

2015 年 SSM 調査による調査不能と項目無回答の基礎分析 *¹

三輪 哲
(東京大学)

前田忠彦
(統計数理研究所)

【論文要旨】

本稿の目的は、2015 年 SSM 調査における調査不能と項目無回答にかんして、基本傾向を整理することである。結果を要約すると、第一に、回収率はだいたい 5 割程度であった。第二に、欠票理由の分布は、前回 2005 年 SSM 日本調査のそれと比べ、概ね変わらなかった。第三に、2005 年に比して、今回はデータの基本属性の分布にかんする偏りが小さくなった。第四に、人口的属性や住居形態など個人特性は接触不能と関連し、地域特性は拒否と関連していた。そして第五に、10 年前よりも、項目無回答率はほとんどの基本質問項目について下がった。

キーワード：調査不能、接触不能率、拒否率、項目無回答

1. はじめに

本稿の目的は、2015 年 SSM 調査における調査不能と項目無回答にかんして、基本傾向を整理することである。社会調査データ分析に基づく学術的研究はもちろん意義あるものだが、その価値は用いたデータの質と量に依存する部分も大きい。ならば、2015 年 SSM 調査にかんしてもまた、社会階層と移動研究の実質的議論を展開する以前に、データそのものの基本的な特徴を把握することが欠かせない。すなわち、回収と無回答の割合と、それにみられるパターンを検討しておくことが重要といえる。

そこで以下では、まず 2015 年 SSM 調査における回収票と欠票理由の分布を確認する。そして母集団からの乖離を検討した後に、調査不能の要因を探索的に分析する。さらに、項目ごとにみられる無回答に注目し、その割合の時点間比較と属性別パターンの検討をおこなう。

2. 2015 年 SSM 調査の回収票と欠票理由

2.1 有効票と無効票

回収率や欠票理由の検討に先立ち、まずは 2015 年 SSM 調査での有効票と無効票について定義をしておこう。ただし、ここで示す内容や用語法は、新しいものでは決してなく、前回 2005 年 SSM 日本調査での三隅・三輪 (2008) の考え方を踏襲したものである。

2015 年 SSM 調査でも、前回同様に、面接票と留置票の 2 種類の方法を組み合わせる形で、実査をおこなった。面接票が回収できない場合は、無効票とした。留置票が回収できずとも、面接票だけ回収できたものは、有効票に含めた。両方回収できれば、もちろん有効票となる。

さらに、有効抽出と無効抽出を区別しておく。ここでいう有効抽出とは、それが有効票で

¹ 本研究は、JSPS 科研費 JP25000001 の助成を受けたものです。

あっても無効票であっても、回収率の算定に含まれるものである。一方、無効抽出とは、回収率の算定からは除外されるものである。具体的には、後に示す欠票理由のうち、死亡、転居、住所不明、長期不在（入所・入院）、対象不適格、障害・認知症、を無効抽出とした。長期不在（それ以外）、短期不在、拒否、その他、を有効抽出とした。回収できた有効票も、もちろんここでいう有効抽出に含まれる。今回の2015年SSM調査においては、正規対象が無効抽出に該当したときには、予備対象へとアタックした*²。なお、調査設計や状況にかんしての詳細は、本報告書所収の白波瀬（2018）を参照されたい。

2.2 回収票数と回収率の算定

以上述べた考え方にしたがって、回収票数と回収率を算定した結果は、次の図1のとおりである。全体の回収率は50.09%である*³。仮に正規対象のみに絞って算定しても49.96%であって、予備対象を用いたことで回収率が上がったとはいえない。面接票と留置票の両方の完了というように条件を厳しく設定しても、回収率はそれほど変わらず、48.97%となる。

全体回収状況

総数 16,100	有効票 7,817	有効抽出 7,817	留置欠票 176	有効票としての全体回収率 =7,817÷(7,817+7,788)×100 = 50.09% (面接票回収率はこれと同じ) 参考) 面接と留置両方の完了に限った回収率 =(7,817-176)÷(7,817+7,788)×100 = 48.97%
	無効票 8,283	有効抽出 7,788	留置完了 0	
		無効抽出 495	留置完了 1	

