

総務省統計委員会担当室 委託業務

賃金関連統計の比較検証に 関する調査研究

報告書

令和3年3月

株式会社メトリクスワークコンサルタンツ

目次

第1章	調査研究の概要	1
1.1.	本調査研究の背景と目的	1
1.1.1.	背景	1
1.1.2.	目的	1
1.1.3.	調査項目	1
1.2.	本調査研究の概要	2
1.2.1.	本調査研究で扱う賃金統計調査の概要	2
1.2.2.	実施体制	3
第2章	調査研究の結果	5
2.1.	調査対象の比較	5
2.1.1.	本調査項目の目的	5
2.1.2.	調査方法・結果	6
2.1.3.	調査結果からの示唆	15
2.2.	調査方法の影響（サンプルの偏りの有無）	17
2.2.1.	本調査項目の目的	17
2.2.2.	調査方法・結果	17
2.2.3.	調査結果からの示唆	23
2.3.	調査対象の回答状況	25
2.3.1.	本調査項目の目的	25
2.3.2.	調査方法・結果	25
2.3.3.	調査結果からの示唆	41
2.4.	行政記録情報の活用可能性の検討	42
2.4.1.	本調査項目の目的	42
2.4.2.	調査方法・結果	42
2.4.3.	調査結果からの示唆	45
2.5.	母集団情報の比較検証	47
2.5.1.	本調査項目の目的	47
2.5.2.	調査方法・結果	47
2.5.3.	調査結果からの示唆	54
2.6.	品質固定された賃金指数の試算	56
2.6.1.	本調査項目の目的	56
2.6.2.	調査方法・結果	56
2.6.3.	調査結果からの示唆	63
第3章	調査研究のまとめ	64
	引用文献	67

第1章 調査研究の概要

1.1. 本調査研究の背景と目的

1.1.1. 背景

公的統計の整備に関しては、新統計の検討や複数統計にまたがる技術的・横断的課題の整理、行政記録情報の活用など、様々な取組が進められている。そうした中、近年の社会構造の変化などを受けて、特に雇用・賃金情勢に対する関心が高まっており、雇用・賃金関連統計を用いた分析に対するニーズが高まっている。

統計ユーザーの利便性を向上させるうえで、類似する統計間の整合性の検証や相違点の明確化が非常に重要となっており、統計委員会においても、2019年に行われた賃金関連統計の一部改正に係る審議の際に委員などから示された指摘などを踏まえ、答申の「今後の課題」の中で統計精度の検証や類似統計との比較などが求められている。また、雇用・賃金関連統計の精度向上や体系整備に向けた課題の整理が求められている。

1.1.2. 目的

上記の背景を踏まえ、本調査研究では毎月勤労統計調査、賃金構造基本統計調査、民間給与実態統計調査の調査票情報などを用いた比較検証や精度検証などを中心に、各統計の類似点・相違点、および利用に当たっての留意点などを整理することにより、統計ユーザーの利便性向上を図ることを目的とする。

1.1.3. 調査項目

上記目的のために、本調査研究では以下の6点についての検討を行う（それぞれが2.1から2.6に対応する）。

- 調査対象の比較：複数の統計調査間の属性情報などの調整を行ったうえで、各統計の比較を行い、差異の有無および要因の検証を行う。
- 調査方法の影響（サンプルの偏りの有無）：調査対象となる事業所内における抽出の偏りの有無の検証を行う。
- 調査対象の回答状況：調査への回答状況について、事業所属性と回答率の関係を検証し、実態を明らかにするとともに、複数回調査対象になった事業所を用いた回答パターンについての検証を行う。
- 行政記録情報の活用可能性の検討：行政記録情報活用の一例である源泉徴収票等情報転記ツールについて、ツールの利用状況の整理、および回答状況に及ぼす影響の検証を行う。
- 母集団情報の比較検証：統計調査の対象（母集団）に含まれない労働者の存在が、賃金統計に及ぼし得る影響について検討を行う。

- 品質固定された賃金指数の試算：労働者の属性を制御した（品質固定した¹）賃金指数の算出を行ったうえで、賃金変動の要因についての分析を行う。

1.2. 本調査研究の概要

1.2.1. 本調査研究で扱う賃金統計調査の概要

本調査研究では、毎月勤労統計調査、賃金構造基本統計調査、民間給与実態統計調査を中

表 1.1.1 本調査研究で扱う賃金関連統計調査の概要

統計調査名	基幹統計調査			一般統計調査
	毎月勤労統計調査	賃金構造基本統計調査	民間給与実態統計調査	(参考) 職種別民間給与実態調査
管轄省庁	厚生労働省	厚生労働省	国税庁	人事院
調査対象給与	毎月分給与	6月分給与	前年年収	4月分給与
調査時期	翌月	7月	1月上旬-2月	4月下旬-6月央
公表時期	速:翌々月初 確:翌々月下旬	概要:翌年3月 詳細:翌年6月	概要:9月 詳細:11月	概要:8月 詳細:11月
母集団	- 5人以上の常用労働者を雇用する事業所 常用労働者	- 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人の事業所に限る) - 10人以上の常用労働者を雇用する公営事業所 全労働者(常用労働者+臨時労働者)	- 源泉徴収義務者(民間事業所)に勤務する給与所得者(所得税納税有無を問わない) - 従事員1人以上の事業所 給与所得者	- 企業規模50人以上かつ事業所規模50人以上の事業所 常時勤務する従業員のうち期間を定めず雇用されている者(公務と類似する職種に該当する者に限る)
母集団フレーム	事業所母集団データベース(約133万事業所)	事業所母集団データベース(約133万事業所、約4,000万人)	K S Kシステム源泉徴収義務者名簿(約350万か所)	独自の母集団名簿(約5.8万事業所、約400万人)
サンプルサイズ	約3.3万か所	約7.8万か所 約165万人	約2.9万か所 約32万人	約1.2万か所 約50万人

¹ ここでいう「品質固定」とは労働者属性を固定することを意味しており、同じ属性の労働者の中での異質性については制御できない点は留意が必要である。

表 1.2.1 本調査研究で扱う賃金関連統計調査の概要（続き）

統計調査名	基幹統計調査			一般統計調査
	毎月勤労統計調査	賃金構造基本統計調査	民間給与実態統計調査	(参考) 職種別民間給与実態調査
抽出率	【事業所】 (30人以上) 業種別 - 30-99人：1/2-1/256 - 100-499人：1/1-1/144 - 500人以上：1/1 (5-29人) 第一段：全国約7万の調査区を5層に分け各層ごとに、所定の抽出率によって抽出。 第二段：産業別に標本事業所を無作為抽出	【事業所】 都道府県、産業、事業所規模別 - 5-9人：3.0% - 10-29人：5.0% - 30-99人：9.2% - 100-499人：19.9% - 500-999人：56.2% - 1,000-4,999人：65.1% - 5,000-14,999人：90.4% - 15,000人以上：100% 【労働者】 - 100人以上：産業、事業所規模別 - 100人未満：事業所規模別	【事業所】 - 1-9人：1/400 - 10-29人：1/200 - 30-99人：1/60 - 100-499人：1/15 - 500-999人：1/3 - 1,000-4,999人：1/1 - 5,000人以上：1/1 - 本社：1/1 【給与所得者】 - 1-9人：1/1 - 10-29人：1/2 - 30-99人：1/6 - 100-499人：1/20 - 500-999人：1/100 - 1,000-4,999人：1/200 - 5,000人以上：1/200（上限100人） - 本社：1/20 - 年間給与額2千万円超は全数調査	【事業所】 規模に関わらず一定

心に用いて分析を行った。加えて、職種別民間給与実態調査、就業構造基本調査を補足的に用いて分析を行った。本調査研究で扱う賃金関連統計調査の概要は表 1.1.1 のとおり。

1.2.2. 実施体制

本調査研究は、川口大司東京大学経済学研究科政策評価研究教育センター長をアドバイザーとし、その助言の下で、委託先である株式会社メトリクスワークコンサルタンツが実施した。実施体制は表 1.1.12 のとおり。

また、本調査研究を進めるにあたって、調査対象となる統計調査を実施する関係府省からの出席者から構成される検討会を開催し、情報提供や助言を得た。検討会への参加府省は表 1.2.3 のとおり。

表 1.2.2 調査研究の実施体制

役割	氏名	所属・役職
アドバイザー	川口 大司	東京大学経済学研究科政策評価研究教育センター長
担当部署	萩野 覚	総務省統計委員会担当室室長
	鈴木 純一	同室次長
	佐藤 正昭	総務省統計局統計情報システム管理官併任政策統括官付
	孕石 真浩	総務省統計委員会担当室室長補佐
	紺野 仁志	同室室長補佐
委託先	青柳 恵太郎	株式会社メトリクスワークコンサルタンツ代表取締役
	西野 宏	同代表取締役
	鳥谷部 貴大	同連携研究員（東京大学経済学研究科）

表 1.2.3 検討会参加府省

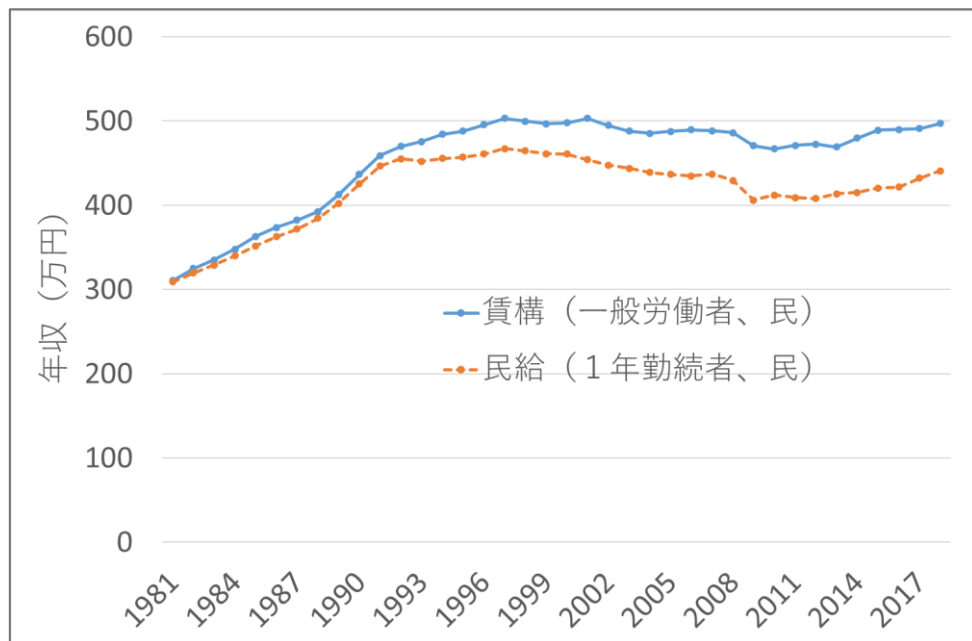
参加府省	担当部署
厚生労働省	雇用・賃金福祉統計室 賃金福祉統計室
国税庁	長官官房企画課
人事院	給与局 給与第一課

第2章 調査研究の結果

2.1. 調査対象の比較

2.1.1. 本調査項目の目的

各統計はそれぞれ集計対象となる給与、業種区分、規模、労働者の年齢・職業などの対象範囲における違いがあることから、集計される結果にも差異が生じ、統計間の直接的な比較を行うことが難しい。例えば、賃金構造基本統計調査と民間給与実態統計調査の1人当たり給与を比較すると、図 2.1.1 のように近年では前者の方が高い数値となっている。これには、賃金構造基本統計調査は常用労働者が5人以上いる事業所を対象とする一方、民間給与実態統計調査は従業員数1人以上の事業所を対象とするなど、統計間の属性の違いが影響している可能性が考えられる。こうした差異は、個票データを用いて統計間の属性情報を調整することで取り除かれることが理論的には期待されるが、抽出の偏りなどの非標本誤差が生じている場合は、属性情報をそろえたとしても統計間に差異が残存することとなる。本調査項目では、賃金構造基本統計調査をベースとして、民間給与実態統計調査と職種別民間給与実態調査とをそれぞれ比較することにより、各統計間の差異についての検証を行う。



注：賃金構造基本統計調査の年収はきまって支払われる現金給与額×12 +特別給与額で計算した。民間給与実態調査の年収は1年間勤続した人を対象に給与所得額を給与所得者数で割ったものである。

出所：川口(2013) p17 を基に総務省作成。

図 2.1.1 賃金構造基本統計調査と民間給与実態調査の年収比較 (万円)

2.1.2. 調査方法・結果

(1) 賃金構造基本統計調査と民間給与実態統計調査の比較

賃金構造基本統計調査と民間給与実態統計調査はそれぞれ以下を調査対象としているため、両統計の比較においてはまず分析対象をそろえる必要がある。

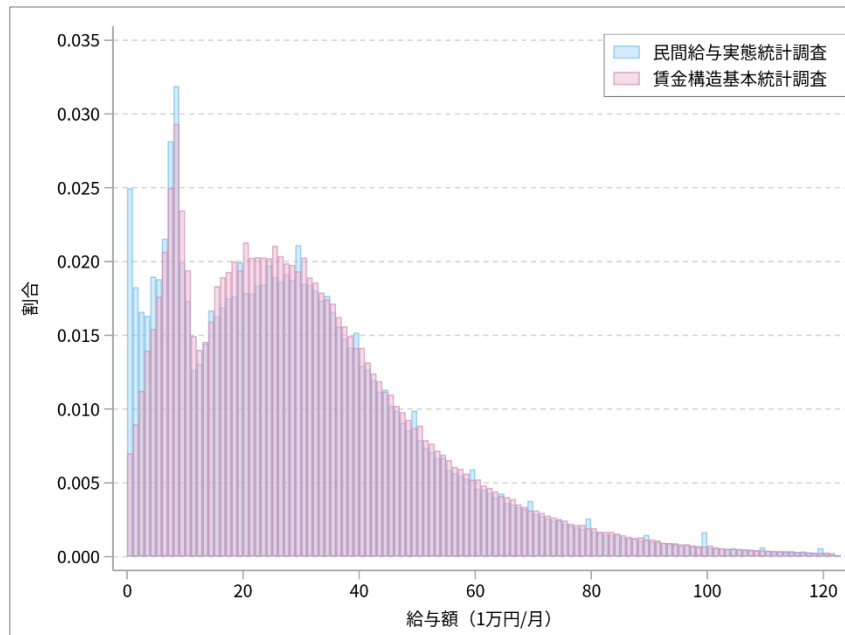
- 賃金構造基本統計調査：5人以上の民営事業所と10人以上の公営事業所の労働者
- 民間給与実態統計調査：民間の源泉徴収義務者に勤務している給与所得者

したがって、賃金構造基本統計調査は民営事業所にサンプルを限定し、民間給与実態統計調査は従業員数5人以上事業所にサンプルを限定した。このサンプルを用いて、労働者の賃金の月当たりの平均値と分布を計算した。なお、賃金構造基本統計調査は調査年の6月の給与額を調査しているのに対し、民間給与実態統計調査は年間の給与額を調査している。そのため、民間給与実態統計調査の年間給与額を12で除し²、月当たりの給与額を算出し比較を行った。ここではいずれの調査も2012年から2017年のデータをプールして使用している。

こうした調整を行ったうえで算出したそれぞれの統計における賃金の分布を図2.1.1に示す。分布からは、民間給与実態統計調査の方が5万円/月以下の労働者の割合が高いが、中所得から高所得層の割合はあまり変わらないことがわかる。民間給与実態統計調査で極端に低い値が多いのは、12か月間働いていない労働者の年間所得も12で割ってしまっていることが原因である可能性が考えられる。なお、賃金分布の形状に山が2つあるのはパートタイム労働者とフルタイム労働者がいることが原因であると考えられる。

また、表2.1.1には月当たりの賃金の平均値、および各分位点の値を示している。賃金構造基本統計調査の平均値は31.659万円/月((1)列)であるのに対し、民間給与実態統計調査の全サンプルを使った平均値は31.595万円/月((2)列)となっている。その差は640円/月程度((4)列)であり、統計的にも有意な差とはなっていない。各分位点の値を比較すると、図2.1.2でも見たように、低分位点の値は差が大きい一方、中央値やより高い分位点の値の差は比較的小さくなっていることがわかる。なお、民間給与実態統計調査においては、1年を通じて勤務した給与所得者の平均賃金などが公表されているが、1年を通じて勤務したサンプルに限定した平均値は35.846万円/月((3)列)となっており、賃金構造基本統計調査との差は4.187万円/月((5)列)に広がる。1年を通じて勤務したサンプルは雇用が安定していると考えられるため、そうでないサンプルを含む値よりも大きな値になることは自然な結果であると考えられる。

² 理想的には実際の所得を得た月数で年間給与額を除すべきだが、データには12か月間所得を得たかどうかの情報しかないので、一律に12で除すという処理をした。約85%の労働者が12か月間所得を得ている。



注：民間給与実態統計調査の分析サンプルは5人以上事業所の労働者、賃金構造基本統計調査の分析サンプルは民営事業所に限定されている。民間給与実態統計調査の所得は年間給与額/12で計算した。賃金構造基本統計調査の所得はきまって支給する現金給与額+特別給与額/12で計算した。データはいずれの調査も2012-2017年。

図 2.1.2 1月当たりの所得の分布 (万円/月)

表 2.1.1 賃金構造基本統計調査と民間給与実態統計調査の比較（万円／月）

	賃金構造基本	民間給与実態統計調査		賃金分布の比較	
	統計調査	全サンプル	1年勤続者	(2)-(1)	(3)-(1)
	(1)	(2)	(3)	(4)	(5)
平均値	31.659	31.595	35.846	-0.064	4.187
	[23.234]	[30.543]	[31.058]	(0.207)	(0.204)
10 パーセンタイル	7.230	5.342	8.850	-1.888	1.620
				(0.082)	(0.076)
25 パーセンタイル	14.993	12.583	18.467	-2.410	3.473
				(0.195)	(0.140)
50 パーセンタイル (中央値)	27.573	26.900	30.675	-0.673	3.102
				(0.175)	(0.163)
75 パーセンタイル	42.431	42.000	45.325	-0.431	2.894
				(0.259)	(0.271)
90 パーセンタイル	60.754	61.233	64.817	0.479	4.062
				(0.482)	(0.475)
観測数	7,803,201	1,771,270	1,496,204		

注：単位は「万円／月」。賃金構造基本統計調査の分析サンプルは民営事業所、民間給与実態統計調査の分析サンプルは5人以上事業所の労働者に限定されている。ブラケット内の値は標準偏差、括弧内の値は標準誤差（事業所レベルでクラスタリング）を表している。各パーセンタイル値の標準誤差は事業所クラスタリングブートストラップを用いて計算した（シミュレーション回数は1000回）。データはいずれの調査も2012-2017年。

(2) 賃金構造基本統計調査と職種別民間給与実態調査の比較

職種別民間給与実態調査の対象は常用労働者数が50人以上の事業所に従事しており、雇用期間に定めのない常勤労働者を対象としている。したがって、賃金構造基本統計調査の個票データを用いて、同様の基準を満たすサンプルに限定したうえで、職種別民間給与実態調査の公表値との比較を行った。比較に当たっては、川口（2013）と同様の調査年（2011年）および職種を用いて分析を行った³。

比較の結果を表 2.1.2 に示す。まず、大学准教授や医師など一部の職種を除き、職種別民間給与実態調査のサンプルの年齢が賃金構造基本統計調査のサンプルよりも高いことがわかる（(5)列）。次に、きまって支給される現金給与月額についても職種別民間給与実態調査の方が高い値となっている（(6)列）。ただし、比較が一部の職種に留まったこと、および対象となる属性（調査対象となる職種カテゴリーの定義⁴、給与支給月の違いなど）を十分に一致させることができなかつたため、結果については留意が必要である。

³ 川口(2013) では、賃金構造基本統計調査の個票データを利用していないため、サンプルの比較可能性を高める工夫として、賃金構造基本統計調査のうち、企業規模が1000人の事業所の集計結果を利用している。

⁴ なお、2019年以前の賃金構造基本統計調査では、一部の職種に該当する労働者のみ職種を記入することとされていた。また、役職のある労働者については職種の調査対象としていなかったため、職種別の集計には役職者は含まれていない。比較の際にはこうした点にも留意が必要である。

