

Discussion Paper Series A No.631

賃金構造基本統計調査に対するデータリンケージの可能性について

村田磨理子
((公財)統計情報研究開発センター)
伊藤伸介
(中央大学経済学部)

2015年12月

Institute of Economic Research
Hitotsubashi University
Kunitachi, Tokyo, 186-8603 Japan

賃金構造基本統計調査に対するデータリンケージの可能性について

村田 磨理子 ((公財)統計情報研究開発センター)

伊藤 伸介 (中央大学経済学部)

概要

本稿は、企業の業績や財務内容が雇用や賃金に及ぼす影響についての実証分析を目指して、異なる統計調査の調査票情報の統合とパネルデータ化の検証を試みたものである。本稿では、最初に、賃金構造基本統計調査について、共通の事業所番号を用いて複数の調査年をリンケージしたパネルデータを作成するだけでなく、共通の事業所番号に基づいて厚生労働省の賃金構造基本統計調査と経済産業省企業活動基本調査をリンケージしたデータの作成を行った。さらに、名称・住所等を利用した賃金構造基本統計調査と法人企業統計調査の名寄せによる完全照合だけでなく、企業規模等の属性を用いた統計的照合を行い、比較・検証を行った。最後に、賃金構造基本統計調査に基づくリンケージデータの可能性を述べた。

賃金構造基本統計調査の調査票情報は、2種類の調査票（事業所票と個人票）に対応した2つのファイルから構成される。同一の調査年の事業所票と個人票は、都道府県番号と事業所一連番号によって照合可能である。しかしながら、事業所一連番号は、調査年ごとに改めて設定されるため、同じ事業所であっても、調査年によって一連番号は異なる。同様に、労働者の一連番号も、調査年ごとに改めて設定されている。なお、事業所の名称及び所在地、労働者の氏名は、調査票情報（以下、個票データと呼称）には収録されていない。

異なる調査年の個票データを照合するために、母集団（抽出枠）である事業所・企業統計調査または経済センサスの事業所番号を利用することができる。ただし、抽出枠は数年おきに切り替わり、抽出枠が同じ場合は、同一の事業所番号によって、直接リンケージを行うことが可能であるが、抽出枠の切り替えの前後の年次においてリンケージを行う場合には、新旧事業所番号の対応情報も併せて必要になる。本稿では、抽出枠が平成18年事業所・企業統計調査及び平成21年経済センサス基礎調査に対応する期間をリンケージしたパネルデータの作成について述べている。賃金構造基本統計調査では、複数年をリンケージできた事業所に関しては、企業規模が大きい事業所が比較的多くなる等の特徴が見られる。

なお、賃金構造基本統計調査の調査票情報は、抽出枠である事業所・企業統計調査及び経済センサス基礎調査と事業所番号によってリンケージされることから、これらの調査結果も併せて利用することができる。したがって、賃金構造基本統計調査の個票データにない項目として、例えば、資本金額、本所・支所の別、本所や親会社に関する情報を得ることが可能である。

次に、企業の収益や財務内容等の企業属性を得るため、本稿では、企業活動基本調査及び法人企業統計調査を利用して、リンケージの可能性について検証を行った。

企業活動基本調査は、独自の永久企業番号を保持していることと、平成23年以降の個票データに経済センサスの事業所番号が収録されていることから、前述の賃金構造基本統計調査との間でリンケージすることが比較的容易である。ただし、両調査の対象の範囲（産業、

規模)に違いがあることから、リンケージできた事業所(企業)は企業活動基本調査の1割程度とあまり多くない。さらに、調査単位に関して事業所と企業という違いがあるため、例えば、複数事業所からなる企業の場合、1つの事業所(貸金構造基本統計調査)と企業全体(企業活動基本調査)がリンケージされることに注意した上で、リンケージされたデータを利用する必要がある。

一方、法人企業統計調査は、経済センサスの事業所番号のような共通の識別番号を持たないため、名称・住所・電話番号等による名寄せによって、貸金構造基本統計調査との間で完全照合を行った。統計的照合では、「本社所在地」、「業種」、「規模」の3つの共通項目を使用しており、名寄せによる完全照合と比較して、正解率を確認した。その結果、企業規模が大きいほど正解率が高くなること、産業大分類別にみると、「電気・ガス・熱供給・水道業」、金融業、保険業」はほぼ半数が正解だが、「宿泊業、飲食サービス業」など4業種はすべて不正解であることなどが明らかになった。

1. はじめに

わが国では、2009年における統計法の全面施行以降、政府統計の個票データを用いた実証研究が広範に行われているが、その一方で、同種の統計調査の個票データにおける縦断的なリンケージや、複数の統計調査の個票データにおけるデータリンケージに関する研究も行われてきた¹。わが国における政府統計のマイクロデータ間の照合(マッチング)に関する研究は、1970年代における松田(1978)の研究に遡ることができるが、その後、総務省統計局が実施する家計調査と貯蓄動向調査におけるリンケージに関する研究が行われただけでなく(美添・荒木(1999))、経済産業省企業活動基本調査(以下「企業活動基本調査」と総務省の科学技術研究調査のリンケージ等、異種のマイクロデータ間のデータリンケージが行われてきた(松田(1999, 200頁))。

ところで、わが国の統計調査は、主に、世帯・人口系の統計調査と事業所・企業系の統計調査に類別されるが、一方では個人や世帯、他方では企業や事業所が調査対象であることから、世帯・人口系の個票データと事業所・企業系の個票データのリンケージは極めて困難である。また、世帯・人口系のデータ同士、ないしは事業所・企業系のデータ同士のリンケージに関しても、リンケージの対象となる統計調査の個票データにおいて共通の直接的な識別子が存在しなければ、データリンケージは容易でないのが現状である。

