

Panel Data Research Center at Keio University
DISCUSSION PAPER SERIES

DP2013-005

February, 2014

「慶應義塾家計パネル調査(KHPS)」と「日本家計パネル調査(JHPS)」における
Longitudinal / Cross-sectional ウェイトおよびパネル統合ウェイトの作成

石井加代子 (慶應義塾大学) *

野崎華世 (慶應義塾大学) *

【要旨】

慶應義塾大学パネルデータ設計・解析センターで調査実施・管理されている「慶應義塾大学家計パネル調査 (Keio Household Panel Survey: KHPS)」および「日本家計パネル調査 (Japan Household Panel Survey: JHPS)」は、古いもので調査開始より 10 年が経過している。パネル調査においてサンプル脱落の問題は不可避であり、KHPS と JHPS も例外ではない。本稿では、サンプル脱落によって生じるバイアスの修正、および、初回調査のサンプリング時点におけるバイアスの修正を行うべく、Longitudinal ウェイトと Cross-sectional ウェイトの作成を試みた。さらに、KHPS における追加サンプルの突合、および KHPS サンプルと JHPS サンプルの統合を実現するべく、それに必要となる統合ウェイトについても作成を試みた。

* 慶應義塾大学大学院商学研究科 特任講師

Panel Data Research Center at Keio University
Keio University

「慶應義塾家計パネル調査(KHPS)」と「日本家計パネル調査(JHPS)」における Longitudinal / Cross-sectional ウェイトおよびパネル統合ウェイトの作成*

石井加代子（慶應義塾大学）

野崎華世（慶應義塾大学）

要約

慶應義塾大学パネルデータ設計・解析センターで調査実施・管理されている「慶應義塾大学家計パネル調査（Keio Household Panel Survey: KHPS）」および「日本家計パネル調査（Japan Household Panel Survey: JHPS）」は、古いもので調査開始より 10 年が経過している。パネル調査においてサンプル脱落の問題は不可避であり、KHPS と JHPS も例外ではない。本稿では、サンプル脱落によって生じるバイアスの修正、および、初回調査のサンプリング時点におけるバイアスの修正を行うべく、Longitudinal ウェイトと Cross-sectional ウェイトの作成を試みた。さらに、KHPS における追加サンプルの突合、および KHPS サンプルと JHPS サンプルの統合を実現するべく、それに必要となる統合ウェイトについても作成を試みた。

* 本稿の分析に際して、慶應義塾大学パネルデータ設計・解析センターより「慶應義塾家計パネル調査」と「日本家計パネル調査」の個票データの提供を受けた。樋口美雄教授および山本勲准教授には折に触れて貴重なコメントをいただいた。当然ながら、本稿の分析と結果の解釈の責任は筆者にのみある。

1. はじめに

慶應義塾大学パネルデータ設計・解析センターでは、「慶應義塾大学家計パネル調査 (Keio Household Panel Survey: KHPS)」および「日本家計パネル調査 (Japan Household Panel Survey: JHPS)」の 2 つのパネル調査の実施・管理を行っている。成人男女を対象に、KHPS は 2004 年より、JHPS は 2009 年より調査が開始され、現在に至るまで毎年調査が実施されている。KHPS と JHPS の 2 つの調査はそれぞれ個別の調査として実施・管理がなされてきたが、2013 年調査より 2 つの調査の調査票を統一化し、データの管理の一元化を図ることとした。これにより、2 つのデータを統合して分析することが可能となり、パネルデータにおいて不可避なサンプル脱落によるサンプルサイズの減少の部分的な解消につながる。なお、KHPS と JHPS の調査項目には、もともと類似の項目が多くあったため、調査票の統一化により調査内容が大幅に変更されることはない。

KHPS においては調査開始より 10 年、JHPS においては調査開始より 5 年が経つが、現在まで、データのサンプリング・バイアスに対する修正や、サンプル脱落によるバイアスの修正について対策が講じられてこなかった。そこで、本稿では、これらのバイアスに対する修正ウェイトの作成、加えて、KHPS と JHPS の 2 つの調査データを統合した際に必要となるサンプルのバイアスを修正するウェイトの作成について検討する。以降では、KHPS および JHPS の調査の概要について説明したうえで、修正ウェイトの作成について検討する。

2. 調査の概要

ここでは、KHPS と JHPS の 2 つの調査の調査設計と現在までの経緯について説明する。

KHPS (「慶應義塾家計パネル調査」: Keio Household Panel Survey) は 2004 年 1 月に第 1 回調査、JHPS (「日本家計パネル調査」: Japan Household Panel Survey) は 2009 年 1 月に第 1 回調査を行い、現在に至るまで同一対象者を追跡調査している。KHPS は調査開始時点で 20 歳以上 69 歳以下の男女 4,000 名を対象に調査を実施、一方、JHPS は調査開始時点で 20 歳以上 (年齢の上限設定なし) の男女 4,000 名を対象に調査を実施した。また、KHPS においてはサンプル脱落への対応として、第 4 回調査 (2007 年) に当時 20 歳以上 69 歳以下の男女 1,400 名を追加、さらに、2012 年に当時 20 歳以上 69 歳以下の男女 1,000 名を追加している。

(1) 調査対象者の抽出方法

上述のとおり、KHPS では初回調査 (2004 年)、第 4 回調査 (2007 年)、第 9 回調査 (2012 年) の 3 時点、JHPS では初回調査 (2009 年) でサンプリングを行っている。KHPS と JHPS のサンプリングにおいて特筆すべき点として、①抽出単位は個人単位であり、②抽出方法として層化 2 段無作為抽出法が用いられており、③予定した標本サイズを確保するため、正規に抽出された対象者 (正規対象者) と同じ調査区・同じ性別・同じ年齢階級 (10 歳刻

み)の予備対象者(正規対象者1人につき10人)を用意している点である。また、調査対象者が有配偶者の場合は、その配偶者に対しても同様の調査を行っている点も特記しておく。

図表 1 KHPS/JHPS のサンプリングの特徴

1. サンプルの抽出単位は個人単位
2. 層化2段無作為抽出法の適用
3. 正規対象者と同じ地域・性・年齢の予備対象者を用意
4. 調査対象者の配偶者に対しても調査

層化2段無作為抽出では、抽出の第1段階で、全国を地域・都市階級により24層(8地域¹×3都市階級²)に層化し、各層に住民基本台帳人口の人口割合に合わせ標本数を配分する。そのうえで、1つの調査区あたりの標本数を10程度(2007、2012年の追加サンプルにおいては5程度)として各層の調査区数を決定し、各層において所定数の調査区を無作為抽出する。調査区は、抽出単位として国勢調査の調査区を使用している。第2段階では、選定された調査区の住民基本台帳を抽出台帳として、調査対象適格者を対象に、指定された起番号、抽出間隔に基づき1調査区について約10人(2007、2012年の追加サンプルにおいては5人)を抽出している。これが正規対象者である。

図表 2 KHPS2004 における標本数の配分方法 (例)

	14大市	その他の市	町村	合計
北海道	61	80	38	179
東北	32	162	101	295
関東	498	693	146	1,337
中部	69	476	181	726
近畿	174	404	81	659
中国	36	142	55	233
四国	-	83	42	125
九州	75	235	136	466
合計	945	2,275	780	4,000

註) まず、全国を地方・都市階級により24層に層化し、各層に2003年3月31日現在の住民基本台帳人口の人口割合で標本数を配分。次に、1つの調査地域あたりの標本数を10程度として各層の調査地域数を決定し、所定数の調査区を無作為抽出した。調査地域は、抽出単位として2000年国勢調査の調査区を使用した。

¹ 8地域とは、北海道、東北、関東、中部、近畿、中国、四国、九州である。

² 3都市階級とは、18大市、その他の市、町村である。

何らかの理由で正規対象者に調査協力を拒否された場合、あらかじめ選定しておいた予備対象者で代替することにより、予定した標本サイズ（初回 4,000、2007 年 1,500、2012 年 1,000）を確保している。前述のとおり、予備対象者は正規対象者 1 人につき 10 人用意され、それぞれ、正規対象者と同じ調査区内に居住し、同じ性・同じ年齢階級（10 歳刻み）であるため、回答者が予備対象者に代替されたとしても、地域・性別・年齢区分においては、無作為抽出の結果を維持すると考えられる。

（2） 調査方法

調査方法については、KHPS と JHPS で若干の違いがある。KHPS においては、すべての調査対象者に対して、各調査区の調査員が調査対象者の居宅に赴き調査票を渡し、後日、調査員が再度訪問し調査対象者が記入した調査票を回収する自計式の留置調査法を用いている。

一方、JHPS においては、調査方法の差異が調査票の回収率や各調査項目の回答率に及ぼす影響を明らかにするために、実験的に（ア）2 つの調査方式を実施し、（イ）Web 回答のオプションを設けた。さらに、調査員のモチベーションが回収率に対する影響を分析するため、（ウ）調査員向け事前説明会の内容や、（エ）調査員への完了報酬にバリエーションを設けた。詳しくは、下記に示すとおりである。