対象別内訳

正規対象	総数 16,100	有効票 6,910	有効抽出 6,910	留置欠票 155	有効票としての正規対象回収率 =6,910÷(6,910+6,920)×100 = 49.96% 参考) 面接と留置両方の完了に限った回収率 =(6,910-155)÷(6,910+6,920)×100 = 48.84%
		無効票 9,190	有効抽出 6,920	留置完了 0	
			無効抽出 2,270		
予備対象	総数 4,025	有効票 907	有効抽出 907	留置欠票 21	有効票としての予備対象回収率 =907÷(907+868)×100 = 51.10% 参考) 面接と留置両方の完了に限った回収率 =(907-21)÷(907+868)×100 = 49.92%
		無効票 1,133	有効抽出 868	留置完了 9	
			無効抽出 265		
	未使用 1,985				

図1 回収状況の全体像

² ただし、抽出数が足りなかった場合には予備対象へのアタックができないので、予備対象のアタック数は無効抽出数とは完全には一致しない。

³ 無効抽出の定義を2005年SSM日本調査と揃えた場合、正規対象回収率は48.76%となる。

2.2 欠票理由の内訳

これまでは、欠票にかんしては、ただ「無効票」としてのみ扱い、その中身を問わなかった。もちろん実際には、中身は多様である。今回の2015年SSM調査において、欠票の中身、すなわち理由がいかなるものであったか、そしてそれは前回の2005年SSM調査の結果とどの程度類似しているのか、これより検討していく。なおここで、三隅・三輪（2008）で示された2005年SSM調査の欠票理由の「転出」は今回では「転居」に、また「入院・入所」は「長期不在（入所、入院）」に、「一時不在」は「短期不在」に、「抽出ミス」は「対象不適格」に、さらに「病気・障害」は「障害・認知症」へと置き換わっている。ただし、高齢者を念頭に置いて設定された「障害・認知症」を除き、それらカテゴリーの中身にはほとんど違いはないので、比較可能なものとみなして結果の解釈をしていく。

表1 欠票理由の分布

	2005SSM	2015SSM		
		正規対象のみ	予備対象のみ	全アタック
死亡	0.1%	0.3%	0.3%	0.3%
転居	6.4%	10.4%	9.0%	10.2%
住所不明	1.2%	1.1%	1.3%	1.1%
長期不在（入所、入院）	0.4%	1.2%	1.0%	1.2%
長期不在（それ以外）	1.2%	0.4%	0.5%	0.4%
短期不在	9.5%	3.1%	7.2%	3.6%
拒否	39.4%	39.2%	34.6%	38.5%
対象不適格	0.2%	0.1%	0.2%	0.1%
障害・認知症 [†]	1.2%	0.9%	1.1%	0.9%
その他	0.0%	0.2%	0.3%	0.2%

注：[†]は、2005年では「病気・障害」というカテゴリで、内容がやや異なる。

欠票理由の分布は、上の表1の通りである*⁴。2015年SSM調査データについては、正規対象、予備対象それぞれの集計結果と、全体での集計結果をともに示した。2015年SSM調査では、正規対象と比べると、予備対象では短期不在の割合が高く、拒否がやや低い傾向にある。ただし正規が大部分を占めるため、正規のみの結果と全アタックの結果は、概ね類似している。

2005年と2015年を比べると、大きくみると、ある程度似た結果が得られたということができよう。たとえば、欠票理由のうち最頻のものは「拒否」であって、およそ4割弱である。

⁴ 列合計で100%に足しあがらないのは、欠票の数を分母にするのではなく、全アタック数を分母にして割合を求めたからである。

それは、2005 年も 2015 年もほぼ同じ水準にある。そこから大きく離れるが、次に相対的に頻度が高めの理由は、「転居」と「短期不在」である。そのほかの理由は 1 パーセントポイント程度かそれ未満に過ぎず、少数派というべきであろう。