そうした状況の中で、周防ほか(2009)及び古隅(2015)は、事業所・企業統計調査と法人企業統計調査の個票データをもとに、住所、企業名や電話番号を用いた名寄せによるデータリンケージの精度を検証し、データリンケージの注意点を指摘している。その一方で、異種の統計調査において識別番号によるリンケージが可能であったとしても、リンケージされたデータの特性を踏まえた上で、実証分析を行うことが必要かと考える。

本稿では、企業の業績や財務内容が雇用や賃金に及ぼす影響についての実証分析を目指して、賃金構造基本調査を例に、異なる統計調査の調査票情報の統合とパネルデータ化の検証を試みる。

2. 賃金構造基本統計調査におけるパネルデータの作成について

本節では、賃金構造基本統計調査におけるパネルデータの作成方法の概要について述べる。最初に、賃金構造基本統計調査の個票データの特徴を概括する。つぎに、賃金構造基本統計調査の個票データをもとにしたパネルデータの作成方法を述べる。最後に、作成された賃金構造基本調査のパネルデータにおけるデータ特性を明らかにする。

¹ マイクロデータ間の照合(マッチング)は、完全照合(exact matching)と統計的照合(statistical matching)に大別される。美添等によれば、完全照合とは、「同じ調査客体(統計単位)を対象にした別々の調査から得られる一連のデータについて、一連世帯番号などの個体識別番号を利用して照合を行い、データファイルを再構築すること(美添・荒木(2000,239頁))」であり、統計的照合は、「異なったファイルに存在する類似の個体(統計単位)に関するデータを連結することによって、結合された情報を含むファイルを作成すること(美添・荒木(1999,11頁))」と定義されている。

(1) 賃金構造基本統計調査の概要

賃金構造基本統計調査は、主要産業に雇用される労働者の賃金の実態について、労働者の性別、年齢、学歴等の人口社会的属性や、雇用形態、就業形態、職種、勤続年数、経験年数といった就業属性別に把握することを目的とした基幹統計である。調査の対象は、全国の5人以上の常用労働者を雇用する民営事業所（5～9人の事業所については、企業規模が5～9人の事業所に限定）、及び10人以上の常用労働者を雇用する公営事業所で日本標準産業分類(平成19年11月改定)に基づく16大産業〔C:鉱業,採石業,砂利採取業、D:建設業、E:製造業、F:電気・ガス・熱供給・水道業、G:情報通信業、H:運輸業,郵便業、I:卸売業,小売業、J:金融業,保険業、K:不動産業,物品賃貸業、L:学術研究,専門・技術サービス業、M:宿泊業,飲食サービス業、N:生活関連サービス業,娯楽業、O:教育,学習支援業、P:医療,福祉、Q:複合サービス事業、R:サービス業(他に分類されないもの)〕である。層化二段抽出法によって、都道府県、産業及び事業所規模別に一定の方法で抽出した事業所が第1次抽出単位、労働者が第2次抽出単位としてそれぞれ設定されている。

賃金構造基本統計調査における調査事項は、事業所の属性、労働者の性別、雇用形態、就業形態、学歴、年齢、勤続年数、労働者の種類、役職、職種、経験年数、実労働日数、所定内実労働時間数、超過実労働時間数、きまって支給する現金給与額、超過労働給与額、調査前年1年間の賞与、期末手当等特別給与額等である。これらの調査事項をもとに、賃金構造の実態を事業所と労働者の両面において捕捉することが可能になっている。

(2) 賃金構造基本統計調査の個票データの特徴

賃金構造基本統計調査の個票データは、2種類の調査票（事業所票と個人票）に対応した2つのファイルから構成される。図1は、賃金構造基本統計調査の個票データのファイル構造に関するイメージ図を示したものである。事業所票データは、1事業所の情報が1レコードに収録され、個人票データは、労働者1個人の情報が1レコードに収録されている。

同一の調査年の事業所票と個人票は、都道府県番号と事業所一連番号によって照合できる。なお、事業所の名称と所在地、及び労働者の氏名は、個票データには収録されていない。

本稿の分析に使用した平成20～25年の個票データのレコード数は、表1に示されている。事業所票のレコード(事業所)数が、約55000～57000レコードであるのに対して、個人票のレコード(個人)数は、約120万～130万レコードとなっている。

(3) 異なる調査年のデータのリンケージについて

事業所一連番号は、調査年ごとに改めて設定されることから、同じ事業所であっても調査年によって異なる番号が付与される。同様に、労働者の一連番号も調査年ごとにレコードに与えられることから、労働者の一連番号に基づいて同一個人を特定することはできない。

一方で、賃金構造基本統計調査は、事業所・企業統計調査または経済センサスをサンプリングフレーム（抽出枠）として抽出されている。抽出枠における事業所の識別番号(事業

所番号)は、賃金構造基本統計調査の個票データに収録されている。ただし、抽出枠は数年ごとに更新される。調査年との対応表は、表2で示されている。

賃金構造基本統計調査独自の事業所一連番号は、毎年改めて設定されるので、複数の調査年にわたって賃金構造基本統計調査の個票データをリンケージするためには利用できない。一方、抽出枠である事業所・企業統計調査または経済センサス基礎調査は、識別番号である事業所番号を保持していることから、抽出枠が同じ場合は、抽出枠における事業所番号によってリンケージを行うことが可能である。例えば、平成20～23年は平成18年事業所・企業統計調査の事業所番号によってリンケージを行うことができ、平成24、25年は平成21年経済センサス基礎調査の事業所番号に基づいて、賃金構造基本統計調査とのリンケージが可能である。

他方、異なる抽出枠においては、事業所番号は必ずしも継続されない。しかしながら、事業所・企業統計調査及び経済センサス基礎調査の個票データには、その時点の識別番号だけでなく、当該事業所の前回調査における識別番号も保持しているので、それを経由してリンケージを行うことが可能になる。図2は、抽出枠における事業所番号の収録状況に関するイメージ図を示したものである。

