

Discussion Paper Series

---

University of Tokyo  
Institute of Social Science  
Panel Survey

---

東京大学社会科学研究所 パネル調査プロジェクト  
ディスカッションペーパーシリーズ

東大社研パネル調査リフレッシュサンプルの基礎分析

Sample Bias of the Japanese Life Course Panel Survey (JLPS) 2019

仲修平 (東京大学社会科学研究所)

三輪哲 (東京大学社会科学研究所)

Shuhei Naka

Satoshi Miwa

March 2020

No.118

## 東大社研パネル調査リフレッシュサンプルの基礎分析

仲修平・三輪哲

(東京大学社会科学研究所)

### 要約

東京大学社会科学研究所では、若年者を対象とする総合的な追跡調査である JLPS (Japanese Life Course Panel Surveys) を 2007 年から実施してきた。JLPS の Wave13 では、「若年調査」の対象者は 32-46 歳となり、20 歳代の若年者が対象外となる。そこで新たにリフレッシュサンプルとして継続「若年調査」より若い層である 20-31 歳 (2019 年時点) を対象として調査を実施した。本稿の目的は、その調査によって得られたデータの代表性を検討することである。主に Wave1 (2007 年調査) の分布と比較したところ、以下の点が明らかとなった。第 1 に、回収率は Wave1 と比べると男女ともに高いことである。第 2 に、層化に用いた性別・年齢・地域ブロックの分布は基準とした値よりも一定程度の偏りが見られた。具体的には、20-25 歳の男性がやや過小に、26-31 歳の女性はやや過大にサンプリングされていた。地域ブロックについては、関東と近畿のサンプル数が基準値と比べると少ないことが示された。市郡規模については妥当な分布となっていた。第 3 に、層化以外の変数では、高学歴層と就業者の比率が高い傾向となっていることがわかった。以上の検討結果から、リフレッシュ調査データは Wave1 の継続調査と比べると、基本属性については同等かより歪みが小さい一方で、層化に用いた以外の変数は相対的に偏りが大きいと結論づける。そのため、JLPS データが代表性を十分に保証しうる貴重な研究資料でありつづけるためには、データの偏りへの対処を視野に入れる必要があることを主張する。

### 謝辞

本研究は、日本学術振興会 (JSPS) 科学研究費補助金・特別推進研究 (25000001, 18H05204)、基盤研究 (S) (18103003, 22223005) の助成を受けたものである。東京大学社会科学研究所 (東大社研) パネル調査の実施にあたっては、社会科学研究所研究資金、株式会社アウトソーシングからの奨学寄付金を受けた。パネル調査データの使用にあたっては東大社研パネル運営委員会の許可を受けた。

## 1. 本稿の目的

本稿の目的は、2019年に20歳から31歳を対象として実施された「働き方とライフスタイルの変化に関するパネル調査」(Japanese Life Course Panel Surveys-JLPS)で得られたデータの代表性を検討することである。東京大学社会科学研究所では、2007年より若年・壮年者に焦点をあてた総合的な追跡調査を継続してきたが(e.g. 石田編 2017; 佐藤・石田編 2019; 石田ほか編 2020), JLPSのWave13では「若年調査」の対象者が32-46歳となり、20歳代が調査対象外となっている<sup>1</sup>。そこで新たにリフレッシュサンプルとして、継続の「若年調査」より若い層である20-31歳(2019年時点)を対象として調査を実施した。

JLPSのリフレッシュサンプル調査(以下、リフレッシュ調査)は2007年の調査と同様に、地域・都市規模・性別・年齢により層化したうえで全国から対象者を抽出し、原則として郵送配布・訪問回収の方法で調査を実施した<sup>2</sup>。JLPS-2007は、「2007年初春における日本社会の若年層の正確な見取り図を提供できる」調査であったが(三輪 2008: 16), JLPS-2019の分布はJLPS-2007の分布に比べてどうなっているのだろうか。リフレッシュ調査によるサンプル(以下、若年リフレッシュサンプル)が、2007年「若年調査」によるサンプル(以下、若年継続サンプル)と比較してどのような特徴を持つのかを、個別の研究に先立って明らかにしておく必要がある。