ただ細かくみると、いくつか重要な違いがあるのも確かである。10 年前に比べ、転居が増加する一方で、短期不在と長期不在が減少した。これは、2015 年 SSM 調査においては、追加調査を 2015 年晩秋まで断続的におこなっていったことがあるだろう。事実上、長期間アタックし続けることによって、当初は短期不在だと思われていた調査対象者が実は転居であったと判明することもあるからだ。さらにまた、標本抽出から実査に至るまでの時間が従来以上に長くなったこと、その期間に年度をまたいだことも、転居の増加を招いた調査設計上の問題点だったと思われる。それから他に増加している理由は、死亡および入院・入所である。これは、2015 年 SSM 調査で初めて 70 代の年代の人たちを調査対象に含めたことが背景にあるのだろう。あくまで局所的だが、調査のやり方や対象次第で調査不能の内容の分布が変わっていることが注目される。

3. 標本抽出と調査不能による母集団との乖離

実際に分析可能なのは有効票からなるデータセットであるが、そこに至るまでに、どのように、そしてどの程度、偏りが生じるのか。それを検討するために、計画標本と母集団^{*5}との基本属性の分布のずれと、有効票の集合である回収標本と計画標本とのずれを、確認したい。

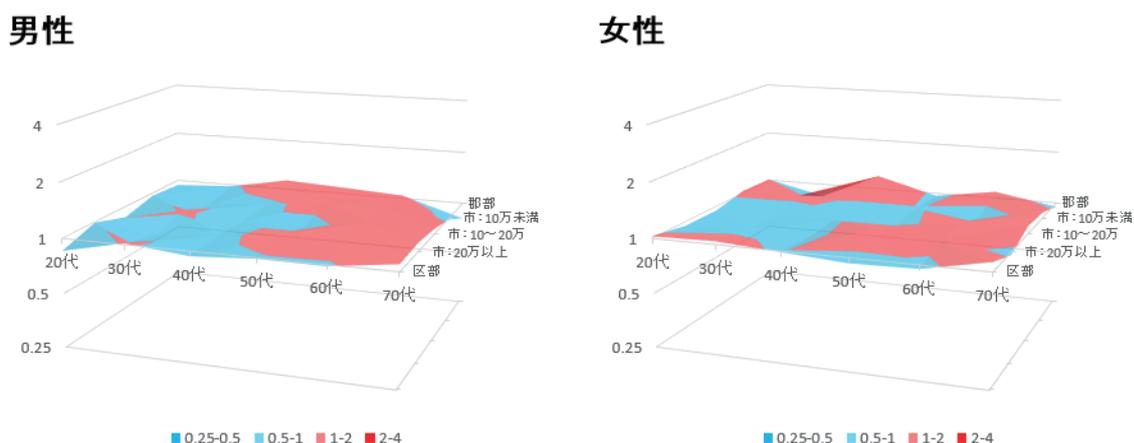


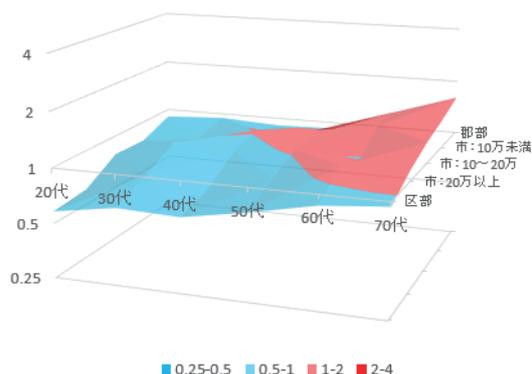
図2 2015年SSM調査における計画標本と母集団との属性分布の乖離

図2は、無作為抽出された標本が、母集団からいかに乖離しているかを示すものである。

⁵ ここで母集団の属性分布としたものは、平成27年1月1日住民基本台帳年齢階級別人口(市区町村別)(日本人住民)から算出した相対度数である。

市郡規模・年代・性別を組み合わせたうえで、計画標本における当該属性の比率を母集団のそれで除した値を縦軸にプロットした。0.9 倍から 1.1 倍の範囲に 8 割以上の値が含まれ、だいたい 1 付近の値をとっている。最小の値でも 0.811（郡部・70 代・男性）、最大は 1.199（郡部・40 代・女性）である。標本抽出による基本属性の偏りは、決して大きくはない。

