

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・高卒パネル調査(JLPS-H)における
初期脱落の影響とそれにともなうバイアスの補正

Exploring the effects of early dropout and the correction for the bias
in JLPS-H

山口泰史

(東京大学社会科学研究所)

Yasufumi YAMAGUCHI

March 2021

No.132

東大社研・高卒パネル調査(JLPS-H)における初期脱落の影響と それにもなうバイアスの補正

山口泰史（東京大学社会科学研究所）

要旨 「東大社研・高卒パネル調査（JLPS-H）」は、高校卒業時から34～35歳時（2021年1月時点）まで、16年間にわたって同一個人を追跡して回答を得ている、日本国内では他に類をみない質問紙調査である。だが、同調査はWave 1を実施する際に学校を通じて対象者に依頼し、それ以降の調査に応諾した対象にのみWave 2以降の回答依頼をおこなう調査設計である。それゆえWave 2時点での大規模脱落が生じ、それにもなうバイアスが懸念される。

本稿では同データについて、性別などの基本属性における脱落の偏りがどの程度生じているのかを確認した上で、そのような偏りを補正するために、IPW（Inverse Probability Weight）法に基づくウェイトの作成を試み、事例的分析を通じて、ウェイトの効果を確認した。その結果、（1）JLPS-Hでは、Wave 16までの各Waveで特定の基本属性に偏った標本脱落が生じていること、（2）Wave 16における世帯の暮らし向きを従属変数とした事例的分析でみる限り、IPW法に基づくウェイトは、ささやかにながら推定結果を変えうるため、ウェイトを用いて分析する方が望ましいと考えられること、の2点が分かった。

謝辞 本研究は、科学研究費補助金基盤研究（S）（22223005）、基盤研究（C）（25381122）、基盤研究（B）（16H03778）および厚生労働科学研究費補助金政策科学推進事業（H16-政策-018）の助成を受けたものである。東京大学社会科学研究所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社会科学研究所パネル調査企画委員会の許可を受けた。

1. 本稿の目的

「東大社研・高卒パネル調査」(以下、**JLPS-H** と表記) は、東京大学社会科学研究所が、2004年3月に高校を卒業した人々を対象とし、ほぼ1年毎に実施している、質問紙によるパネル調査である。

この調査の特徴の1つは、調査を開始するにあたって、所属高校を通じて対象者にアクセスして**Wave 1**の回答を得た上で、追跡調査の許諾を本人から得て**Wave 2**以降の調査を実施していることである。加えて、**Wave 1**調査と並行して、調査対象者の所属高校についての調査も実施しており、調査対象者が高校生時点から所属していた学校についての細かな情報が得られている。これによって、高校生時点から34~35歳時点(2021年1月現在)までのキャリアや意識の変遷を追うことができるだけでなく、どのような高校出身であったのか、また高校生の頃にどのような友人たちと高校生活を送っていたのか、といった経験がその後のキャリアや意識にもたらす作用についても検討することが可能となっている。

JLPS-Hが、高校を通して調査対象にアクセスするような調査設計であることは、上記のような利点をもたらす。一方で、このような調査設計は、調査データの取得や分析における問題ももたらすことになる。

代表的な問題として挙げることができるのが、いわゆる標本脱落の問題である。学校を通じて調査対象者にアクセスすることで、たとえば学校の授業のなかで実施されることもあってか、**Wave 1**の回収率は高い水準になり、多くの回収数を得られることになる。しかしながら、**Wave 2**以降の追跡調査については、**Wave 1**実施前に調査対象者から応諾を得ているわけではなく、**Wave 1**に回答してくれたからといって、**Wave 2**についても回答してもらえとは限らない。まして、高校卒業時は、親元を離れて一人暮らしを始めたり、就職や進学によって、生活リズムが大きく変わったりする時期でもあり、追跡調査に応諾していたとしても、調査への回答が得られないこともしばしば見られる。これらの影響によって、**Wave 2**以降の回収率は、**Wave 1**に比べて大きく下がりうる。