（ア） 調査方式

調査方式による回収率への影響を把握するため、調査全地点を二分して（A 地点群、B 地点群）、それぞれ異なる調査方式を実施した。ただし、2013 年調査以降は、B 地点群の調査方式は廃止され、すべて A 地点群の調査方式に統一された。

《A 地点群》

調査員が調査対象者の居宅に赴き調査票を渡し、後日、調査員が再度訪問し調査対象者が記入した調査票を回収する自計式の留置調査法を実施。

《B 地点群》

質問項目を分割し、調査員が調査対象者の居宅に赴き調査票を渡し、後日、調査員が再度訪問し調査対象者が記入した調査票を回収する自計式の留置調査と、調査員が口頭で対象者に質問して回答してもらう面接調査を併用。

（イ） Web 回答の選択肢の用意（対象者の任意）

インターネット環境が整っている調査対象者には、留置法による調査項目について、Web での回答オプションも用意。

（ウ） 事前説明会への研究メンバーの参加の有無

調査員が所属する調査支社局を 2 つのグループに分け、1 つのグループでは研究メンバー（大学教員）が調査員の事前説明会に参加し、JHPS 調査の主旨およびその実験の意義などについて説明。もう 1 つのグループは通常の事前説明会を実施。研究メンバーが直接働きかけることによって調査員の熱意や、調査対象者に対する説明度合に差があるかどうかを

検討する目的。

(エ) 調査員への完了報酬の方式

調査員は、調査が完了すると完了報酬を受け取る。この完了報酬の差が調査の回収率に影響するか確認するため、調査員を二分して、一方のグループには完了報酬を正規対象者と予備対象者で差をつけ、もう一方のグループには差をつけないよう設定。具体的には、前者のグループにおいては、正規対象者の完了報酬を3,300円、予備対象者の完了報酬を2,600円とし、後者のグループでは、正規対象者でも予備対象者でも完了報酬を一律2,800円とした。

3. 回収率

(1) 初回調査の疑似回収率（参考値）

通常の調査では、抽出された対象者に対して調査依頼を行い、調査に協力してもらえなかった場合は未回収となる。しかしながら、KHPSとJHPSでは、上述のとおり、所定の標本サイズを確保するため、正規対象者に協力を拒否された場合、予備対象者に協力を依頼している。そのため、通常の調査と比較可能な回収率を算出することができない。

そこで、調査員が各調査につき記入する調査員確認票を用い、各調査対象者との接触状況を把握し、調査員が接触可能であった対象者のうち、実際に回答に協力した対象者の割合を「疑似回収率（参考値）」として図表3に掲載する。

図表3 初回調査の疑似回収率（参考値）

	KHPS			JHPS
	2004年 wave1	2007年 新規サンプル	2012年 新規サンプル	2009年 wave1
調査対象者数	13,430	5,868	3,183	12,549
有効回答数 *1	4,005	1,419	1,012	4,022
接触可能対象数	9,737 *3	4,256	2,331	9,654
接触可能率(%) *2	72.5 *3	72.5	73.2	76.9
疑似回収率(%) *4	41.1	33.3	43.4	41.7

*1 予備対象として調査した人も含む

*2 (接触可能対象数/調査対象者数)×100

*3 KHPS 2004は、調査員確認票を導入していなかったため、接触可能率はKHPS2007 新規サンプルの値と同じであったと仮定し、接触可能対象数を算出している。

*4 (有効回答数/接触可能対象者数)×100

(2) 継続回収率

KHPS/JHPSそれぞれの第2回調査以降の回答状況と継続回収率（(有効回答数-復活サンプル)/前年度完了数×100）を図表4に示す。KHPSとJHPSのいずれにおいても、おおよそ、調査を重ねるごとに回収率が高まっており、9割前後の回収率を確保している。

図表 4 継続回収率

KHPS									
	2005年	2006年	2007年	2008年	2009年	2010年	2011年	2012年	2013年
	wave2	wave3	wave4	wave5	wave6	wave7	wave8	wave9	wave10
調査対象者数	4,005	3,342	2,894	4,067	3,706	3,448	3,232	3,041	3,888
うち、前年度完了数	4,005	3,342	2,887	4,062	3,691	3,422	3,207	3,030	3,877
有効回答数	3,314	2,887	2,643	3,691	3,422	3,207	3,030	2,865	3,568
うち、復活サンプル	-	0	3	0	4	7	10	10	11
欠票	691	455	251	371	273	222	187	175	320
継続回収率(%)*1	82.7	86.4	91.4	90.9	92.6	93.5	94.2	94.2	91.7

JHPS				
	2010年	2011年	2012年	2013年
	wave2	wave3	wave4	wave5
調査対象者数	4,022	3,490	3,170	2,839
うち、前年度完了数	-	3,470	3,160	2,821
有効回答数	3,470	3,160	2,821	2,581
うち、復活サンプル	-	6	4	8
欠票	552	316	342	248
継続回収率(%)*1	86.3	90.9	89.1	91.2

*1 (有効回答数－復活サンプル)/前年度完了数×100

4. 修正ウェイト作成の手順

諸外国のパネルデータにおけるサンプルバイアスの修正ウェイトの作成は、一般的に図表 5 の手順を踏む。KHPS/JHPS のサンプル作成においても、この手順を参考にする。なお、図表 5 の手順に加えて、KHPS/JHPS では、追加サンプルの統合、および、KHPS サンプルと JHPS サンプルの統合のためのウェイト作成を行う必要がある。

図表 5 ウェイト作成における一般的な手順

- ① ウェイトをつけるサンプルの選定
- ② 選択確率の逆数としてのinitial ウェイトの算出
- ③ 回答傾向同一グループ (response homogenous groups) の作成もしくは回答傾向に関する (response propensities) モデルの作成による無回答バイアスの修正
- ④ 上記の手順で作成されたある変数 (性、年齢、学歴など) のウェイト付き値が母集団値に一致するように、センサス統計などをもとに再度ウェイトを調整 (calibration)

(参考) Watson(2012) p.8

作成するウェイトは、初回調査における Cross-sectional ウェイト³と、2回目以降のサン

³ 同一個人を追跡するパネルデータでは、初回調査以降の人口変動 (特に流入) を把握しきれないため、2回目以降の調査を用いて横断面の集計をするのは望ましくない。そのため、このペーパーでは2回目以降のデータに対する Cross-sectional ウェイトの作成は行わない。しかしながら、諸外国のパネルデータでは、このような欠点を踏まえつつ、2回目以降のデータに対しても Cross-sectional ウェイトを提供してい

プル脱落を修正した **Longitudinal** ウェイトである。**Cross-sectional** ウェイトは、当該時点における母集団分布とのかい離を修正するウェイトであり、**Longitudinal** ウェイトは、サンプル脱落によるバイアスを修正し、初回調査の分布に調整するウェイトである。

ウェイトをつける単位としては、KHPS/JHPS ではサンプルの抽出単位が個人のため、個人単位のウェイトのみ作成する⁴。一方、諸外国のパネルデータの多くはサンプルの抽出単位が世帯であり、**Cross-sectional** ウェイトについては個人単位のウェイトと世帯単位のウェイトが作成されている。ただし、世帯は時間を通じて変形する可能性があり、パネル分析の対象にはならないため、**Longitudinal** ウェイトについては個人単位のウェイトのみが用意されている。

なお、図表 5 の①のウェイトをつけるサンプルの選定については、KHPS も JHPS はサンプル抽出の単位が個人であるため、ウェイトを付すサンプルは、単純に調査に協力した回答者に限定することができる。一方、諸外国のパネルデータの多くは、サンプルの抽出単位が世帯であり、世帯員全員（しばしば 15 歳未満を除く）を調査対象としているため、ウェイトを付すサンプルの選定作業がやや複雑である。

5. 初回調査における **Cross-sectional** ウェイト

ここでは、パネルの初回調査時点の **Cross-sectional** ウェイトとして、KHPS2004 の wave1 と JHPS2009 の wave1 に加え、KHPS2007 新規サンプルの初回時点、および KHPS2012 新規サンプルの初回時点の **Cross-sectional** ウェイトを作成する。

初回調査における **Cross-sectional** ウェイトの目的は、サンプリングから生じるバイアスの修正と、調査対象者の調査協力拒否、すなわち無回答から生じるバイアスの修正を行うことである。前者については、選択確率の逆数からなる **initial** ウェイトを作成することで対処するのが一般的であるが、KHPS と JHPS の正規対象者については、層化 2 段無作為抽出法により正規対象者が抽出されており、各対象者の選択確率はすべて等しいため、この範囲においては **initial** ウェイトによる調整は必要ない。また、正規対象者から協力が得られない場合に代わって回答する予備対象者も、該当する正規対象者と同じ居住地・性別・年齢階級という条件のもと無作為に抽出されているため、**initial** ウェイトによりサンプリング・バイアスを修正する必要はない。

一方、予備対象者を用意して予定標本数を確保しているといえども、得られたデータには正規対象者と予備対象者が回答拒否したことで生じたバイアスがあるため、この部分についてはウェイトによる修正が必要である⁵。正規対象者も予備対象者も無作為に抽出されて