そこで本稿では、JLPSのリフレッシュ調査データの代表性を次の手順で検討する。第2節では、リフレッシュ調査の概要、標本抽出と実査などの調査方法を説明する。第3節では、回収率および調査不能に関する情報を記述する。第4節では、計画標本と回収標本の基本属性分布を主にJLPS-2007の分布と比較しながら記述する。第5節では、就業状況や学歴等の主要変数について、回収標本と公的統計との乖離を検証する。第6節ではそれらを踏まえて、リフレッシュ調査データのもつ偏りを評価する。

## 2. 調査方法

### (1) 継続調査とリフレッシュ調査

JLPSのWave13は、2007年から継続している調査と対象外となった20歳から31歳を対象とした調査を2019年に実施した。リフレッシュ調査は、継続調査の調査設計や質問項目

---

<sup>1</sup> 東大社研パネル調査プロジェクトは、若年・壮年調査のほかに高卒パネル調査を実施している。高卒パネル調査については、佐藤編(2017)を参照されたい。

<sup>2</sup> JLPS-2007の調査設計や基礎的な分布については、三輪(2008)を参照されたい。また、JLPS-2019の調査結果に関する速報は、石田ほか(2020)でも確認することができる。

において互換性を保証するデザインとなっている。それによって、リフレッシュ調査は継続してきた若年調査データに対してコーホート間の比較が可能となるデータである。見方を変えれば、新たに20歳代が加わったことにより、JLPSは20歳代から40歳代の働き方とライフスタイルの変化を追跡するという総合的な調査の役割を維持している。

表1 リフレッシュサンプル調査の標本設計

		リフレッシュサンプル調査				
		割当 地点数	母集団人口	割当 標本規模	層の 重み	標本の 重み
全国計		264	14,814,974	6,600	1.000	1.000
北海道	21大市	4	242,586	108	.016	.016
	20万以上市	1	59,241	26	.004	.004
	その他の市	4	185,621	83	.013	.013
	町村	2	85,016	38	.006	.006
東北	21大市	3	145,359	65	.010	.010
	20万以上市	5	257,329	115	.017	.017
	その他の市	7	389,687	174	.026	.026
	町村	3	151,277	67	.010	.010
関東	21大市	7	370,204	166	.025	.025
	20万以上市	21	1,228,101	547	.083	.083
	その他の市	27	1,591,135	709	.107	.107
	町村	4	195,475	87	.013	.013
京浜	21大市	34	2,001,853	892	.135	.135
北陸	21大市	2	90,457	40	.006	.006
	20万以上市	3	160,390	71	.011	.011
	その他の市	5	276,113	123	.019	.019
	町村	1	41,723	19	.003	.003
東山	20万以上市	2	113,419	50	.008	.008
	その他の市	6	329,717	147	.022	.022
	町村	2	87,971	39	.006	.006
東海	21大市	8	462,808	206	.031	.031
	20万以上市	6	331,130	147	.022	.022
	その他の市	11	658,065	293	.044	.044
	町村	2	94,289	42	.006	.006
近畿	21大市	14	814,947	363	.055	.055
	20万以上市	12	717,693	320	.048	.048
	その他の市	14	790,682	352	.053	.053
	町村	2	118,211	53	.008	.008
中国	21大市	5	235,738	105	.016	.016
	20万以上市	4	186,428	83	.013	.013
	その他の市	6	336,063	150	.023	.023
	町村	1	53,097	24	.004	.004
四国	20万以上市	3	173,440	77	.012	.012
	その他の市	3	165,998	74	.011	.011
	町村	1	55,433	25	.004	.004
九州	21大市	8	416,787	185	.028	.028
	20万以上市	6	348,921	155	.024	.023
	その他の市	11	646,724	288	.044	.044
	町村	4	205,846	92	.014	.014

注：母集団人口は、2018年1月1日現在の住民基本台帳人口の数値である。

## (2) 標本抽出

標本抽出は、2007年調査と同様に層化二段無作為抽出によって行われた。具体的には地域・都市規模・性別・年齢により層化したうえで、対象者を全国から抽出し、原則として郵送配布・訪問回収の方法で調査を実施した<sup>3</sup>。リフレッシュ調査における調査対象の母集団は、2018年1月1日現在で満20歳から31歳（1987年～1998年出生）の男女個人である<sup>4</sup>。

事前の層化に用いた地域は、北海道・東北・関東・京浜・北陸・東山・東海・近畿・中国・四国・九州の11に分けられた<sup>5</sup>。都市規模は、21大都市・人口20万人以上の市・その他の市・町村の4つである。これらの組み合わせから44の層ができあがるが、実際には21大都市が存在しない地域もあるため、39層となった<sup>6</sup>。