(4) 事業所票のパネルデータ化

図2で示されるデータの構造を踏まえて、事業所票のパネルデータ化を実際に試みる。ここでは、平成20～25年のそれぞれについて、同一事業所について当年と前年のデータをリンケージした。抽出枠が同一か異なるかの違いから、次のような処理が必要になる。

- 平成25年と24年のリンケージには、平成21年経済センサス基礎調査の事業所番号を利用する
- 平成24年と23年のリンケージに当たっては、まず平成21年経済センサス基礎調査の個票データを使って、平成21年事業所番号と平成18年事業所番号の対応表を作成する。作成した対応表によって、事業所番号を変換してリンケージする。
- 平成23年と22年、22年と21年、21年と20年は、それぞれ、平成18年事業所・企業統計調査の事業所番号を利用する

表3は、リンケージできたデータの事業所数と、その事業所に含まれる個人票の人数を示したものである。

(5) 事業所の属性の追加

前述のとおり、賃金構造基本統計調査の調査票情報は、抽出枠である事業所・企業統計調査及び経済センサス基礎調査と事業所番号を介してリンケージを行うことが可能である。それによって、事業所・企業や経済センサスの個票データに含まれる調査事項も併せて利用することができる。具体的には、経済センサス基礎調査には以下のような調査事項が含まれている(甲調査の場合)。

ア 事業所に関する事項

(ア) 名称

(イ) 電話番号

- (ウ) 所在地
- (エ) 開設時期
- (オ) 従業者数
- (カ) 事業の種類
- (キ) 業態
- イ 企業に関する事項
 - (ア) 経営組織
 - (イ) 資本金等の額
 - (ウ) 外国資本比率
 - (エ) 決算月
 - (オ) 持株会社か否か
 - (カ) 親会社の有無
 - (キ) 親会社の名称
 - (ク) 親会社の所在地及び電話番号
 - (ケ) 子会社の有無及び子会社の数
 - (コ) 法人全体の常用雇用者数
 - (サ) 法人全体の主な事業の種類
 - (シ) 国内及び海外の支所等の有無及び支所等の数
 - (ス) 本所の名称
 - (セ) 本所の所在地及び電話番号

したがって、賃金構造基本統計調査と経済センサス基礎調査のリンケージを行うことによって、賃金構造基本統計調査の個票データには含まれない調査事項として、例えば、資本金額、本所・支所の別、本所や親会社に関する情報が得られる。

(6) リンケージされた事業所の分布特性

本節の最後に、リンケージされた事業所の分布特性を見ていくことにしたい。表4と表5はそれぞれ、平成25年と24年をリンケージしたパネルデータに含まれる事業所について分布特性を見たものである。各調査年の対象事業所全体と比較すると、企業規模の大きい事業所の割合が高く、企業規模の小さい事業所はかなり少ない(表4)。同様に、産業大分類別に比較すると、C(鉱業、採石業、砂利採取業)、F(電気・ガス・熱供給・水道業)、J(金融業、保険業)の割合が相対的に大きく、逆に、P(医療、福祉)についてその割合が小さくなっていることが確認できる(表5)。

3. 賃金構造基本統計調査と企業活動基本調査に関するリンケージデータの作成

賃金構造基本統計調査には、企業の業績や財務内容に関する情報が捕捉されていない。したがって、企業業績や財務内容が雇用や賃金に及ぼす影響についての実証分析を行う場合、企業の収益や財務情報に関する企業属性を得るためには、経済産業省企業活動基本調

査及び法人企業統計調査の個票データを利用してリンケージを行うことが求められる。そこで、本節では、賃金構造基本統計調査と企業活動基本調査におけるリンケージデータの特徴について述べる。

(1) 企業活動基本調査の概要

最初に、企業活動基本調査の概要を述べる。企業活動基本調査は、日本標準産業分類に掲げる大分類C－鉱業、採石業、砂利採取業、大分類E－製造業、大分類F－電気・ガス・熱供給・水道業（中分類35－熱供給業及び中分類36－水道業を除く。）、大分類G－情報通信業の一部、大分類I－卸売業、小売業、大分類J－金融業、保険業の一部、大分類K－不動産業、物品賃貸業の一部、大分類L－学術研究、専門・技術サービス業の一部、大分類M－宿泊業、飲食サービス業の一部、大分類N－生活関連サービス業、娯楽業の一部、大分類O－教育、学習支援業の一部及び大分類R－サービス業（他に分類されないもの）の一部に属する事業所を有する企業のうち、従業者50人以上かつ資本金額又は出資金額3,000万円以上の会社を対象としている。調査事項は、次のとおりである。

- 1)企業の名称及び所在地
- 2)資本金額又は出資金額
- 3)企業の設立形態及び設立時期
- 4)直近1年間の組織再編行為の状況
- 5)企業の決算月
- 6)事業組織及び従業者数
- 7)親会社、子会社・関連会社の状況
- 8)資産・負債及び純資産並びに投資
- 9)事業内容
- 10)取引状況
- 11)事業の外部委託の状況
- 12)研究開発、能力開発
- 13)技術の所有及び取引状況
- 14)企業経営の方向

企業活動基本調査における調査対象の企業には、独自の永久企業番号が割り当てられている。さらに、平成23年以降は経済センサス基礎調査の事業所番号が収録されている。したがって、企業活動基本調査の個票データと経済センサスの個票データのリンケージを行うことは容易である。本研究では、平成24年の企業活動基本調査の個票データを用いて、賃金構造基本統計調査とのリンケージの検証を行っている。

(2) 賃金構造基本統計調査と企業活動基本調査のリンケージの方法

賃金構造基本統計調査と企業活動基本調査のリンケージにおいて留意すべき点としては、この2調査における調査対象の範囲（産業、規模）の違いや、調査の単位（事業所か企業か）の相違が指摘できる。特に、調査の単位の違いについては、例えば、複数事業所

からなる企業の場合、1つの事業所（賃金構造基本統計調査）と企業全体（企業活動基本調査）がリンケージされることから、リンケージデータの利用に当たっては、その点に注意する必要がある。