表 2.1.2 賃金構造基本統計調査と職種別民間給与実態調査の比較

	賃金構造基本統計調査		職種別民間給与実態調査		差	
	年齢 (歳)	きまって支給 する現金給与額 (万円/月)	年齢 (歳)	きまって支給 する現金給与額 (万円/月)	(3)-(1)	(4)-(2)
	(1)	(2)	(3)	(4)	(5)	(6)
大学教授	56.6 [6.8]	68.1 [12.4]	56.5	74.2	-0.06 (0.14)	6.09 (0.26)
大学准教授	45.8 [6.9]	54.0 [8.6]	47.0	58.7	1.21 (0.19)	4.78 (0.23)
大学講師	43.3 [9.4]	46.1 [13.3]	42.3	50.5	-1.03 (0.39)	4.42 (0.55)
医師	41.6 [12.6]	78.1 [46.3]	42.4	93.4	0.82 (0.39)	15.32 (1.43)
歯科医師	43.4 [11.7]	61.7 [33.7]	39.4	73.6	-3.96 (2.11)	11.96 (6.05)
薬剤師	36.8 [10.1]	33.4 [10.2]	34.7	34.3	-2.11 (0.47)	0.95 (0.48)
看護師	36.3 [10.7]	32.2 [8.9]	36.6	34.5	0.25 (0.12)	2.33 (0.10)
自家用乗用 自動車運転手	55.2 [11.4]	21.9 [11.9]	52.0	41.6	-3.18 (0.66)	19.69 (0.69)
守衛	57.9 [11.4]	22.0 [11.8]	52.7	39.5	-5.18 (0.63)	17.52 (0.65)
用務員	56.0 [11.6]	16.5 [9.6]	52.4	33.6	-3.56 (0.58)	17.12 (0.48)

注：賃金構造基本統計調査の分析サンプルは事業所規模50人以上の民営事業所に限定。職種別民間給与実態調査の値は平成23年職種別民間給与実態調査の第5表（人事院）から抜粋した。賃金構造基本統計調査の値も同じ2011年のデータを用いている。ブラケット内の値は標準偏差、括弧内の値は標準誤差（事業所レベルでクラスターリング）を表している。標準誤差の計算の際、職種別民間給与実態調査の値は定数として扱った。

(3) 非正規労働者の抽出割合について

賃金構造基本統計調査と民間給与実態統計調査のそれぞれにおける非正規労働者の割合を算出し、比較を行った。なお、両調査で「非正規労働者」の定義に違いがあるため、解釈に当たっては留意が必要である。本分析では、それぞれの調査において、以下に該当する者を「非正規労働者」として計算した。

賃金構造基本統計調査：以下 A)-C)の3つの異なる定義を用いた。

- A) 「非正規労働者（正社員・正職員以外の労働者）」とされる者
- B) 「非正規労働者」もしくは「臨時労働者（常用労働者以外の労働者）」とされる者
- C) 「非正規労働者」もしくは「臨時労働者」もしくは「短時間労働者」とされる者

※ 正社員・正職員：常用労働者のうち、事業所で正社員、正職員とする者をいう。

※ 常用労働者：1又は2に該当する労働者をいう。

- 1. 期間を定めずに雇われている労働者
- 2. 1か月以上の期間を定めて雇われている労働者

※ 短時間労働者：同一事業所の一般の労働者より1日の所定労働時間が短い又は1日の所定労働時間が同じでも1週の所定労働日数が少ない労働者。

民間給与実態統計調査：「非正規の給与所得者」とされている者。

※ パートタイマー、アルバイト、派遣社員、契約社員、嘱託等をいう。

なお、賃金構造基本統計調査では「正社員・正職員」を先に定義し、それに当てはまらない者を「非正規（正社員・正職員以外）」と定義しているが、民間給与実態統計調査では先に「非正規」を定義し、「正規」を「役員、青色事業専従者及び非正規を除く給与所得者」と定義している。

分析結果を図 2.1.3 に示す。2017年の非正規労働者の割合を見ると、賃金構造基本統計調査では、定義によって数値が異なるが約36-38%となっている。他方、民間給与実態統計調査では、割合は約30%となっている。ただし、前述のとおり、「非正規労働者」をどう定義するかによって数値が変わりうるため、割合自体の比較結果の解釈は難しい。非正規労働者の割合が2012年から2016年にかけて緩やかに上昇しており、2017年には若干減っているという傾向は2つの調査で共通している。

図 2.1.4 と図 2.1.5 はそれぞれ男女別、事業所規模別の分析結果を示している。男性の賃金構造基本統計調査と民間給与実態統計調査の差は男性で約5-6%（約22-23%と約17%）、女性で約3-8%（約51%-56%と約48%）と差の水準としては大きな差はない。ただし、男性の非正規労働者の割合は女性の半分以下であるため、ずれとしては男性の方が大きいことになる。事業所規模別では、300人未満事業所では民間給与実態統計調査の方が非正規労働者の比率が低く、300-499人事業所では差はあまりない。500人以上事業所では、賃金構

造基本統計調査の方が低い比率となっている。1,000人以上事業所では差がさらに大きくなっている。事業所規模で非正規率の水準の差が異なることについては、上述の非正規労働者の定義の問題に加え、賃金構造基本統計調査と民間給与実態統計調査における事業所規模の捉え方の違いが考えられる。民間給与実態統計調査の場合、給与等の支払事務を取り扱う事業所が単位となるため、本社で一括して事務を行っている場合は、複数の事業所で経済活動が行われていても調査上の単位は1つの事業所となる。したがって、経済活動が行われている事業所単位で調査上の単位が定義される賃金構造基本統計調査と比べて、民間給与実態統計調査の方が規模の大きい事業所の数が相対的に多くなる傾向がある（表2.1.3を参照）⁵。その結果、民間給与実態統計調査の大規模事業所では規模の大きい企業の影響が強く出ているという可能性がある一方で、賃金構造基本統計調査では、規模の大きい企業も事業所ごとに細分化され、中・小規模に割り振られることが多いため、大規模の企業の影響が平準化され、事業所規模別に見てもその影響が強く表れないという可能性が考えられる。

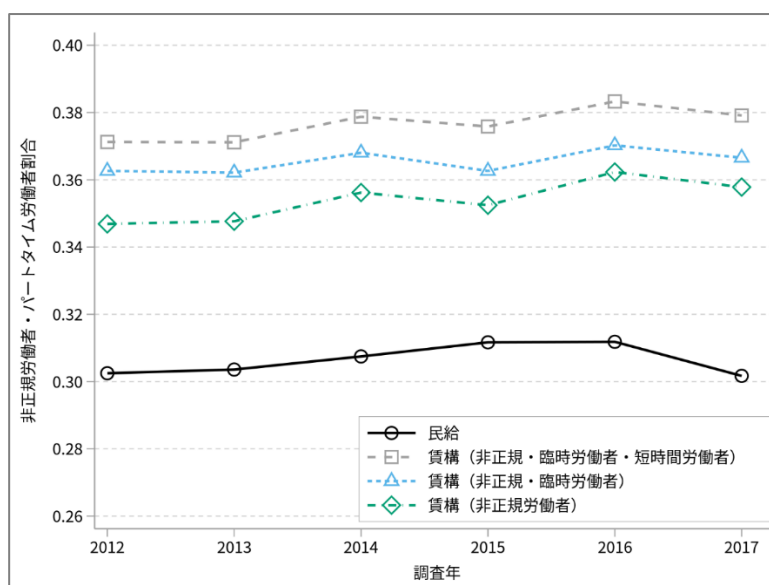
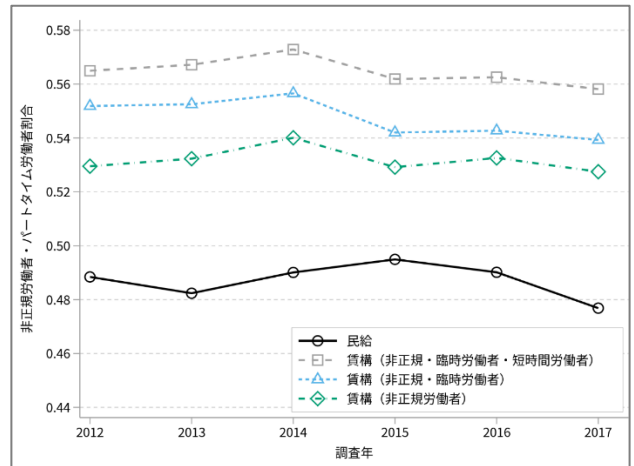
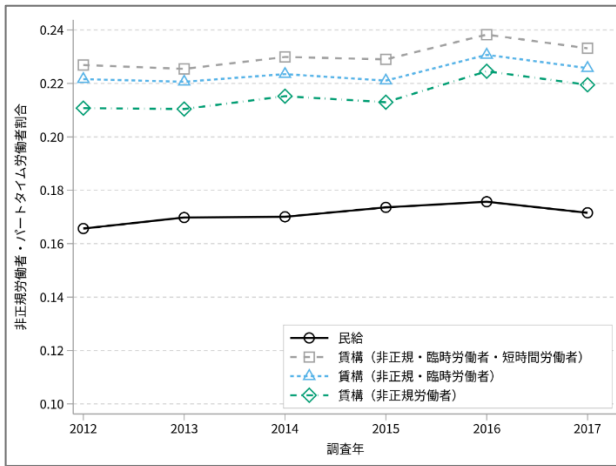


図 2.1.3 賃金構造基本統計調査と民間給与実態統計調査の非正規労働者割合の比較

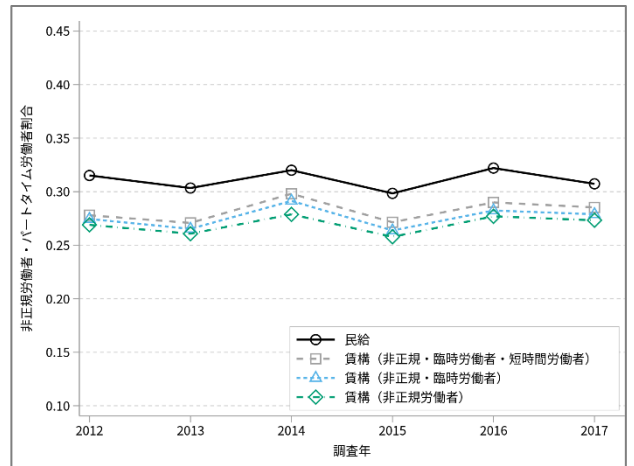
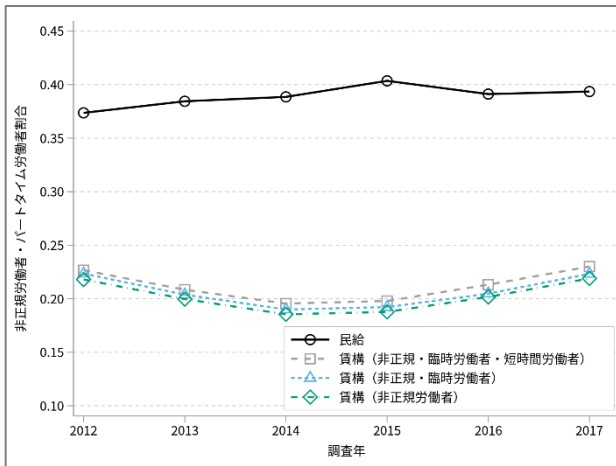
⁵ 例えば、コンビニエンス・ストアなどの場合、賃金構造基本統計調査では個々の店舗が調査の対象となる事業所として定義される。他方、仮に給与等の支払事務が個々の店舗ではなく本社などで一括して行われている場合、民間給与実態統計調査では本社が調査対象となる1つの事業所として定義される。その結果、民間給与実態統計調査の方が事業所規模が大きくなる傾向がある。



(a) 男性

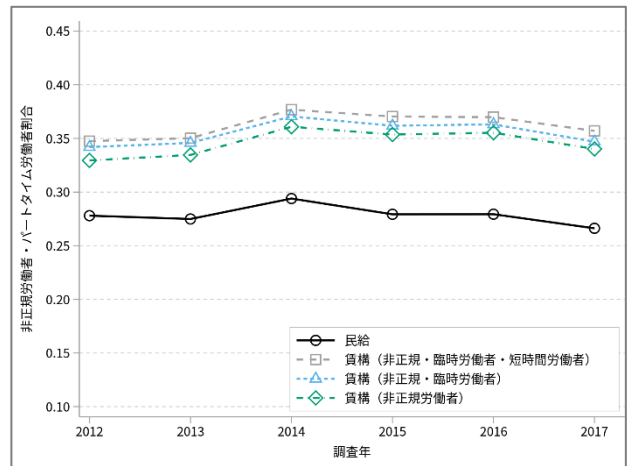
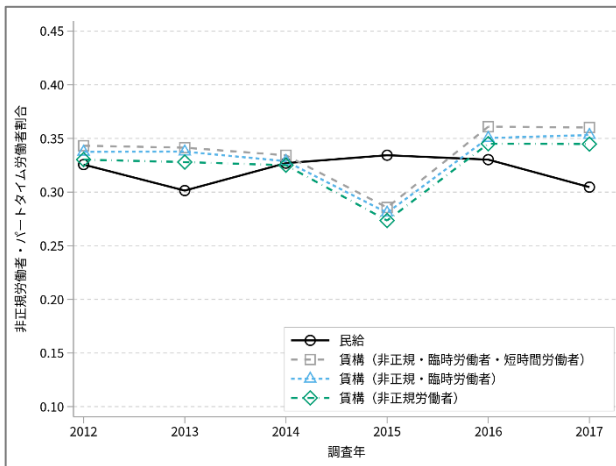
(b) 女性

図 2.1.4 賃金構造基本統計調査と民間給与実態統計調査の非正規労働者割合 (男女別)



(a) 1000人以上

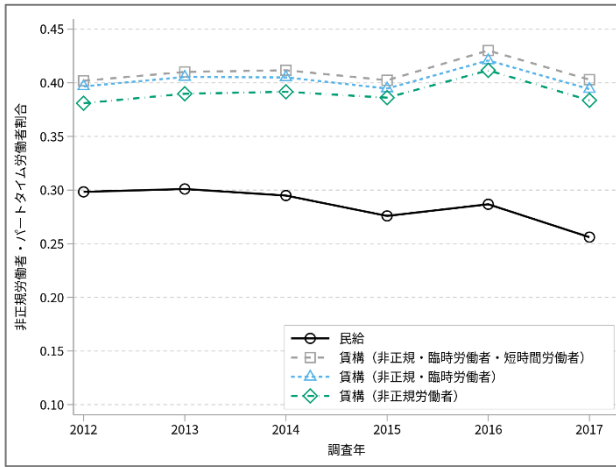
(b) 500-999人



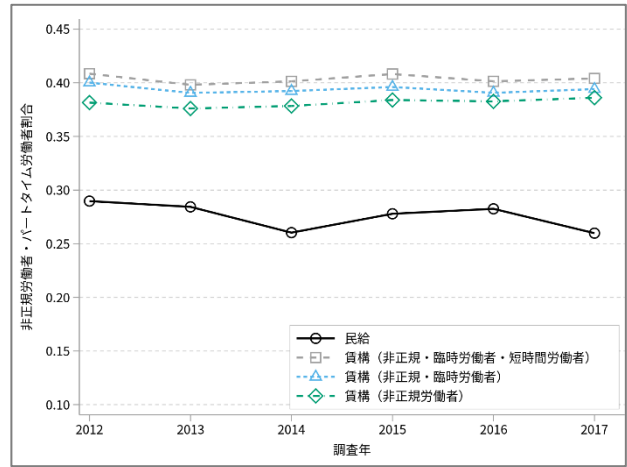
(c) 300-499人

(d) 100-299人

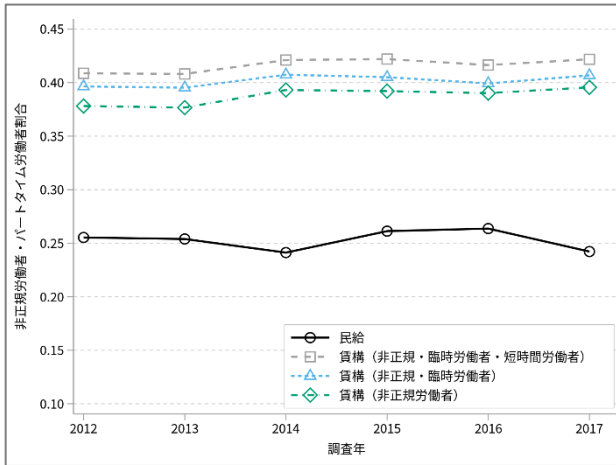
図 2.1.5 賃金構造基本統計調査と民間給与実態統計調査の非正規労働者割合の比較 (事業規模別)



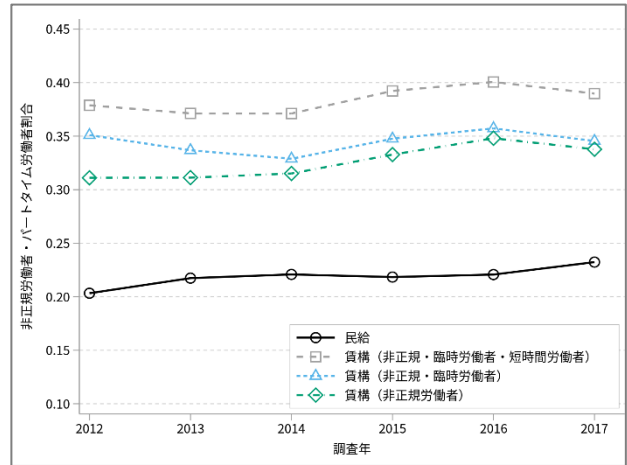
(e) 50-99 人



(f) 30-49 人



(g) 10-29 人



(h) 5-9 人

図 2.1.6 賃金構造基本統計調査と民間給与実態統計調査の非正規労働者割合の比較
(事業規模別) 続き

表 2.1.3 事業所規模ごとの調査対象事業所数の内訳

事業所規模	賃金構造基本 統計調査	全体に占める 割合	民間給与実態 統計調査	全体に占める 割合
1,000人以上	2,571	0.009	5,347	0.085
500-999人	5,179	0.019	5,655	0.090
300-499人	3,411	0.012	3,489	0.055
100-299人	30,197	0.108	15,229	0.241
50-99人	28,834	0.103	6,857	0.109
30-49人	47,584	0.171	7,759	0.123
10-29人	124,264	0.446	12,664	0.201
5-9人	36,723	0.132	6,076	0.096
合計	278,763		63,076	

注：データはいずれの調査も 2012-2017 年。

2.1.3. 調査結果からの示唆

上記分析結果のとおり、賃金構造基本統計調査と民間給与実態統計調査は、公表値における数値の乖離はあるものの、個票データを用いて各調査の対象の属性を揃えるための調整を行った結果、両調査の差はほぼ解消された。したがって、両統計調査の公表値における差異は調査対象および集計に用いているサンプルの違いに起因するものであると考えられる。

他方、賃金構造基本統計調査と職種別民間給与実態調査とに差がみられたことについては、後者は集計された公表値を用いていることもあり、対象となるサンプルの属性を十分に一致させることができなかつたことが一因であると考えられる。分析に当たっては、対象となる事業所の規模や分析に用いる職種などを可能な限り揃えたものの、同じ職種のカテゴリであっても、その詳細な定義には多少の違いがあることや、賃金構造基本統計調査は毎年6月の、職種別民間給与実態調査は毎年4月の給与を対象として調査をしていることなど、調整できていない属性が多く残されている。

いずれの比較からも、異なる統計調査の結果を比較する際には、調査対象や集計方法の差異を適切に考慮することの重要性が示唆される。そうした差異に対して適切な考慮をせずに集計値の比較をすることは誤った結論につながってしまう可能性があるため、十分な留意が必要である。また、各統計調査はそれぞれの目的のために実施されているため、調査対象や集計方法が異なり、その結果として集計値に違いが出ることは自然なことである。調査結果の利用者がそうした調査対象や集計方法の違いを正しく認識したうえで、調査結果を解釈することができるよう、調査方法や集計方法に関する正確な情報をわかりやすく公表することも重要である。さらに、非正規労働者の比較においては、調査間の定義が異なるため比較が困難であったが、可能な範囲で統計調査間の共通の定義や調査方法を用いる

ことも重要であると考えられる⁶。

⁶ 今後は、民間給与実態統計調査における非正規の定義が変更される予定であり、その場合、非正規労働者の比率は変わる（賃金構造基本統計調査の数字に近づく）可能性がある。

2.2. 調査方法の影響（サンプルの偏りの有無）

2.2.1. 本調査項目の目的

調査対象事業所や調査対象労働者の属性などに加えて、実査における調査方法がサンプルの偏りをもたらす可能性も考えられる。例えば、毎月勤労統計調査では事業所ごとの支給給与総額を労働者数で除すことによって 1 人当たり賃金を算出している。他方、賃金構造基本統計調査においては、調査対象事業所の抽出は調査実施者が行うが、事業所内での労働者の抽出は調査対象事業所に委ねられており、こうした抽出方法の違いがもたらす非標本誤差が生じている可能性がある。