男性



女性

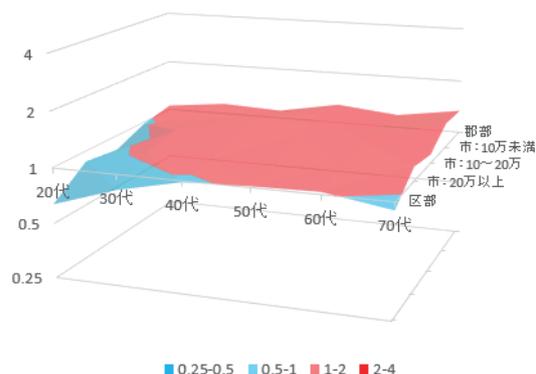
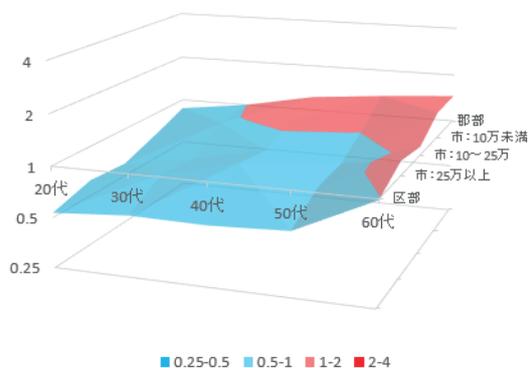


図3 2015年SSM調査における回収標本と計画標本との属性分布の乖離

男性



女性

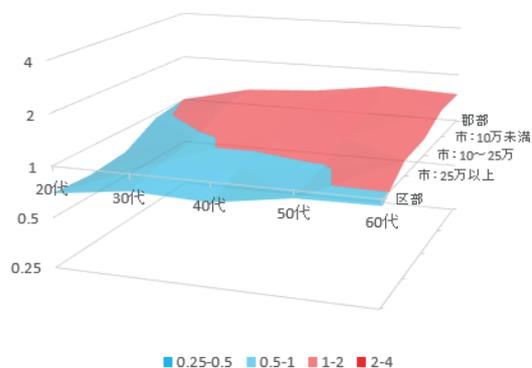


図4 2005年SSM日本調査における回収標本と計画標本との属性分布の乖離

それが、回収標本となると、やはり計画標本からの乖離が大きくなる。図3では、全60セルのうちおよそ17%が0.8倍以下と過小になっている。その反面、約22%が1.2倍以上と過大になっている。回収率に、性差、年代差、地域差があるゆえ、予想通りではある。

ただし2005年の結果を示す図4と比べると、まだ2015年SSM調査では回収標本の偏りは比較的小さい。同様に、0.8倍以下を過小、1.2倍以上を過大として、2005年SSM日本調査での計画標本の偏りをみると、過小のセルも過大のセルも、それぞれ30%ずつ存在する。よって、2015年でも調査不能による属性の偏りは避けられなかったが、若干改善されたといえ

る。

4. 調査不能の要因分析

ここでは、調査不能の要因を検討する。調査不能に至る過程を、第1段階の接触不能⁶、その後にくる第2段階の拒否⁷という2段階に分けて検討した保田（2007）に倣い、同様のモデルを適用する。拒否率の分析が、接触できた条件下での拒否という、条件付き確率となっている点に留意されたい。このようにすることによって、調査不能に与える諸要因が、どの段階で影響するかを峻別でき、単に調査不能の予測分析をするよりも深い洞察が得られる。

表2 接触不能と拒否を予測するロジットモデル推計結果

	接触不能			拒否		
	Coef.	Std. Err.	z	Coef.	Std. Err.	z
市郡規模: ref 町村						
区部	.184	.091	2.02 *	.308	.067	4.58 **
20万以上市	-.150	.094	-1.60	.221	.066	3.35 **
10-20万市	-.064	.098	-0.65	.050	.069	0.72
10万未満市	.116	.092	1.26	-.074	.066	-1.12
地域ブロック: ref 北海道・東北						
関東	.030	.078	0.38	.217	.057	3.77 **
中部	-.039	.086	-0.45	-.105	.062	-1.68
近畿	-.075	.088	-0.85	.218	.064	3.39 **
中四国	.103	.098	1.04	-.160	.073	-2.19 *
九州・沖縄	.048	.093	0.52	-.043	.070	-0.62
性別: ref 男性						
女性	-.332	.044	-7.54 **	-.198	.033	-5.94 **
年代: ref 20-34歳						
35-49歳	-.724	.056	-12.93 **	.048	.052	0.93
50-64歳	-1.162	.064	-18.16 **	-.013	.052	-0.26
65歳以上	-1.098	.064	-17.21 **	-.152	.052	-2.91 **
住居形態: ref 集合住宅						
一戸建て	-.554	.053	-10.47 **	-.229	.043	-5.27 **
オートロック	.282	.072	3.92 **	-.016	.070	-0.22
その他	1.282	.182	7.06 **	-.172	.235	-0.73
調査時期: ref 第I期						
第II期	.037	.054	0.69	-.097	.040	-2.42 *
第III期	.225	.053	4.24 **	.041	.041	1.00
定数項	-.692	.116	-5.98 **	.009	.091	0.10