るものがあるため、これについては今後の課題とする。

⁴ KHPS/JHPS においても、世帯単位の **Cross-sectional** ウェイトの作成は可能であるため、今後の課題とする。

⁵ KHPS・JHPS では、正規対象者と同じ居住地・性・年齢階級の予備対象者が用意されているため、回答者が予備対象者に代替されたとしても、地域・性別・年齢においては、無作為抽出の結果を維持することができる。しかし、学歴や所得水準、世帯類型など、その他の条件については統御されていないため、極端な例として、対象者全体で学歴の高い人が調査に協力する確率が 0 という場合、得られたサンプルは、

おり、回答行動に違いがあるとは考え難いため、得られたデータ全体で無回答バイアスを修正するウェイトを作成する。なお、回答拒否した正規対象者や予備対象者の情報はほとんどないため⁶、図表 5 の③にあるようなモデルを用いたバイアスの修正を行うことはできない。そこで、Calibration (④) によって対処することとする。

Calibration では、Post-stratification の手法を用いて、学歴（中学卒／高校卒／短大・高専卒／大学・大学院卒／在学中）・就業状態（主に仕事／通学のかたわら仕事／家事のかたわら仕事／休職中／求職中／通学・家事）・男女別配偶状態（男性有配偶／女性有配偶／男性無配偶／女性無配偶）・年齢階級（5 歳刻み）・地域（8 ブロック）の変数を軸に、母集団とみなされるベンチマーク統計（主に国勢調査）に合わせてウェイトを作成する。Post-stratification にかんする簡単な例を図表 6 に示す。

図表 6 Post-stratification の例

	母集団の分布	サンプルの分布	母集団 / サンプル	ウェイト
年齢18-34×男性	17%	13%	0.17/0.13	1.308
年齢18-34×女性	18%	12%	0.18/0.12	1.500
年齢35-64×男性	16%	15%	0.16/0.15	1.067
年齢35-64×女性	17%	15%	0.17/0.15	1.133
年齢65以上×男性	16%	22%	0.16/0.22	0.727
年齢65以上×女性	16%	23%	0.16/0.23	0.696
合計	100%	100%		

註：各カテゴリーにかんする母集団における分布とサンプルにおける分布の比から、ウェイトを作成する。

図表 7 Iterative proportional fitting “raking” algorithm の例

変数A,B,Cを軸にPost-stratificationのウェイトを作成する
1. 変数Aを軸に、データをベンチマーク統計に合わせてウェイトWAを作成。 ・ウェイトWAで修正したデータを用いて、Bにかんする度数分布表を作成。
2. 変数Bを軸に、データをベンチマーク統計に合わせてウェイトWA*WBを作成 ・ウェイトWA * WB で修正したデータを用いて、Cにかんする度数分布表を作成。
3. 変数Cを軸に、データをベンチマーク統計に合わせてウェイトWA*WB*WCを作成。 ・ウェイトWA * WB * WC で修正したデータを用いて、Aにかんする度数分布表を作成。
4. 変数Aを軸に、データをベンチマーク統計に合わせてウェイトWA*WB*WC * WA2を作成。 ・ウェイトWA * WB * WC * WA2で修正したデータを用いて、Bにかんする度数分布表を作成。
5. 変数Bを軸に、データをベンチマーク統計に合わせてウェイトWA*WB*WC * WA2 * WB2を作成。 ・ウェイトWA * WB * WC * WA2 * WB2で修正したデータを用いて、Cにかんする度数分布表を作成。
⋮
⋮
⋮
これを繰り返す行い、ウェイト付きデータの度数分布とベンチマーク統計の度数分布が変わらなくなるまで行う。

図表 6 では、年齢階級と性別を軸に、サンプルのバイアスを修正するウェイトを作成して

学歴の高い人がいないというバイアスを持ったデータになる。

⁶ KHPS/JHPS には、対象者 1 人ずつにつき調査員が記入する「調査員確認票」なるものが存在するが、訪問回数や拒否の理由など、回答拒否者について得られる情報が少なく、これを用いて無回答によるバイアスを修正するモデルを作成することはできない。

いる。このように、軸となる変数が少ない場合は、手計算により Post-stratification のウェイトを作成することができるが、今回のように軸とする変数が複数ある場合、ベンチマーク統計における多次元のクロス表を入手することができなかつたり、できたとしても各セルの値が小さすぎて、手計算による Post-stratification が現実的でない。そこで、今回はレイキング (iterative proportional fitting “raking” algorithm) による Post-stratification のウェイトを作成する⁷。作業内容としては、図表 7 に示すとおりである。

図表 8 利用するベンチマーク統計一覧

変数	ベンチマーク統計	KHPS2004	KHPS2007	KHPS2012	JHPS2009
学歴	総務省「就業構造基本調査」	2002年	2007年	2012年	2007年
就業状態	総務省「国勢調査」	2005年	2005年	2010年	2010年
配偶状態(男女別)	総務省「国勢調査」	2005年	2005年	2010年	2010年
年齢階級	総務省「人口推計」	2003年10月	2006年10月	2011年10月	2008年10月
居住地域	KHPS/JHPS	KHPS2004	KHPS2007	KHPS2012	JHPS2009

ベンチマークとして用いる統計を図表 8 にまとめる。居住地域・年齢・性別については、サンプリングの際に、母集団と一致するよう設計されているが、他の変数を軸に Post-stratification を行うと、居住地域・年齢・性別の分布が影響を受けてしまうため、これらも考慮して Post-stratification を行った。なお、「国勢調査」は 5 年に 1 度（大規模調査は 10 年に 1 度）、「就業構造基本調査」も 5 年に 1 度であり、残念ながら KHPS/JHPS の各パネルのサンプル抽出時期と合わせるできない。「人口推計」は毎年の推計値が公表されているため、KHPS/JHPS のサンプル抽出時期（調査開始年の 1 月）にもっとも近い前年 10 月の値を用いることとした。

図表 9 Post-stratification により作成されたウェイトの分布

	KHPS2004	KHPS2007	KHPS2012	JHPS2009
1パーセンタイル値	0.295	0.205	0.220	0.362
25パーセンタイル値	0.645	0.553	0.576	0.701
50パーセンタイル値	0.796	0.795	0.830	0.859
75パーセンタイル値	1.225	1.222	1.144	1.143
99パーセンタイル値	3.195	3.813	4.231	3.097
平均値	1	1	1	1

Post-stratification による Cross-sectional ウェイトの結果を示すため、軸とした変数にかんするベンチマーク統計の分布とともに、各パネル(KHPS2004 サンプル, KHPS2007 追加サンプル, KHPS2012 追加サンプル, JHPS2009 サンプル)のウェイトなし集計値の分布と、Cross-sectional ウェイト付き集計値の分布を付表 1 に掲載している。Raking による Post-stratification により、すべての項目でベンチマーク統計の分布とほぼ一致する形になるようウェイトが作成されている。図表 9 には、サンプル抽出時点の Cross-sectional ウェ

⁷ Stata の ipfweight コマンドによる推計。

イトの分布にかんする集計を表す。

6. Longitudinal ウェイト

第 2 回目調査以降には、サンプル脱落によるバイアスを修正する Longitudinal ウェイトを作成する必要がある。ここでは、初回調査における Cross-sectional ウェイトをもとに各期のウェイトを作成していく⁸。

【例】 wave5 における Longitudinal ウェイトの作成

wave1 の Cross-sectional ウェイトをもとに、wave1 から wave5 まで回答継続しているか否かについて、wave1 時点の情報を用い回答継続確率モデルを推計する。推計された回答継続確率の逆数を用い、wave2 から wave5 までのサンプル脱落によるバイアスを修正する Longitudinal ウェイトを作成する。

$$W_{adjust, wave5} = W_{adjust, wave1} \frac{1}{P_{response, wave2 to wave5}}$$

- $P_{response, wave2 to wave5}$: wave1 から wave5 までの継続回答確率。

実際の分析方法としては、Logit モデル（回答継続=1，その他=0）を用いて、回答継続確率を推計し、その回答継続確率の逆数をサンプル脱落によるバイアスを修正する Longitudinal ウェイトとする。回答行動を説明する変数として用いた変数は、居住地域、世帯人数、性別、婚姻状況、持家状況、学歴、就業状況、年齢 10 歳階級、各年齢階級における後半 5 歳ダミー、世帯所得といった調査対象者の属性を表す変数である。さらに、正規対象者ダミーと、JHPS サンプルにおいては、面接調査ダミー、ウェブ回答ダミーといった調査方法にかんする変数も用いた。