したがって、本調査項目では、賃金構造基本統計調査と民間給与実態統計調査のそれぞれについて、事業所内におけるサンプルの偏りの有無について検証を行う。

2.2.2. 調査方法・結果

(1) 賃金構造基本統計調査

賃金構造基本統計調査と毎月勤労統計調査を共通の事業所コードによって事業所レベルで結合できることに着目し、2つの調査それぞれから計算される 1 人当たり賃金が一致するか確認する。毎月勤労統計調査では、常用労働者、実労働時間、きまって支給する給与の事業所の合計値が調査されているため、ここから事業所の総労働時間や給与総額を常用労働者数で割ることで、常用労働者 1 人当たりの労働時間数や給与額を計算できる。もし、賃金構造基本統計調査において、個人票対象の労働者が事業所内で無作為に抽出されているならば、個人票から計算される常用労働者 1 人当たりの労働時間数や給与額は平均的には毎月勤労統計調査から計算されるものと一致するはずである。したがって、事業所ごとに賃金構造基本統計調査から得られる値と毎月勤労統計調査から得られる値の差を計算して、その傾向を見ることで労働時間数や給与額が高いもしくは低い労働者が抽出されやすいのか知ることができる。

分析は、毎月勤労統計調査は規模が 500 人以上の事業所については原則全数調査を実施していることから、500 人以上の事業所に限定して行う。まず賃金構造基本統計調査（2016-17 年）から 500 人以上の事業所を抽出し、各事業所の常用労働者レベルのデータを事業所レベルのデータとして集計した（5037 事業所）。そのうえで、共通事業所コードを用いて毎月勤労統計調査（2016-17）の事業所とのマッチングを行った（マッチングできた企業数は 2193 事業所）。さらに、特別給与額の調整を行ったうえで、賃金構造基本統計調査と毎月勤労統計調査のデータを比較した。

分析結果を表 2.2.1 に示す。まず、賃金構造基本統計調査の 500 人以上の事業所（(2) 列）のうち、毎月勤労統計調査とマッチできた分析に用いるサンプル（比較サンプル（(3)

表 2.2.1 賃金構造基本統計調査と毎月勤労統計調査の比較

	賃金構造基本統計調査			毎月勤労 統計調査	毎勤との 比較	参考
	全事業所 (1)	500 人以 上事業所 (2)	比較 サンプル (3)	(4)	(3)-(4) (5)	(3)-(2) (6)
常用労働者数 (人)	33.94 [109.80]	962.34 [1101.64]	1147.16 [1209.31]	1153.22 [1167.76]	-6.06 (16.65)	184.82 (21.19)
労働日数 (日)	20.00 [5.25]	20.05 [3.58]	20.13 [3.43]	19.59 [2.22]	0.54 (0.06)	0.08 (0.04)
所定労働時間 (時間)	140.14 [50.27]	149.46 [33.02]	151.10 [31.50]	147.26 [20.52]	3.84 (0.48)	1.64 (0.40)
所定外労働時間 (時間)	7.55 [15.57]	12.88 [17.95]	14.66 [18.34]	15.57 [9.81]	-0.91 (0.18)	1.78 (0.18)
総労働時間 (時間)	147.70 [56.46]	162.34 [40.56]	165.76 [39.57]	162.83 [25.05]	2.93 (0.53)	3.42 (0.48)
きまって支給する 現金給与額 (万円/月)	22.93 [14.76]	36.70 [21.20]	36.88 [19.62]	36.67 [10.27]	0.21 (0.13)	0.18 (0.23)
特別給与額 (万円/月)	3.58 [5.92]	9.94 [10.75]	10.81 [9.95]	11.42 [6.25]	-0.61 (0.07)	0.86 (0.14)
総給与額 (万円/月)	26.51 [18.86]	46.65 [29.20]	47.69 [27.39]	48.09 [15.46]	-0.40 (0.18)	1.04 (0.34)
事業所数	114,410	5,037	2,193			

注：サンプルは 500 人以上の事業所に限定。「常用労働者数」以外は全て「常用労働者 1 人当たり」の値。ブラケット内の値は標準偏差、括弧内の値は標準誤差（事業所レベルでクラスタリング）を表している。2016-17 年の数値を比較しているが、賃金構造基本統計調査においては調査前年の特別給与額を聞いているため、特別給与額のみ 2015-16 年の値となっている。

列) は、事業所規模が大きい、労働者 1 人当たり労働時間が長い、労働者 1 人当たり給与支給額が高いという傾向がみられる。本分析はこうしたサンプルを用いたものであることには、解釈において留意が必要である。

賃金構造基本統計調査と毎月勤労統計調査の比較 ((5)列) においては、常用労働者数には特に差は見られない (統計的にも有意ではない)。労働日数や労働時間については、賃金構造基本統計調査の方が 2-3%程度長いことが分かる。また、給与額については、きまって支給される給与には差は無いが、特別支給額 (賞与) が毎月勤労統計調査の方が若干多く、総給与額も毎月勤労統計調査の方が約 4,000 円程度高い数字になっている。他方、総給与額の差は平均値の 1%にも満たない差であり、統計的に有意な差とはなっているものの、実質的には大きな乖離は無いと考えることができる。

表 2.2.2 賃金構造基本統計調査と毎月勤労統計調査の比較（男性）

	賃金構造基本統計調査			毎月勤労 統計調査	毎勤との 比較	参考
	全事業所	500 人以 上事業所	比較 サンプル		(3)-(4)	(3)-(2)
	(1)	(2)	(3)	(4)	(5)	(6)
常用労働者数（人）	18.77 [76.14]	610.56 [821.34]	798.86 [964.70]	803.39 [1004.17]	-4.52 (12.44)	188.30 (16.34)
労働日数（日）	20.63 [5.05]	20.24 [3.62]	20.29 [3.45]	19.83 [2.36]	0.46 (0.06)	0.05 (0.04)
所定労働時間（時間）	150.26 [47.15]	152.92 [31.67]	154.16 [29.99]	150.85 [20.78]	3.30 (0.52)	1.23 (0.40)
所定外労働時間（時間）	9.51 [17.58]	14.66 [19.66]	16.58 [19.68]	17.66 [10.62]	-1.08 (0.22)	1.92 (0.20)
総労働時間（時間）	159.76 [53.83]	167.58 [39.87]	170.74 [38.50]	168.51 [25.25]	2.21 (0.59)	3.15 (0.49)
きまって支給する 現金給与額（万円／月）	27.96 [17.40]	42.94 [25.34]	42.33 [22.93]	42.75 [12.80]	-0.43 (0.22)	-0.61 (0.29)
特別給与額（万円／月）	4.58 [6.91]	11.53 [11.85]	12.22 [10.50]	12.98 [6.72]	-0.76 (0.09)	0.69 (0.15)
総給与額（万円／月）	32.54 [21.87]	54.47 [33.03]	54.55 [30.08]	55.73 [17.34]	-1.19 (0.27)	0.08 (0.38)
事業所数	114,410	5,037	2,193			

注：サンプルは 500 人以上の事業所に限定。「常用労働者数」以外は全て「常用労働者 1 人当たり」の値。ブラケット内の値は標準偏差、括弧内の値は標準誤差（事業所レベルでクラスタリング）を表している。2016-17 年の数値を比較しているが、賃金構造基本統計調査においては調査前年の特別給与額を聞いているため、特別給与額のみ 2015-16 年の値となっている。

表 2.2.2 と表 2.2.3 にそれぞれ男女別の分析結果を示す。男性のどの給与項目も毎月勤労統計調査の方が高い。例えば、給与総額の場合、毎月勤労統計調査の方が賃金構造基本統計調査の値よりも 11,900 円（2.1%）程度高いことが分かる。女性については、きまって支給する現金給与額は賃金構造基本統計調査の方が高く、特別給与額は毎月勤労統計調査の方が高い。給与総額では賃金構造基本統計調査の方が 1.4%高くなっている。

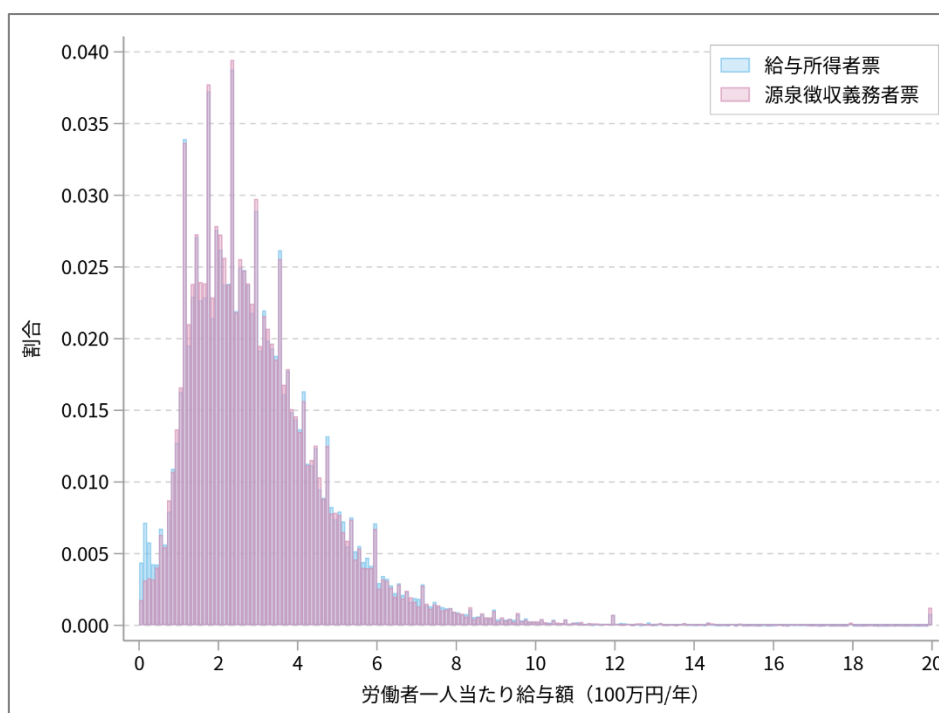
表 2.2.3 賃金構造基本統計調査と毎月勤労統計調査の比較（女性）

	賃金構造基本統計調査			毎月勤労 統計調査	毎勤との 比較	参考
	全事業所 (1)	500 人以 上事業所 (2)	比較 サンプル (3)	(4)	(3)-(4) (5)	(3)-(2) (6)
常用労働者数（人）	15.17 [52.87]	351.78 [588.32]	348.29 [550.82]	349.83 [433.77]	-1.54 (6.37)	-3.48 (11.13)
労働日数（日）	19.36 [5.32]	19.80 [3.61]	19.87 [3.49]	19.07 [2.28]	0.80 (0.06)	0.07 (0.04)
所定労働時間（時間）	131.09 [50.88]	145.40 [34.33]	147.19 [33.02]	140.97 [21.87]	6.22 (0.52)	1.79 (0.46)
所定外労働時間（時間）	4.36 [10.39]	9.66 [14.27]	9.92 [14.19]	10.18 [6.86]	-0.25 (0.19)	0.26 (0.17)
総労働時間（時間）	135.45 [54.50]	155.06 [39.96]	157.11 [38.85]	151.15 [24.40]	5.97 (0.57)	2.05 (0.51)
きまって支給する 現金給与額（万円／月）	17.57 [10.75]	28.35 [15.30]	27.68 [13.62]	26.75 [7.58]	0.93 (0.13)	-0.67 (0.20)
特別給与額（万円／月）	2.58 [4.37]	6.95 [7.29]	7.38 [6.76]	7.82 [4.35]	-0.45 (0.08)	0.43 (0.10)
総給与額（万円／月）	20.15 [13.79]	35.30 [20.81]	35.06 [19.04]	34.58 [11.13]	0.48 (0.18)	-0.24 (0.28)
事業所数	114410	5037	2193			

注：サンプルは 500 人以上の事業所に限定。「常用労働者数」以外は全て「常用労働者 1 人当たり」の値。ブラケット内の値は標準偏差、括弧内の値は標準誤差（事業所レベルでクラスタリング）を表している。2016-17 年の数値を比較しているが、賃金構造基本統計調査においては調査前年の特別給与額を聞いているため、特別給与額のみ 2015-16 年の値となっている。

(2) 民間給与実態統計調査

民間給与実態統計調査の調査項目は、源泉徴収義務者（事業所）に関する調査項目と、事業所内で抽出された個々の給与所得者に関する調査項目の 2 つに大別することができる。源泉徴収義務者票（事業所票）から給与所得者数と総給与支払額がわかるため、給与所得者 1 人当たりの給与支払額がわかる。一方、給与所得者票（個人票）からは、事業所においてサンプリングされた給与所得者の所得がわかる。抽出に偏りが無ければ両者の分布は一致することが期待されるため、源泉徴収義務者票と給与所得者票を事業所レベルでマッチして、給与所得者票から得られる事業所平均所得の分布が源泉徴収義務者票から得られるものと一致するか検証した。

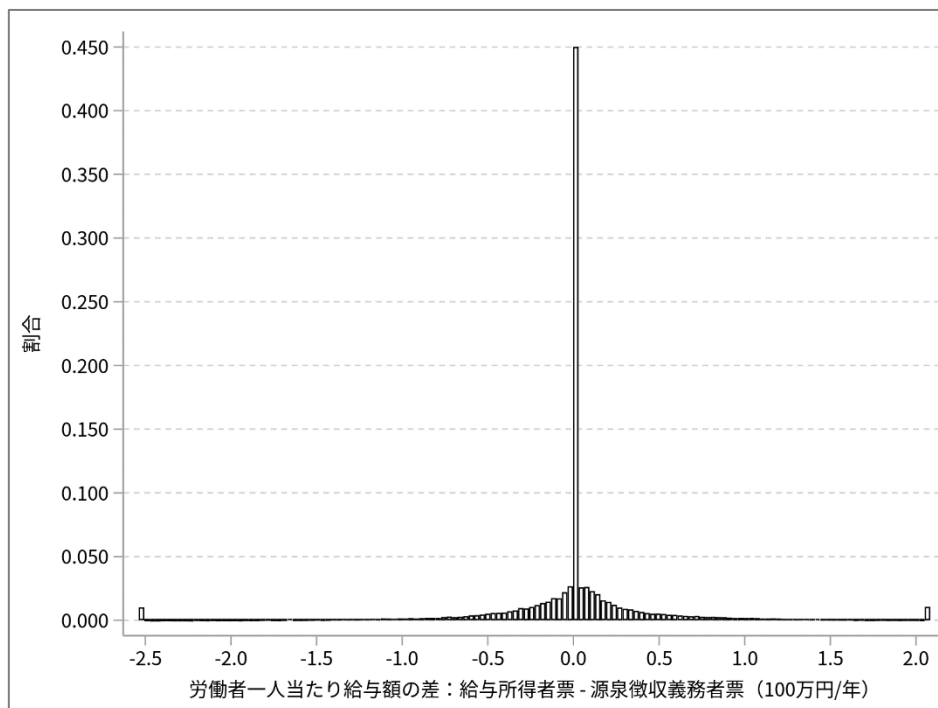


注：グラフを見やすくするために、プロットする際に所得額を 2000 万円でキャップしている。データは 2012-2019 年。

図 2.2.1 民間給与実態統計調査の源泉徴収義務者票と給与所得者票の年間給与総額の分布

図 2.2.1 は 2012 年から 2019 年のデータをプールしそれぞれの調査票から得られる所得の分布をプロットしたものである。源泉徴収義務者票の値は総給与支払額を 3 月、6 月、9 月、12 月の従業員数の平均値で割り、計算した。図を見ると、給与所得者票から計算した分布では平均年間給与額が 50 万円以下の層が多い。この点については、1 年のうち数か月間しか働いていない給与所得者がサンプリングされたことに起因すると考えられる。他方、源泉徴収義務者票のデータから計算した分布では 2000 万円以上の層がやや多くなっている（図ではグラフを見やすくするために、所得額を 2000 万円でキャップしている。また、源泉徴収義務者票のデータでは、極端に高い値を取ることがあり、その原因の一つとして、従業員 1 人当たり給与を計算するために、総給与支払額を「平均」従業員数で割っていることが考えられる）。

表 2.2.4 は給与所得者票と源泉徴収義務者票それぞれから計算される労働者 1 人当たり所得の事業所ごとの差の平均値と分布を表している。平均値では源泉徴収義務者票から計算した値の方が 15 万円/年高くなっている（(3)列）。一方、分位点の値を比較すると、低分位点の値は源泉徴収義務者票の方が高く、高分位点の値は給与所得者票の方が高い。ただし、平均の差と比べると、分位点の値の差は限定的である。平均値では差が大きく、分布では差が小さいのは、上述のとおり、源泉徴収義務者票から計算された給与所得者 1 人当た



注：グラフを見やすくするために、プロットする際に差の値を 1 パーセントの値と 99 パーセントの値でキャップしている。データは 2012-2019 年。

図 2.2.2 民間給与実態統計調査の源泉徴収義務者票と給与所得者票の
年間給与総額の差の分布

り所得は上方バイアスがかかっている可能性があり、特に数レコードの値が平均値に大きな影響を与えていることがわかった。実際、上位下位 1%をトリムして分析を行うと、平均値の差は小さくなった ((6)列)。統計的には有意に給与所得者票の値の方が高いが、差は 9000 円/年 (平均値の 0.3%) と極めて小さい値になっている。図 2.2.2 には、二つの調査票の差の分布を示しているが、この図からも多くの事業所で差は小さいが、左側の裾が長い
ため、平均値が外れ値に左右されやすいことが確認できる。また、分位点の値は外れ値の影響を受けにくい
ため、フルサンプルと上位下位 1%をトリムしたサンプルとでほぼ同じ結果が得られた。

表 2.2.4 民間給与実態統計調査の給与所得者 1 人当たり年間給与総額 (100 万円/年)

	フルサンプル			上下 1 % をトリムしたサンプル		
	給与所 得者票	源泉徴収 義務者票	差 (1)-(2)	給与 所得者票	源泉徴収 義務者票	差 (4)-(5)
	(1)	(2)	(3)	(4)	(5)	(6)
平均値	3.049	3.202	-0.153	3.020	3.011	0.009
	[1.994]	[20.007]	(0.090)	[1.887]	[1.857]	(0.002)
10 パーセンタイル	1.181	1.200	-0.019	1.200	1.200	0.000
			(0.007)			(0.005)
25 パーセンタイル	1.774	1.795	-0.021	1.777	1.790	-0.013
			(0.006)			(0.005)
50 パーセンタイル(中央値)	2.690	2.681	0.009	2.675	2.662	0.013
			(0.004)			(0.004)
75 パーセンタイル	3.881	3.849	0.032	3.845	3.809	0.036
			(0.006)			(0.006)
90 パーセンタイル	5.289	5.223	0.067	5.209	5.138	0.071
			(0.011)			(0.009)
観測数	163,993	163,993	163,993	157,683	157,683	157,683

注：ブラケット内の値は標準偏差、括弧内の値は標準誤差（事業所レベルでクラスタリング）を表している。各パーセンタイル値の標準誤差は事業所クラスタリングブートストラップを用いて計算した（シミュレーション回数は 1000 回）。データは 2012-2019 年。

2.2.3. 調査結果からの示唆

上記分析結果のとおり、賃金構造基本統計調査では男女ともに労働時間が長いサンプルが抽出されていることが分かった。給与額については、男性については給与が若干低いサンプルが、女性については給与が高いサンプルが抽出されている傾向が伺われる（男女合わせた全体では、偏りが打ち消し合った結果大きな差とはなっていない）。データからはその偏りの原因については検証することができないが、総じて、女性の方が男性よりも差が大きく、中核的な労働者が抽出されている可能性が推察される。

民間給与実態統計調査については、上位下位 1% をトリムしたうえで比較を行ったところ、給与額に大きな差は無く、偏りなく抽出されていると考えることができる。

調査間でこのような違いが出る理由については、データから実証的に検証することはできないが、1 つの可能性として、調査の記入要項における説明方法の違いが考えられる。賃金構造基本統計調査では、事業所内での抽出方法として等間隔抽出法が示されているが、あくまで抽出方法の一例としての提示であり、無作為抽出であれば他の方法を用いても良いとの記載になっている。この場合、抽出の担当者が無作為抽出に関する正確な理解を持って

いない場合、適切な無作為抽出が行われない可能性がある。他方、民間給与実態統計調査の記入要領においては、抽出方法として等間隔抽出法を用いることを所与としており、記入対象者を決めるための方法についても丁寧な説明がされている。さらに、フローチャートに数字を記入していくことで、何人の対象者をどの間隔で抽出すべきかを自動で計算できる Excel ファイルも提供しており⁷、抽出の担当者が無作為抽出に関する正確な理解を持っているかどうかに関わらず、手順を追っていただけで適切な無作為抽出ができるようになっていく。したがって、民間給与実態統計調査以外の調査においても、抽出方法を統一し、そのためのツールを提供するなどの方策を行うことが抽出の無作為性を担保するうえで有効である可能性がある。