このような回収率の低下は、調査データの分析において、バイアスを生むことにつながる。とくに、回収率の低下が何らかの属性の影響下で生じている場合、データやその分析結果に歪みが生じることは避けられない。たとえば、女性よりも男性のほうが、**Wave 2**以降の調査に対して回答してくれなくなる傾向が強い場合、**Wave 2**以降のデータは、当初想定した男女比率よりも、女性比率が高いデータになってしまう。すなわち、女性の回答傾向がより強く反映されたデータになってしまうことになる。

以上の問題意識を踏まえて、本稿では**JLPS-H**データについて、以下の2つの分析をおこなう。

第一に、回答者の基本属性として、性別/出身高校のタイプ/高3時校内成績をピック

アップし、各 Wave のデータでこれらの属性の分布がどのように変わるのかを検討する。たとえば、上述のように男性の方が調査から脱落しやすいのであれば、Wave 2 以降の調査では男性比率が下がり、女性比率が上がるということになっているだろう。

第二に、そのように検討した標本脱落の状況を踏まえて、標本脱落にともなうバイアスを補正する方策を検討する。具体的には、パネル調査に限らず一般的に標本脱落の補正に用いられることの多い、Inverse Probability Weighting 法 (IPW 法) による補正可能性を検討する。IPW 法は、標本への残りやすさを確率として算出して、逆確率で重み付けをおこなうことでデータのバイアスを補正する、という方法である。また、この方法を使用することの意義を検討する上で、IPW 法使用の有無による回答分布の比較、重回帰分析の分析結果の比較をおこなう。

なお、本稿と同様の問題関心から JLPS-H における標本脱落の検討をおこなったものとして、山口 (2018)、山口 (2019) がある。本稿は、これらの先行研究を土台に、Wave 15、Wave 16 のデータを追加した上で、より積極的に IPW 法を使用することの意義を探ろうとする論考だと位置づけられる。

2. JLPS-H の調査設計と標本脱落

(1) JLPS-H の調査設計

ここであらためて、JLPS-H の調査設計を簡単にまとめておく¹。JLPS-H は、無業率と進学率を踏まえて 47 都道府県の中から抽出された、神奈川県・石川県・宮城県・秋田県の 4 県の、県内に所在する公立高校計 162 校に調査を依頼している。このうち、101 校の 7,563 名から Wave 1 の回答を得ている。なお、Wave 1 の調査は 2004 年 1~3 月に実施された。

追跡調査については、応諾する場合は調査票郵送先住所を記述してもらうという形式で、Wave 1 調査票の末尾で依頼をおこなっている。追跡調査に応諾があったのは 2,036 名であり、このうち 501 名より Wave 2 回答を得ている。

Wave 3 以降の調査は、郵送法で実施された。なお、回答については、調査票への記入・返送に加えて、Web での入力・提出も可能であった。調査はおよそ 1 年毎に実施され、回収数は多少の変動はありつつも、おおよそ 500 前後で推移してきた。2021 年 3 月現在で、Wave17 まで実施されており、分析に使用できるのは Wave16 までのデータである。

なお、JLPS-H の回収状況については、百瀬・山口・鈴木 (2020) を参照されたい。

ここまでの話をまとめれば、Wave 1 における回収数を 100%としたとき、追跡調査への応諾者比率は 26.9%、Wave 2 回収者の比率は 6.6%ということになる。Wave3 以降につい

¹ 調査設計の詳細については、石田編著 (2006) や佐藤編 (2017) の序章を参照されたい。

でも、おおよそ Wave 2 と同等の比率で推移している。このことから、JLPS-H においては、Wave 1 から Wave 2 にかけて、追跡調査に応諾するか、また、実際に Wave 2 に回答を得られるか、という 2 つのプロセスを経て大規模な標本脱落が生じ、その後はさほど標本が減少すること無く推移しているとまとめることができる。