注: * $p < .05$ ** $p < .01$ (両側検定)

⁶ 死亡、転居、住所不明、長期不在、短期不在、障害・認知症を接触不能に、回収完了、拒否を接触できたものとしてまとめて、接触不能か否かを判別する分析をおこなった。

⁷ 注5で述べた接触できた者だけに絞って、拒否か否かを判別する分析をおこなった。

4.1 接触不能の要因分析

まずは、表2の左側に示した接触不能の要因から検討したい。調査不能にかんして、地域による差異がみられることは、既にみた通りである。だが、調査不能までの過程を分けて、接触不能の段階に限ってしてみると、思いのほか地域差は大きくはない。町村部と政令指定都市（区部）のあいだでこそ5%水準で統計的有意差がみられるが、それを除くと、都市規模についても地域ブロックについても顕著な差はみられなかった。

他方で、個人属性による差異は明瞭にあらわれている。女性のほうが接触しやすい傾向が観察される。そして、年代の影響はいっそう大きく、20-34歳の若年層よりも35-49歳の中年層のほうが接触しやすく、そのうえの高年層になるとさらに接触しやすくなっていたようだ。

住居形態もまた、接触不能と強く関連することが見いだされた。よく知られる、一戸建て居住者は集合住宅居住者よりも接触しやすいとする結果は、ここでも観察された。オートロック集合住宅の居住者が、相対的に接触しにくかったという結果も、よくみられる結果であろう。

また、調査時期については、第Ⅲ期のみが、他よりも接触不能になりやすかったことがうかがえる。この理由は、追加調査に用いることのできた時間でいくらか説明できよう。2015年SSM調査では、2015年の1-3月（第Ⅰ期）、4-5月（第Ⅱ期）、6-8月（第Ⅲ期）に分けて実査がなされたが、その期間中に接触できなかった対象者には、その後晩秋まで追加調査を随時おこなっていった。そうすると最終的に追加調査を打ち切るまでの時間が異なるゆえ、第Ⅲ期の対象者が、もっとも接触するチャンスが低くなったものと考えられる。

4.2 拒否の要因分析

次に、接触できた条件のもとでの拒否の起きやすさについて検討しよう。先にみた接触不能の分析結果とは異なり、地域による影響が色濃く出ている。対象者が居住しているのが、都市規模がより大きな地域ほど、また関東や近畿といった大都市圏であると、面接調査に対して拒否しやすい傾向があることがわかった。回収率の地域差の存在は調査現場で広く知られた事実であるが、それは接触した後に調査を応諾するか拒否するかの段階において顕現することが明らかとなった。

個人属性は拒否の段階でも重要な予測要因足りうるが、その影響の程度は、接触不能の段階に比して、いくらか小さなものであるようだ。女性のほうが拒否されにくいのは結果から確かそうといえる。しかしながら、その係数推定値の絶対値をみると、接触不能の段階のそれと比して明らかに小さい。さらにまた、年代による違いも、65歳以上の高齢層のみ他よりも拒否率が低い傾向こそあるが、差異はそれくらいで違いの程度もそれほど大きくない。

住居形態については、一戸建て居住者のみ拒否されにくい傾向がみられた。一戸建て居住者は接触不能にもなりにくいわけであるので、二つの段階を重ねて、いっそう回答票が得や