⁷ <https://www.nta.go.jp/publication/statistics/kokuzeicho/r02minkyu/excel/kimekata.xlsm>

2.3. 調査対象の回答状況

2.3.1. 本調査項目の目的

統計調査では、一般的に、業種・地域・事業所規模などによって層化したうえで標本抽出を行うが、その際には回収率 100%を想定して標本誤差を計算している。したがって、回答率が低下すると統計精度が低下してしまう。また、調査への未回答が事業所属性に関わらず無作為に生じているのなら、調査から得られたデータは母集団の代表性を有するが、調査回答確率が事業所属性に依存している場合、代表的なデータを得られず、一般にそのデータを用いた分析にはバイアスが生じる。他方、もし観察可能な属性だけに依存しているなら、その属性をサンプリングウエイトや回帰分析によって制御することで、バイアスを緩和もしくは解消できる。そのため、調査の回答状況と事業所属性の関係を分析しデータの代表性を検証することは重要であり、もし代表的でないとしてもバイアスを緩和するためにはどのような事業所属性を用いればよいか明らかにする必要がある。

したがって、本調査項目では、賃金構造基本統計調査と民間給与実態統計調査のそれぞれについて、事業所属性と回答率の関係を検証し、実態を明らかにする。また、事業所レベルのパネルデータを作成し、複数回調査対象になった事業所を用いた分析を行うことで、回答パターンの検証を行う。

2.3.2. 調査方法・結果

(1) 調査への回答状況と事業所属性の関係

賃金構造基本統計調査と民間給与実態統計調査のそれぞれについて、名簿情報に記載されている事業所のうち、実際に調査に回答している事業所を調査回答事業所とし、事業所属性（賃金構造基本統計調査：事業所規模・産業・都道府県・都市規模、民間給与実態統計調査：事業所規模・産業）ごとに回収率を計算した。

【賃金構造基本統計調査】

2012年から2017年のデータを用いて、事業所規模別、産業別、都道府県別、都市規模別で回答状況を分析した。表 2.3.1 は事業所規模別の回答状況を示している。規模別では小規模事業所（特に 5-9 人事業所と 10-29 人事業所）では回収率が低くなっている。産業別で見ると（表 2.3.2）、「卸売業、小売業」、「不動産業、物品賃貸業」、「宿泊業、飲食サービス業」、「生活関連サービス業、娯楽業」、「教育、学習支援事業」といった産業での回答率が比較的低い数字となっている。都道府県別では（表 2.3.3）、東京都、大阪府といった都市圏に加え、茨城県、群馬県、徳島県、奈良県などでも低い数字となっている。都市規模別では（表 2.3.5）政令指定都市や中核市で、他の市町村よりも回答率が低くなっている。

図 2.3.1 から図 2.3.5 には、調査回答有無を被説明変数とし、調査年（2012 年）、産業（製造業）、事業所規模（100-299 人）都道府県（東京都）都市規模（その他の市）の各属性が

表 2.3.1 賃金構造基本統計調査回答状況：事業所規模別（割合）

	2012年	2013年	2014年	2015年	2016年	2017年
全て	0.69	0.69	0.70	0.71	0.70	0.69
15,000人以上	1.00	1.00	1.00	1.00	1.00	1.00
5,000-14,999人	0.92	0.81	0.84	0.82	0.94	0.86
1,000-4,999人	0.88	0.85	0.87	0.85	0.84	0.81
500-999人	0.87	0.84	0.84	0.84	0.83	0.81
300-499人	0.85	0.82	0.81	0.79	0.83	0.83
100-299人	0.83	0.81	0.82	0.82	0.80	0.80
50-99人	0.79	0.79	0.80	0.81	0.79	0.77
30-49人	0.77	0.76	0.77	0.77	0.77	0.75
10-29人	0.67	0.67	0.68	0.69	0.69	0.67
5-9人	0.45	0.48	0.49	0.50	0.51	0.48

ミー変数を説明変数とした回帰分析結果を示しているが（カッコ内は参照カテゴリ）、各図からも上記の傾向を確認することができる⁸。

⁸ ここでの推定式は以下のとおり。

$$y_{ijt} = x'_{ijt} \beta + \eta_j + \pi_t + u_{ijt}$$

i, j, t はそれぞれ事業所、都道府県、調査年を表しており、 η_j と π_t はそれぞれ都道府県、調査年固定効果である。 x'_{ijt} は産業、事業所規模、都市規模の各属性ダミー変数を含む説明変数のベクトルである。被説明変数の y_{ijt} は調査回答を表すダミー変数である。

表 2.3.2 賃金構造基本統計調査回答状況：産業別（割合）

	2012年	2013年	2014年	2015年	2016年	2017年
全て	0.69	0.69	0.70	0.71	0.70	0.69
鉱業、採石業、砂利採取業	0.70	0.76	0.74	0.74	0.72	0.71
建設業	0.71	0.71	0.72	0.75	0.74	0.71
製造業	0.76	0.75	0.76	0.76	0.75	0.74
電気・ガス・熱供給・水道業	0.87	0.85	0.90	0.88	0.87	0.87
譲歩通信業	0.69	0.70	0.69	0.70	0.70	0.68
運送業、郵便業	0.73	0.75	0.73	0.75	0.72	0.71
卸売業、小売業	0.69	0.69	0.72	0.71	0.68	0.67
金融業、保険業	0.80	0.81	0.82	0.83	0.83	0.81
不動産業、物品賃貸業	0.61	0.62	0.62	0.63	0.66	0.67
学術研究、専門・技術サービス業	0.70	0.72	0.74	0.73	0.72	0.72
宿泊業、飲食サービス業	0.59	0.57	0.58	0.59	0.54	0.50
生活関連サービス業、娯楽業	0.56	0.57	0.59	0.59	0.58	0.56
教育、学習支援業	0.61	0.63	0.62	0.67	0.70	0.69
医療、福祉業	0.80	0.78	0.81	0.81	0.79	0.75
複合サービス業	0.85	0.87	0.84	0.87	0.85	0.87
サービス業（他に分類されないもの）	0.70	0.70	0.72	0.71	0.70	0.69

表 2.3.3 賃金構造基本統計調査回答状況：都道府県別（割合）

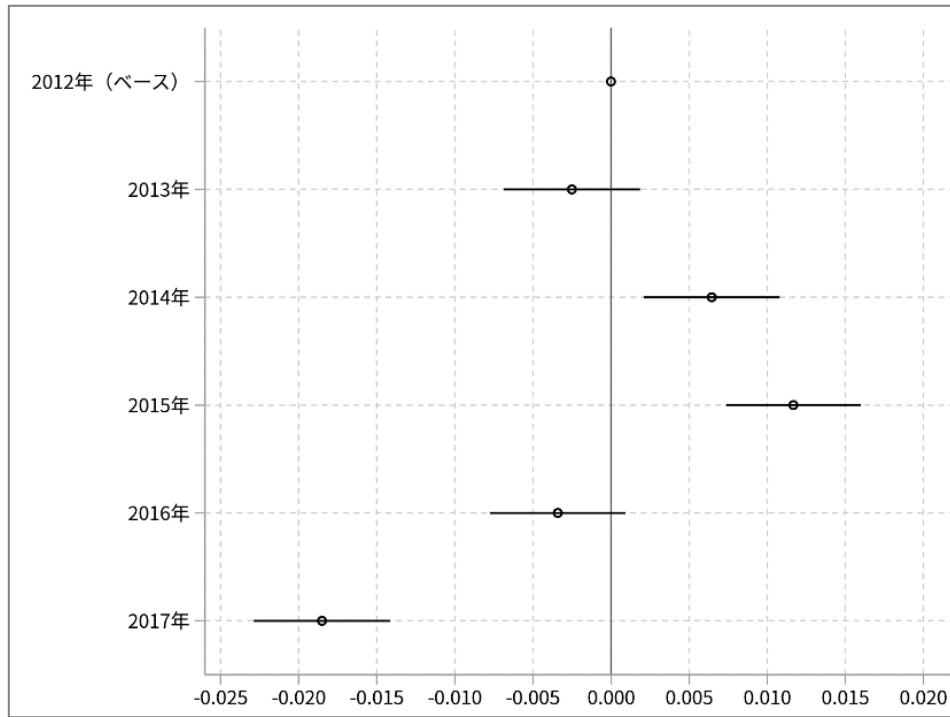
	2012年	2013年	2014年	2015年	2016年	2017年
全て	0.69	0.69	0.70	0.71	0.70	0.69
北海道	0.74	0.72	0.77	0.73	0.74	0.72
青森県	0.70	0.75	0.73	0.74	0.75	0.70
岩手県	0.76	0.81	0.81	0.83	0.83	0.80
宮城県	0.65	0.64	0.69	0.66	0.69	0.65
秋田県	0.78	0.79	0.79	0.82	0.81	0.84
山形県	0.74	0.70	0.73	0.73	0.77	0.77
福島県	0.64	0.67	0.72	0.75	0.70	0.72
茨城県	0.61	0.58	0.63	0.61	0.60	0.59
栃木県	0.71	0.70	0.76	0.79	0.77	0.70
群馬県	0.63	0.61	0.67	0.68	0.56	0.62
埼玉県	0.70	0.73	0.72	0.72	0.73	0.72
千葉県	0.70	0.69	0.74	0.63	0.65	0.66
東京都	0.66	0.59	0.58	0.62	0.62	0.61
神奈川県	0.67	0.63	0.62	0.64	0.65	0.65
新潟県	0.76	0.78	0.81	0.82	0.82	0.80
富山県	0.68	0.73	0.76	0.76	0.76	0.73
石川県	0.73	0.76	0.74	0.78	0.77	0.78
福井県	0.71	0.73	0.74	0.78	0.77	0.75
山梨県	0.69	0.73	0.70	0.70	0.63	0.64
長野県	0.71	0.72	0.71	0.73	0.76	0.76
岐阜県	0.73	0.68	0.70	0.68	0.67	0.68
静岡県	0.67	0.67	0.65	0.69	0.64	0.69
愛知県	0.68	0.70	0.72	0.71	0.73	0.70
三重県	0.72	0.69	0.74	0.81	0.74	0.68
滋賀県	0.60	0.62	0.69	0.68	0.69	0.68
京都府	0.64	0.63	0.65	0.62	0.63	0.62
大阪府	0.62	0.64	0.61	0.57	0.57	0.55
兵庫県	0.66	0.65	0.67	0.68	0.75	0.70
奈良県	0.62	0.61	0.69	0.70	0.64	0.57

表 2.3.4 賃金構造基本統計調査回答状況：都道府県別（割合）続き

	2012年	2013年	2014年	2015年	2016年	2017年
和歌山県	0.67	0.64	0.63	0.65	0.66	0.65
鳥取県	0.69	0.72	0.71	0.71	0.71	0.70
島根県	0.77	0.77	0.80	0.82	0.78	0.79
岡山県	0.68	0.69	0.68	0.69	0.74	0.71
広島県	0.70	0.72	0.74	0.72	0.68	0.64
山口県	0.75	0.79	0.78	0.81	0.78	0.74
徳島県	0.59	0.57	0.56	0.59	0.66	0.62
香川県	0.63	0.64	0.70	0.73	0.72	0.70
愛媛県	0.74	0.72	0.77	0.76	0.72	0.74
高知県	0.65	0.65	0.68	0.73	0.66	0.66
福岡県	0.76	0.73	0.79	0.78	0.79	0.79
佐賀県	0.72	0.77	0.74	0.73	0.74	0.74
長崎県	0.76	0.77	0.81	0.82	0.79	0.76
熊本県	0.78	0.79	0.79	0.85	0.82	0.82
大分県	0.69	0.74	0.69	0.76	0.74	0.73
宮崎県	0.60	0.66	0.67	0.67	0.70	0.68
鹿児島県	0.73	0.74	0.74	0.75	0.76	0.69
沖縄県	0.80	0.81	0.78	0.71	0.65	0.66

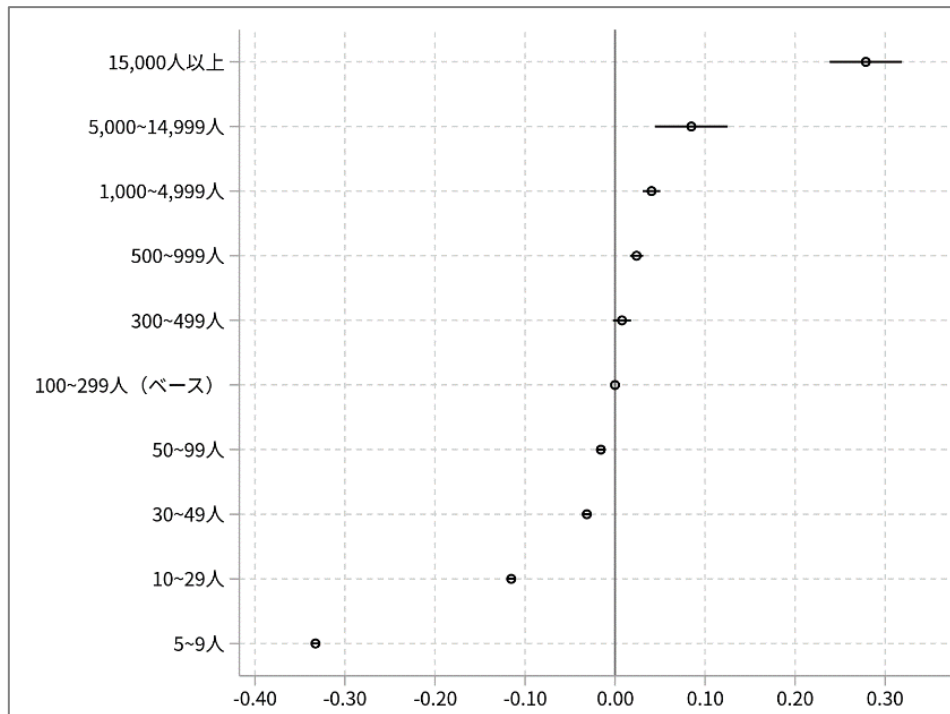
表 2.3.5 賃金構造基本統計調査回答状況：都市規模別（割合）

	2012年	2013年	2014年	2015年	2016年	2017年
全て	0.69	0.69	0.70	0.71	0.70	0.69
政令指定都市	0.67	0.66	0.66	0.67	0.67	0.66
中核市	0.69	0.69	0.72	0.73	0.71	0.70
その他の市	0.70	0.70	0.72	0.72	0.71	0.70
町村	0.70	0.71	0.71	0.71	0.72	0.69



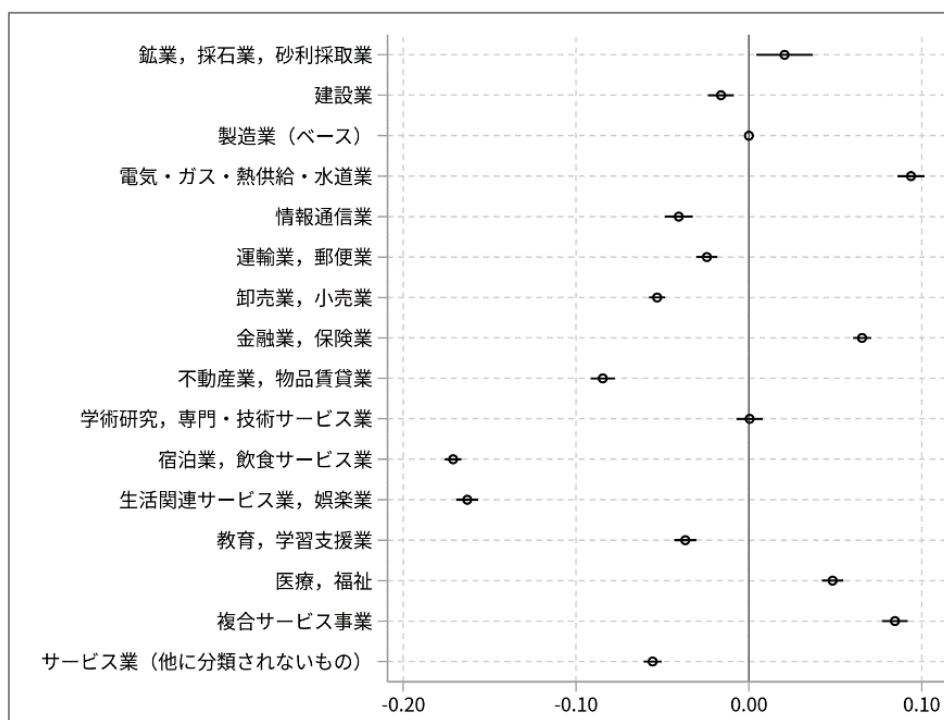
注：データは2012-2017年。

図 2.3.1 賃金構造基本統計調査の回収率の回帰分析結果（年ダミーの係数）



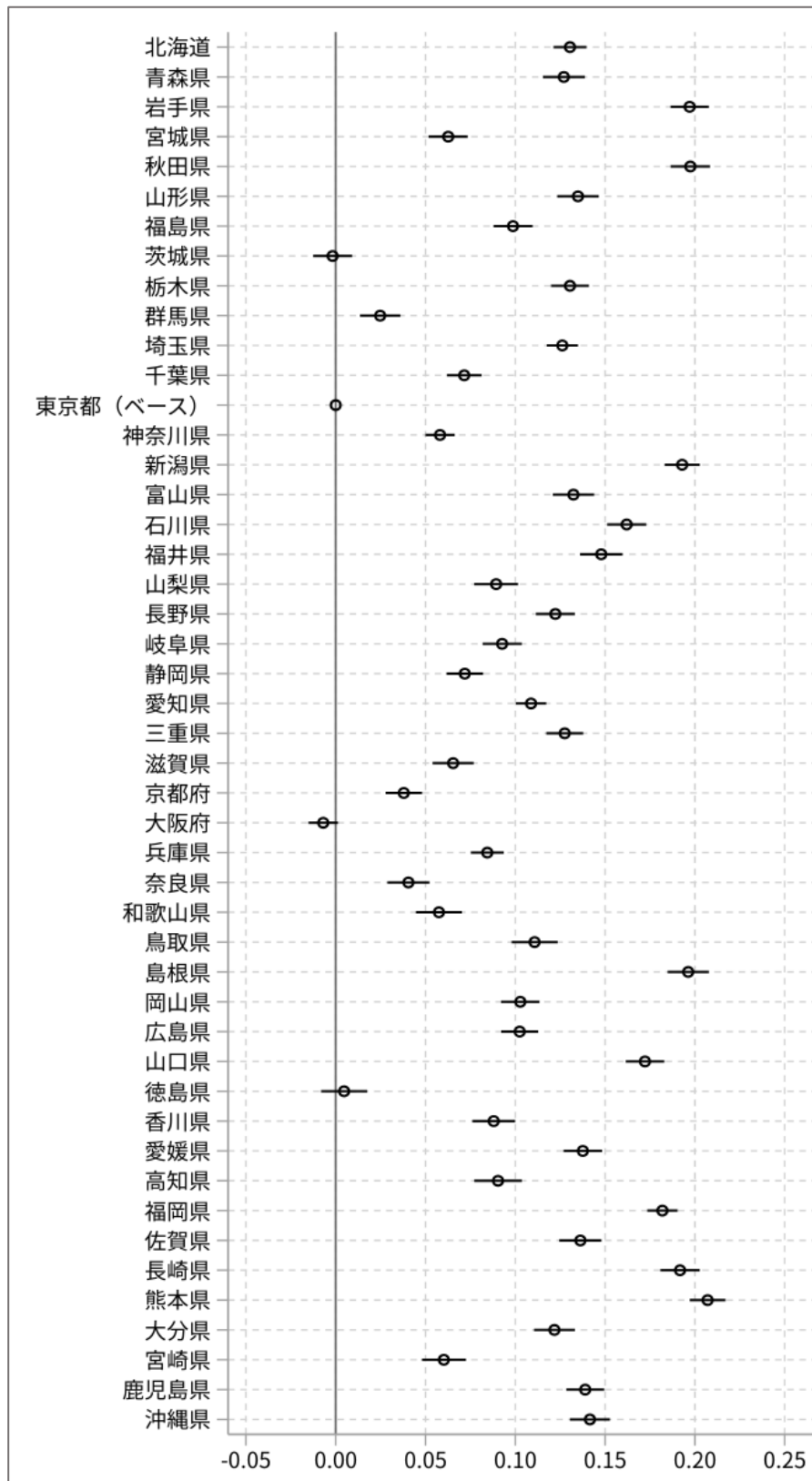
注：データは2012-2017年。

図 2.3.2 賃金構造基本統計調査の回収率の回帰分析結果（事業所規模ダミーの係数）



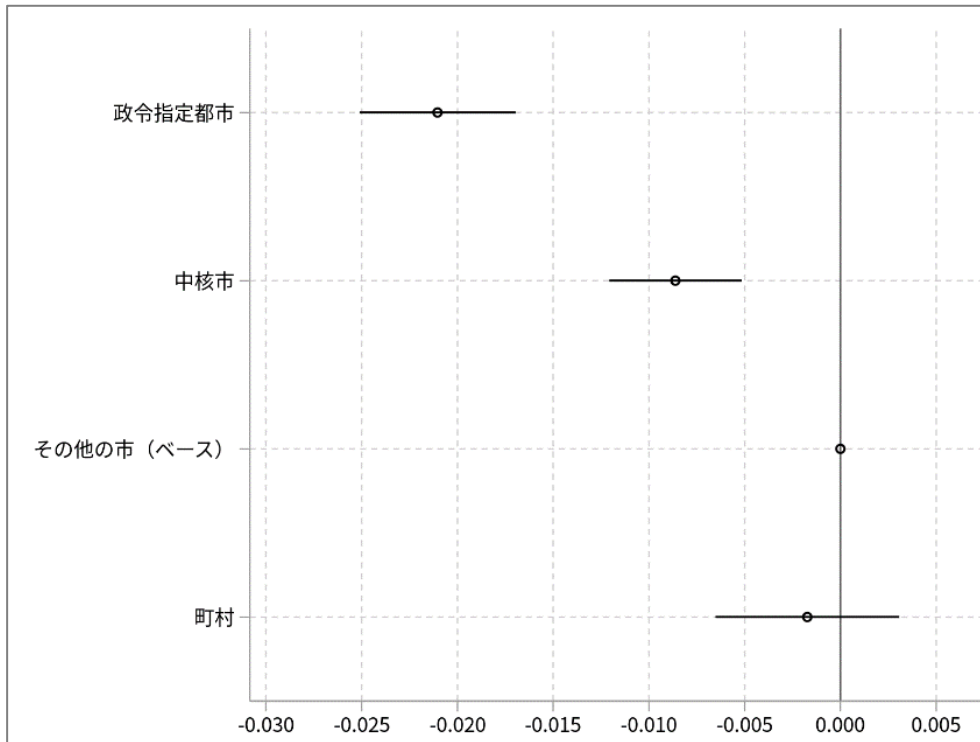
注：データは 2012-2017 年。

図 2.3.3 賃金構造基本統計調査の回収率の回帰分析結果 (産業ダミーの係数)