(2)標本脱落の問題と対策

標本脱落は、特定の属性に偏ってそれが生じた場合にデータにおけるバイアスを生じさせうるといふ問題に加えて、そもそもある程度のサンプルサイズを確保しなければ多変量分析をはじめとする統計手法の適用に耐えないという問題がある。それゆえ、仮にデータに偏りを生じさせないとしても、標本脱落それ自体が改善されるべき問題だといえる。

そのことを踏まえれば、標本脱落への対応として、標本脱落が生じにくい調査設計をとったり、あらかじめサンプルサイズを多めに確保したり、という対応が考えられる。また、リフレッシュサンプルとして、追加の調査対象者を抽出し、元々の調査対象者とのデータの差異を丁寧に見つつ、両者を結合してサンプルサイズを確保する、という対応もありうる。

だが、JLPS-H に限っていえば、既に Wave 16 まで実施されている調査であり、特定の高校（前述の 4 県 101 校）出身者であり、ここ 16 年にわたって質問紙調査に回答している、ないしは少なくとも回答を依頼されているというやや特殊な調査対象となっているといえる。そのため、JLPS-H における標本脱落への対処として、上述のような方法はかならずしも適切ではない。

そこで、本稿では IPW 法を用いて標本脱落に対処する方法を検討することになる。IPW 法では、本稿が対象とするようなパネル調査における標本脱落を事例に説明すれば、次のようなプロセスで脱落の補正を試みることになる。①標本への残存を“1”、脱落を“0”として、二項ロジスティック回帰分析を用いて、標本残存の予測確率を推定する。②その予測確率の逆数を取る。③予測確率の逆数を、分析における確率ウェイトとして用いることでバイアスを補正する。

パネル調査における標本脱落に対して、このような IPW 法による補正を試みた事例としては、坂本（2006）や三輪（2014）が挙げられる。

坂本（2006）、三輪（2014）はそれぞれ、「消費生活に関するパネル調査」（財団法人家計経済研究所）、「全国家族調査パネルスタディ」（日本家族社会学会）というパネル調査データを対象に、標本脱落にともなって生じたデータのバイアスを IPW 法で補正できるか検討をおこなっている。坂本（2006）は、IPW 法を用いない場合と用いた場合の分析結果の比較を通じて、IPW 法を用いることで分析結果が変化しうることを示している。また、三輪（2014）は、「2 つの度数分布の違いをみるための統計指標で、何パーセント分だけ相対度数をシフトさせれば両者が完全に一致するかを示す」（三輪 2014:171）ものである非類似指数

を指標とした分析を通じて、IPW 法を用いることで、サンプルに特定の属性への偏りを生む標本脱落の影響を軽減しうることを示唆している。

もっとも、坂本（2006）や三輪（2014）が IPW 法の適用対象とした調査の対象者は成人男女である。一方で、JLPS-H は学校在学時を起点とするようなパネル調査であり、井上（2009）も示唆するように、そのような調査設計の調査特有の標本脱落が生じている可能性もある。そのため、JLPS-H における標本脱落に対しても、IPW 法が効果を持つかどうかは明らかではない。

そこで、次節では、まず JLPS-H における標本脱落の状況として、各 Wave における後述する 4 つの属性の分布の変化を確認した上で、それを踏まえて IPW 法を使用した場合のバイアスの補正効果を確認してゆく。

3. 標本脱落の状況とその補正

(1) JLPS-H における標本脱落と属性の関連

まずは、JLPS-H データにおける標本脱落がどのように生じているのかを確認してゆく。具体的には、先に述べたように、回答者の基本属性である性別／出身高校のタイプ／高 3 時校内成績（いずれも Wave 1 での回答）それぞれの分布が、Wave によってどう変わってゆくかをみる。ここで確認しておきたいのは、標本脱落の原因というより、標本脱落によって、データにどのような変化が生じたのか、ということである。