リンケージの手順は、以下のとおりである。最初に、平成21年経済センサス基礎調査の事業所番号をリンケージのキー変数として用いることによって、賃金構造基本統計調査の事業所票データ、経済センサス基礎調査の個票データ、及び企業活動基本調査の個票データをリンケージする。つぎに、賃金構造基本統計調査の事業所・企業単位のリンケージデータを賃金構造基本統計調査の個人票データにリンケージすることによって、賃金構造基本統計調査における個人に関する属性を追加的に付与する。なお、リンケージにおいては、①平成24年賃金構造基本統計調査の個票データに対するデータリンケージ、及び②平成24年と平成25年の賃金構造基本統計調査のパネルデータに対するデータリンケージが行われた。

（3）リンケージデータの特徴

リンケージデータの事業所・企業の分布特性は、表6～表9で示されている。表6を見ると、平成24年賃金構造基本統計調査と平成24年企業活動基本調査においてリンケージができた事業所・企業の本数は3,166レコードであって、企業活動基本調査全体の1割程度と少ないことがわかる。さらに、平成24年と平成25年の賃金構造基本統計調査のパネルデータにおいて、平成24年企業活動基本調査とリンケージできた事業所・企業の本数は、1,313レコードであることが確認された。また、リンケージデータにおける勤め先企業の規模は、100～299人が最も多くなっている（表7）。産業別に見ると、G情報通信業に該当する企業の割合が相対的に大きくなっている（表8）。また、単一事業所企業が少ないことも確認できる（表9）。

さらに、賃金構造基本統計調査における現金給与額の分布（図4、表10）をみると、平成24年と平成25年の賃金構造基本統計調査のそれぞれの分布特性や、平成24年と平成25年の2時点間のパネルデータにおける分布において、2つのピークが存在する。それに対して、企業活動基本調査とリンケージした事業所に限定した分布については、下側のピークがないことから、分布特性について違いが見られる。

（4）リンケージの課題

本稿でのリンケージの方法では、複数事業所企業の場合、原則として賃金構造基本統計調査における本所事業所のデータのみが、企業活動基本調査における企業全体のデータとリンケージされることになる。実際には、同一企業の本所と支所の両方に賃金構造基本統計調査のレコードが含まれる場合や、支所みのレコードが存在することが考えられる。したがって、リンケージの前に本所・支所の統合を行うことで、リンケージできる範囲が拡大する可能性が考えられる。

また、複数の調査間で重複している項目（例えば、産業、資本金額、従業員数）において、定義の違い等に起因するずれが生じていることが確認された。分析目的にあった選択・統合が必要である。

4. 賃金構造基本統計調査と法人企業統計調査の名寄せによる完全照合の結果と統計的照合の評価

企業の収益や財務内容等の企業属性に関しては、法人企業統計調査を利用して分析を行うことも考えられる。賃金構造基本統計調査とのリンケージにおいては、経済センサスの事業所番号のような共通の識別番号を持たないため、当初は統計的照合を試みた。その後、名称、所在地の文字情報や本所・支所の情報を用いて、個票データの完全照合を行うことができた。本節では、賃金構造基本統計調査と法人企業統計調査の名寄せによる完全照合の状況と、統計的照合に関する評価結果について述べる。

(1) 法人企業統計調査の概要

法人企業統計調査は、法人の企業活動の実態を明らかにし、併せて法人を対象とする各種統計調査のための基礎となる法人名簿を整備することを目的に、財務省が統計法に基づく基幹統計調査として実施された調査である。営利法人等を対象とする標本調査（年次別調査と四半期別調査）であって、平成 20 年度調査からは、「金融業、保険業」が調査対象に含まれている。なお、調査事項は、資産・負債・純資産の状況や損益等について確定した決算の計数である。

資本金 1 億円以上の管理法人は、基本的に、調査年が変わっても同一法人には同じ法人番号が割り当てられており、パネルデータ化が比較的容易である。しかしながら、他の統計調査とのリンケージについては、経済センサス等の共通の識別番号を持たないため、名称、所在地等による名寄せが必要となる。

(2) 賃金構造基本統計調査と法人企業統計調査の個票データとの完全照合

本研究では、平成 23 年（2011 年）のデータを使用して、賃金構造基本統計調査と法人企業統計調査の個票データの完全照合に関する検証を行う。前述のとおり、2 つの調査に共通の ID は存在しない。また、賃金構造基本統計調査の個票データには事業所の名称、住所がないことから、抽出枠である平成 18 年事業所・企業統計調査における企業（ないしは事業所）の名称と住所を用いる。その際、支所・支社・支店に関しては、本社の名称・住所を用いた。また、名寄せによる完全照合を行う上で、名称は「(株)」・「株式会社」を除去し、さらに余分な空白の除去やカナ・英数字を全角に変換する等の処理（正規化）が適用され、住所は市区町村レベルまでの一致で判定を行うために用いられた。

表 11 と表 12 は、名寄せによる賃金構造基本統計調査と法人企業統計調査の個票データの完全照合の結果を示したものである。検証結果から、賃金構造基本統計調査の 9,337 事業所が法人企業統計調査の 3,191 企業とリンケージすることが明らかになった。法人企業統計調査における企業規模、業種をみると、約 8 割は企業規模 100 人以上の企業であり

（表 11）、企業の産業大分類別で見ると、「E 製造業」、「I 卸売業、小売業」、「J 金融業、保険業」に該当する企業が多いことが明らかになった（表 12）。

(3) 賃金構造基本統計調査と法人企業統計調査における統計的照合の評価

本研究では、賃金構造基本統計調査と法人企業統計調査を対象に、完全照合ではなく、統計的照合の評価も行うことによって、賃金構造基本統計調査と法人企業統計調査の統計的照合の精度検証も行っている。本研究では、2調査に共通して利用できる3項目、すなわち、本社所在地、業種、企業規模によるマッチングによって統計的照合を行った。本研究においては、本社所在地については市区町村までを使い、業種に関しては原則として法人企業統計の業種の分類を用いた。さらに、企業規模については原則として賃金構造基本統計調査の企業常用労働者数の階級区分を使用している。なお、賃金構造基本統計調査において検証対象となるレコードについては、民営事業所のうち、単独事業所・本社事業所に限定した。