注：データは2012-2017年。

図 2.3.4 賃金構造基本統計調査の回収率の回帰分析結果（都道府県ダミーの係数）



注：データは2012-2017年。

図 2.3.5 賃金構造基本統計調査の回収率の回帰分析結果（都市規模ダミーの係数）

【民間給与実態統計調査】

賃金構造基本統計調査と同様に、民間給与実態統計調査についても、事業所属性ごとに回収率を計算した。民間給与実態統計調査で利用可能な事業所属性は、事業所規模と産業のみであるため、その2つの属性に絞って分析を行った。ここでは2012年から2019年にかけてのデータを用いている。

年ごとの回収率について、2012年から18年にかけての回収率は約70%であるが、2019年のみ63%と低い数字となっている。この点については、新型コロナウイルス感染症の流行に伴い回収率が低調だったことに加え、緊急事態宣言の発令により、事業所に対する催促や照会を停止していたことが影響していると考えられる。事業所規模別で見ると（表2.3.6）、賃金構造基本統計調査と同様に小規模事業所の回収率の低さが顕著である。その一方、100人以上の事業所に関しては、事業所規模と回収率の間に明確な関係は認められない。産業別（表2.3.7）では、「小売業」、「料理・旅館・飲食店業」、「製造小売業」、「サービス業」などの回収率が低い。また、法人と比べて個人事業主の回収率は顕著に低い数字となっている。

図2.3.6から図2.3.8には、賃金構造基本統計調査と同様に調査回答有無を被説明変数とし、調査年（2012年）、産業（製造業）、事業規模（100-299人）の各属性ダミー変数を説

表 2.3.6 民間給与実態統計調査の回答状況：事業所規模別（割合）

	2012年	2013年	2014年	2015年	2016年	2017年	2018年	2019年
全て	0.73	0.71	0.71	0.70	0.70	0.71	0.73	0.63
500人未満で資本金10億円以上の株式会社の本社	0.84	0.83	0.82	0.83	0.83	0.82	0.84	0.77
5000人以上	0.84	0.82	0.82	0.82	0.82	0.84	0.83	0.73
1000-4999人	0.85	0.84	0.86	0.84	0.84	0.84	0.85	0.72
500-999人	0.85	0.85	0.85	0.85	0.84	0.85	0.85	0.68
100-499人	0.82	0.81	0.82	0.81	0.81	0.81	0.82	0.70
30-99人	0.78	0.79	0.79	0.79	0.79	0.80	0.79	0.66
10-29人	0.76	0.73	0.73	0.73	0.72	0.74	0.76	0.61
1-9人	0.54	0.51	0.48	0.48	0.48	0.49	0.52	0.49

明変数とした回帰分析結果を示しているが（カッコ内は参照カテゴリ）、各図からも上記の傾向を確認することができる⁹。

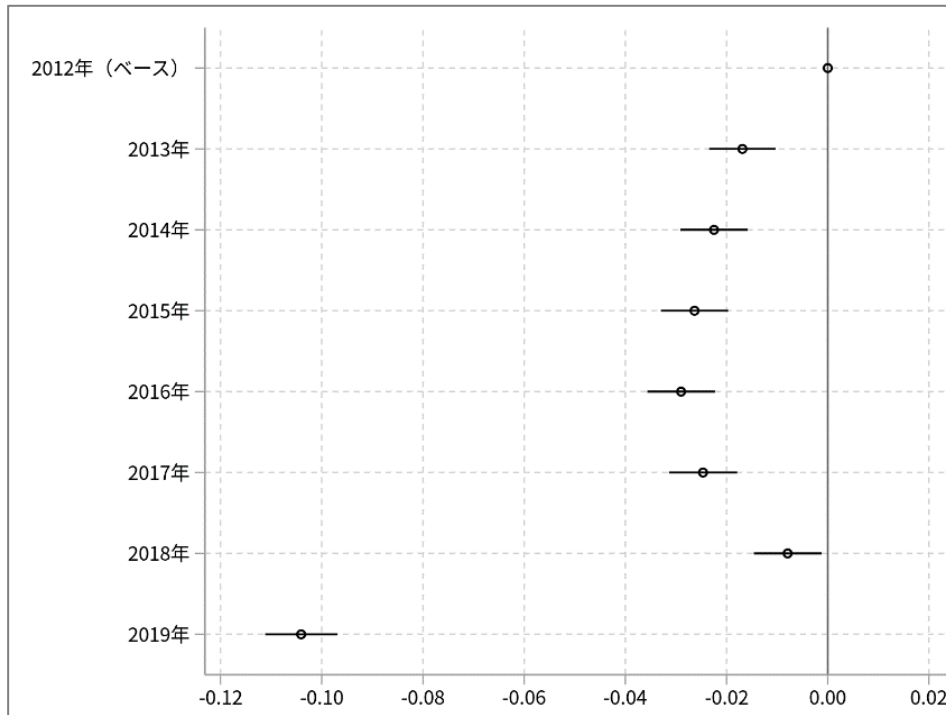
⁹ ここでの推定式は以下のとおり。

$$y_{it} = x'_{it} \beta + \pi_t + u_{it}$$

それぞれの記号などは脚注8と同様だが、民間給与実態統計調査には都道府県や都市規模のデータが含まれていないため、 x'_{it} は産業と事業所規模の各属性ダミーのみを含んでいる。

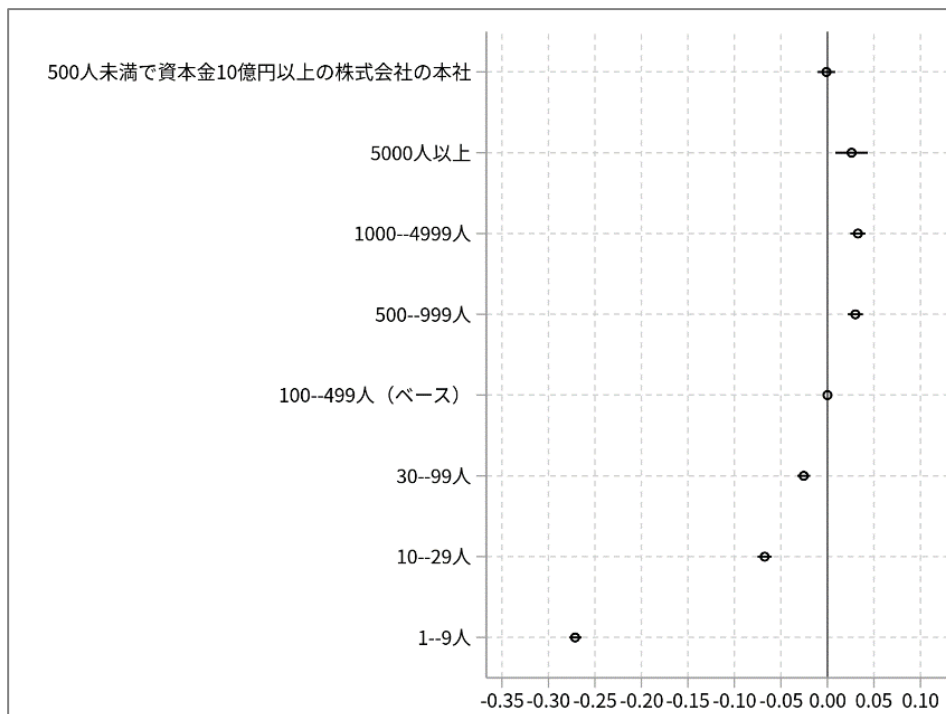
表 2.3.7 民間給与実態統計調査の回答状況：産業別（割合）

	2012年	2013年	2014年	2015年	2016年	2017年	2018年	2019年
法人：製造業	0.82	0.80	0.80	0.81	0.80	0.82	0.83	0.73
法人：卸売業	0.76	0.76	0.76	0.74	0.74	0.75	0.77	0.64
法人：小売業	0.73	0.73	0.71	0.69	0.72	0.70	0.74	0.63
法人：建設業	0.73	0.70	0.69	0.69	0.68	0.68	0.71	0.63
法人：運送業	0.84	0.82	0.82	0.82	0.83	0.83	0.84	0.73
法人：サービス業	0.75	0.73	0.73	0.73	0.72	0.73	0.74	0.62
法人：料理・旅館・飲食店業	0.67	0.61	0.63	0.60	0.62	0.62	0.63	0.46
法人：その他	0.80	0.78	0.78	0.77	0.78	0.78	0.80	0.72
個人：小売業	0.40	0.38	0.37	0.39	0.36	0.37	0.41	0.39
個人：卸売業	0.52	0.52	0.44	0.38	0.34	0.52	0.43	0.50
個人：製造小売業	0.48	0.38	0.44	0.37	0.41	0.44	0.42	0.40
個人：製造卸売業	0.44	0.48	0.42	0.51	0.42	0.47	0.36	0.59
個人：受託加工業	0.40	0.52	0.46	0.54	0.49	0.42	0.53	0.49
個人：修理業	0.52	0.49	0.49	0.46	0.45	0.52	0.47	0.43
個人：サービス業	0.46	0.47	0.43	0.36	0.38	0.43	0.40	0.43
個人：建設業	0.50	0.48	0.48	0.46	0.45	0.47	0.46	0.47
個人：その他の営業	0.44	0.37	0.39	0.40	0.39	0.45	0.48	0.38
個人：農業	0.44	0.46	0.44	0.47	0.43	0.45	0.43	0.47
個人：その他の事業	0.63	0.58	0.56	0.54	0.53	0.53	0.58	0.53
個人：その他	0.47	0.47	0.38	0.43	0.38	0.39	0.45	0.43



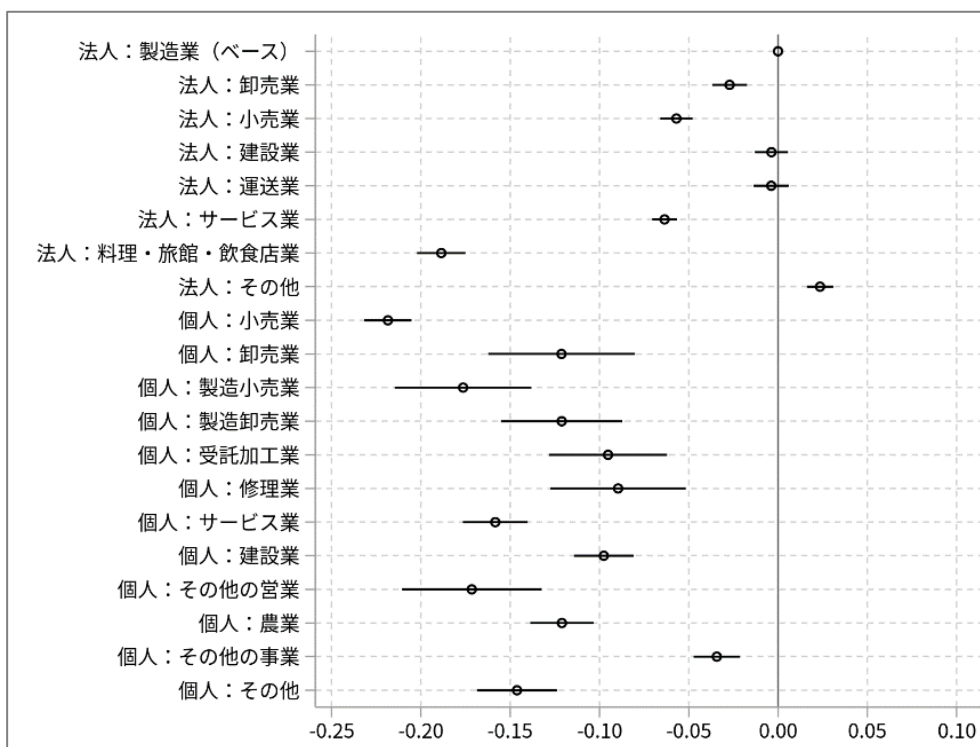
注：データは2012-2019年。

図 2.3.6 民間給与実態統計調査の回収率の回帰分析結果（年ダミーの係数）



注：データは2012-2019年。

図 2.3.7 民間給与実態統計調査の回収率の回帰分析結果（事業所規模ダミーの係数）



注：データは2012-2019年。

図 2.3.8 民間給与実態統計調査の回収率の回帰分析結果（産業ダミーの係数）

(2) パネル構造を用いた分析

調査においては、同じ事業所が複数年にわたって調査対象となることは珍しくない。したがって、事業所の ID 情報を用いて事業所パネルデータを作成することで、同一の事業所の回答パターンの分析を行うことが可能となる。以下では、賃金構造基本統計調査と民間給与実態統計調査のそれぞれについて、複数回調査の対象となった事業所のサンプルを用いて、回答状況の詳細な分析を行う。

【賃金構造基本統計調査】

表 2.3.8 は賃金構造基本統計調査の 2012 年から 2017 年における抽出回数と回答回数の分布を表している。表からは抽出回数に関わらず、毎回回答する事業所の割合が最も高い (5 回以上対象となっている事業所のうち、全回回答している割合は 7 割超) ことがわかる。その一方で、複数年の調査対象になっているにも関わらず、1 度も回答しない事業所も一定数 (同、2-3%) 存在することもわかる。

こうした回答状況の構造を詳しく検証するために、遷移行列の推定を行った。表 2.3.9 には、前回抽出時の回答状況別に次回抽出時の回答確率を示している。それによると、前回抽出時に回答した事業所の約 90%が次回の抽出時も回答している一方で、前回抽出時に回答

表 2.3.8 賃金構造基本統計調査の抽出回数と回答回数の分布

調査回答回数	調査対象回数					
	1回	2回	3回	4回	5回	6回
0回	102,099	8,236	862	166	55	9
1回	172,323	11,276	978	136	50	6
2回	0	38,875	2,155	251	50	7
3回	0	0	8,298	594	89	12
4回	0	0	0	2917	284	6
5回	0	0	0	0	2067	37
6回	0	0	0	0	0	216
事業所数合計	274,422	58,387	12,293	4,064	2,595	293
全回答の割合	0.628	0.666	0.675	0.718	0.797	0.737
回答0回の割合	0.372	0.141	0.070	0.041	0.021	0.031

注：データは2012-2017年。

表 2.3.9 賃金構造基本統計調査の調査回答の遷移確率：前回の回答状況別

前回の回答状況	今期の回答状況（割合）		観測数	事業所数
	回答	未回答		
回答	0.876	0.124	88,487	64,136
未回答	0.346	0.654	18,523	16,338

注：データは2012-2017年。

表 2.3.10 賃金構造基本統計調査の調査回答の遷移確率：前々回と前回の回答状況別

前々回 → 前回の回答状況	今期の回答状況（割合）		観測数	事業所数
	回答	未回答		
回答 → 回答	0.920	0.080	24,082	15,688
未回答 → 回答	0.715	0.285	1,600	1,588
回答 → 未回答	0.524	0.476	1,620	1,613
未回答 → 未回答	0.224	0.776	2,076	1,659

注：データは2012-2017年。

しなかった事業所が次回回答する確率は約35%にとどまっている。また、表2.3.10は前々回抽出時と前回抽出時の回答状況別に次回抽出時の回答確率を計算した結果を示している。その結果からは、前回抽出時に回答している事業所でも、前々回の回答状況によって、次回抽出時の回答確率は大きく異なることが見て取れる（前回と前々回に回答した事業所が次

表 2.3.11 賃金構造基本統計調査の事業所規模別抽出回数（割合）

事業所規模	抽出回数		
	1回以上	2回以上	3回以上
15,000人以上	0.0000	0.0001	0.0002
5,000-14,999人	0.0005	0.0012	0.0028
1,000-4,999人	0.0110	0.0241	0.0499
500-999人	0.0223	0.0462	0.0884
300-499人	0.0125	0.0202	0.0293
100-299人	0.1134	0.1880	0.2770
50-99人	0.0969	0.1192	0.1214
30-49人	0.1620	0.1822	0.1660
10-29人	0.4235	0.3172	0.1922
5-9人	0.1577	0.1017	0.0729
観測数	459,064	184,642	67,868

注：データは2012-2017年。

回答する確率は 92%なのに対し、前回回答していても前々回答していない事業所では次の回答確率が約 72%にとどまる)。こうしたことから、特定の事業所に未回答が生じがちな傾向があることがわかる。

なお、表 2.3.9 と表 2.3.10 の分析対象は 2012 年から 2017 年の間にそれぞれ 2 回、3 回以上抽出された事業所であり、母集団よりも大規模な事業所が多く含まれている点には留意が必要である（表 2.3.11）¹⁰。

【民間給与実態統計調査】

民間給与実態統計調査を用いた同様の分析結果を表 2.3.14 に示す。ここでは 2012 年から 2019 年のデータを用いている。賃金構造基本統計調査と同様に、抽出回数に関わらず、毎回回答する事業所の割合が最も高い（5 回以上対象となっている事業所のうち、全回回答している割合はおよそ 5 割）ことがわかるが、複数年の調査対象となっているものの、全て未回答としている事業所も一定数（同、2-4%）存在している。

前々回・前回の回答状況を踏まえた分析についても同様に行い、その結果を表 2.3.14 から表 2.3.12 に示す。分析結果はほぼ賃金構造基本統計調査と同様であり、前回・前々回の回答状況が次の回答に大きく影響を及ぼすこと、以前は調査協力していた事業所でも、一度未回答となると、その後の調査でも未回答となる傾向があるという結果が得られてい

¹⁰ なお、規模の大きい事業所は抽出確率が高いため、本分析に規模の大きい事業所が多く含まれること自体は自然なことである。

表 2.3.14 民間給与実態統計調査の抽出回数と回答回数の分布

調査回答回数	調査対象回数							
	1回	2回	3回	4回	5回	6回	7回	8回
0回	52,638	1,136	229	80	41	35	15	95
1回	86,892	2,034	292	104	65	22	22	73
2回	0	6,524	624	174	84	35	18	104
3回	0	0	1,942	386	113	74	36	128
4回	0	0	0	1,031	277	77	50	176
5回	0	0	0	0	607	186	85	233
6回	0	0	0	0	0	420	159	428
7回	0	0	0	0	0	0	434	897
8回	0	0	0	0	0	0	0	2,996
事業所数合計	139,530	9,694	3,087	1,775	1,187	849	819	5,130
全回答の割合	0.623	0.673	0.629	0.581	0.511	0.495	0.530	0.584
回答0回の割合	0.377	0.117	0.074	0.045	0.035	0.041	0.018	0.019