表1に示したように、それら264層に地点を割り当てた。さらにそれらの地点の中で、性別・年代（20-25歳、26-31歳）を組み合わせた4層に分け、抽出すべき個人の数を割り当てた。実際の標本抽出のためには、住民基本台帳をサンプリング台帳とした。選ばれた地点は全部で264地点であり、抽出対象となる個人は6,600人であった。

## (3) 実査

実査は、2019年2月から3月にかけて行われた。調査票は郵送により配布され、回収は調査員が訪問することによってなされた。ただし、転送不要の調査票が戻った146人に限り、訪問回収から郵送回収に切り替えて調査票を投函した。調査票への回答の記入は対象者自身が行う「自記式」の留置調査である。なお、実査に際しては、東京大学社会科学研究所が一般社団法人中央調査社に委託して行われた。

回収票は2,380人に対して欠票は4,206人であった<sup>7</sup>。回答票のうち次年度Wave2の調査に協力することに同意したのは2,049ケース（リフレッシュサンプル）である<sup>8</sup>。本稿ではこれらのケースを分析対象として、2007年若年調査のうち20歳から31歳を取り出した2,081ケース（継続サンプル）の分布を比較する。

回収に際して、リフレッシュ調査は継続調査とは異なり一次調査と二次調査の区別をしていない。この点は回収後の分布に影響する重要な点であるために、以下に説明を加える。継続調査は、正規対象者だけではなく第1の予備対象と第2の予備対象までは、すべての

<sup>3</sup> JLPS2007の調査設計については、三輪（2008）を参照されたい。

<sup>4</sup> 本調査の対象者にはJLPS2007と同様に外国籍の人は含まれない。

<sup>5</sup> 北陸は新潟県・富山県・石川県・福井県から構成されるのに対して、東山は山梨県・長野県・岐阜県を統合したカテゴリである。

<sup>6</sup> JLPS-2007で層化に用いた地域は「16大都市」であったが、2007年以降に新潟市、浜松市、岡山市、相模原市、熊本市が政令指定都市に加わるにより「21大都市」となっている。

<sup>7</sup> 本稿の分析対象のデータは、「2019年10月31日配布バージョン」である。なお、6,600人のうち14人は、調査票の事前にあいさつ状を送付した段階で調査を拒否したために調査を終了した。

<sup>8</sup> 調査の継続に関する同意は、継続調査では「口頭」により同意を得ているのに対して、リフレッシュ調査では「書面」で同意を得ている。

対象者に対して調査票の回収を試みている（三輪 2008: 3-4）。ここまでを一次調査と呼ぶのに対して、二次調査は一次調査で完了できない層にのみ絞って行われる調査である。そのため、二次調査は「回収され難かった層」に対して集中的に行われるものであるため、一次調査と二次調査は性質が異なるのである。継続調査では、「回収された後のレベルでも妥当な属性分布が得られるようにと、回収手続きに工夫が施された」（三輪 2008: 4）。そのため、次節以降でリフレッシュ調査データの分布と比較する際は、継続調査のうち一次調査の回収票を取り出して検討する。

### 3. 回収率

継続調査とリフレッシュ調査の粗回収率は（計画標本のうち回収された票の割合）、それぞれ 30.6%と 36.1%である。リフレッシュ調査（JLPS-2019）は、前回調査（JLPS-2007）に比べると 6 ポイント近く上昇していることがわかる。ただし、粗回収率には転居や長期不在などが含まれているために、その値は検討の余地が残される。そこで、三輪（2008）と同様の方法を用いて「純」回収率を算出しておきたい。粗回収率の分母は、「総アタック数」であるのに対して、純回収率ではそれが「有効アタック数」となっている点に特徴がある。

$$\text{回収率} = \frac{\text{有効回収票}}{\text{有効アタック数}} = \frac{\text{有効回収票}}{\text{総アタック数} - \text{転居} - \text{長期不在} - \text{住所不明} - \text{その他}}$$

表 2 JLPS の回収率

	有効回収票数(A)	総アタック数(B)	有効アタック数(C)	回収率(A/C)
<b>JLPS-2007</b>				
合計	2,081	6,802	5,659	36.8
男性	924	3,471	2,780	33.2
女性	1,157	3,331	2,879	40.2
<b>JLPS-2019</b>				
合計	2,380	6,600	5,763	41.3
男性	1,114	3,373	2,903	38.4
女性	1,266	3,227	2,860	44.3