図5は、本社所在地、業種、企業規模の組み合わせが同一のグループ別の企業数の分布を示したものである。本社所在地、業種と企業規模別にグループ分けを行うと、法人企業統計調査の対象となるレコードは18,772グループ、賃金構造基本統計調査の対象となるレコードは21,374グループにそれぞれ分けることができる。一方、1グループ当たりの企業数の分布を確認すると、それぞれの調査においてグループの中の8割は1企業のみで構成されていることがわかる。

つぎに、賃金構造基本統計調査と法人企業統計調査におけるグループ別企業数の対応関係を見たのが、表13である。表13では、同じ属性のグループがマッチングされている。表13を見ると、一方しか該当企業がないグループが多く、「1企業」対「1企業」でマッチングしたグループに含まれる企業数は、約2,000と少ない。

そこで、「1企業」対「1企業」でマッチングした約2,000組について、完全照合の結果と比較し、統計的照合の「正解率」を確認したのが、図6である。図6からは、企業規模が大きいほど正解率が高くなることがわかる。また、産業大分類別にみると、「電気・ガス・熱供給・水道業」、「金融業、保険業」においては、マッチングされた組のほぼ半数が正解だが、「宿泊業、飲食サービス業」など5業種では、それらのすべてが不正解であることが確認できた。

5. おわりに

本稿では、賃金構造基本統計調査の個票データにおけるデータリンケージの可能性を追究した。本研究から明らかになったように、事業所・企業統計調査の事業所番号を介して、賃金構造基本統計調査と企業活動基本調査のリンケージが可能である。本稿において議論された賃金構造基本統計調査と企業活動基本調査のリンケージデータをもとに、企業業績や財務情報が企業の個別労働者における賃金率に及ぼす影響が実証的に明らかにされている(伊藤・出島(2015))。その一方で、賃金構造基本統計調査と企業活動基本調査のリンケージデータについては、企業規模の大きな企業が照合されやすい傾向になるため、リンケージデータによる実証分析を行う場合には、セレクションバイアスの可能性について留意する必要があると思われる。

他方、本研究では、賃金構造基本統計調査と法人企業統計調査を対象に、名称・住所・電話番号等による名寄せによる完全照合を行うだけでなく、「本社所在地」、「業種」と「規模」に基づく統計的照合を行い、完全照合と比較することによって、統計的照合の正解率を確認

した。その結果、企業規模が大きいほど正解率が高くなること、産業大分類別にみると、産業によって正解率が大きく異なることがわかった。

謝辞

本研究は、平成 26 年度一橋大学経済研究所共同利用共同研究拠点事業プロジェクト研究「企業の業績と雇用政策が家計の就業・資産選択に及ぼす影響に関する計量分析」（研究代表者：北九州市立大学 林田実）及び平成 27 年度日本学術振興会科学研究費補助金「家計の資産選択や労働供給に関する政策評価のためのマイクロシミュレーション分析」〔基盤研究(C)(課題番号 15K03399)〕（研究代表者：中央大学 伊藤伸介）における研究成果の一部を発表するものである。

本研究において使用した「賃金構造基本統計調査」、「事業所・企業統計調査」、「経済センサス基礎調査」、「企業活動基本調査」及び「法人企業統計調査」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ます。

参考文献

- 伊藤伸介・出島敬久(2015)「企業業績が個別労働者の賃金に与える効果に関するマイクロデータ分析—企活と賃金センサスのデータリンケージをもとにして—」, 『経済学論纂』(中央大学)56 巻, 1・2 合併号, 13~37 頁
- 周防節雄・古隅弘樹・宮内環(2009)「法人企業統計調査と事業所・企業統計調査の統合データによる企業データベース：1983~2005 年」, 『統計数理』, Vol.57, No.2, 277~303 頁
- 古隅弘樹(2015)「公的産業統計を組み合わせた遡及的データベースの構築」, 『産業構造の変容と公的統計の利用（中間報告）』, 統計数理研究所共同研究レポート 337, 5~23 頁
- 松田芳郎(1978)『データの理論』岩波書店
- 松田芳郎(1999)『マイクロ統計データの描く社会経済像』日本評論社
- 美添泰人・荒木万寿夫(1999)「家計調査と貯蓄動向調査の統計的マッチング」(財)統計研究会『平成 10 年度 統計的マッチングにより発生する誤差の要因等の検証に関する調査研究報告書』, 9~97 頁
- 美添泰人・荒木万寿夫(2000)「完全照合」松田芳郎・伴金美・美添泰人編『講座マイクロ統計分析②：マイクロ統計の集計解析と技法』日本評論社, 239~250 頁

図1 賃金構造基本統計調査の個票データのファイル構造に関するイメージ図
事業所票のデータファイル

都道府 県番号	事業所一連番 号	産業分 類	規模	...
01	00001			
01	00002			
01	00003			
01	00004			
01	00005			
01	00006			
01	00007			

個人票データファイル

都道府 県番号	事業所一 連番号	労働者一 連番号	性 別	雇用形態	...
01	00001	01			
01	00001	02			
01	00001	03			
01	00001	04			
01	00002	01			
01	00002	02			
01	00002	03			

表1 賃金構造基本統計調査の個票データのレコード数

	事業所票のレコード (事業所)数	個人票のレコード (個人)数
平成 25 年	57,744	1,297,471
平成 24 年	57,426	1,285,162
平成 23 年	54,779	1,202,220
平成 22 年	55,338	1,225,079
平成 21 年	55,775	1,220,418
平成 20 年	53,434	1,205,421