注：データは2012-2019年。

表 2.3.13 民間給与実態統計調査の調査回答の遷移確率：前回の回答状況別

前回の回答状況	今期の回答状況（割合）		観測数	事業所数
	回答	未回答		
回答	0.898	0.102	59,986	19,967
未回答	0.424	0.576	11,024	6,335

注：データは2012-2019年。

表 2.3.12 民間給与実態統計調査の調査回答の遷移確率：前々回と前回の回答状況別

前々回 → 前回の回答状況	今期の回答状況（割合）		観測数	事業所数
	回答	未回答		
回答 → 回答	0.920	0.080	38,185	11,055
未回答 → 回答	0.732	0.268	3,269	2,887
回答 → 未回答	0.573	0.427	3,131	2,775
未回答 → 未回答	0.288	0.712	3,884	1,880

注：データは2012-2019年。

表 2.3.15 民間給与実態統計調査の事業所規模別抽出回数（割合）

事業所規模	抽出回数		
	1回以上	2回以上	3回以上
500人未満で資本金10億円以上の株式会社の本社	0.1066	0.2588	0.3138
5000人以上	0.0230	0.0533	0.0654
1000-4999人	0.1435	0.3517	0.4314
500-999人	0.0764	0.1654	0.1556
100-499人	0.1367	0.1210	0.0313
30-99人	0.1046	0.0264	0.0022
10-29人	0.0921	0.0083	0.0001
1-9人	0.3171	0.0150	0.0001
観測数	233,081	93,551	74,163

注：データは2012-2019年。

る。なお、分析対象に母集団よりも大規模な事業所が多く含まれている点についても、賃金構造基本統計調査を用いた分析と同様である（表 2.3.15）。

2.3.3. 調査結果からの示唆

調査への回答状況の分析を行った結果、賃金構造基本統計調査と民間給与実態統計調査ともに、共通の属性を持つ事業所における回答率が低い傾向にあることが明らかになった。まず、事業所規模では小規模事業所での回答率が低く、産業別で見ると、小売業やサービス業、宿泊・飲食サービス業などの産業で回答率が低い傾向がみられる。また、賃金構造基本統計調査の結果からは、東京都や大阪府といった地域や、政令指定都市や中核市といった都市部での回答率が他の地域よりも低いことが明らかになった。また、複数回調査対象となった事業所のパネルデータを用いた分析からは、賃金構造基本統計調査と民間給与実態統計調査に共通して、一度も回答していない事業所が一定数存在すること、および前回・前々回の回答状況が次の回答有無に大きく影響を及ぼすことが明らかになった。

本分析では、こうした傾向を実証するにとどまっているが、今後はこうした傾向が生じる要因について定性・定量の両面から分析を行い、回答率を改善するための取組を検討することが重要である。その際には、ナッジ¹¹などの行動経済学の理論を活用した取組を検討することも有効であると考えられるとともに、そうした取組の効果を適切な方法で検証し、そのエビデンスに基づく政策立案を行うことが重要であると考えられる。

¹¹ 直訳すると「ひじで軽くつつく」の意で、多額の経済的インセンティブや罰則といった手段を用いるのではなく、人が意思決定する際の環境をデザインすることで、自発的な行動変容を促すアプローチ。例えば、調査依頼のメッセージを工夫するなどの方法が考えられる。

2.4. 行政記録情報の活用可能性の検討

2.4.1. 本調査項目の目的

2.3でも上述したように、調査への回答率を高めることは統計の精度を高めるうえで非常に重要な点となる。回答率の向上のためには回答者の負担を軽減することが重要であり、そのための取組として行政記録情報の活用が進められている。また、行政記録情報の活用は調査の正確性を高めることに寄与することも期待される。したがって、本調査項目においては、行政記録情報活用の取組として導入が進められている源泉徴収票等情報転記ツール¹²（以下、「転記ツール」という）について、転記ツールの利用状況の整理、および回答状況に及ぼす影響の検証を行う。

2.4.2. 調査方法・結果

(1) 転記ツールの利用状況

転記ツールの利用に関する情報は、2020年1月10日から2月5日までに民間給与実態統計調査にオンラインで回答した事業所を対象としたアンケート調査によって得られたものである。アンケート調査に回答した1086事業所のうち、344事業所（31.7%）が転記ツールを利用していると回答し、742事業所（68.3%）が未利用と回答した。

事業所の規模別では、5000人以上の大規模事業所で利用率が高い一方で、中規模事業所における利用率は低い傾向がある（表 2.4.3）。また、産業別に見ると、製造業、サービス業、運送業で高く、建設業では低い利用率となっている（表 2.4.3）。

表 2.4.1 転記ツールの利用状況

	事業所数	割合
調査対象	1,839	-
回答	1,086	0.591
未回答	753	0.409
利用	344	0.317
未利用	742	0.683

注：データは2020年に実施のアンケート調査結果。

¹² 事業所が民間給与実態統計調査の調査票を記入する段階において、オンラインなどで税務署に提出した源泉徴収票等のデータを取り込み、共通する項目を自動的に調査票に転記するツール。

表 2.4.2 転記ツールの利用状況：事業所規模別（割合）

事業所規模	利用有
500 人未満で資本金 10 億円以上の株式会社の本社	0.456
5000 人以上	0.849
1000--4999 人	0.318
500--999 人	0.487
100--499 人	0.358
30--99 人	0.273
10--29 人	0.334
1--9 人	0.429

注：利用割合はアンケート回答事業所に占める利用事業所の割合。データは 2020 年に実施のアンケート調査結果。

表 2.4.3 転記ツールの利用状況：産業別（割合）

産業	利用率
法人：製造業	0.400
法人：卸売業	0.380
法人：小売業	0.425
法人：建設業	0.332
法人：運送業	0.537
法人：サービス業	0.402
法人：料理・旅館・飲食店業	0.407
法人：その他	0.310

注：利用割合はアンケート回答事業所に占める利用事業所の割合。データは 2020 年に実施のアンケート調査結果。

(2) 転記ツールの回答状況への影響

転記ツールの利用が回答状況に及ぼす影響を検証するために、ここでは税額率（年間税額/年間給与支払額）に対する影響について分析した。分析に当たっては、転記ツールの利用有無別に源泉徴収義務者のデータを用いた固定効果モデル（2.5.1）と、転記ツールの利用有無に回答した源泉徴収義務者のデータを用いた差の差モデル¹³（2.5.2）を用いて推定を行っ

¹³ 転記ツールの利用事業所における利用前（t-1 期）と後（t 期）の差と、未利用事業所の同時期（t-1 期, t 期）の差を比較することで転記ツール利用の効果を検証する分析方法。t 期と t-1 期の差について、利用事業所と未利用事業所の差を取ることから「差の差」と呼ばれる。

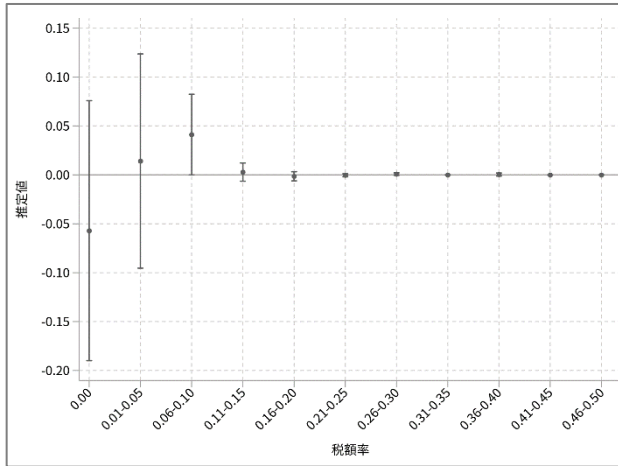
た。なお、この分析では民間給与実態統計調査の 2018 年と 2019 年のデータを利用している。

$$y_{ijt}^k = \gamma After_t + x'_{ijt}\beta + \eta_j + \epsilon_{ijt} \quad (2.5.1)$$

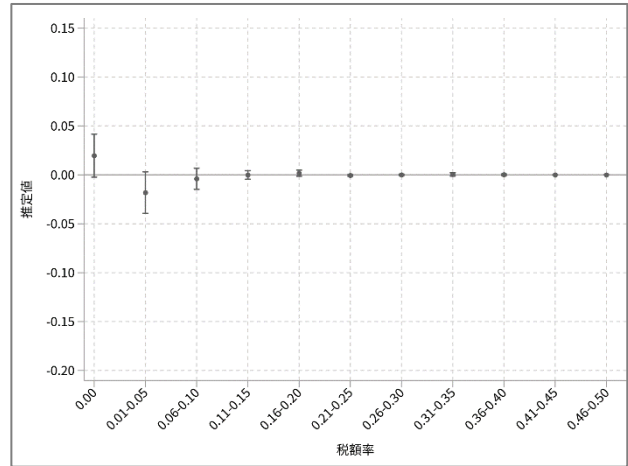
$$y_{ijt}^k = \delta After_t \cdot Treat_j + \gamma After_t + x'_{ijt}\beta + \eta_j + \epsilon_{ijt} \quad (2.5.2)$$

ここでは、 i, j, k はそれぞれ労働者、源泉徴収義務者、調査年を示す添え字である。 y は給与所得者票から計算される労働者の税率額が k 番目の層に入っているか否かを示すダミー変数、 $After_t$ は 2019 年のデータであれば 1 を示すダミー変数である。 $Treat_j$ は 2019 年に転記ツールを利用した場合に 1 を、そうでない場合に 0 を取るダミー変数である。 x_{ijt} は労働者の属性を示す変数群（性別、年齢、勤続年数、12 カ月間給与支払い）である。 η_j は源泉徴収義務者の固定効果を示している。

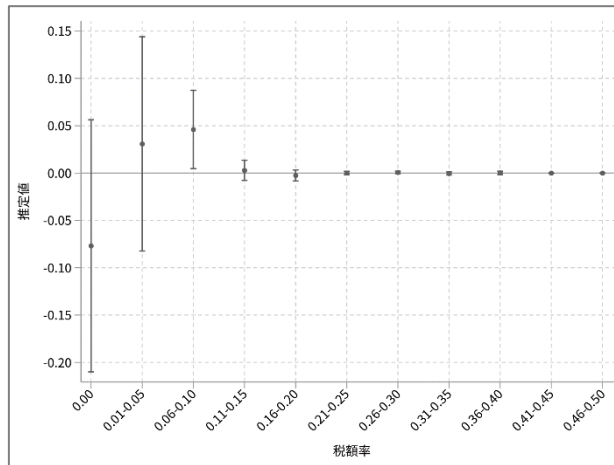
図 2.4.1 はそれぞれのモデルの推定結果を示している。分析の結果、固定効果モデル、差の差モデルのいずれにおいても、転記ツールを利用した場合、統計的には有意ではないが、税額ゼロの回答が減少し、税額率 1-5% の回答が増加する傾向がみられた。また、税額率 6-10% の回答も増えており、こちらは統計的にも有意な結果となっている。なお、転記ツールを利用しなかった源泉徴収義務者には上記の傾向が見られず（図 2.4.1(a)）、また転記ツールを利用した源泉徴収義務者の人員数、支払い給与総額、源泉徴収総額についても 2018 年から 2019 年にかけての変化が無いことも確認できていることから、こうした変化が単純な時系列トレンドでは説明できないと考えられる。したがって、これらの結果を踏まえると、転記ツールを活用した事業所では申告税額に変化が生じており（税額ゼロが減り、税額率 1-10% の割合が増えた）、それらは転記ツールによる効果であることが示唆される。



(a) 固定効果モデル：転記ツール利用有



(b) 固定効果モデル：転記ツール利用無



(c) 差の差モデル

注：(a)と(b)は式(2.5.1)の推定値と95%信頼区間、(c)は式(2.5.2)の推定値と95%信頼区間を表している。標準誤差は源泉徴収義務者レベルでクラスタリングした。データは転記ツールの利用に関しては2020年に実施のアンケート調査結果、税額率などについては2018-19年の民間給与実態統計調査。

図 2.4.1 転記ツールの税額率への影響

2.4.3. 調査結果からの示唆

分析の結果、転記ツールの利用についてはアンケート調査に回答した企業の30%程度にとどまっており、規模や産業によって利用率に差が見られることが明らかになった。転記ツールは回答者の負担軽減につながると考えられることから、ツールの活用率を高める取組を進めることが重要であると考えられる。

また、転記ツールを利用している事業所では、転記ツールの利用によって申告税額が増えたことを示唆する結果が得られている。サンプルサイズや利用できるデータの期間が限られており、さらなる分析が必要となるが、転記ツールの利用により回答の正確性向上も期待

できる可能性がある。したがって、回答者の負担軽減に加え、正確性向上の観点からも転記ツールの活用を促していくことは重要であると考えられる。また、転記ツールに限らず、他の行政記録情報の活用を積極的に検討していくことが必要であると考えられる。

2.5. 母集団情報の比較検証

2.5.1. 本調査項目の目的

賃金構造基本統計調査では 5 人以上の事業所が母集団となっている一方で、民間給与実態統計調査では従業員 1 人以上の源泉徴収義務者が対象となっており、賃金構造基本統計調査ではカバーされていない 1-4 人事業所の労働者も調査対象となっている。また、フリーランスなどの雇人のいない個人事業主については、源泉徴収義務者に該当しないことが多いため、その場合は民間給与実態統計調査でもカバーされない。また、賃金構造基本統計調査の対象は労働者に限られており、事業主や役員などは含まれていない。本調査項目では、こうした調査の母集団にそもそも含まれない層の存在が賃金統計に及ぼす影響について検討を行う。

2.5.2. 調査方法・結果

(1) 従業員 1-4 人事業所について

上述のとおり、民間給与実態統計調査では、従業員数 1 人以上の源泉徴収義務者が調査対象となっており、賃金構造基本統計調査の対象とはならない 1-4 人事業所の労働者も調査対象となっている¹⁴。ここでは、民間給与実態統計調査の 2012 年から 2019 年のデータを用いて、従業員数 4 人以下の源泉徴収義務者に雇用されている労働者の属性や所得を明らかにする。

まず、源泉徴収義務者全体に占める従業員数 4 人以下の源泉徴収義務者の割合は 83%を占めている。その一方で、従業員数 4 人以下の源泉徴収義務者に従事している労働者は労働者全体の 19%である。

そのうえで、従業員数 1-4 人の事業所と従業員数 5 人以上の事業所に勤務している労働者の比較を行った結果を表 2.5.1 に示す。分析からは、従業員数 1-4 人の事業所に勤務している労働者は、女性の比率が高い (4.2 ポイント)、年齢が高い (7.5 歳)、勤続年数が長い (3.7 歳) という傾向がみられた。また、所得に関しては、年齢が高い・勤続年数が長いという傾向があるにも関わらず、給料・手当等が 9.5%、賞与等が 69.2%低く、給与総額では 19.4%ほど低い数字になっている。従業員数 4 人以下の事業所を含めた平均給与総額は 377 万円であることから、従業員数 4 人以下の事業所を含めずに平均給与総額を算出した値と比べ、4%ほど低い値になっている。給与額 (対数値) に関する回帰分析結果 (表 2.5.2) をみても、11.1% (給料・手当等) ないし 18.8% (給与総額) 低いとの結果が得られ ((1)列、(3)列)、労働者属性 (性別・年齢・勤続年数) を制御すると給与額の差は、17.8% (給料・手当等)

¹⁴ 厳密には、民間給与実態統計調査では従業員数なのに対して、賃金構造基本統計調査では常用労働者数によって調査対象が定義されている。また、調査対象も源泉徴収義務者と事業所で若干の違いがあることに注意が必要である。

表 2.5.1 従業員数4人以下の源泉徴収義務者に従事している労働者の属性

	従業員数 4 人以下	従業員数5人以上	差：(1)-(2)
	(1)	(2)	(3)
女性（割合）	0.470 [0.499]	0.427 [0.495]	0.042 (0.003)
年齢（歳）	50.902 [15.201]	43.439 [13.815]	7.463 (0.076)
勤続年数（年）	13.212 [13.175]	9.524 [10.295]	3.688 (0.077)
給料・手当等（100万円/年）	2.952 [3.022]	3.263 [3.190]	-0.311 (0.019)
賞与等（100万円/年）	0.201 [0.542]	0.652 [1.142]	-0.451 (0.008)
合計給与（100万円/年）	3.153 [3.174]	3.914 [3.849]	-0.761 (0.025)
観測数	163,051	2,239,078	

注：ブラケット内の値は標準偏差、括弧内の値は標準誤差（事業所レベルでクラスタリング）を表している。データは民間給与実態統計調査の2012-2019年。

表 2.5.2 従業員数 4 人以下の源泉徴収義務者に従事している労働者の属性
(回帰分析結果)

被説明変数	ln(給料・手当等)		ln(給与総額)	
	(1)	(2)	(3)	(4)
従業員数 4 人以下事業所	-0.111 (0.008)	-0.178 (0.005)	-0.188 (0.009)	-0.252 (0.005)
労働者属性	制御無	制御有	制御無	制御有
観測数	2,402,114	2,402,114	2,402,114	2,402,114

注：括弧内の値は標準誤差（事業所レベルでクラスタリング）を表している。給与額が 0 のレコードは分析から除外した（15 レコード）。制御した労働者属性は性別、年齢、勤続年数。データは民間給与実態統計調査の 2012-2019 年。

ないし 25.2%（給与総額）とさらに広がる（(2)列、(4)列）¹⁵。

¹⁵ ここでの推定式は以下のとおり。

$$\ln(\text{earnings}_{it}) = x'_{it} \beta + u_{it}$$

i, t はそれぞれ労働者と調査年を表している。説明変数の x'_{it} は労働者属性（性別・年齢・勤続年数）のベク

(2) フリーランス労働者について

賃金構造基本統計調査は従業員 5 人以上の事業所を対象としており、民間給与実態統計調査でも源泉徴収義務者ではない自営業者が含まれていないため、いわゆるフリーランス労働者はいずれの調査にもカバーされていない、したがって、ここでは就業構造基本調査を用いて、同調査における雇人の無い自営業者を本分析におけるフリーランスと定義し¹⁶、その割合や所得分布などを検証した。

雇人の無い自営業者の割合は、図 2.5.1 に示すように一貫して減少傾向にあり、1977 年の約 13%から 2017 年では就業者の約 6%となっている。他方、自営業者に占める雇人の無い自営業者の割合は 70%程度と高いうえ、近年は増加傾向にある。

2017 年の就業構造基本調査における所得分布（表 2.5.3）を見ると、雇人の無い自営業者は低所得である傾向が強く、約 60%が 200 万円未満となっている。雇人の無い自営業者を除く就業者（フリーランス以外の就業者）における 200 万円未満労働者が 33%程度、雇人のある自営業者における同割合が 30%程度であることと比べると、非常に高い数字となっている。また、各所得階級の中央値を用いて計算した「平均所得」でも、雇人の無い自営業

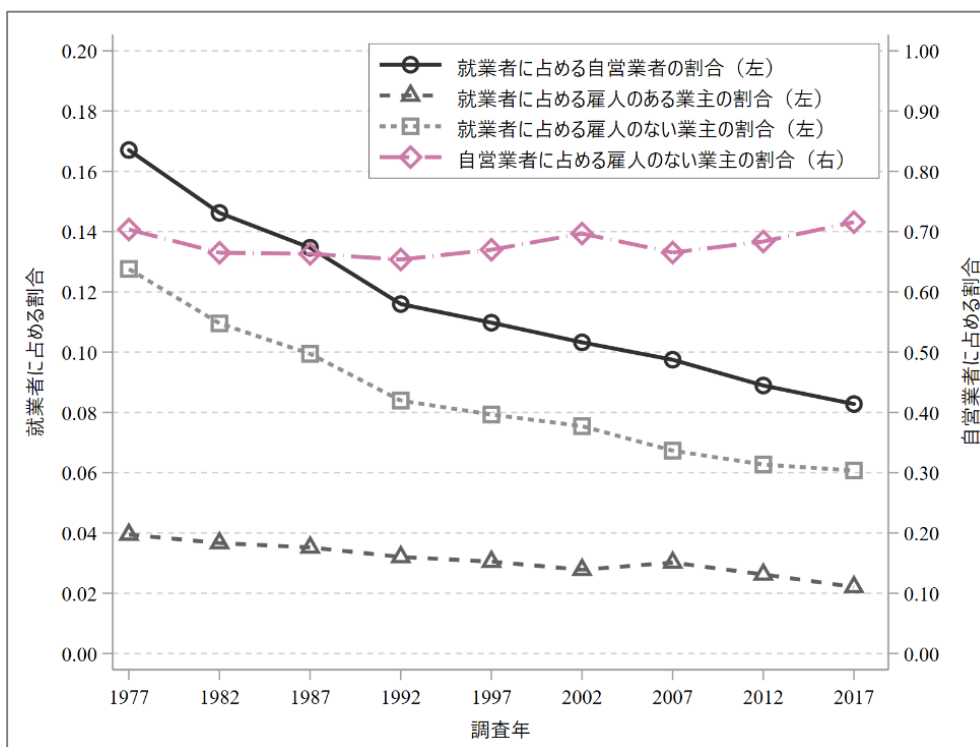


図 2.5.1 就業構造基本調査における自営業者の割合と構成の推移

トルで、被説明変数は給料・手当等、および給与総額の自然対数値である。

¹⁶ 内閣府が実施している調査などにおけるフリーランスの定義では、「店舗のある個人店主」や「農林漁業従事者」が除外されていることも多く、それらの調査との定義の違いについては留意が必要である。