表 1 に、Wave 1 から Wave 16 までの各 Wave の有効回答者における、Wave 1 で回答された性別／出身高校タイプ／高 3 時校内成績の分布を示した。

まず性別をみると、Wave 1 では男性が 48.7%、女性が 51.1%（無回答は省略）とおおよそ 1 対 1 となっているのに対し、Wave 2 では男性の比率が 42.7%と大きく下がり、女性が 57.3%となっている。このような有効回答者が女性に偏る傾向は、その後の Wave でいっそう顕著になっていく。たとえば Wave 3 では男性が 38.3%、女性が 61.7%と女性比率が 6 割を超え、直近の Wave 16 に至っては、男性が 31.4%、女性が 68.6%と、女性比率が 7 割に迫る勢いとなっている。この結果から、JLPS-H では男性の方が女性よりも脱落しやすく、Wave が進むにつれて女性の回答が過大に評価されるデータとなっていくことが分かる。

次に出身高校タイプについてみると、Wave 1 では進学校 41.1%、進路多様校 38.4%、専門高校・学科 20.5%という比率であるのに対し、Wave 2 では進学校出身者の比率が大きくなり、54.8%を占めるようになる。すなわち、進学校出身者よりも進路多様校や専門高校・学科の出身者の方がパネル調査から脱落しやすく、進学校出身者の回答が過大に評価されるデータとなっていることになる。ただし、Wave 4 を過ぎると、大学進学者も卒業する時期であるためか、このような進学校出身者比率の高まりはやや改善される。直近の Wave 16

のデータでは、進学校出身者比率は5割を切る48.4%となっており、Wave 1の41.1%にまでは至らないものの、Wave 2の54.8%に比べれば、Wave 1における比率に近づいていると言える。

最後に高3時校内成績についてみると、Wave 1では上位33.1%、中位31.7%、下位33.1%、無回答2.1%であるが、Wave 2では上位層の比率が大きくなり、41.5%を占めるようになる。中位層の比率はほぼ変わらないものの、下位層の比率は23.0%と10ポイント近く低くなっている。すなわち、高3時の校内成績が低かった層は高かった層に比べてパネル調査から脱落しやすく、成績上位層の回答が過大に評価されるデータとなることが分かる。もっとも、出身高校タイプと同様に、成績上位層の比率がWave 1より大きくなる傾向は変わらないものの、そのようなバイアスはWaveが進むにつれて多少改善される様子もうかがえる。

表1 各Waveの有効回答者における性別、出身高校タイプ、高3時校内成績の分布

	N	性別		出身高校タイプ			高3時校内成績			
		男性	女性	進学校	進路 多様校	専門高校 ・学科	上	中	下	無回答
Wave1	7,563	48.7	51.1	41.1	38.4	20.5	33.1	31.7	33.1	2.1
Wave2	482	42.7	57.3	54.8	29.9	15.4	41.5	31.1	23.0	4.4
Wave3	645	38.3	61.7	51.8	30.5	17.7	41.6	29.6	24.6	4.2
Wave4	526	38.8	61.2	54.6	27.6	17.9	41.0	29.6	25.1	4.4
Wave5	509	39.5	60.5	49.3	32.4	18.3	37.9	31.9	25.9	4.3
Wave6	449	38.8	61.3	47.0	35.0	18.0	40.9	30.5	24.5	4.1
Wave7	498	40.2	59.8	48.0	32.3	19.7	34.9	31.8	29.5	3.9
Wave8	487	36.8	63.2	46.0	34.3	19.7	38.1	30.6	27.3	4.0
Wave9	496	38.7	61.3	45.2	35.7	19.2	35.6	30.4	30.2	3.9
Wave10	468	36.8	63.3	48.3	32.9	18.8	34.7	32.2	29.3	3.7
Wave11	451	34.2	65.9	48.1	34.4	17.5	35.9	31.2	29.3	3.7
Wave12	448	34.6	65.4	48.4	33.5	18.1	36.9	30.7	28.7	3.7
Wave13	435	35.9	64.1	47.8	32.4	19.8	35.6	32.2	28.4	3.8
Wave14	472	35.4	64.6	48.3	32.0	19.7	37.8	30.2	28.0	4.1
Wave15	426	33.3	66.7	49.5	32.9	17.6	37.6	28.5	29.9	4.1
Wave16	405	31.4	68.6	48.4	33.3	18.3	37.8	29.1	28.4	4.7