表2 賃金構造基本統計調査の調査年とサンプリングフレーム(抽出枠)の対応

調査年	抽出枠
平成 16 年	平成 13 年事業所・企業統計調査
平成 17 年	
平成 18 年	平成 16 年事業所・企業統計調査
平成 19 年	
平成 20 年	平成 18 年事業所・企業統計調査
平成 21 年	
平成 22 年	
平成 23 年	
平成 24 年	平成 21 年経済センサス基礎調査
平成 25 年	

図2 抽出枠における事業所番号の収録状況（イメージ）

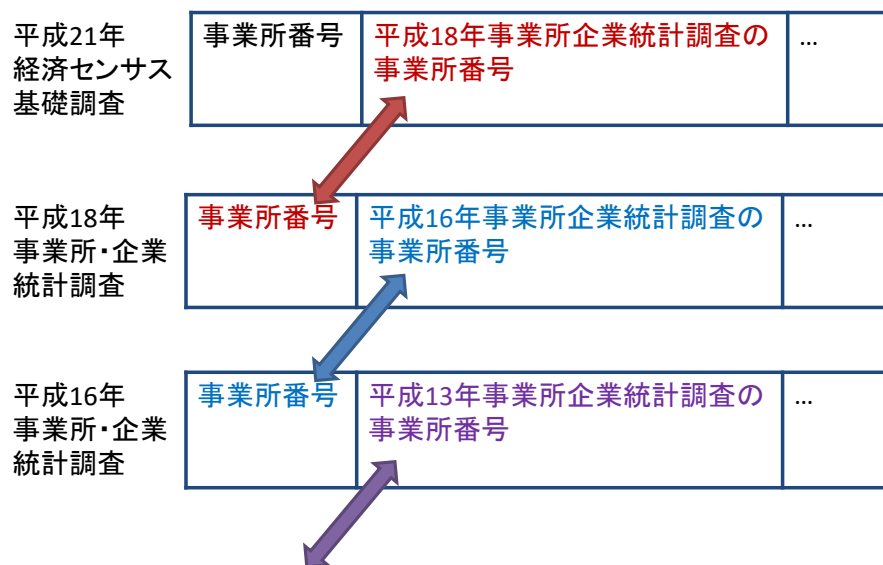


図3 抽出枠が異なる場合のリンケージの例

平成24年のデータ				平成23年のデータ			
都道府県 番号	事業所 一連番 号	...	平成21年経済セン サス基礎調査の 事業所番号	都道府県 番号	事業所 一連番 号	...	平成18年事業所・ 企業統計調査の 事業所番号
01	00001		0000000000001	01	00001		0000000000001
01	00002		0000000000003	01	00002		0000000000004
01	00003		0000000000007	01	00003		0000000000006

事業所番号対応 表（平成21年経済 センサス基礎調査 の調査票情報 データから作成）	平成21年経済セン サス基礎調査の 事業所番号	平成18年事業所・ 企業統計調査の 事業所番号
	0000000000001	0000000000003
	0000000000002	0000000000005
	0000000000003	0000000000006

事業所番号の
組み合わせに
従って接続す
る

3

表3 賃金構造基本統計調査のパネルデータ化における同一事業所のリンケージ状況

リンケージした 調査年	リンケージされた事 業所数	リンケージされた事業所に 含まれる個人票の人数
平成 25, 24 年	12,758	358,185
平成 24, 23 年	8,377	242,817
平成 23, 22 年	12,899	350,020
平成 22, 21 年	12,303	345,451
平成 21, 20 年	10,193	296,684

表4 賃金構造基本統計調査のパネルデータにおける企業規模別事業所数

	平成 25 年	平成 24 年	平成 25, 24 年パネ ル
5,000 人以上	8,297	8,819	2,323
1,000～4,999 人	7,211	7,757	2,133
500～999 人	3,760	3,613	1,026
300～499 人	2,881	2,841	703
100～299 人	7,978	7,653	2,250
30～99 人	9,414	9,010	1,867
10～29 人	10,243	9,854	1,274
5～9 人	7,960	7,879	1,182
計	57,744	57,426	12,758