図表 10 KHPS2004 パネルの Longitudinal weight にかんする分布

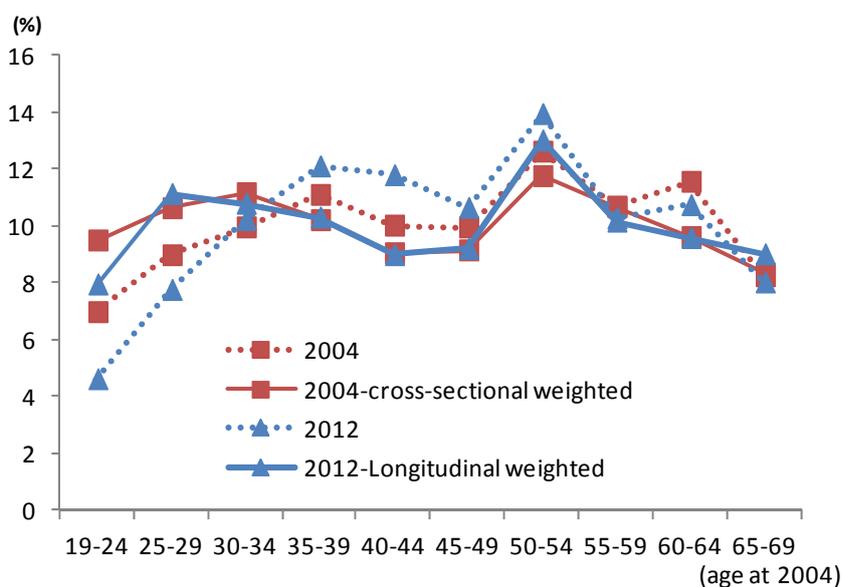
	KHPS2004 パネル								
	2005	2006	2007	2008	2009	2010	2011	2012	2013
1パーセンタイル値	0.34	0.39	0.42	0.44	0.46	0.49	0.51	0.54	0.58
25パーセンタイル値	0.75	0.82	0.89	0.94	1.00	1.05	1.10	1.15	1.22
50パーセンタイル値	0.94	1.06	1.16	1.25	1.32	1.40	1.46	1.56	1.66
75パーセンタイル値	1.49	1.71	1.86	2.00	2.14	2.26	2.38	2.50	2.68
99パーセンタイル値	4.25	5.09	5.71	6.40	7.04	7.38	7.78	8.52	8.62
平均値	1.21	1.38	1.52	1.63	1.75	1.85	1.95	2.06	2.19

⁸ Watson (2012)によると、Longitudinal ウェイトの作成方法は 2 つあるとされている。まず 1 つは、前期のウェイトと前期の情報をもとに今期のウェイトを作成する方法（回答確率モデル）、もう 1 つは、初回のウェイトと初回の情報をもとに今期のウェイトを作成する方法（回答継続確率モデル）である。前者の場合、最新の情報により無回答行動を説明することができるが、復活サンプルに対してウェイトを作成することができない。一方、後者の場合、無回答行動の説明にはならないが、推計に必要な情報をそろえやすく、初回データの分布に戻ることができる。カナダの家計パネル Survey of Labour and Income Dynamics (SLID)、オーストラリアの家計パネル Household, Income and Labour Dynamics (HILDA) においても、後者の方法が用いられており、KHPS/JHPS においてもこの方法を踏襲する。

回答継続確率の推計結果については、付表 2～付表 4 に示す。参考までに、KHPS2004 サンプルについて作成された Longitudinal weight の分布にかんする集計を図表 10 に表す。サンプル脱落を補うために、調査回を重ねるごとに全体的にウェイトの値が大きくなっていくことがわかる。

Longitudinal ウェイトの効果を確認するために、KHPS2004 サンプルにおける 2004 年時点の Cross-sectional ウェイトによる修正済み分布と、2012 年時点の Longitudinal ウェイトによる修正済み分布を比較する（図表 11）。図表から明らかなおとおり、Longitudinal ウェイトにより、2004 年時点の年齢分布を着実に再現するよう修正がなされていることがわかる。

図表 11 Longitudinal ウェイトによる分布の修正
(KHPS2004 サンプルの 2004 年と 2012 年の比較)



7. 複数パネルの統合ウェイト

同一主体を長期間追跡調査するパネル調査では、サンプル脱落の問題が不可避であり、これにより標本調査の根幹となるサンプルの代表性が損なわれてしまう。これに対処するため、多くのパネル調査ではサンプルの追加を行い、脱落により失われた部分を補完したり、さらには、若年層や移民サンプルを追加することにより年々老朽化するパネルの補正を行ったりしている。

2004 年から調査が開始され、今年で 10 年目を迎える KHPS も例外ではなく、2007 年と 2012 年にサンプルを追加している。また、質問項目や調査方法が類似している JHPS も KHPS の追加サンプルとみなすことで、サンプルを拡充することができる。そこで、ここでは複数のパネル(KHPS2004 サンプル(K04)、KHPS2007 新規サンプル(K07)、KHPS2012

新規サンプル(K12)、JHPS2009 サンプル(J09) の統合方法について検討する。

ここでは単純な方法として、各サンプルの母集団の年齢分布の違いに着目したうえで、統合の方法を検討する。例えば、2007年に追加された K07 を K04 と統合することを考える(図表 12)。K04、K07 ともにサンプリング時点で 20 歳から 69 歳を対象にしており、2007 年時点において、K04 の年齢幅は 23 歳から 73 歳、K07 の年齢幅は 20 歳から 69 歳である。2 つのサンプルがオーバーラップしているのは 23 歳から 69 歳の部分であり、この部分に比重を減らすウェイトを付すことで統合が可能となる。

図表 12 統合のイメージ (K04 と K07 を 2007 年時点で統合)



- K07 の 23~69 歳にウェイトをつける
- K04 の 23~69 歳にウェイトをつける

より複雑な例として、K04 と K07、J09、K12 の統合を考える。4 つのパネルが揃う 2012 年時点で、K04 の年齢幅は 28 歳から 77 歳、K07 の年齢幅は 25 歳から 74 歳、J09 の年齢幅は 23 歳から上限なし、K12 の年齢幅は 20 歳から 69 歳である。パネルがオーバーラップする部分は、全部で 5 か所あり、①すべてのパネルがオーバーラップする部分は 28 歳から 69 歳、②K07、J09、K12 のみがオーバーラップする部分は 25 歳から 27 歳、③J09 と K12 のみがオーバーラップする部分は 23 歳から 27 歳、④K04、K07、J09 のみがオーバーラップする部分は 70 歳から 74 歳、⑤K04 と J09 のみがオーバーラップする部分は 70 歳から 77 歳の部分であり、それぞれに比重を減らすウェイトを作成することで、4 つのパネルを統合することができる。

図表 13 統合のイメージ (K04、K07、K12、J09 を 2012 年時点で統合)



- すべてのパネルの 28~69 歳にウェイトをつける
- K07、J09、K12 の 25~27 歳にウェイトをつける
- J09 と K12 の 23~27 歳にウェイトをつける
- K04、K07、J09 の 70~74 歳にウェイトをつける
- K04 と J09 の 70~77 歳にウェイトをつける

それでは、どのようにパネル統合用のウェイトを作成するのか。ここでは 2 つの方法があ

るとされている (Watson, 2012)。1 つは、By combining the estimates と呼ばれる方法で、各々のパネルのサンプルサイズとサンプリングのデザイン効果 (design effect) からウェイトを作成し、計算されたウェイトをつけて各パネルを足し合わせることで、母集団を推計するというものである。もう 1 つは、By pooling the samples と呼ばれるもので、あらかじめ統合する予定の複数のパネルを 1 つのパネルとしてプーリングしたうえで、各回答者の抽出確率と回答継続確率によって、統合ウェイトを作成する方法である。こちらの場合、プーリングしたデータにこのウェイトをつけて集計することで、母集団を推計することができる。2 つの方法は以下のそれぞれの式で表すことができる。なお、本稿では【方法 1】を踏襲し、統合ウェイトを算出する。

【例】KHPS2004 サンプルと KHPS2007 サンプルを統合する。

(方法 1)

$$\hat{Y}_{combined} = \theta Y_{K04, wave4} + (1 - \theta) Y_{K07, wave1}$$

$$\theta = \frac{n_{K04}}{n_{K04} + n_{K07} \left(\frac{d_{K04}}{d_{K07}} \right)}$$

$$W_{combined} = \begin{cases} \theta \times W_{K04, adjust\ w6} : K04 \text{ でオーバーラップする部分の回答者に。} \\ (1 - \theta) \times W_{K07, design} : K07 \text{ でオーバーラップする部分の回答者に。} \end{cases}$$

- θ は統合ウェイトとなる Panel allocation factor
- n は統合時における各パネルのサンプルサイズ。

- K04 と K07 のサンプリングデザインは同一と仮定すると、design effect の比である $\left(\frac{d_{K04}}{d_{K07}} \right)$ は 1。⇒ θ はサンプルサイズ比で決定。

(方法 2)

$$W_{pooled, i} = \frac{1}{p_{i, K04} + p_{i, K07}}$$

$$W_{pooled, i} = \begin{cases} \frac{1}{p_{i, K04} + \hat{p}_{i, K07}} & \text{if } i \in S_{K04} \\ \frac{1}{\hat{p}_{i, K04} + p_{i, K07}} & \text{if } i \in S_{K07} \end{cases}$$