注: 第1波の性別「無回答」0.25%は上表から省略した

(2) JLPS-Hにおける標本脱落の推定

前項において、JLPS-Hにおいても、特定の属性に偏って標本脱落が生じていることが分かり、それにとまって、Wave 2以降のデータにWave 1標本からのズレが生じていることが示唆された。そこで、IPW法を同データに用いるために、本項でウェイトを作成する。

もっとも、どのWaveにおける回答を分析に使用するかによって、用いるべきウェイトは

変わってくる。たとえば、Wave 2 時点の生活満足度を Wave 1 と Wave 2 に含まれる変数で回帰したいのであれば、Wave 2 における回答有無を Wave 1 の変数で予測し、予測確率の逆数を作成することになる。もし、Wave 16 時点の生活満足度を Wave 16 時点の変数と Wave 1、Wave 2 に含まれる変数で回帰するのであれば、分析には Wave 1、Wave 2、Wave 16 の 3 つの Wave に回答したケースが用いられることになるため、Wave 2 と Wave 16 のどちらも回答したかそうでないかを Wave 1 の変数で予測し、予測確率の逆数を作成する必要がある。ここでは、Wave 2 における回答有無を従属変数として、IPW 法のためのウェイトを作成してみることにしたい。上に述べたように、従属変数を変更すれば、個々の分析に合ったウェイトを作成することが可能である。

表 2 Wave 2 における回答の有無の予測（二項ロジスティック回帰分析）

	Wave 2 回答（あり=1/なし=0）		
	Exp(B)	B	S.E.
性別（基準：男性・無回答）			
女性	1.213	0.193†	0.099
高校所在県（基準：A県）			
B県	1.399	0.336*	0.142
C県	1.328	0.284*	0.132
D県	1.245	0.219	0.161
Wave 1 時年齢（基準：17歳）			
18歳	1.006	0.006	0.136
19歳以上・無回答	1.313	0.272	0.353
学校タイプ（基準：進路多様校）			
進学校	1.542	0.433***	0.116
専門高校・専門学科	1.207	0.188	0.150
高卒後希望／予定進路（基準：就職）			
大学・短大	2.049	0.717***	0.170
専門学校	2.074	0.730***	0.178
未定・その他・無回答	1.843	0.612*	0.250
高3時校内成績（基準：中位）			
上位	1.302	0.264*	0.111
下位	0.727	-0.319*	0.128
無回答	0.401	-0.915	0.592
切片	0.021	-3.857***	0.226

注) n=7,563, ***p<0.000 **p<0.001 *p<0.05 †p<0.1

Wave 2 における回答有無を予測してウェイトを作成するにあたって、説明変数には Wave 1 で得た以下の変数を使用している。性別、高校所在県、Wave 1 時年齢、出身高校の学校タイプ、高卒後希望／予定進路、高3時校内成績、である。学卒時の標本脱落が離家や生活リズムの変化などによって生じやすいと想定されることを踏まえると、本当であれ

ば離家や転居、生活リズムの変化についての情報を用いることが理想的ではある。しかしながら、パネル調査からの脱落以前にこれらの情報を得ることは困難であり、JLPS-HのWave 1でもこれらの情報は得られていない。そこで、表1で使用した変数に加えて、離家や転居、生活リズムの変化との関連が大きいと思われる高卒後希望／予定進路、また高校所在県、Wave 1時年齢についてもウェイト作成に使用している。