注 1：有効アタック数とは、総アタック数から「転居」・「長期不在」・「住所不明」・「その他」（死亡、病気、障害など）を除いた数である。

注 2：JLPS-2007 は一次調査の値である。

注 3：JLPS-2007 と JLPS-2019 の粗回収率（合計）はそれぞれ 30.6%と 36.1%である。

つまり、「有効アタック数」とは、転居・長期不在・住所不明・その他（入院や死亡等）までを外したものである。この式によって、最終的な純回収率を計算した結果を表 2 に示し

た。リフレッシュ調査と継続調査の純回収率は、それぞれ 41.3%と 36.8%である。性別ごとに見ると、リフレッシュ調査の男性と女性の純回収率は、それぞれ 38.4%と 44.3%である。それに対して継続調査の純回収率は、それぞれ 33.2%と 40.2%である。リフレッシュ調査の回収率は前回調査に比べると、男女ともに高いと判断できる。

続いて、回収率の背景にある調査不能の理由を確認する。どのような理由により票を回収できていないのかを知ることによって、リフレッシュ調査と継続調査の回収率に生じた差異に迫ることができる（表 3）。

表 3 調査不能理由の分布

	JLPS-2007	JLPS-2019
【有効回収】	30.6	36.1
【調査不能】		
転居	10.5	8.5
長期不在	1.4	2.2
一時不在	22.8	12.9
住所不明（宛先不明含む）	3.4	1.0
拒否（受け取り拒否含む）	27.4	21.7
それ以外合計	4.1	17.6
それ以外の内訳		
その他	0.6	1.0
受取拒否	0.1	0.0
記入漏れ	0.1	0.0
白票	0.1	0.0
郵送返送依頼	2.3	16.2
あて名不明	0.9	0.4
実数	6,802	6,600

表 3 を見ると、リフレッシュ調査は継続調査に比べると、一時不在と拒否の比率が 10 ポイントと 6 ポイント近く下がっている。逆にいえば、リフレッシュ調査はこれらの層から票を回収できたことによって回収率が上昇したと考えられる。もちろん、調査不能の記述のみだけで回収率の上昇を説明することは難しく、安易な判断をすることはできない。例えば、調査における回収不能率を「接触率」と「接触できた条件下での拒否率」に分解することが考えられる（保田 2008）。

そのアイデアに基づいて継続調査を検討した結果、回収率は地域特性（市郡規模と地域ブロック）・個人属性（性別と年齢）・住居（一軒家と集合住宅等）によって異なることがわかっている（三輪 2008）。未回収あるいは調査不能となる人々はリフレッシュ調査と継続調査でどこが異なるのか／異なるらないのか、に関する分析は別途検討する予定である。

#### 4. 計画標本と回収標本の基本属性

本節では、先述した標本抽出の方法にもとづいて得られた計画標本と回収標本のうち基本属性（年齢・性別・地域ブロック・市郡規模）の分布を検討する（表 4 から表 6）。計画標本とは、正規対象として抽出された対象者の集合を意味する。比較のための参照基準として、2007 年に実施された継続調査のうち「一次調査」における属性の分布を用いる。継続調査は上述したようにすべての対象者にアタックする「一次調査」に加えて、一次調査で完了できない層のみに絞って行われた「二次調査」からなる。それによって、回収された後のレベルでも妥当な属性分布が得られている（三輪 2008）。

それに対して、リフレッシュ調査では一次調査と二次調査の区別をしていないために、継続調査に比べると分布の偏りがより大きくなっていることが予想される。そこでより厳密に分布を比較するために、継続調査の一次調査の回収票を用いた。したがって、現時点で公開されているデータの分布とは若干異なる点には留意が必要である<sup>9</sup>。また、2015 年 10 月に実施された国勢調査における属性の分布も一部に用いるが、本節でもリフレッシュ調査（JLPS-2019）と継続調査（JLPS-2007）の比較に主眼を置く。

表 4 から表 6 の「計画標本」の列を見ると、層化に用いた性別、年齢、地域ブロックおよび市郡規模の 4 変数は、標本抽出の際に大きな偏りが生じていないことがわかる。ところが、「回収標本」については市郡規模の分布を除いて一定程度の偏りが見られる。