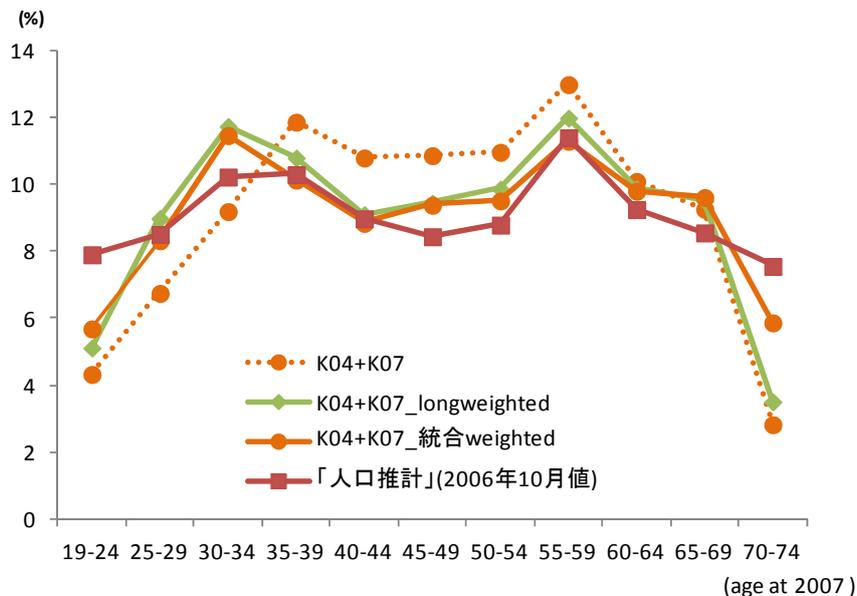
- p は probability of selection and response

- K04 回答者の実際の probability of selection and response と、K04 回答者が K07 サンプルにいた場合の推定 probability of selection and response.
- K07 回答者の実際の probability of selection and response と、K07 回答者が K04 サンプルにいた場合の推定 probability of selection and response.

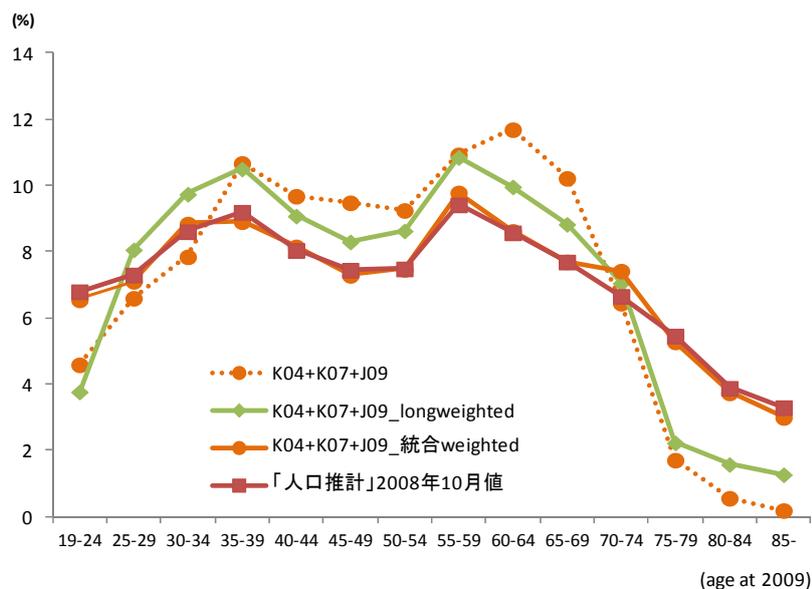
実際に統合した結果を図表 14～図表 16 に示す。K04 サンプルと K07 サンプルの統合の結果をみると (図表 14)、ウェイトをつけずに単純に 2 つのパネルを統合しただけでは、母集団分布から顕著なかい離がみられるが、統合ウェイトを用いて修正すると、母集団分

布に近づくことがわかる。ただし、2つのパネルの開始時点が3年しか空いていないため、Longitudinal ウェイトによる修正と統合ウェイトによる修正で分布に大差がないことがわかる。

図表 14 KHPS2004 サンプルと KHPS2007 サンプルの統合後の年齢分布



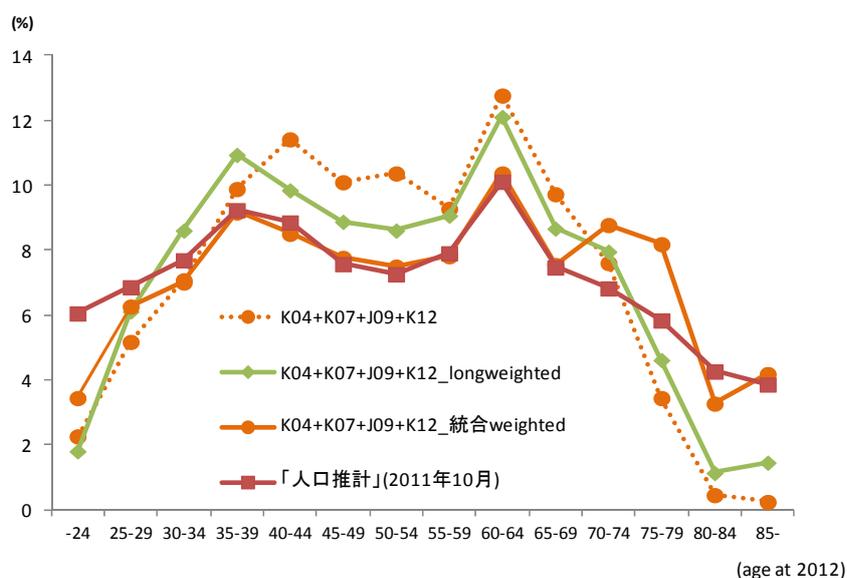
図表 15 KHPS2004 サンプルおよび KHPS2007 サンプル、JHPS2009 サンプルの統合後の年齢分布



次に、K04 サンプルと K07 サンプル、J09 サンプルの3つのパネルの統合の結果について

てみていく。JHPS ではサンプリングの際に年齢の上限を設定していないため、ウェイトなしで統合すると K04 サンプルと K07 サンプルには存在しない 70 代後半層の割合が少ない。また、20 代前半の層にかんしても、K04 サンプルと K07 サンプルでは少なくなっているため、この部分でも割合が少ない。しかしながら、統合ウェイトを用い、この部分のサンプルの比重を高めることで、母集団の分布に飛躍的に近づくことがわかる。

図表 16 KHPS2004 サンプルおよび KHPS2007 サンプル、JHPS2009 サンプル、KHPS2012 サンプルの統合後の年齢分布



最後に、4つのパネルを統合した結果についても確認しておく。先ほどの結果同様に、ウェイトなしの集計では若年層と後期高齢層で母集団よりも割合が低くなっている。年齢に上限を設定しない J09 サンプルが加わっていても、後期高齢層の割合は低くなっており、Longitudinal ウェイトによってもこの部分の乖離は修正されない。統合ウェイトを用いることで、母集団の分布に近づくことが確認できる。

8. 今後の課題

KHPS および JHPS のウェイトについては、定期的にその内容を更新・改良していく予定である。今回は初の試みということで、もっともベーシックな方法によるウェイトの作成を行った。今後さらに改良を加える必要があるが、特に統合ウェイトに関しては、年齢分布以外の項目にも着目して、ウェイトの割り当てを行うよう改良を加える必要がある。

日本を代表するパネルデータであるよう、調査票や調査方法における工夫によりサンプル脱落の軽減を試みるとともに、ウェイトの作成により補助的にデータの信頼性を高めていく。

