため、本コンバータ上では山梨県上九一色村を甲府市側に含めているこ

```

*****
** 【Stata でのコンバータ使用例のイメージ】
** データファイル名: muni_pop[yyyy].dta ([yyyy] は調査年を表す)
** (1980年から2010年まで5年毎の市区町村データ)
** データにおける変数: year id_muni pop
** year: 調査年
** id_muni: 調査年時点の市区町村コード
** pop: 人口総数
*****

=====
** 各年データの再集計
=====
** 5年毎のループ処理をする
forvalues i = 1980(5)2010 {
    ** 市区町村データ
    use "muni_pop`i`.dta", clear

    ** コンバータの接続キーを作成する
    gen merge_id_muni = id_muni

    ** 接続キーを用いてコンバータ内の2015年時点の市区町村コードと市区町村名を追加
    merge 1:1 merge_id_muni using "Converter.dta", /*
        */ keepusing(id_muni2015 name_muni2015)

    ** 原データ側にしかなかった不要な情報を削除(接続漏れがないか念のため確認)
    drop if _merge == 1

    ** コンバータ側にしかなかった不要な情報を削除
    drop if id_muni2015 == .

    ** 2015年時点の市区町村コードによって再集計する
    by id_muni2015, sort: egen totalpop = total(pop)

    ** 重複を削除する
    duplicates drop id_muni2015, force

    ** 必要な変数のみを保存
    keep year id_muni2015 name_muni2015 totalpop

    ** データの整形
    replace year = `i`
    sort id_muni2015
    order year id_muni2015 name_muni2015 totalpop

    ** 2015年基準で再集計した過去データの保存
    save "muni_pop`i`_base2015.dta", replace
}

=====
** 過去データを統合し, long型としてパネルデータを構築
=====
** メモリ上のデータを削除
clear

** 5年毎のループ処理で, long型のパネルデータを構築
forvalues i = 1980(5)2010 {
    append using "muni_pop`i`_base2015.dta.dta"
}

** パネルデータとして保存する
save "muni_pop_panel1980-2010.dta", replace

```

図 5: 市区町村コンバータを用いた市区町村パネルデータの作成例 (Stata のコード例)

注) 著者作成。コードの書き方の一例であり, 各自の状況に応じて適宜修正が必要である。

表 5 コンバータの正確性を検証する際に使用するデータ

項目	検証するデータ：「国勢調査」(総務省統計局)，人口総数
対象期間	1980–2010 年 (5 年毎)
解答データ出所	内閣府，「選択する未来」委員会，「市区町村別 人口・経済関係データ」のウェブページ上で公開されている市区町村パネルデータ (内閣府, 2015)。
原データ出所	総務省統計局 (2019) の e-Stat における国勢調査 (第 1 次基本集計，都道府県編，都道府県・市部・郡部・市区町村)。
検証結果	山梨県上九一色村の合併の扱いにより内閣府データとコンバータ再集計で差が生じる。それ以外は全ての年次で完全に一致する。

項目	検証するデータ：「工業統計調査」(経済産業省)，製造品出荷額等
対象期間	2001–2012 年 (毎年)
解答データ出所	内閣府，「選択する未来」委員会，「市区町村別 人口・経済関係データ」のウェブページ上で公開されている市区町村パネルデータ (内閣府, 2015)。
原データ出所	経済産業省 (2019) のウェブサイト上で公開されている各年の工業統計調査の市区町村編の集計結果。
検証結果	内閣府データでは千葉県大網白里市が非掲載である一方で，コンバータ再集計では千葉県大網白里町の市制施行 (2013 年 1 月 1 日) の情報を反映しているため，2012 年工業統計調査の千葉県大網白里町のデータを引き継ぐ。それ以外の差は金額の単位の丸めの誤差から生じていると思われる。内閣府データは単位が百万円，原データの単位は万円である。

注) 上段は国勢調査の人口総数に関する検証内容，下段は工業統計調査の製品出荷額等に関する検証内容を示している。検証に用いた Stata コード及びデータはオンライン上で公開している。

とが違いの要因である⁶⁾。

次に工業統計調査の製造品出荷額等の検証結果について述べる。唯一の大きな違いは，2012 年の千葉県大網白里市の扱いである。内閣府データの 2012 年を見ると千葉県大網白里市のデータが非掲載になっているが，経済産業省に掲載されている 2012 年工業統計調査の市区町村編の集計結果では千葉県大網白里町としてデータが含まれている。2013 年 1 月 1 日に千葉県大網白里町は千葉県大網白里市へ市制変更しているため，市区町村コードの変更が行われているが，本コンバータでは市制施行後の接続に対応していることから違いが生じている。他に製造品出荷額等の数値に小さな違いが生じているが，金額の単位の丸めの誤差から結果の違いが生じていると考えられる。内閣府データは単位が百万円として四捨五入されて掲載されており，経済産業省の原データの集計単位は万円に掲載されており，集計の段階の四捨五入が原因と思われる⁷⁾。

以上より，本稿で作成した市区町村コンバータを用いることで，市町村合併前後でも同一の地理単位に基づいた信頼性の高い市区町村パネルデータが作成できると考えられる。既に内閣府 (2015) でも市区町村パネルデータが提供されているため，利用したい項目があれば改めてパネル化する必要はない。しかしながら，状況に応じて新たな項目が必要になる場合，市区町村コンバータを利用することで新たな変数についても容易に市区町村パネルデー

⁶⁾ ただし現在公開されているバージョン (平成 27 年 1 月 22 日更新) の内閣府データでは山梨県上九一色村の人口が山梨県富士河口湖町にも含まれておらず，集計時に除かれている可能性がある。

⁷⁾ なお工業統計調査の製造業出荷額等では，山梨県上九一色村がそもそも経済産業省の原データに含まれていないため人口総数と同様の問題は生じていない。

タを構築できるようになり、分析の幅を広げることができる。

6 結語

本稿では、1980年から2020年までの市区町村パネルデータを作成するためのコンバータの作成方法について解説した。政府統計の集計データの公表は市区町村単位で公表されることが多いが、市区町村合併により地理的単位が変化するため、長期的に同一の地理単位でデータの時系列比較をすることが困難になっていた。市区町村コンバータを各種の市区町村データに適用することにより、1980年から2020年までの地理単位を統一した市区町村パネルデータを容易に作成できるようになる。また、個票データを用いた分析や市区町村合併を自然実験とした実証分析においても適用することができる。市区町村コンバータが地域経済分析の今後の発展に大きく寄与することが期待される。

参考文献

経済産業省 (2019) 「工業統計調査」。

<http://www.meti.go.jp/statistics/tyo/kougyo/index.html>。

総務省統計局 (2019) 「e-Stat 政府統計の総合窓口」。

<https://www.e-stat.go.jp/> (2019年2月20日確認)。

内閣官房まち・ひと・しごと創生本部 (2019) 「RESAS 地域経済分析システム」。

<https://resas.go.jp/> (2019年2月20日確認)。

内閣府 (2015) 「『選択する未来』委員会 市区町村別 人口・経済関係データ」。

https://www5.cao.go.jp/keizai-shimon/kaigi/special/future/keizai-jinkou_data.html (2019年2月20日確認)。

総務省 (2007) 「全国地方公共団体コード仕様」, 4月。

http://www.soumu.go.jp/main_content/000137948.pdf (2019年2月20日確認)。