表 2.5.3 所得分布：2017年就業構造基本調査（割合）

	全労働者 (1)	雇用者または雇用のある業主 (2)	雇用者 (3)	自営業者		雇用のない自営業者の累積割合 (6)
				雇用あり (4)	雇用なし (5)	
就業者に占める割合	1.000	0.939	0.917	0.022	0.061	
所得分布						
50万円未満	0.059	0.047	0.046	0.083	0.244	0.244
50-99万円	0.114	0.112	0.113	0.066	0.147	0.391
100-149万円	0.104	0.102	0.103	0.084	0.127	0.518
150-199万円	0.075	0.074	0.074	0.080	0.087	0.605
200-249万円	0.107	0.108	0.108	0.102	0.092	0.697
250-299万円	0.080	0.081	0.081	0.078	0.067	0.764
300-399万円	0.139	0.141	0.142	0.133	0.097	0.861
400-499万円	0.100	0.103	0.103	0.092	0.052	0.913
500-599万円	0.070	0.073	0.073	0.069	0.029	0.942
600-699万円	0.048	0.050	0.050	0.039	0.017	0.959
700-799万円	0.035	0.037	0.037	0.027	0.012	0.971
800-899万円	0.022	0.023	0.023	0.023	0.007	0.978
900-999万円	0.014	0.014	0.014	0.020	0.005	0.983
1000-1249万円	0.020	0.020	0.020	0.039	0.007	0.99
1250-1499万円	0.005	0.006	0.005	0.015	0.002	0.992
1500万円以上	0.009	0.009	0.008	0.051	0.004	0.996
平均年間所得（万円）	345.6	354.5	352.8	425.5	209.7	

注：端数処理の関係で、(6)の累計は1にならない。データは就業構造基本調査の2017年。

者を除く就業者が 355 万円、雇人のある自営業者が 426 万円なのに対して、雇人の無い自営業者は 210 万円と極めて低い数値となっている。

次に、フリーランス労働者（雇人の無い自営業者）の属性情報を集計した。表 2.5.4 からわかる雇人の無い自営業者の特徴として、高年齢、男性の割合が高い、婚姻率が高い、学歴が低いことが挙げられる。表 2.5.5 の現在の仕事について理由を見てみると、知識や技能を活かすためにフリーランスとして働いている割合が高い。その一方で、学校を卒業した、より良い労働条件が見つかった、という理由でフリーランスとして働いている人は少ない。また、表 2.5.6 からわかるように就業継続希望については、転職希望は低く、現在の仕事を続けたいもしくは辞めたいと考えている割合が高くなっている。上述のように、フリーランス労働者はその他の労働者とは異なる特徴を持っており、また景気変動が所得へ与え

表 2.5.4 雇人の無い自営業者（フリーランス）の属性（割合）

属性	雇人の無い 自営業者 (1)	その他の 労働者 (2)	差 (1) - (2) (3)
年齢	57.321 [15.132]	45.369 [14.623]	11.952 (0.135)
女性	0.294 [0.456]	0.446 [0.497]	-0.151 (0.004)
未婚	0.196 [0.397]	0.300 [0.458]	-0.103 (0.004)
在学中	0.003 [0.050]	0.027 [0.162]	-0.025 (0.001)
小学・中学	0.177 [0.382]	0.064 [0.245]	0.113 (0.003)
高校・旧制中	0.403 [0.491]	0.370 [0.483]	0.034 (0.004)
専門学校 (1年以上2年未満)	0.076 [0.265]	0.053 [0.224]	0.023 (0.002)
専門学校 (2年以上4年未満)	0.069 [0.253]	0.086 [0.281]	-0.017 (0.002)
専門学校(4年以上)	0.000 [0.017]	0.002 [0.048]	-0.002 (0.000)
短大	0.050 [0.218]	0.083 [0.277]	-0.033 (0.002)
高専	0.010 [0.097]	0.011 [0.105]	-0.001 (0.001)
大学	0.198 [0.399]	0.296 [0.456]	-0.097 (0.004)
大学院	0.016 [0.126]	0.034 [0.182]	-0.018 (0.001)
観測数	28,649	476,678	

注：ブラケット内の値は標準偏差、括弧内の値は標準誤差を表している。データは就業構造基本調査の2017年。

る影響も異なる可能性があるため、賃金動向を見る際にはこれらの点に留意する必要がある。

表 2.5.5 現在の仕事についての理由（割合）

理由	雇人の無い 自営業者 (1)	その他の 労働者 (2)	差 (1) - (2) (3)
失業していた	0.065 [0.247]	0.103 [0.304]	-0.038 (0.002)
学校を卒業した	0.093 [0.290]	0.219 [0.414]	-0.126 (0.002)
収入を得る必要が生じた	0.156 [0.363]	0.163 [0.369]	-0.007 (0.003)
知識や技能を生かしたかった	0.276 [0.447]	0.123 [0.328]	0.153 (0.004)
社会に出たかった	0.021 [0.143]	0.030 [0.171]	-0.009 (0.001)
時間に余裕ができた	0.032 [0.176]	0.037 [0.188]	-0.005 (0.001)
健康を維持したい	0.032 [0.175]	0.009 [0.096]	0.022 (0.001)
よりよい条件の仕事が見つかった	0.040 [0.196]	0.129 [0.336]	-0.089 (0.002)
その他	0.286 [0.452]	0.186 [0.389]	0.100 (0.004)
観測数	28,649	476,678	

注：ブラケット内の値は標準偏差、括弧内の値は標準誤差を表している。データは就業構造基本調査の2017年。

表 2.5.6 現在の仕事の継続希望（割合）

	雇人の無い 自営業者 (1)	その他の 労働者 (2)	差 (1) - (2) (3)
続けたい	0.822 [0.383]	0.791 [0.406]	0.030 (0.003)
別の仕事もしたい	0.071 [0.257]	0.064 [0.245]	0.007 (0.002)
変わりたい	0.035 [0.184]	0.106 [0.308]	-0.071 (0.002)
やめたい	0.072 [0.259]	0.039 [0.193]	0.034 (0.002)
観測数	28,649	476,678	

注：ブラケット内の値は標準偏差、括弧内の値は標準誤差を表している。データは就業構造基本調査の2017年。

(3) 高所得層について

賃金構造基本統計調査では調査対象を労働者としているため、役員などの比較的高所得である可能性が高いと考えられる層が対象に含まれていない¹⁷。他方、民間給与実態統計調査は給与所得者を対象としており、役員であっても給与所得者であれば調査の対象となる（さらに、調査対象事業所において、給与の年額が2000万円を超える給与所得者は全て調査対象となる）。こうした構造を活用して、高所得層における所得分布の裾の大きさの比較を行い、賃金構造基本統計調査の母集団に含まれない高所得層についての検討を行う。

所得分布の右裾は以下のパレート分布 $F(y)$ で良く近似できることが知られている。

$$F(y) = 1 - \left(\frac{y_m}{y}\right)^\alpha \quad (y \geq y_m)$$

α は所得分布の裾の大きさの特徴づけるパラメータとなっており、 α の値が小さいほど裾の大きい分布となる。つまり α の値が小さいほど高所得者の間の不平等が大きいことになる。特に、 $\alpha \leq 1$ のとき平均が無く、 $\alpha \leq 2$ のとき分散が無い。Atkinson et. al. (2018) にしたがうと、上式は以下のように書ける。

$$\ln y = C + \frac{1}{\alpha} \ln \frac{1}{1 - F(y)}$$

ここで、 C は α と y_m からなる定数である。この式を年間所得が900万円（所得分布の約95

¹⁷ 賃金構造基本統計調査においても、一般の従業者と同様の基準で給与が支払われている者は役員であっても労働者として調査対象に含まれるが、その場合でも調査票に記入される金額は役員報酬を除く給与部分のみである。

表 2.5.7 パレート係数の推定結果

調査年	賃金構造基本統計調査			民間給与実態統計調査		
	男女計	男	女	男女計	男	女
2012年	5.000 (0.023)	5.004 (0.024)	4.934 (0.046)	3.270 (0.002)	3.269 (0.002)	3.282 (0.006)
2013年	4.774 (0.021)	4.771 (0.020)	4.811 (0.074)	2.985 (0.005)	2.983 (0.005)	3.006 (0.013)
2014年	4.714 (0.012)	4.716 (0.013)	4.684 (0.019)	3.083 (0.005)	3.082 (0.005)	3.098 (0.017)
2015年	4.625 (0.009)	4.628 (0.009)	4.597 (0.017)	3.048 (0.006)	3.047 (0.006)	3.054 (0.018)
2016年	4.654 (0.010)	4.655 (0.010)	4.644 (0.018)	3.032 (0.005)	3.029 (0.006)	3.064 (0.011)
2017年	4.767 (0.010)	4.766 (0.010)	4.767 (0.024)	2.942 (0.005)	2.940 (0.005)	2.959 (0.018)

注：括弧内の値は標準誤差（事業所レベルでクラスタリング）を表している。賃金構造基本統計調査の年間所得はきまって支給する現金給与額×12 + 特別支給額で計算した。

パーセンタイル) 以上の層を用いて推定した。表 2.5.7 に示す推定結果を見ると 2012 年から 2017 年の期間において、賃金構造基本統計調査から得られたパレート係数は民間給与実態調査よりも大きなものとなっている。全数調査を行っている民間給与実態調査を基準として考えると、賃金構造基本統計調査では高所得層のカバレッジが十分でないことがデータからも確認することができる。さらに、標準誤差も民間給与実態調査の方が小さくなっており、より正確な推定値が得られていることがわかる。一方、男女差については両調査で共通して限定的なものとなっている。

2.5.3. 調査結果からの示唆

本調査項目では、賃金構造基本統計調査ではカバーされていない従業員 1-4 人事業所の労働者、およびいずれの賃金統計にも含まれていないフリーランス（雇人のいない自営業者）の割合や所得分布を検証した。その結果、1-4 人事業所の労働者は従業員 5 人以上の事業所の労働者と比べ給料・手当等が 9.5% 低く、雇人の無い自営業者（フリーランス）は 2017 年では約 60% の所得が 200 万円未満となっているなど、いずれの労働者も賃金が低い傾向が見られた。

こうした結果からは、小規模事業所の労働者やフリーランスを含む全労働者の賃金水準は、賃金構造基本統計調査や民間給与実態統計調査が示す賃金水準よりも低い水準である

と考えられる。就業構造が大きく変化している中で、特に賃金の低い層がシステムティックに調査対象から抜け落ちてしまうと、労働者全体の賃金水準を正確に把握することができなくなってしまうため、将来的にはこうした層を含めた調査を新規調査も含め検討する必要があると考えられる。

また、賃金構造基本統計調査はその調査の性質上、高所得層のカバレッジについては十分でないことがデータからも明らかになり、高所得層も含めた所得水準を把握するうえでは、留意が必要であると考えられる。

2.6. 品質固定された賃金指数の試算

2.6.1. 本調査項目の目的

高齢化や女性の労働市場への参入、大学進学率の上昇などを通じて労働者の人口構成は刻々と変化しており、平均賃金の変化は必ずしも経済動向と一致しない可能性がある。例えば、女性がパートタイムで労働することが増えると、パートタイム労働者の賃金はフルタイム労働者の賃金より低い傾向にあるため、他の労働者の賃金に変化がなくても平均賃金は下がる。したがって、賃金変動から経済動向を読み取ろうとする場合には、労働者構成の変化に伴う賃金変動と同一属性の労働者の賃金変動を切り分けて考えることが重要である。そこで、本分析では賃金構造基本統計調査のデータを用いて労働者の属性を制御した（品質固定した）賃金指数の算出を行ったうえで、賃金変動の要因についての分析を行う。

2.6.2. 調査方法・結果

(1) 労働者属性を制御した賃金指数の算出

まず、労働者属性を制御した賃金指数の算出を行う。ここでは回帰分析を用いて、以下の3つの指数の算出を行う。

【モデル1：ヘドニック指数】

1989年から2017年の賃金構造基本統計調査のデータをプールしたうえで、労働者の時間当たり賃金（対数値）を被説明変数とし、労働者属性と年ダミーを説明変数とした回帰分析を行った。

$$\ln(\text{wage}_{it}) = x'_{it} \beta + \tau_t + u_{it}$$

ここで得られた $\hat{\tau}_t$ を用いて、以下のとおり賃金指数（ \tilde{w}_t ）を算出する。

$$\tilde{w}_t = \exp(\hat{\tau}_t) - 1$$

時間当たり賃金は（きまって支給する現金給与額＋特別給与額／12）／（所定労働時間＋所定外労働時間）として計算している。説明変数として用いた労働者属性は、年齢（3次項まで）、性別、都道府県、学歴¹⁸である。 τ_t が労働者の属性を制御したうえで t 年の賃金が1989年より何%高いかを示している。

【モデル2：ラスパイレス指数】

¹⁸ 学歴は、中学、高校、高専・短大、大学・大学院、不明の5つのカテゴリとしている。短時間労働者（2004年まではパートタイム労働者）の学歴は調査されていないため、学歴カテゴリは「不明」とした。

以下の手順で、ラスパイレス指数を作成した。

1. 調査年ごとに時間当たり賃金（対数値）を労働者属性に回帰する。したがって推計式は

$$\ln(\text{wage}_{it}) = x'_{it} \beta_t + \tau_t + u_{it}$$

であり、用いた変数はモデル 1 と同じである。）

2. 1. で得られた賃金モデルを 1989 年の労働者属性の分布に当てはめる。つまり、ラスパイレス指数（ \tilde{w}_t ）は以下のように計算される。

$$\tilde{w}_t = \frac{\sum_{i \in I_0} \exp(x'_{i0} \hat{\beta}_t)}{\sum_{i \in I_0} \exp(x'_{i0} \hat{\beta}_0)} - 1$$

ただし、 I_0 は 1989 年の労働者の集合であり、モデル 1 との比較のために、ラスパイレス指数から 1 を引いてある。

モデル 1 では、年齢や教育などの労働者属性の価格が時間を通じて一定であり、賃金変動はどの属性の労働者でも同じであるような状況を想定している。一方、モデル 2 では、労働者属性の価格が時間を通じて変化することを許しており、労働者属性ごとに異なる賃金変動を経験するという、より現実的なモデルとなっている。

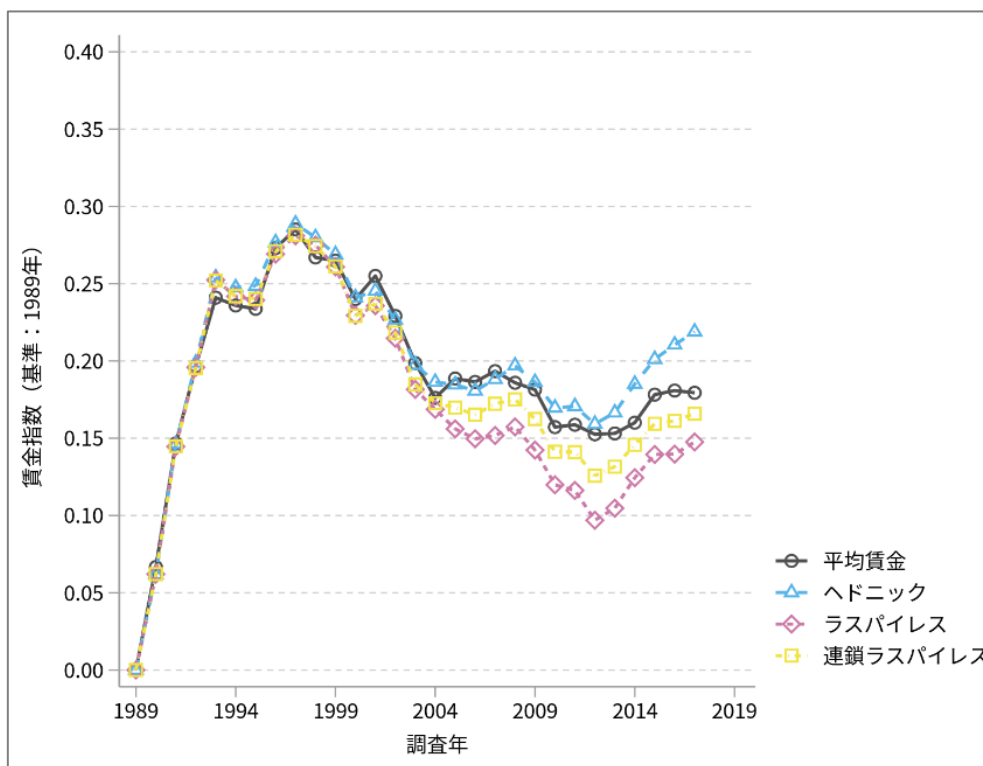
【モデル 3：連鎖ラスパイレス指数】

モデル 2 と同様に調査年ごとに賃金方程式を推定し、以下のように連鎖方式のラスパイレス指数（ \tilde{w}_t ）を作成した（モデル 1 との比較のために、ラスパイレス指数から 1 を引いている）。

$$\tilde{w}_t = \frac{\sum_{i \in I_0} \exp(x'_{i0} \hat{\beta}_1)}{\sum_{i \in I_0} \exp(x'_{i0} \hat{\beta}_0)} \times \frac{\sum_{i \in I_1} \exp(x'_{i1} \hat{\beta}_2)}{\sum_{i \in I_1} \exp(x'_{i1} \hat{\beta}_1)} \times \dots \times \frac{\sum_{i \in I_{t-1}} \exp(x'_{it-1} \hat{\beta}_t)}{\sum_{i \in I_{t-1}} \exp(x'_{it-1} \hat{\beta}_{t-1})} - 1$$

モデル 3 の特徴として、ベースとなる労働者構成が毎年変化していることが挙げられる。これは、労働者構成が初期時点で固定されているモデル 2 のラスパイレス指数と対照的である。指数の計算は複雑となるが、毎年ベースをシフトさせることで、より現実を反映した指数となっている。

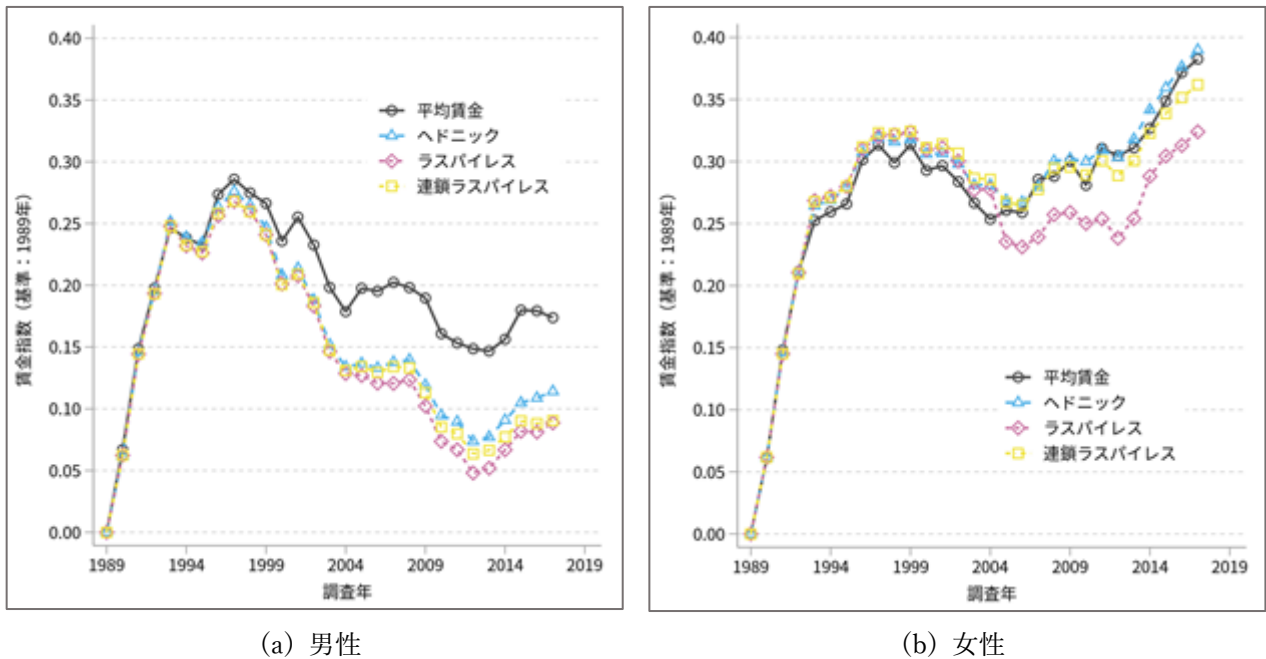
図 2.6.1 に属性の制御を行っていない平均賃金およびモデル 1-3 から作成された賃金指数の推移をプロットしている。モデル 1、2 とともに 2000 年代前半までは平均賃金と似たような変化を示している。モデル 1 は 2000 年代以降も平均賃金と同様の推移を示しているが、2014 年以降は平均賃金よりも高くなっている。このことは、モデル 1 では人的資本（労働者属性）の価格が年によって異なることを許していないため、賃金変化を十分に捉えられて