企業規模

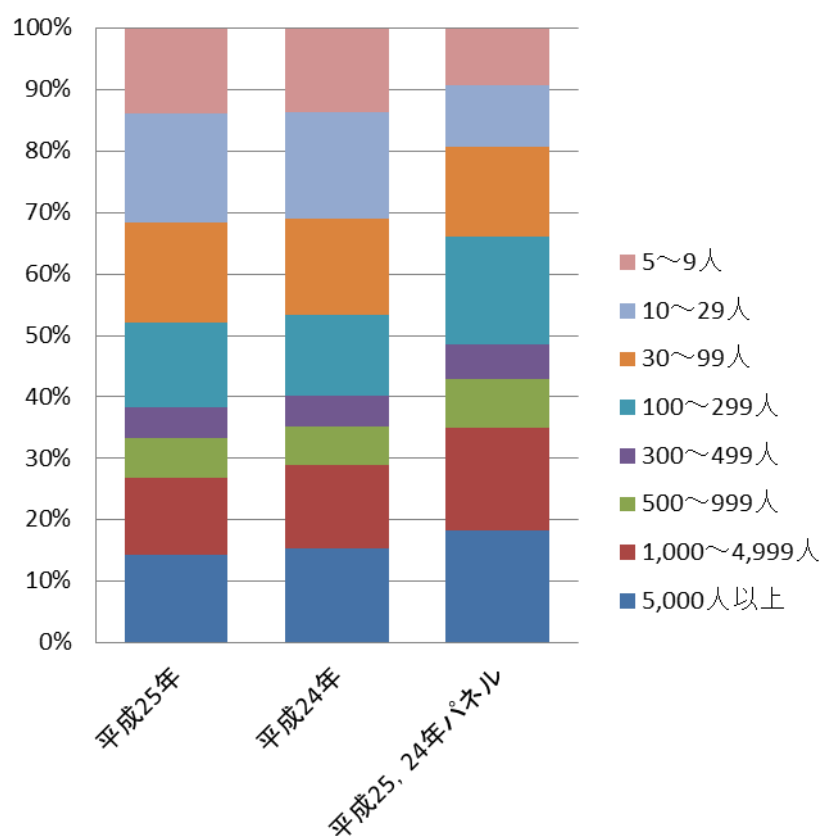
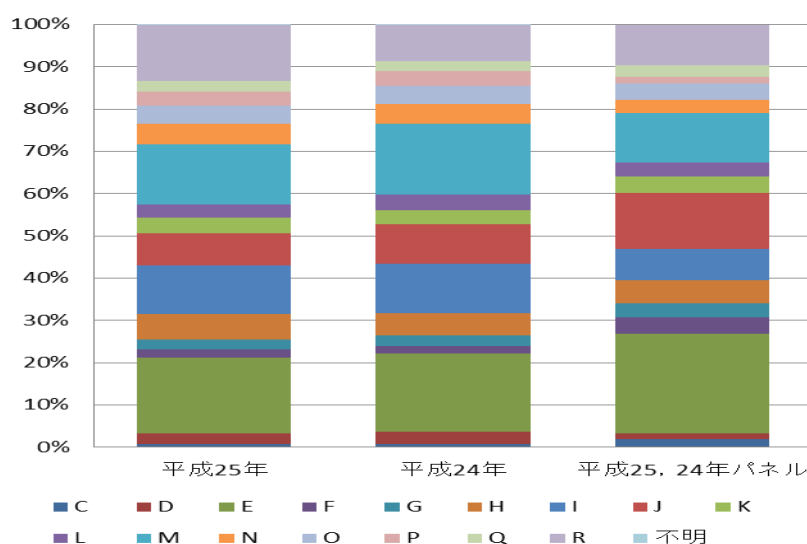


表5 賃金構造基本統計調査のパネルデータにおける産業大分類別事業所数

	平成25年	平成24年	平成25, 24年パネル
C	426	400	238
D	1,496	1,671	193
E	10,348	10,619	2,987
F	1,062	1,111	513
G	1,444	1,398	415
H	3,439	2,995	685
I	6,681	6,730	948
J	4,376	5,350	1,704
K	2,139	1,932	498
L	1,709	2,119	422
M	8,280	9,664	1,482
N	2,765	2,615	394
O	2,466	2,471	502
P	1,935	2,030	193
Q	1,501	1,319	358
R	7,665	4,978	1,226
不明	12	24	0
計	57,744	57,426	12,758

産業大分類



注：C: 鉱業，採石業，砂利採取業 D: 建設業 E: 製造業 F: 電気・ガス・熱供給・水道業
 G: 情報通信業，H: 運輸業，郵便業，I: 卸売業，小売業，J: 金融業，保険業 K: 不動産業，物品賃貸業
 L: 学術研究，専門・技術サービス業，M: 宿泊業，飲食サービス業，N: 生活関連サービス業，娯楽業
 O: 教育，学習支援業，P: 医療，福祉，Q: 複合サービス事業，R: サービス業(他に分類されないもの)

表6 データリンケージの状況

データの種類	リンケージされた事業所数(企業数)	リンケージされた事業所に含まれる個人票の人数
企活(平成24年)	30,647	-
賃金(平成24年)+企活(平成24年)	3,166	121,366
賃金(平成25,24年パネル)+企活(平成24年)	1,313	53,626

注 「賃金」は賃金構造基本統計調査、「企活」は企業活動基本調査をそれぞれ表している。以下同様。

表7 賃金構造基本統計調査と企業活動統計調査のリンケージデータの特徴
企業規模別企業数

	賃金(平成25年)	賃金(平成24年)+企活(平成24年)	賃金(平成25,24年パネル)+企活(平成24年)
5,000人以上	8,297	124	68
1,000~4,999人	7,211	372	201
500~999人	3,760	396	190
300~499人	2,881	315	125
100~299人	7,978	1,359	565
30~99人	9,414	590	161
10~29人	10,243	10	3
5~9人	7,960	0	0
計	57,744	3,166	1,313

表8 賃金構造基本統計調査と企業活動統計調査のリンケージデータの特徴
産業大分類別企業数

	賃金（平成 25 年）	賃金（平成 24 年）+ 企活（平成 24 年）	賃金（平成 25, 24 年パネル）+ 企活（平成 24 年）
C	426	20	15
D	1,496	47	17
E	10,348	1,677	672
F	1,062	61	50
G	1,444	204	103
H	3,439	13	9
I	6,681	566	197
J	4,376	39	24
K	2,139	55	26
L	1,709	99	39
M	8,280	143	55
N	2,765	106	43
O	2,466	12	6
P	1,935	0	0
Q	1,501	1	1
R	7,665	123	56
不明	12	0	0
計	57,744	3,166	1,313

表9 賃金構造基本統計調査と企業活動統計調査のリンケージデータの特徴
単一事業所企業・複数事業所企業の別企業数

	賃金（平成 24 年）+ 企活（平成 24 年）	賃金（平成 25, 24 年パネル）+ 企活（平成 24 年）
単一事業所企業	653	268
複数事業所企業	2,513	1,045
計	3,166	1,313