付表 1 Raking による Post-stratification の結果

	KHPS2004			KHPS2007			KHPS2012			JHPS2009		
	K04 (1)	ベンチマーク (2)	K04 weighted (3)	K07 (1)	ベンチマーク (2)	K07 weighted (3)	K12 (1)	ベンチマーク (2)	K12 weighted (3)	J09 (1)	ベンチマーク (2)	J09 weighted (3)
年齢 20 ~ 24	6.6%	9.2%	9.2%	7.8%	8.6%	8.6%	8.1%	7.7%	7.7%	7.1%	6.8%	6.8%
25 ~ 29	8.9%	10.6%	10.6%	7.4%	9.4%	9.4%	8.8%	8.7%	8.7%	7.3%	7.3%	7.3%
30 ~ 34	10.0%	11.3%	11.3%	8.3%	11.4%	11.4%	10.4%	9.7%	9.7%	7.7%	8.6%	8.6%
35 ~ 39	10.8%	9.9%	9.9%	13.3%	10.9%	10.9%	11.4%	11.7%	11.7%	10.3%	9.2%	9.2%
40 ~ 44	10.1%	9.2%	9.2%	10.2%	9.4%	9.4%	11.9%	11.2%	11.2%	8.2%	8.1%	8.1%
45 ~ 49	10.0%	9.2%	9.2%	10.4%	9.1%	9.1%	8.7%	9.6%	9.6%	7.9%	7.5%	7.5%
50 ~ 54	12.5%	11.7%	11.7%	10.9%	9.9%	9.9%	10.9%	9.2%	9.2%	8.3%	7.5%	7.5%
55 ~ 59	10.7%	10.7%	10.7%	13.1%	12.7%	12.7%	8.1%	10.0%	10.0%	8.9%	9.4%	9.4%
60 ~ 64	11.7%	9.7%	9.7%	9.8%	9.6%	9.6%	12.9%	12.8%	12.8%	12.1%	8.6%	8.6%
65 ~ 69	8.8%	8.6%	8.6%	8.7%	9.0%	9.0%	8.5%	9.5%	9.5%	10.4%	7.7%	7.7%
70 ~ 74	-	-	-	-	-	-	-	-	-	6.9%	6.7%	6.7%
75 ~ 79	-	-	-	-	-	-	-	-	-	3.4%	5.5%	5.5%
80 ~ 84	-	-	-	-	-	-	-	-	-	1.1%	3.9%	3.9%
85 ~ 89	-	-	-	-	-	-	-	-	-	0.3%	2.1%	2.1%
90 ~ 94	-	-	-	-	-	-	-	-	-	0.1%	0.9%	0.9%
95 ~ 99	-	-	-	-	-	-	-	-	-	0.0%	0.3%	0.0%
100歳以上	-	-	-	-	-	-	-	-	-	0.0%	0.0%	0.0%
学歴												
中学卒	13.6%	23.8%	23.8%	8.8%	19.1%	19.1%	7.4%	17.4%	17.4%	12.0%	19.6%	19.1%
高校卒(含む専門)	52.5%	43.6%	43.6%	50.4%	51.1%	51.1%	49.8%	50.5%	50.5%	50.4%	52.5%	51.1%
短大・高専卒	11.8%	13.7%	13.7%	15.7%	8.2%	8.2%	14.5%	8.4%	8.4%	11.9%	8.5%	8.3%
大学・院卒	19.4%	16.2%	16.2%	22.2%	18.9%	18.9%	24.4%	21.1%	21.1%	22.8%	19.4%	18.9%
通学中	2.7%	2.7%	2.7%	3.1%	2.7%	2.7%	4.0%	2.6%	2.6%	3.0%	2.8%	2.7%
就業状態												
主に仕事	59.0%	50.5%	50.5%	58.3%	50.5%	50.5%	60.5%	50.1%	50.1%	53.8%	50.1%	50.1%
通学のかたわらに仕事	1.6%	0.6%	0.6%	1.8%	0.6%	0.6%	2.6%	0.7%	0.7%	1.9%	0.7%	0.7%
家事などのかたわらに仕事	11.4%	8.5%	8.5%	14.7%	8.5%	8.5%	10.9%	8.0%	8.0%	12.2%	8.0%	8.0%
仕事を休んでいた	1.6%	1.0%	1.0%	1.3%	1.0%	1.0%	1.4%	1.1%	1.1%	1.6%	1.1%	1.1%
仕事を探していた	3.0%	3.7%	3.7%	2.6%	3.7%	3.7%	3.4%	4.0%	4.0%	2.5%	4.0%	4.0%
通学・家事	23.4%	35.7%	35.7%	21.4%	35.7%	35.7%	21.3%	36.1%	36.1%	27.9%	36.1%	36.1%
男女別配偶状態												
有配偶男性	36.8%	31.0%	31.0%	33.6%	31.0%	31.0%	33.8%	30.6%	30.6%	35.3%	30.6%	30.6%
有配偶女性	36.1%	33.0%	33.0%	39.9%	33.0%	33.0%	35.4%	30.6%	30.6%	36.6%	30.6%	30.6%
無配偶男性	13.2%	18.8%	18.8%	12.1%	18.8%	18.8%	16.5%	17.5%	17.5%	13.2%	17.5%	17.5%
無配偶女性	14.0%	17.1%	17.1%	14.5%	17.1%	17.1%	14.3%	21.4%	21.4%	14.9%	21.4%	21.4%
居住地												
北海道	4.5%	4.5%	4.5%	4.4%	4.4%	4.4%	4.5%	4.5%	4.5%	4.4%	4.4%	4.4%
東北	7.4%	7.4%	7.4%	7.2%	7.2%	7.2%	7.4%	7.4%	7.4%	7.5%	7.5%	7.5%
関東	33.4%	33.4%	33.4%	33.4%	33.4%	33.4%	34.4%	34.4%	34.4%	32.8%	32.8%	32.8%
東海	18.1%	18.1%	18.1%	19.0%	19.0%	19.0%	17.9%	17.9%	17.9%	18.5%	18.5%	18.5%
近畿	16.5%	16.5%	16.5%	16.3%	16.3%	16.3%	16.4%	16.4%	16.4%	16.2%	16.2%	16.2%
中国	5.8%	5.8%	5.8%	5.8%	5.8%	5.8%	5.4%	5.4%	5.4%	6.0%	6.0%	6.0%
四国	3.1%	3.1%	3.1%	3.0%	3.0%	3.0%	3.0%	3.0%	3.0%	3.3%	3.3%	3.3%
九州	11.1%	11.1%	11.1%	10.9%	10.9%	10.9%	11.1%	11.1%	11.1%	11.4%	11.4%	11.4%

註:

- (1): ウェイトなしで集計した各データの集計値。
- (2): ベンチマーク統計(国勢調査、人口推計、就業構造基本調査など)の値。
- (3): Cross-sectionalウェイトを掛けた各データの集計値。

付表 2 回答継続確率モデル 推計結果 (KHPS2004 パネル)