表 4 を見ると、男性の標準化残差がマイナス（計画標本に対して少ない傾向）となっている。とくに、20-25 歳の男性サンプルサイズが相対的に小さくなっていることがわかる。それに対して女性の標準化残差はプラス（計画標本に対して多い傾向）となっており、26-31 歳のサンプルサイズが大きいことが見て取れる。それに対して継続調査では、26-31 歳の男性サンプルサイズが小さく、20-25 歳の女性サンプルサイズが大きい。このように内実の相違は見られるが、JLSP は男性サンプル数が少ないのに対して女性サンプル数が多いという類似の傾向を持つことが示された。

一方、地域ブロックを見ると（表 5）、東北と中部のサンプルは回収されやすいのに対して、関東と近畿のサンプルは回収されにくい傾向である。そうした傾向はリフレッシュ調査と同様に継続調査においても見られた。それ以外の地域では、中国と九州のサンプルはやや多いものの、基準となる計画標本と大差は見られない。

市郡規模はカイ 2 乗統計量が統計的に有意ではないことからわかるように、より適切な分布に近い値となっている（表 6）。全般的に分布を見ると、リフレッシュ調査の標準化

<sup>9</sup> SSJDA から公開されているデータは、一次調査と二次調査を統合した「最終回収標本」である。両調査の分布については三輪（2008）を参照されたい。



残差は継続調査に比べると同等ないし小さく、より妥当な分布が得られていると判断できる<sup>10</sup>。ただし詳細に見ると、リフレッシュ調査では 21 大市のサンプル数が相対的に少ないのに対して、20 万未満市のサンプル数が多くなっている。そうした傾向は継続調査においても類似していることがわかる。

表 4 計画標本と回収標本における性別・年齢の分布

	JLPS-2019 計画標本			JLPS-2019 回収標本			JLPS-2007 回収標本			国勢調査 (2015年)
	度数	構成割合	標準化残差	度数	構成割合	標準化残差	度数	構成割合	標準化残差	構成割合
男性										
20-25歳	1,629	24.7	1.172	426	20.8	-2.944	434	20.9	-0.204	24.0
26-31歳	1,749	26.5	-0.546	509	24.8	-1.753	490	23.5	-5.283	26.8
合計	3,378	51.2		935	45.6		924			50.8
女性										
20-25歳	1,555	23.6	0.768	509	24.8	1.634	542	26.0	6.050	23.1
26-31歳	1,667	25.3	-1.292	605	29.5	3.065	615	29.6	0.500	26.1
合計	3,222	48.8		1,114			1,157			49.2
総計	6,600	100.0		2,049	100.0		2,081	100.0		100.0
$\chi^2$ (df=3)	5.60	n.s		23.80	***		64.81	***		

注 1： $\chi^2$ の理論分布としては、国勢調査における構成割合を採用した。

注 2：JLPS-2007 は一次調査の値である。

表 5 計画標本と回収標本における地域ブロックの分布

	JLPS-2019 計画標本			JLPS-2019 回収標本			JLPS-2007 回収標本			国勢調査 (2015年)
	度数	構成割合	標準化残差	度数	構成割合	標準化残差	度数	構成割合	標準化残差	構成割合
地域ブロック										
北海道	255	3.9	-0.562	88	4.3	0.662	95	4.6	0.814	4.0
東北	421	6.4	-0.097	158	7.7	2.328	155	7.4	1.214	6.4
関東	2,401	36.4	-0.246	662	32.3	-3.184	652	31.3	-3.159	36.6
中部	1,177	17.8	0.278	417	20.4	2.865	455	21.9	4.376	17.7
近畿	1,088	16.5	0.215	300	14.6	-1.942	308	14.8	-2.034	16.4
中国	362	5.5	0.069	130	6.3	1.703	140	6.7	2.441	5.5
四国	176	2.7	0.554	49	2.4	-0.470	70	3.4	1.610	2.6
九州	720	10.9	-0.068	245	12.0	1.397	206	9.9	-1.391	10.9
総計	6,600	100.0		2,049	100.0		2,081	100.0		100.0
$\chi^2$ (df=7)	0.83	n.s		34.05	***		45.88	***		