注：データは賃金構造基本統計調査の1989年から2017年。

図 2.6.1 賃金指数の推移

いない可能性を示している。モデル2は2000年代後半以降、平均賃金よりも大きく減少している。モデル3もモデル2と同様に2000年代後半以降に平均賃金よりも減少しているが、減少幅はモデル2よりも小さい。モデル2とモデル3の乖離の原因の一つとして、男性の賃金低下が考えられる。モデル2は1989年の労働者構成を基準としたラスパイレス指数であるため、男性労働者により大きなウェイトがかかっている。一方、モデル3の連鎖ラスパイレス指数は基準となる労働者構成を毎年シフトさせているため、女性の労働参加率の上昇にともなって、男性労働者へのウェイトは小さくなる。その結果、モデル3と比べて、男性労働者の賃金下落の影響がモデル2に強く反映されていると考えられる。



(a) 男性

(b) 女性

注：データは賃金構造基本統計調査の1989年から2017年。

図 2.6.2 賃金指数の推移（男女別）

図 2.6.2 は男女別の平均賃金と賃金指数をプロットしている。2000 年以降、男性の賃金指数はどのモデルでも平均賃金を大きく下回っている。これは労働人口の高齢化によって、平均賃金を引き上げられているためだと考えられる。女性の賃金指数はモデル 1 とモデル 3 では平均賃金とほぼ同様の動きをしている。モデル 2 では 2005 年以降平均賃金よりやや低い。ラスパイレス指数と連鎖ラスパイレス指数の乖離が男性よりも女性で顕著な理由として、2000 代から 2010 年代にかけての M 字カーブの緩和が考えられる。上述のとおり、ラスパイレス指数は 1989 年時点の労働者構成を基準としているため、M 字カーブの山にあたる 20 代前半の女性や 40-50 歳の女性労働者に大きなウェイトがかかっている。M 字カーブの緩和に伴い、20 代後半から 30 代の女性の労働者の重要性が増す中で、（連鎖方式ではない）ラスパイレス指数では現実の労働市場をうまく反映できていない可能性がある。

(2) 平均賃金と賃金指数の乖離について

2000 年代後半からラスパイレス指数や連鎖方式ラスパイレス指数が平均賃金を下回っているということは、この期間に賃金の高い傾向にある労働者が増加したことを意味している。ここでは、各属性がそれぞれ平均賃金と賃金指数の乖離にどの程度寄与しているのかについて分析を行う。

各属性の寄与度の分析に当たっては以下の式を推定した。

$$wage_{it} = \beta_{0t} + x_{1it}\beta_{1t} + \dots + x_{Kit}\beta_{Kt} + u_{it}$$

表 2.6.1 平均賃金と賃金指数の乖離の要因分解

	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015
合計	0.0123	0.0137	0.0068	-0.0009	0.0047
年齢	0.0073	0.0112	0.0081	0.0059	0.0066
女性	-0.0046	-0.0013	-0.0081	-0.0058	-0.0062
教育	0.0126	0.0057	0.0007	0.0039	0.0039
都道府県	-0.0031	-0.0018	0.0061	-0.0048	0.0004

注：データは賃金構造基本統計調査の1990-2015年。

この式から得られる連鎖方式ではないラスパイレス指数 (\tilde{w}_t) は

$$\tilde{w}_t = \hat{\beta}_{ot} + \bar{x}_{10}\hat{\beta}_{1t} + \dots + \bar{x}_{K0}\hat{\beta}_{Kt}$$

となり、したがって平均賃金 (\bar{w}_t) と賃金指数 (\tilde{w}_t) との乖離は、

$$\frac{\bar{w}_t - \tilde{w}_t}{\bar{w}_t} = \sum_{k=1}^K \underbrace{\frac{(\bar{x}_{kt} - \bar{x}_{k0})\hat{\beta}_{kt}}{\bar{w}_t}}_{\text{属性}x_k\text{の構成変化による寄与}}$$

とすることができる。

表 2.6.1 は、労働者の年齢、性別、学歴、都道府県の構成変化による賃金変動の寄与度を表している。例えば、1990年から1995年の間に、属性を制御しない賃金指数（平均賃金）は属性を制御した賃金指数（ここではラスパイレス指数）と比べ1.23%高くなっている。その要因を分解すると、年齢構成の変化と教育水準の変化がそれぞれ0.73%、1.26%平均賃金を押し上げていることがわかる。他方、女性の就業割合の増加や就業地の変化がそれぞれ0.46%、0.31%平均賃金を下げる方向に働いていることがわかる。

分析結果からは、どの時期でも、年齢構成の変化が平均賃金を上昇させていることがわかる。これは労働人口の高齢化によるものだと考えられる。また、女性の就業率の上昇によって、平均賃金は減少している。特に2000年以降は、女性の就業率の上昇による影響が労働人口の高齢化による影響を相殺している。学歴の構成変化も平均賃金の上昇に大きく寄与している。これは、若いコホートの方が学歴は高い傾向にあるため、若いコホートの労働市場への参入によって学歴の高い労働者が増えたことによるものだと考えられる。

(3) 賃金変動の要因分解

2.6.2(2)では、賃金変動のうち、労働者属性を一定にした時の賃金変動に着目しているが、賃金変動を労働者属性の分布の変化と労働者属性を所与としたときの賃金の変化に分解で

きる。賃金変動を労働者属性の分布の変化と t 期における属性 j の労働者のシェアと平均賃金をそれぞれ s_{jt} と w_{jt} とおき、労働者全体の平均賃金を w_t とおく。変数 x_t の前期との差分を $\Delta x_t (=x_t - x_{t-1})$ で表せば、平均賃金の変化は以下のように分解される。

$$\frac{\Delta \bar{w}_t}{\bar{w}_{t-1}} = \underbrace{\sum_{j \in J} \Delta s_{jt} \frac{w_{jt-1}}{\bar{w}_{t-1}}}_{\text{労働者属性の分布の変化に起因する賃金変動}} + \underbrace{\frac{\sum_{j \in J} s_{jt-1} \Delta w_{jt}}{\sum_{j \in J} s_{jt-1} w_{jt-1}}}_{\text{労働者属性の分布を一定にしたときの賃金変動 (ラスパイレス指数)}} + \underbrace{\sum_{j \in J} \Delta s_{jt} \frac{\Delta w_{jt}}{\bar{w}_{t-1}}}_{\text{労働者属性の分布の変化と労働者属性ごとの賃金変化の相関}}$$

ここでは、右辺第一項は、労働者構成の変化による賃金変化率を表し、右辺第二項は労働者属性を所与としたうえでの実質的な賃金の変化率を示す（すなわち、 $t-1$ 期を基準時点としたラスパイレス指数（厳密にはラスパイレス指数から 1 を引いたもの）となる）。例えば、年齢の高い労働者は比較的賃金が高い傾向があるが、高齢化により年齢の高い労働者が増えた（労働者構成が変わる）ことによる賃金変化率は右辺第一項で捉えられることとなる。他方、同じ年齢（属性）の労働者における賃金上がることによる賃金増加率は右辺第二項で捉えられることになる。右辺第三項は労働者構成の変化と賃金変化の相関関係を表しており、シェアの大きい労働者属性の賃金が増加していれば正となり、反対にシェアの小さい労働者属性の賃金が増加していれば負となる。

分析においては、1990 年から 2015 年までの 5 年ごとに賃金変動を上記の式にしたがって分解した。分析に用いた労働者属性は、年齢カテゴリ（5 歳刻み、75 歳以上は 1 つのカテゴリとした）、性別、大学卒ダミー、一般労働者ダミーである。

分析結果を表 2.6.2 に示す。1990 年から 1995 年にかけては、賃金変動のほとんどが労働者属性を所与としたときの賃金の変化によって説明することができる。つまり、賃金変動のほとんどが実質的な賃金の上昇であるということがわかる。1995 年から 2000 年にかけて

表 2.6.2 賃金変化の要因分解

	1990– 1995	1995– 2000	2000– 2005	2005– 2010	2010– 2015
全体変動	0.157	0.005	-0.041	-0.026	0.018
労働者構成の変化による変動	0.005	0.010	-0.004	-0.008	0.004
労働者属性を所与としたときの賃金変化による変動	0.153	-0.004	-0.043	-0.019	0.015
相関	-0.001	-0.001	0.006	0.001	0.000
基準年の平均賃金 (円/時間)	1,869	2,161	2,173	2,082	2,027

注：データは賃金構造基本統計調査の 1990-2015 年。

は、賃金の高かった労働者属性のシェアが増加している一方で、労働者属性を所与としたときの賃金は減少している。その結果、全体的な賃金変動は小さくなっている。2000年から2005年の間は、賃金変動のほとんどが、労働者属性を所与としたときの賃金の変化によって説明され、実質的な賃金下がっていることが見て取れる。2005年から2010年については、全体的にはマイナスとなっており、その変化には労働者属性を所与としたときの賃金変化と労働者属性のシェアの変化のいずれもが寄与している。2010年から2015年は、全体的には賃金は上昇しており、労働者属性を所与としたときの賃金の変化が寄与する部分が多いが、労働者属性のシェアの変化の寄与も相対的に大きい。また、どの期間も労働者属性の分布の変化と労働者属性ごとの賃金の変化の相関はほとんどないことがわかる。なお、2005年以降に労働者属性の分布の変化の寄与度が大きくなっていることは、図2.6.1で2005年から、平均値とモデル2やモデル3の値が乖離し始めていることと整合的な結果である（労働者属性の分布の変化が全くなければ、平均賃金とラスパイレス指数は一致するはずである。）

さらに、男女別に同様の分析を行った結果を表2.6.3と表2.6.4に示す。男性については、労働者属性の構成変化による影響が賃金上昇に寄与している一方で、1995年以降は各グループ内では賃金は下降傾向にあるのが特徴的である。女性については、労働者属性の構成変化による影響が男性と比べて限定的である。これは女性の賃金の年齢プロファイルが男性よりもフラットであり、高齢化の影響が限定的であることが原因だと考えられる。また、グループ内の賃金は2000年から2005年の期間を除き、上昇傾向にあるのも特徴的である。

表 2.6.3 賃金変化の要因分解（男性）

	1990- 1995	1995- 2000	2000- 2005	2005- 2010	2010- 2015
全体変動	0.155	0.003	-0.031	-0.031	0.018
労働者構成の変化による変動	0.014	0.020	0.012	0.002	0.004
労働者属性を所与としたときの賃金変化による変動	0.140	-0.016	-0.049	-0.032	0.015
相関	0.000	-0.002	0.006	-0.001	0.000
基準年の平均賃金（円/時間）	2,234	2,580	2,587	2,507	2,027

注：データは賃金構造基本統計調査の1990-2015年。

表 2.6.4 賃金変化の要因分解（女性）

	1990- 1995	1995- 2000	2000- 2005	2005- 2010	2010- 2015
全体変動	0.193	0.021	-0.025	0.016	0.053
労働者構成の変化による変動	0.002	-0.010	-0.002	0.000	0.012
労働者属性を所与としたとき の賃金変化による変動	0.196	0.032	-0.027	0.016	0.042
相関	-0.005	0.000	0.004	0.000	-0.001
基準年の平均賃金（円/時間）	1,215	1,450	1,481	1,444	1,467

注：データは賃金構造基本統計調査の1990-2015年。

2.6.3. 調査結果からの示唆

上記分析からは、労働者の属性を制御した賃金指数と属性を制御しない賃金指数（平均賃金）との間に乖離が生じていることが明らかになった。そうした乖離は賃金の高い傾向にある労働者が増加したことを意味しているが、年齢構成の変化や教育水準の変化がそうした乖離の要因となっていることも分析から明らかになっている。また、賃金変化の要因を分解すると賃金変化全体において労働者構成の変化が占める割合も一定程度となっている。

こうした結果は、賃金変化を議論するにあたっては、労働者の属性を制御した賃金指数に基づくことの重要性を示唆している。とりわけ、少子高齢化や女性の就業率の上昇といった影響により労働市場の構成が大きく変化している中においては、そうした労働者の構成変化による賃金への影響と同一の属性を持つ労働者の実質的な賃金変化による影響を切り分けて考えなければ、経済動向に関する誤った認識に至ってしまう可能性もあるため、属性を制御した賃金指数の算出は極めて重要となると考えられる。

第3章 調査研究のまとめ

本調査研究における分析結果を、「統計調査間の比較（ユーザー利便性の向上）」、「統計精度の向上（非標本誤差の縮小）」、「賃金関連統計の体系（経済実態の把握）」という3つの観点からまとめる。

【統計調査間の比較（ユーザー利便性の向上）】

「2.1 調査対象の比較」では、複数の統計調査間の比較を通じて、統計間の水準の違いについて検証を行った。賃金構造基本統計調査と民間給与実態統計調査を用いた分析の結果、個票データを用いて統計調査間での対象の属性を揃えるための調整を適切に行うことにより、両統計は比較可能な水準の値となることが明らかになった。他方で、賃金構造基本統計調査と職種別民間給与実態調査の比較では、可能な範囲で両統計の調整を行ったものの、後方で調整できていない属性情報が残ったため、統計量にも乖離が残ることとなった。これらの結果からは、それぞれの調査の対象や調査項目の定義、集計方法が統計間の差異の大きな原因となっていることが伺われる。

また、賃金構造基本統計調査と民間給与実態統計調査の非正規労働者の分布の違いについての分析においても、両統計間で非正規労働者の分布が異なっていることが示されたが、これについても「非正規労働者」の定義が異なることが原因の1つと考えられる。

こうした点からは、3つの示唆を得ることができる。1つは、統計間の異同を議論する際に単純な集計結果を用いることへの留意である。上述のとおり、各統計では対象や調査項目、集計方法が異なるため、単純に公表されている集計値を用いて比較をすることは誤った議論につながりかねないため、留意が必要である。統計間の比較を行う際には、調査対象や集計方法の差異を適切に考慮することが重要である。

2つ目は、調査方法や集計方法に関する詳細な情報を公開することの重要性である。上述のように、調査間の比較に当たっては、調査間の違いを調整する必要があるが、調査対象や調査項目、集計方法などが明らかになっていないと、調整を行うこと自体がそもそも不可能となる。

3つ目に、それぞれの統計調査はそれぞれの目的に基づき最適に設計することが優先されるべきではあるが、可能な限り統計間で共通の調査方法や集計方法を取ることが重要であると考えられる。例えば、非正規労働者の分類などにおいては、共通の分類を用いることで、調査間の比較可能性を担保することが可能となる。

【統計精度の向上（非標本誤差の縮小）】

統計調査の設計に当たっては、無作為抽出によるサンプルから100%の回答率で回答が得られることを前提として標本誤差などが計算されている。したがって、様々な非標本誤差を原因としてこうした前提からの乖離が生じてしまうと、統計の精度低下やバイアスを引き

起こすことにつながってしまう。本調査分析では、こうした非標本誤差について、事業所内での抽出の偏りの有無、および調査への回答状況の観点から分析を行った。

事業所内での抽出（「**2.2 調査方法の影響（サンプルの偏りの有無）**」）については、民間給与実態統計調査では偏りが確認されなかったものの、賃金構造基本統計調査では労働時間や給与額について若干の偏りが見られた。こうした偏りを是正するためには、事業所内での適切な無作為抽出を徹底する以外に方策はないため、各事業所において抽出を担当する者が適切な抽出を実施できるようなツールを提供するといった方策が考えられる。

調査の回答状況の分析（「**2.3 調査対象の回答状況**」）においては、賃金構造基本統計調査と民間給与実態統計調査ともに、小規模事業所や小売業やサービス業、宿泊・飲食サービス業などの産業で回答率が低いという共通の結果が得られた。また、複数回の調査対象になっている事業所の回答パターンの分析からは、前回・前々回の回答状況が次の回答有無に大きく影響を及ぼすことが明らかになった。また関連して、回答者の負担軽減・回答率の向上に寄与すると期待される行政記録情報活用の一例として、源泉徴収票等情報転記ツールに関する分析（「**2.4 行政記録情報の活用可能性の検討**」）も行った。転記ツールの利用状況はアンケート回答企業の30%程度にとどまっており、さらなる活用推進が必要と考えられるものの、転記ツールの利用は回答の正確性向上に寄与している可能性が示唆された。回答率の改善のためには、転記ツールをはじめとする行政記録情報の活用を検討するとともに、上記傾向が生じる要因について定性・定量の両面から分析を行い、回答を促進するための取組を検討することが重要である。その際には、ナッジなどの行動経済学の知見を踏まえた取組の検討が有効であると考えられるとともに、そうした取組の効果を適切な方法で検証し、そのエビデンスに基づく意思決定を行っていくことが重要であると考えられる。

【賃金関連統計の体系（経済実態の把握）】

賃金統計調査は、調査を通じて賃金動向・経済動向を明らかにし、それに基づいた適切な政策立案を行っていくことを目的の1つとしている。統計調査を通じて経済動向を把握していくうえでの留意点について、本調査研究では母集団のカバレッジ、および労働者の属性を制御した（品質固定した）労働指数の試算という2つの観点から分析を行った。

母集団のカバレッジに関する分析（「**2.5 母集団情報の比較検証**」）においては賃金構造基本統計調査ではカバーされていない従業員1-4人事業所の労働者、およびいずれの賃金統計にも含まれていないいわゆるフリーランス（本分析では就業構造基本調査における「雇人のいない自営業者」）の割合や所得分布を検証した。その結果、調査の対象に含まれていないいずれの層においても賃金が低い傾向が見られた。したがって、小規模事業所の労働者やフリーランスを含む全労働者の賃金水準は、賃金構造基本統計調査や民間給与実態統計調査が示す賃金水準よりも低い水準であると考えられる。就業構造が大きく変化している中で、特に賃金の低い層がシステマティックに調査対象から抜け落ちてしまうと、労働者全体の賃金水準を正確に把握することができなくなってしまうため、将来的にはこうした層を

含めた調査を新規調査も含め検討する必要があると考えられる。

「2.6 品質固定された賃金指数の試算」では労働者の属性を制御した（品質固定した）賃金指数の算出を行い、実態としての賃金変動を検証したうえで、賃金変動の要因についての分析を行った。分析からは、労働者の属性を制御した賃金指数と属性を制御しない賃金指数との間に乖離が生じていることが明らかになった。そうした結果は賃金の高い傾向にある労働者が増加したことを意味しているが、年齢構成の変化や教育水準の変化が乖離の要因となっていることも分析から明らかになった。また、賃金変化の要因を分解した結果、労働者構成の変化が賃金変化に一定程度の影響を及ぼしていることが明らかになった。

これらの結果からは、賃金変化の議論において労働者の属性を制御した賃金指数を用いることの重要性が示唆される。特に、少子高齢化や女性の就業率の上昇といった要因が労働市場の構成に大きく影響を及ぼしている中においては、労働者の構成変化による賃金への影響と実質的な賃金の変化を切り分けて考えなければ、経済動向に関する誤った認識に至ってしまう可能性もある。労働者の属性を制御した（品質固定した）賃金指数のこうした重要性を踏まえると、定期的に賃金指数を算出し、それに基づいた賃金変動の議論を行っていくことが望ましいと考えられる。

引用文献

川口大司 (2013) 「賃金」、『日本労働研究雑誌』、633 : 14-17.

Atkinson, A.B., Casarico, A. & Voitchovsky, S. (2018) Top incomes and the gender divide. *Journal of Economic Inequality*, 16, 225–256.

賃金関連統計の比較検証に関する調査研究 報告書

令和3年(2021年)3月

委託者 総務省統計委員会担当室 〒162-8668 東京都新宿区若松町19番1号

受託機関 株式会社メトリクスワークコンサルタンツ

〒160-0004 東京都新宿区四谷2-11-8 オフィスコート四谷 4F