	KHPS 2004								
	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011	2011-2012	2012-2013
Region									
Hokkaido	-0.117 [0.207]	-0.0377 [0.180]	0.0887 [0.176]	0.146 [0.171]	0.143 [0.169]	0.208 [0.168]	0.137 [0.167]	0.197 [0.166]	0.138 [0.165]
Tohoku	0.17 [0.179]	-0.0659 [0.145]	-0.127 [0.138]	-0.145 [0.136]	-0.0565 [0.136]	-0.0135 [0.136]	-0.0371 [0.135]	0.0731 [0.135]	0.0561 [0.135]
Chubu	-0.154 [0.121]	-0.239** [0.103]	-0.230** [0.0985]	-0.213** [0.0968]	-0.237** [0.0957]	-0.202** [0.0955]	-0.183* [0.0955]	-0.164* [0.0958]	-0.179* [0.0963]
Kinki	0.227* [0.135]	0.356*** [0.115]	0.386*** [0.108]	0.420*** [0.104]	0.397*** [0.101]	0.370*** [0.0993]	0.312*** [0.0978]	0.358*** [0.0971]	0.344*** [0.0968]
Chugoku	0.0825 [0.195]	-0.0683 [0.161]	-0.164 [0.152]	-0.209 [0.146]	-0.252* [0.145]	-0.208 [0.146]	-0.181 [0.147]	-0.19 [0.148]	-0.117 [0.148]
Shikoku	0.394 [0.281]	0.266 [0.229]	0.269 [0.212]	0.247 [0.203]	0.238 [0.199]	0.0727 [0.197]	-0.0295 [0.197]	-0.0271 [0.196]	-0.0112 [0.196]
Kyusyu	0.0435 [0.146]	0.0574 [0.125]	-0.0416 [0.117]	-0.0861 [0.114]	-0.0673 [0.113]	-0.125 [0.112]	-0.119 [0.113]	-0.118 [0.113]	-0.0934 [0.113]
Number of household members									
2 people	0.331* [0.186]	0.395** [0.162]	0.392** [0.156]	0.353** [0.153]	0.375** [0.153]	0.399*** [0.154]	0.326** [0.155]	0.273* [0.155]	0.216 [0.156]
3 people	0.410** [0.185]	0.452*** [0.160]	0.390** [0.154]	0.353** [0.151]	0.481*** [0.151]	0.484*** [0.151]	0.458*** [0.152]	0.359** [0.152]	0.356** [0.153]
4 people	0.510*** [0.196]	0.601*** [0.168]	0.507*** [0.161]	0.418*** [0.158]	0.512*** [0.157]	0.504*** [0.157]	0.430*** [0.157]	0.347** [0.158]	0.286* [0.158]
More than 5 people	0.424** [0.203]	0.542*** [0.176]	0.564*** [0.170]	0.488*** [0.166]	0.481*** [0.165]	0.484*** [0.165]	0.428*** [0.165]	0.364** [0.166]	0.323* [0.167]
Male	-0.0625 [0.0917]	-0.0587 [0.0782]	-0.132* [0.0744]	-0.158** [0.0726]	-0.167** [0.0716]	-0.180** [0.0712]	-0.151** [0.0711]	-0.182** [0.0711]	-0.146** [0.0712]
Original subject	-0.168 [0.107]	-0.134 [0.0912]	-0.159* [0.0866]	-0.0986 [0.0847]	-0.103 [0.0838]	-0.118 [0.0834]	-0.111 [0.0831]	-0.111 [0.0833]	-0.113 [0.0835]
Married	0.183 [0.118]	0.194* [0.102]	0.193** [0.0971]	0.198** [0.0949]	0.185** [0.0944]	0.183* [0.0941]	0.228** [0.0945]	0.209** [0.0950]	0.164* [0.0955]
Owned house	-0.171 [0.111]	-0.287*** [0.0946]	-0.181** [0.0879]	-0.174** [0.0854]	-0.201** [0.0845]	-0.257*** [0.0840]	-0.256*** [0.0836]	-0.236*** [0.0831]	-0.221*** [0.0832]
Education									
High School	0.161 [0.138]	0.176 [0.119]	0.318*** [0.114]	0.329*** [0.111]	0.406*** [0.111]	0.381*** [0.111]	0.397*** [0.111]	0.355*** [0.112]	0.321*** [0.113]
Junior college or Higher professional school	0.565*** [0.192]	0.449*** [0.159]	0.481*** [0.150]	0.476*** [0.146]	0.590*** [0.145]	0.512*** [0.144]	0.553*** [0.144]	0.512*** [0.145]	0.460*** [0.145]
University or Graduate school	0.345** [0.163]	0.315** [0.141]	0.458*** [0.134]	0.453*** [0.131]	0.510*** [0.131]	0.433*** [0.130]	0.435*** [0.130]	0.405*** [0.131]	0.383*** [0.132]
Other or Nonresponse	0.332 [0.243]	0.329 [0.206]	0.359* [0.192]	0.376** [0.188]	0.429** [0.187]	0.293 [0.185]	0.315* [0.185]	0.338* [0.185]	0.376** [0.184]
Worker	-0.0829 [0.103]	-0.118 [0.0885]	-0.0853 [0.0839]	-0.0705 [0.0817]	-0.148* [0.0809]	-0.129 [0.0805]	-0.138* [0.0802]	-0.157* [0.0805]	-0.118 [0.0807]
Age									
30-39	0.121 [0.144]	0.257** [0.124]	0.256** [0.118]	0.345*** [0.115]	0.347*** [0.114]	0.456*** [0.114]	0.415*** [0.114]	0.449*** [0.115]	0.449*** [0.116]
40-49	0.440*** [0.156]	0.444*** [0.131]	0.407*** [0.124]	0.522*** [0.121]	0.535*** [0.120]	0.615*** [0.120]	0.633*** [0.120]	0.614*** [0.121]	0.660*** [0.121]
50-59	0.198 [0.151]	0.332** [0.131]	0.406*** [0.126]	0.544*** [0.124]	0.605*** [0.123]	0.643*** [0.122]	0.583*** [0.123]	0.550*** [0.123]	0.526*** [0.124]
60-69	0.392** [0.174]	0.449*** [0.148]	0.349** [0.141]	0.447*** [0.137]	0.520*** [0.136]	0.527*** [0.136]	0.538*** [0.136]	0.515*** [0.136]	0.460*** [0.137]
Latter half of 10 age group	-0.0262 [0.0859]	0.01 [0.0726]	-0.0461 [0.0689]	-0.018 [0.0672]	-0.0431 [0.0662]	-0.0176 [0.0658]	0.00413 [0.0656]	-0.00174 [0.0656]	-0.0167 [0.0657]
Household Income (ten thousand yen)									
360-559	0.161 [0.126]	0.148 [0.107]	0.0804 [0.101]	0.0263 [0.0972]	0.001 [0.0961]	0.0544 [0.0952]	0.0535 [0.0946]	0.0316 [0.0944]	-0.0173 [0.0944]
560-849	0.16 [0.147]	0.0121 [0.122]	-0.0481 [0.115]	-0.0654 [0.112]	-0.153 [0.111]	-0.0835 [0.110]	-0.0822 [0.110]	-0.0953 [0.110]	-0.172 [0.110]
More than 850	0.328** [0.136]	0.190* [0.113]	0.0685 [0.106]	0.0685 [0.103]	-0.0799 [0.101]	-0.0627 [0.0996]	-0.0531 [0.0993]	-0.0888 [0.0992]	-0.115 [0.0992]
0 or Nonresponse	-0.327** [0.136]	-0.313*** [0.120]	-0.304*** [0.116]	-0.247** [0.115]	-0.319*** [0.115]	-0.291** [0.115]	-0.328*** [0.116]	-0.395*** [0.117]	-0.384*** [0.118]
Constant	0.712*** [0.237]	0.073 [0.203]	-0.248 [0.195]	-0.525*** [0.191]	-0.726*** [0.190]	-0.871*** [0.189]	-0.972*** [0.191]	-0.953*** [0.191]	-0.965*** [0.192]
Number of observations	4,005	4,005	4,005	4,005	4,005	4,005	4,005	4,005	4,005
Log Likelihood	-1797	-2311	-2506	-2607	-2661	-2686	-2698	-2699	-2694

Robust standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

付表 3 回答継続確率モデル 推計結果 (KHPS2007 パネル)

	KHPS 2007					
	2007-2008	2008-2009	2009-2010	2010-2011	2011-2012	2012-2013
Region						
Hokkaido	0.386 [0.442]	0.143 [0.354]	0.218 [0.336]	0.157 [0.313]	0.281 [0.311]	0.419 [0.307]
Tohoku	0.349 [0.344]	0.122 [0.273]	0.147 [0.259]	-0.0431 [0.245]	0.275 [0.251]	0.168 [0.240]
Chubu	0.306 [0.243]	0.18 [0.198]	-0.017 [0.179]	-0.0769 [0.170]	-0.106 [0.165]	-0.0161 [0.163]
Kinki	0.538** [0.266]	0.398* [0.213]	0.301 [0.195]	0.302 [0.185]	0.401** [0.180]	0.490*** [0.175]
Chugoku	0.207 [0.383]	0.0417 [0.306]	-0.0785 [0.284]	0.0186 [0.277]	-0.157 [0.261]	0.0802 [0.262]
Shikoku	-0.128 [0.440]	-0.0832 [0.385]	-0.219 [0.363]	-0.273 [0.346]	-0.424 [0.337]	-0.378 [0.335]
Kyusyu	0.127 [0.275]	0.27 [0.245]	0.211 [0.223]	0.205 [0.209]	0.227 [0.202]	0.355* [0.199]
Number of household members						
2 people	-0.08 [0.359]	0.129 [0.304]	0.2 [0.285]	0.0976 [0.279]	0.214 [0.274]	0.0385 [0.274]
3 people	0.235 [0.358]	0.316 [0.299]	0.138 [0.279]	-0.0648 [0.274]	0.224 [0.271]	0.112 [0.272]
4 people	0.347 [0.367]	0.303 [0.305]	0.406 [0.286]	0.0216 [0.279]	0.218 [0.274]	0.0481 [0.273]
More than 5 people	0.245 [0.378]	0.402 [0.316]	0.359 [0.295]	-0.0624 [0.289]	0.147 [0.285]	-0.0231 [0.285]
Male	-0.0253 [0.175]	0.134 [0.147]	0.175 [0.135]	0.0796 [0.127]	0.04 [0.125]	0.123 [0.123]
Original subject	-0.0566 [0.214]	-0.255 [0.175]	-0.333** [0.158]	-0.194 [0.152]	-0.272* [0.149]	-0.21 [0.148]
Married	-0.0101 [0.231]	0.0859 [0.195]	0.152 [0.178]	0.301* [0.172]	0.234 [0.172]	0.340** [0.170]
Owned house	0.275 [0.209]	0.145 [0.170]	0.132 [0.160]	0.188 [0.154]	0.0653 [0.150]	0.0538 [0.148]
Education						
High School	-0.0276 [0.339]	-0.329 [0.295]	-0.152 [0.263]	0.0392 [0.247]	0.219 [0.236]	0.184 [0.235]
Junior college or Higher professional school	-0.222 [0.380]	-0.42 [0.332]	-0.149 [0.302]	0.0489 [0.284]	0.233 [0.276]	0.253 [0.275]
University or Graduate school	0.237 [0.375]	-0.0108 [0.328]	0.00113 [0.292]	0.146 [0.274]	0.366 [0.265]	0.266 [0.263]
Other or Nonresponse	-0.146 [0.473]	-0.418 [0.394]	-0.24 [0.362]	-0.209 [0.342]	0.0912 [0.337]	0.0351 [0.335]
Worker	-0.0161 [0.204]	-0.0164 [0.168]	0.0881 [0.157]	0.0248 [0.150]	-0.0938 [0.146]	-0.0595 [0.143]
Age						
30-39	0.434 [0.293]	0.338 [0.230]	0.492** [0.213]	0.612*** [0.206]	0.766*** [0.204]	0.711*** [0.201]
40-49	0.359 [0.300]	0.666*** [0.258]	0.664*** [0.230]	0.694*** [0.221]	0.847*** [0.218]	0.748*** [0.216]
50-59	0.0896 [0.271]	0.176 [0.233]	0.337 [0.218]	0.284 [0.211]	0.540** [0.210]	0.484** [0.210]
60-69	0.5 [0.340]	0.411 [0.280]	0.474* [0.257]	0.337 [0.250]	0.597** [0.247]	0.487** [0.246]
Latter half of 10 age group	0.19 [0.165]	0.178 [0.138]	0.112 [0.126]	-0.079 [0.120]	-0.0804 [0.118]	-0.166 [0.116]
Household Income (ten thousand yen)						
360-559	0.314 [0.266]	0.265 [0.217]	0.237 [0.200]	0.315 [0.193]	0.346* [0.190]	0.296 [0.186]
560-849	0.654** [0.285]	0.241 [0.222]	0.315 [0.209]	0.303 [0.199]	0.2 [0.193]	0.126 [0.190]
More than 850	0.0495 [0.260]	0.14 [0.225]	0.088 [0.212]	-0.0545 [0.202]	-0.147 [0.199]	-0.173 [0.199]
0 or Nonresponse	-0.508* [0.264]	-0.317 [0.228]	-0.298 [0.215]	-0.183 [0.210]	-0.281 [0.207]	-0.462** [0.205]
Constant	0.922* [0.496]	0.545 [0.419]	0.03 [0.386]	-0.111 [0.370]	-0.611* [0.359]	-0.614* [0.360]
Number of observations	1,419	1,419	1,419	1,419	1,419	1,419
Log Likelihood	-518	-694.4	-786.1	-850.9	-880.3	-905.6