注 1： $\chi^2$ の理論分布としては、国勢調査における構成割合を採用した。

注 2：JLPS-2007 は一次調査の値である。

<sup>10</sup> 継続調査については、二次調査をした後の方が分布の偏りが大きくなっていった（三輪 2008: 12）。その理由として、二次調査が 16 大都市を中心に行われることで、16 大都市居住の人々が過大に代表されていた。この点に鑑みると、リフレッシュ調査は一次/二次調査の区別をしていないものの、より適切に回収できたといえる。

表 6 計画標本と回収標本における市郡規模の分布

	JLPS-2019 計画標本			JLPS-2019 回収標本			JLPS-2007 回収標本		
	度数	構成割合	標準化残差	度数	構成割合	標準化残差	度数	構成割合	標準化残差
市郡規模									
21大市	2,130	32.3	0.004	623	30.4	-1.486	505	24.3	-2.651
20万以上市	1,591	24.1	-0.053	476	23.2	-0.836	590	28.4	0.889
20万未満市	2,393	36.3	0.016	797	38.9	1.993	764	36.7	0.351
町村	486	7.4	0.052	153	7.5	0.202	222	10.7	2.350
総計	6,600	100.0		2,049	100.0		2,081	100.0	
$\chi^2$ (df=3)	0.01	n.s		6.92	n.s		13.46	*	

注1： $\chi^2$ の理論分布としては、住民基本台帳における構成割合を採用した。

注2：JLPS-2007は一次調査の値である。なお、JLPS-2007の「21大市」はサンプリング時の層化に合わせて「16大市」の値となっている。

## 5. 層化要因以外の代表性

本節では、層化要因以外の変数について継続調査に加えて労働力調査を基準にしてリフレッシュ調査の回収標本特性を検討する。労働力調査は、国勢調査とは異なり標本調査である。ただし、個人数では11万人ほどの大規模データであるため、それを比較対象として検討を進めることには一定の意味がある。なお、ここでは2019年1-3月の四半期平均の値を基準とし、表記する場合はリフレッシュ調査と同じ年齢区分にすることが難しいため、「20-29歳」と「20-34歳」を併記することにした。

表7で取り上げる変数は、すべて層化に用いていないものである。そのため、表4から表6でみた分布よりも大きな偏りが観察される。

働き方に関する変数を見ると、就業者の割合が労働力調査と継続調査に比べて高い値となっていることがわかる。その割合は同時期の労働力調査よりも約10ポイント上回る値となっている。従業上の地位はほぼ妥当な分布が得られているが、家族従業者についてはやや多いことがわかる。ただし、偏りの程度は継続調査に比べると小さくなっている。

産業と職業は、自由回答をもとにコーディングした結果を大分類にまとめている。産業の分布を見ると、第2次産業の値が標準化残差の絶対値が3を超えていることからわかるように、リフレッシュ調査では過小になっている。それに対して、第3次産業の分布はやや過大になっていることがわかる。むしろ、継続調査においても産業分布の偏りは見られるものの、その偏りはリフレッシュ調査の方が大きいことがわかる。

他方、職業の分布を見ると、標準化残差の絶対値が相対的に大きいカテゴリーが散見される。例えば、専門職と事務職が過大になっているのに対して、生産工程・労務職と運輸通信・

保安職が過小となっていることがわかる。これらの偏りは継続調査においても見られるが、リフレッシュ調査の絶対値は継続調査と同等かやや大きいことがわかる。

働き方以外の変数として、婚姻状態と学歴の回答分布を確認しておきたい。婚姻状態の標準化残差はすべて3以内におさまっており、より適切な分布となっていることがわかる。ただし、それらの分布は継続調査の標準化残差と比べると大きな値となっている<sup>11</sup>。それに対して、学歴の分布は労働力調査に比べると、高学歴へ顕著に偏っている。回答の選択肢が異なるために「大学・大学院」の卒業者割合のみを取り上げたが、継続調査とリフレッシュ調査の標準化残差はそれぞれ4.696と5.740となっている。つまり、JLPSは総じて高学歴層による回答が多いといえる。