ウェイトを作成するために、Wave 2における回答有無を二項ロジスティック回帰分析で予測したのが表2である。

上で述べたように従属変数を変えた場合には、当然ながら対数オッズ比やオッズ比、標準誤差の値が変わってくるため、ここではそれらの値については議論しない。この推定によって得られた推定値を元にして、Wave 2に回答した各ケースに対して、表2の独立変数から予測した「回答あり」確率が付与される。その逆数を各ケースの確率ウェイトとして重みづけることで、Wave 2への回答者の属性の偏りを補正することができる、と考えられる。

(3)IPW 法によるバイアスの補正と分析

IPW 法を用いることによって、特定の属性にかたよった標本脱落にもなうバイアスが補正されうることは、山口(2019)が非類似指数を用いた分析によって明らかにしている。ここでは、実際に JLPS-H のデータを用いた分析をおこなうにあたって、そのようにバイアスを補正する IPW 法を用いたら、分析結果がどのように変わるのかを、世帯のくらしむき(Wave 16)を従属変数とした事例分析によって確認してみたい。

世帯のくらしむきは、「現在のあなたの世帯のくらしむきは、この中のどれにあたりますか」と尋ね、豊か／やや豊か／どちらかといえば豊か／どちらかといえば苦しい／やや苦しい／苦しい、の6つの選択肢のなかから解答してもらう設問である。表2にみるように、標本脱落にもなって性別や高校所在県、高卒後希望／予定進路(≒学歴)の分布に偏りが生じているのであれば、それらの変数との相関が大きいとみられる世帯のくらしむきを従属変数とした分析結果にも一定のバイアスが生じている可能性がある。そのようなバイアスを IPW 法で補正できるのであれば、IPW 法使用の有無によって、分析結果も変わると考えられる。なお、この分析において使用するウェイトは、Wave 16における回答の有無を、表2と同じモデルで推定して作成したものである。

分析結果を表3に示した。IPW 法を使用しない場合と使用した場合で、偏回帰係数や頑健標準誤差の推定値には違いがみられることが分かる。ただし、その違いは十分に大きいとはいえない。分析結果を解釈する上で違いがあるといえるのは、性別(男性)と高校所在県(D県)のみである。

世帯のくらしむきを従属変数として事例的に分析をおこない、IPW 法の効果を検討した結果は、次のとおりにまとめられる。JLPS-H に対して IPW 法を用いることは、分析結果

をある程度変化させるため、山口（2019）が示した IPW 法のバイアス補正効果を併せて考えれば、JLPS-H における標本脱落を考慮する上で、IPW 法を用いて分析をおこなう方が望ましい。しかし一方で、少なくとも表 2 に示したようなモデルで作成したウェイトでの IPW 法で世帯のくらしむきを従属変数とした分析を補正する場合には、IPW 法によって大きく分析結果が変わる訳ではない。

表 3 世帯のくらしむき（Wave 16）を従属変数とした重回帰分析（OLS）

	IPW法不使用		IPW法使用	
	B	S. E.	B	S. E.
性別 ref. 女性				
男性	-0.170	(0.120)	-0.282*	(0.128)
高校所在県 ref. A県				
B県	0.122	(0.177)	0.071	(0.180)
C県	0.005	(0.162)	0.051	(0.167)
D県	0.253	(0.173)	0.309†	(0.184)
現在の状況 ref. 正社員・公務員				
非正社員	-0.427**	(0.137)	-0.453**	(0.152)
自営業主・家族従業	-0.262	(0.290)	-0.150	(0.220)
無職、出産・育休中	-0.195	(0.193)	-0.127	(0.250)
対数世帯年収	0.026	(0.027)	0.058	(0.035)
切片	3.805***	(0.202)	3.636***	(0.236)
R-squared	0.042		0.055	

注：N=404, S. E. には頑健標準誤差を示した。*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$