Robust standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

付表 4 回答継続確率モデル 推計結果 (JHPS2009 パネルおよび KHPS2012 パネル)

	JHPS 2009				KHPS2012
	2009-2010	2010-2011	2011-2012	2012-2013	2012-2013
Region					
Hokkaido	-0.00163 [0.233]	0.133 [0.203]	0.0731 [0.177]	0.139 [0.172]	0.0216 [0.486]
Tohoku	0.151 [0.193]	-0.218 [0.153]	-0.122 [0.140]	-0.368*** [0.135]	-0.261 [0.351]
Chubu	0.0441 [0.139]	0.0625 [0.117]	0.11 [0.104]	0.0358 [0.0996]	-0.187 [0.275]
Kinki	0.0513 [0.144]	0.0876 [0.121]	0.173 [0.109]	0.0842 [0.103]	0.102 [0.294]
Chugoku	0.0684 [0.218]	0.297 [0.192]	0.340** [0.167]	0.291* [0.157]	1.057* [0.609]
Shikoku	0.16 [0.286]	0.134 [0.241]	0.203 [0.215]	0.129 [0.204]	0.014 [0.577]
Kyusyu	-0.183 [0.154]	-0.106 [0.132]	-0.00895 [0.120]	0.0133 [0.116]	-0.292 [0.304]
Number of household members					
2 people	0.148 [0.201]	0.0604 [0.177]	-0.0409 [0.163]	0.06 [0.156]	-0.0856 [0.405]
3 people	0.331 [0.206]	0.0974 [0.178]	0.017 [0.164]	0.122 [0.158]	0.261 [0.423]
4 people	0.27 [0.216]	0.0865 [0.188]	-0.0179 [0.173]	0.124 [0.166]	-0.145 [0.421]
More than 5 people	0.321 [0.224]	0.0527 [0.195]	-0.0823 [0.179]	-0.0909 [0.172]	0.648 [0.508]
Male	0.099 [0.0996]	0.0104 [0.0842]	-0.0112 [0.0760]	0.00986 [0.0730]	0.329 [0.208]
Original subject	-0.0488 [0.0993]	-0.0442 [0.0840]	-0.132* [0.0750]	-0.132* [0.0719]	-0.379* [0.195]
Married	0.0564 [0.133]	0.17 [0.113]	0.133 [0.103]	0.138 [0.0993]	0.102 [0.259]
Owned house	0.0634 [0.120]	0.145 [0.103]	0.159* [0.0933]	0.162* [0.0901]	0.328 [0.238]
Education					
High School	0.466*** [0.155]	0.389*** [0.133]	0.407*** [0.122]	0.430*** [0.120]	0.525 [0.372]
Junior college or Higher professional school	0.527*** [0.204]	0.423** [0.172]	0.388** [0.156]	0.480*** [0.152]	0.904** [0.451]
University or Graduate school	0.470*** [0.180]	0.535*** [0.154]	0.483*** [0.140]	0.504*** [0.136]	0.749* [0.411]
Other or Nonresponse	0.585** [0.239]	0.476** [0.203]	0.492*** [0.186]	0.524*** [0.178]	0.509 [0.466]
Worker	-0.00138 [0.109]	-0.0972 [0.0936]	-0.0831 [0.0849]	-0.106 [0.0820]	-0.0969 [0.241]
Age					
30-39	0.461*** [0.163]	0.636*** [0.142]	0.665*** [0.130]	0.666*** [0.124]	-0.102 [0.325]
40-49	0.742*** [0.179]	0.663*** [0.152]	0.638*** [0.138]	0.629*** [0.132]	-0.25 [0.325]
50-59	0.476*** [0.171]	0.355** [0.146]	0.342** [0.133]	0.390*** [0.130]	0.226 [0.355]
60-69	0.745*** [0.179]	0.686*** [0.152]	0.493*** [0.136]	0.532*** [0.132]	-0.14 [0.372]
More than 70	0.660*** [0.213]	0.374** [0.177]	0.266 [0.163]	0.248 [0.158]	
Latter half of 10 age group	0.193** [0.0960]	0.155* [0.0809]	0.126* [0.0722]	0.0313 [0.0689]	-0.256 [0.187]
Household Income (ten thousand yen)					
360-559	-0.0867 [0.150]	-0.136 [0.126]	-0.226** [0.113]	-0.213** [0.108]	-0.0336 [0.311]
560-849	0.0155 [0.160]	0.0333 [0.135]	0.00249 [0.120]	0.00403 [0.115]	0.158 [0.329]
More than 850	0.101 [0.168]	-0.0249 [0.139]	-0.0703 [0.124]	-0.0937 [0.119]	-0.266 [0.334]
0 or Nonresponse	-0.790*** [0.149]	-0.723*** [0.130]	-0.707*** [0.121]	-0.667*** [0.119]	-0.474 [0.319]
Interview Survey	-0.145 [0.0929]	-0.0864 [0.0784]	-0.0757 [0.0704]	-0.130* [0.0674]	
Web Survey	0.0787 [0.312]	-0.153 [0.250]	-0.157 [0.231]	-0.237 [0.224]	
Constant	0.667*** [0.256]	0.298 [0.226]	0.0369 [0.208]	-0.266 [0.201]	1.115** [0.531]
Number of observations	4,022	4,022	4,022	4,022	1,012
Log Likelihood	-1554	-2016	-2375	-2535	-395.8

Robust standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

参考文献

European Commission (2011), 'Description of target variables: Cross-sectional and Longitudinal', EU-SILC 065 (2011 operation).

LaRoche, S. (2003), 'Longitudinal and cross-sectional weighting of the Survey of Labour and Income Dynamics', Income Research Paper Series, Cat.No. 75F0002MIE, No.007, Statistics Canada.

Watson, N. (2012), 'Longitudinal and cross-sectional weighting methodology for the HILDA Survey', *HILDA Project Technical Paper Series*, No.2/12, Melbourne Institute of Applied Economic and Social Research.

石黒格 (2008) 『Stata による社会調査データの分析——入門から応用まで』, 北大路書房.

木村正一 (2005) 「2004 年慶應義塾大学家計パネル調査の標本特性」, 『日本の家計行動のダイナミズム I』慶應義塾大学出版会.

木村正一 (2006) 「KHPS における回答率の変化とその影響」, 『日本の家計行動のダイナミズム II』慶應義塾大学出版会.

直井道生 (2008) 「KHPS2007 新規対象サンプルの標本特性」, 『日本の家計行動のダイナミズム IV』慶應義塾大学出版会.

直井道生 (2009) 「項目回答率とパネル調査回答継続」, 『日本の家計行動のダイナミズム V』慶應義塾大学出版会.

直井道生・山本耕資 (2010) 「日本家計パネル調査の標本設計と代表性」, 『貧困のダイナミズム——日本の税社会保障・雇用政策と家計行動』慶應義塾大学出版会.

直井道生・山本耕資・宮内環 (2010) 「JHPS 調査票回収状況および回答状況における調査実施方法のパフォーマンス」, 『貧困のダイナミズム——日本の税社会保障・雇用政策と家計行動』慶應義塾大学出版会.

直井道生・山本耕資 (2010) 「パネル調査における回答継続と調査実施方法—JHPS2010 を用いた検証」, 『教育・健康と貧困のダイナミズム——所得格差に与える税社会保障制度の効果』慶應義塾大学出版会.

宮内環・McKenzie.C.R.・木村正一 (2005) 「回答行動の分析—調査受諾と拒否の選択行動—」, 『日本の家計行動のダイナミズム I』慶應義塾大学出版会.

宮内環・McKenzie.C.R.・木村正一 (2006) 「パネルデータ継続と回答行動の分析」, 『日本の家計行動のダイナミズム II』慶應義塾大学出版会.

宮内環・McKenzie.C.R.・直井道生 (2008) 「パネルデータ継続調査標本と新規調査標本との比較」, 『日本の家計行動のダイナミズム IV』慶應義塾大学出版会.