表7 回収標本における層化要因以外の分布（男女合計）

	JLPS-2007 回収標本			JLPS-2019 回収標本			労働力調査 (2019年)	労働力調査 (2019年)
	度数	構成割合	標準化残差	度数	構成割合	標準化残差	20-29歳	20-34歳
就業者	1,706	82.0	2.143	1,802	87.9	4.074	77.8	79.9
従業上の地位								
自営業主	38	2.5	0.254	32	2.1	-0.119	1.4	2.1
家族従業者	42	2.8	4.406	17	1.1	1.850	0.6	0.7
雇用者	1,411	94.6	-0.581	1,474	96.8	-0.141	97.9	97.1
産業								
第1次産業	13	0.8	-0.952	15	0.9	-0.946	1.2	1.1
第2次産業	403	24.8	-0.237	318	18.3	-3.393	20.6	22.1
第3次産業	1,208	74.4	0.251	1,407	80.9	1.933	78.2	76.8
職業								
専門	359	21.3	4.295	501	28.0	6.770	20.2	20.7
管理	1	0.1	-1.775	3	0.2	-0.840	0.1	0.3
事務	425	25.3	2.752	417	23.3	3.822	18.2	19.3
販売	225	13.4	-2.444	255	14.3	-1.239	16.2	15.4
サービス	219	13.0	2.068	264	14.8	0.914	15.5	13.9
農業	19	1.1	0.165	12	0.7	-2.480	1.2	1.4
運輸通信・保安	51	3.0	-2.737	54	3.0	-2.646	4.1	4.3
生産工程・労務	384	22.8	-4.035	283	15.8	-7.523	24.5	24.6
婚姻状態								
未婚	1,282	77.0	-0.720	1,294	79.9	-1.488	83.3	74.3
有配偶	364	21.8	1.163	299	18.5	2.701	15.8	24.3
死別・離別	20	1.2	1.153	26	1.6	2.994	0.9	1.4
大学・大学院卒業者	417	36.1	4.696	574	52.3	5.740	-	41.2

注1：2019年の労働力調査は、20-29歳と20-34歳までの年齢階級における構成割合の値を用いた。標準化残差の算出に際しては、20-34歳の値を用いた。

注2：婚姻状態と大学・大学院の度数はそれぞれ20-29歳と25-31歳の値を用いた。

注3：産業と職業の分布は2015年国勢調査の値を用いた。

注4：JLPS-2007は一次調査の値である。

<sup>11</sup> JLPS-2007の最終回収標本（二次調査を含めた20-40歳のデータセット）では、未婚のサンプル数が少ないのに対して、有配偶のサンプル数が多い値となっている（三輪 2008）。

## 6. まとめ

本稿の目的は、2019年に20歳から31歳を対象として実施された「働き方とライフスタイルの変化に関するパネル調査」(Japanese Life Course Panel Surveys-JLPS)で得られたデータの代表性を検討することであった。本調査は、同様の調査設計で2007年から実施されている継続調査において調査対象外となる若年者を新たにリフレッシュサンプルとして回収するものである。リフレッシュ調査の分布を検討する際は、2007年の継続調査と公的統計を主な参照軸とした。検討により得られたデータの性質に関する知見をまとめておきたい。

リフレッシュ調査の回収率は継続調査の回収率よりも5ポイント近く上昇していた。しかしながら、層化する際に用いた基本属性の分布は基準とした値よりも一定程度の偏りが見られた。例えば、20-25歳の男性がやや過小にサンプリングされているのに対して、26-31歳の女性はやや過大に回収されている。ただし、男性サンプル数が少なく、女性サンプル数が多いという傾向は継続調査と類似していることがわかった。地域ブロックについては、東北と中部のサンプル数が相対的に多いのに対して、関東と近畿のサンプル数が基準値と比べると少ないことが示された。市郡規模については、21大市に居住するサンプル数はやや少ないのに対して、20万未満市に居住するサンプル数がわずかに多いものの、ほぼ基準値通りの値となっていた。

層化する際に用いていない学歴や仕事に関する諸変数の分布を見ると、基準とした労働力調査の値からはある程度乖離していることが明らかとなった。偏りの方向性は、高学歴の人が多めになっていること、およびより積極的に働いている人が多いことである。これらの特徴は継続調査においても観察されている(三輪 2008)。他方、仕事に関する変数の分布を確認すると、従業上の地位や産業についてはほぼ基準値に近い値となっていた。ところが、職業の比率を見ると、専門職・事務職の比率が相対的に高いのに対して、生産工程・労務職や運輸通信・保安職の比率が相対的に低いことがわかった。