図4 賃金構造基本統計調査と企業活動統計調査のリンケージデータにおいて「決まって支給する現金給与額」の分布

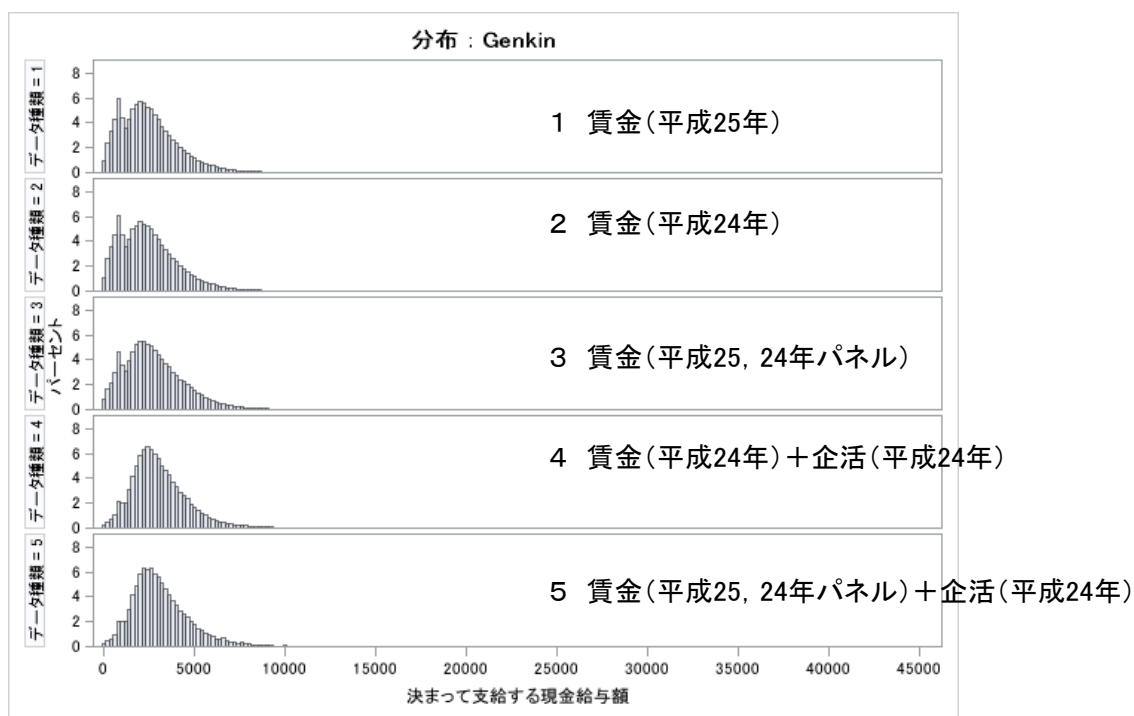


表 10 賃金構造基本統計調査と企業活動統計調査のリンケージデータにおいて「決まって支給する現金給与額」の基本統計量

	平均	標準偏差	歪度	尖度	第1四分位数	中央値	第3四分位数
賃金(平成25年)	2,499	1,654	1.513	6.801	1,300	2,259	3,356
賃金(平成24年)	2,494	1,676	1.461	6.044	1,250	2,250	3,365
賃金(平成25, 24年パネル)	2,797	1,732	1.330	5.071	1,592	2,531	3,725
賃金(平成24年)+企活(平成24年)	3,084	1,596	1.470	5.627	2,038	2,825	3,873
賃金(平成25, 24年パネル)+企活(平成24年)	3,146	1,625	1.253	3.272	2,058	2,868	3,950

表 11 賃金構造基本統計調査と法人企業統計調査のリンケージデータの特徴
企業規模別企業数

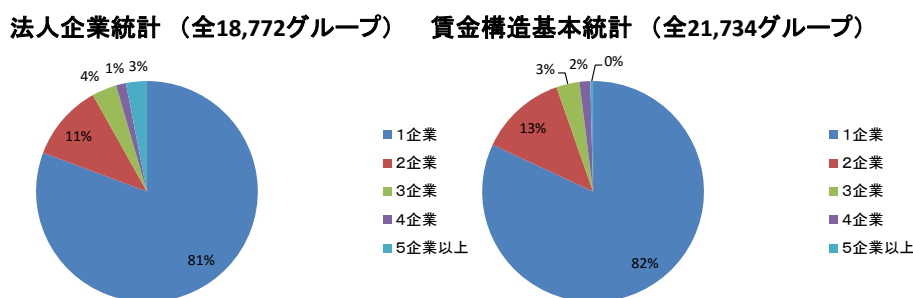
	賃金（平成 23 年）＋法企（平成 23 年）
5,000 人以上	229
1,000～4,999 人	746
500～999 人	467
300～499 人	349
100～299 人	691
30～99 人	382
10～29 人	202
5～9 人	94
4 人以下、不詳	31
計	3,191

注 「法企」は法人企業統計調査を表している。以下同様。

表 12 賃金構造基本統計調査と法人企業統計調査のリンケージデータの特徴
産業大分類別企業数

	賃金（平成 23 年）＋法企（平成 23 年）
A,B 農林水産業	3
C 鉱業，採石業，砂利採取業	23
D 建設業	154
E 製造業	1,098
F 電気・ガス・熱供給・水道業	69
G 情報通信業	128
H 運輸業，郵便業	203
I 卸売業，小売業	457
J 金融業，保険業	448
K 不動産業，物品賃貸業	179
L 学術研究，専門・技術サービス業	53
M 宿泊業，飲食サービス業	149
N 生活関連サービス業，娯楽業	95
O 教育，学習支援業	18
P 医療，福祉	5
R サービス業（他に分類されないもの）	109
合計	3,191

図5 本社所在地、業種、企業規模の組み合わせが同一のグループ別の企業数の分布



7

表13 賃金構造基本統計調査と法人企業統計調査におけるグループ別企業数の対応関係

		賃金構造基本統計調査 1グループ当たり企業数						
		なし	1企業	2企業	3企業	4企業	5企業以上	合計
法人企業統計調査 1グループ当たり 企業数	なし	-	15,020	2,110	502	208	183	18,023
	1企業	12,482	<u>1,981</u>	385	136	70	96	15,150
	2企業	1,566	336	113	26	26	33	2,100
	3企業	461	120	43	21	14	21	680
	4企業	177	47	17	16	6	7	270
	5企業以上	375	89	58	25	13	12	572
	合計	15,061	17,593	2,726	726	337	352	36,795

図6 統計的照合の正解率