4. 議論

本稿では、JLPS-H における標本脱落の状況と、それにもなつて生じるバイアスの補正可能性について検討を進めてきた。具体的には、各 Wave における性別、出身高校タイプ、高 3 時校内成績の分布の違いをみることで、標本脱落の状況を概観し、IPW 法に用いるウェイトの作成を試み、Wave 16 の世帯のくらしむきを従属変数として、実際に IPW 法を用いた場合の分析結果の違いをみてきた。

これらの分析から、以下のことを導けるだろう。

まず、JLPS-H においては、学校を通して実施する調査設計の影響もあつて、調査初期の標本脱落が大きく、特定の属性に偏つた標本脱落が生じている。これはデータの偏りを生んで分析にバイアスを生じさせかねない点で問題だといえる。ただし、そのようなデータの偏りは、IPW 法を用いることで補正できると考えられ、実際に事例的に分析をおこなつてみる限り、JLPS-H に対しても、IPW 法はささやかながら効果を持ちうる。

もっとも、本稿で示したような方法で JLPS-H に対して IPW 法を用いるべきかどうか

は、母集団をどのように想定してどんな分析をおこなうかによっても異なる。今回の分析においては、Wave 1 が想定する母集団の理想的なサンプルだということを暗黙の前提としていたが、実際のところ、回収率は比較的高いとはいえ、Wave 1 でさえも当初予定していた対象者全員から回答を得られたわけではまったくない。だとすれば、分析する Wave ないしはデータにおける標本脱落の状況を確認した上で、そのようなデータの偏りを前提に分析結果の解釈をおこなっていくことも考えられるだろう。もちろん、Wave 1 にできるだけ近づける形で Wave 2 以降のデータを分析したいのであれば、IPW 法を用いることは有用であり、むしろ用いるべきだということもできる。

最後に本稿の限界について述べておく。本稿で IPW 法のウェイトを作成するにあたっては、あくまで表 2 に示した変数のみを用いており、Wave 1 の他の変数、また Wave 1 で尋ねることのできていない個人間の異質性については考慮することができていない。また、上記の議論と関わる論点として、JLPS-H ではサンプリングの過程での対象者の選択、調査を依頼したなかでの Wave 1 回答の有無、追跡調査への応諾の有無、各 Wave に対する回答の有無、と調査からの脱落が生じる段階が複数に渡っている。本稿ではこのうち、最後の各 Wave に対する回答の有無についてのみ分析をおこなっており、それ以外の脱落とそれともなって生じるバイアスについては検討をおこなうことができていない。

パネル調査において、調査からの脱落が生じることは避けられない。そのような脱落について、どんな脱落が生じており、それはどのようなバイアスを生じさせる可能性があり、どうすればそれを回避ないしは補正することができるのかを考えて分析をおこなっていく姿勢が求められているといえよう。

引用文献

- 石田浩編著, 2006, 『高校生進路選択と意識変容』東京大学社会科学研究所.
- 三輪哲, 2014, 「NFRJ-08Panel におけるウェイトによる脱落への対処」『家族社会学研究』26(2): 169-178.
- 百瀬由璃絵・山口泰史・鈴木富美子, 2020, 「高卒 15 年目の、学校教育の効果と奨学金返済、夫婦ペアデータによる夫の子育てと認識のズレ: 高卒パネル調査 wave16 の結果から」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
- 坂本和靖, 2006, 「サンプル脱落に関する分析——『消費生活に関するパネル調査』を用いた脱落の規定要因と推計バイアスの検証」『日本労働研究雑誌』551: 55-70.
- 佐藤香編著, 2017, 『格差の連鎖と若者 3 ライフデザインと希望』勁草書房.
- 山口泰史, 2018, 「東大社研・高卒パネル調査 (JLPS-H) における標本脱落状況の検討」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

ズ No.107

山口泰史, 2019, 「高校在学時を起点とするパネル調査における初期標本脱落とバイアスの補正—『東大社研・高卒パネル調査』を事例として—」『社会と調査』 23: 66-70.

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/>