以上の検討結果を踏まえると、リフレッシュ調査データは2019年初春における若年層の諸相を厳密に再現しているとはまでは言えないと結論づける。だからといって、悲観的に捉えることはないと考えている。今回実施したリフレッシュ調査は、2007年から継続してきたパネル調査の対象を拡げるという重要な役割を担っている。そのことは、若年・壮年者に焦点をあてた総合的な追跡調査の厚みがより増したといえる。ただし、データの利用に際しては、分布の偏りに対して十分な目配りが必要であることを強調しておきたい。

分布の偏りを適切に対処するためには、リフレッシュ調査において回収不能であった層がどのような特性を持っているのか、を理解することは不可欠であろう。さらにいえば、事後的なデータ調整の代表的な作業である「ウェイトの作成」は避けて通れないと考えている。

母集団と回収標本の間において基本属性（性別・年齢・居住地等）の分布の偏りを調整するウェイト値は、パネル調査の場合ではとくに脱落の問題と絡んで重要になる（保田 2016）。JLPS ではウェイト値の算出までは今のところ行っていない。しかしながら、JLPS データが代表性を十分に保証しうる貴重な研究資料であり続けるためには、データの偏りへの対処を含めて検討していく必要があるだろう。

## 参考文献・引用文献

- 石田浩編, 2017, 『教育とキャリア』 勁草書房.
- 石田浩・有田伸・藤原翔編, 2020, 『人生の歩みを追跡する——東大社研パネル調査でみる現代日本社会』 勁草書房.
- 石田浩・石田賢示・大久保将貴, 2020, 「働き方とライフスタイルの変化に関する全国調査 2019」分析結果報告——パネル調査から見る初職への移行、職業キャリア、介護問題」 (<https://csrda.iss.u-tokyo.ac.jp/panel/PR/20PressRelease.pdf>)
- 三輪哲, 2008, 「働き方とライフスタイルの変化に関する全国調査 2007 における標本特性と欠票についての基礎分析」『東京大学社会科学研究所パネル調査プロジェクトディスカッションペーパーシリーズ』 No.10.
- 佐藤香編, 2017, 『ライフデザインと希望』 勁草書房.
- 佐藤博樹・石田浩編, 2019, 『出会いと結婚』 勁草書房.
- 保田時男, 2008, 「低下する回収率と回収不能の要因」 谷岡一郎・仁田道夫・岩井紀子編『日本人の意識と行動——日本版総合的社会調査 JGSS による分析』 東京大学出版会: 447-458.
- 保田時男, 2016, 「パネル調査法」 筒井淳也・水落正明・保田時男編『パネルデータの調査と分析・入門』 ナカニシヤ出版: 11-19.

## 東京大学社会科学研究所パネル調査プロジェクトについて

労働市場の構造変動、急激な少子高齢化、グローバル化の進展などにもない、日本社会における就業、結婚、家族、教育、意識、ライフスタイルのあり方は大きく変化を遂げようとしている。これからの日本社会がどのような方向に進むのかを考える上で、現在生じている変化がどのような原因によるものなのか、あるいはどこが変化してどこが変化していないのかを明確にすることはきわめて重要である。

本プロジェクトは、こうした問題をパネル調査の手法を用いることによって、実証的に解明することを研究課題とするものである。このため社会科学研究所では、若年パネル調査、壮年パネル調査、高卒パネル調査、中学生親子パネル調査の4つのパネル調査を実施している。

本プロジェクトの推進にあたり、以下の資金提供を受けた。記して感謝したい。

文部科学省・独立行政法人日本学術振興会科学研究費補助金

基盤研究 S：2006 年度～2009 年度、2010 年度～2014 年度 基盤研究 C：2013 年度～2016 年度 特別推進研究：2015 年度～2017 年度 若手研究 A：2015 年度～2018 年度  
基盤研究 B：2016 年度～2020 年度 特別推進研究：2018 年度～2024 年度

厚生労働科学研究費補助金

政策科学推進研究：2004 年度～2006 年度

奨学寄付金

株式会社アウトソーシング（代表取締役社長・土井春彦、本社・静岡市）：2006 年度～2008 年度

## 東京大学社会科学研究所パネル調査プロジェクト ディスカッションペーパーシリーズについて

東京大学社会科学研究所パネル調査プロジェクトディスカッションペーパーシリーズは、東京大学社会科学研究所におけるパネル調査プロジェクト関連の研究成果を、速報性を重視し暫定的にまとめたものである。



東京大学社会科学研究所 パネル調査プロジェクト  
<https://csrda.iss.u-tokyo.ac.jp/panel/>