すい層となっているようだ。

最後に、調査時期の影響である。他の要因を統制した後でも、第Ⅱ期だけ拒否が生じにくいことが推定結果からみてとれる。

4.3 調査員の観察されない異質性を考慮した追加分析

興味深いことに、調査不能にかんして、調査時期の影響がみられた。しかもそれは、第Ⅱ期に拒否率が下がり、第Ⅲ期に接触不能率が上がるという、時期と段階にかんして異なる様相をみせていた。そこで、これが調査員の交替によるものなのか、あるいは調査員特性を統制しても残るものなのかなど、もう少し追究してみたい。

調査員のなかには、第Ⅰ期のみ実査にかかわった方もいれば、3期すべてにかかわった方も、他のパターンの方もいた。つまり、調査員が複数回担当していたわけである。それを活かして、調査員と調査時期の組み合わせをケースとしてデータを再構成して、調査不能率を分析するためのパネルデータを作成した。そうした調査員パネルデータを用い、接触不能率と拒否率を従属変数とした、固定効果モデルを推定した。その結果が次の表3である。

表3 接触不能率と拒否率を予測する固定効果モデル推計結果

	接触不能率			拒否率		
	Coef.	Std. Err.	t	Coef.	Std. Err.	t
地点共変量						
女性率	-.173	.060	-2.89 **	-.061	.074	-0.83
高齢率	-.151	.048	-3.15 **	-.033	.058	-0.58
一戸建て率	-.029	.025	-1.19	-.025	.030	-0.83
都市規模平均	.003	.006	0.55	-.012	.008	-1.58
調査時期: ref 第Ⅰ期						
第Ⅱ期	.002	.011	0.18	-.029	.013	-2.23 *
第Ⅲ期	.022	.011	2.01 *	-.010	.013	-0.74
定数項	.279	.036	7.84 **	.571	.044	13.10 **

注: * $p < .05$ ** $p < .01$ (両側検定)

調査員レベルのパネルデータとして分析しても、調査時期と調査不能との関連については、似たような結果が得られた。すなわち、第Ⅱ期には拒否率が低くなり、第Ⅲ期には接触不能率が高くなるようだ。固定効果モデルによる推定で、これらが再確認されたことの意義は大きい。なぜなら、調査員レベルでは観察されない異質性を統制しているのので、たとえば調査員の属性とか能力などとは別に、時期による調査不能の差異をみることができているからである。また、調査員の入れ替わりの影響も、この結果からは排除されている。よって、同程度の調査員が同じような条件の調査地点に赴いたとしても、第Ⅱ期だけは拒否率を低くする

ことができたとみられる。春ゆえの温暖な季節性や、調査や個人情報をめぐる逆風的な社会事象の打撃を受けなかったなど、さまざまな社会レベルの外的条件が良好だったとみることができるともかもしれない。第Ⅲ期の接触不能率の高まりについては、やはり追加調査打ち切りまでに許された時間の少なさと、標本抽出された時期からもっとも時間的に離れたために対象者名簿と実態とのあいだで乖離が大きくなったことによるダメージだと考えられる。

5. 項目無回答の傾向

ここからは、項目すなわち各質問の無回答にかんして分析をおこなう。これまでは調査不能すなわち一票がまるごと無回答となる場合を扱ってきた。一方で、調査票は回収できたのだけど、質問によっては回答してくれないようなこともしばしばある。こちらは、項目無回答 (item nonresponse) と呼ばれる。そのような項目ごとに発生する無回答がどのような項目で、どれくらい生じて、それがどのような層で多く生じるのかも、調査データのクオリティを判断するうえで重要な情報となる。それゆえ、ここでは、SSM 調査における基本的な項目に限定して、項目無回答を検討していく。

5.1 項目無回答の分布比較

扱う基本的な項目としては、(1)現職：従業上の地位、(2)現職：仕事内容、(3)初職：仕事内容*⁸、(4)最終学歴、(5)父学歴、(6)父主職：従業上の地位、(7)父主職：仕事内容、(8)母学歴、(9)婚姻状況、(10)配偶者学歴、(11)15歳時キョウダイ数、(12)住居形態、(13)生活満足感、(14)主観的地位 (10段階階層帰属意識)、(15)個人年収、(16)配偶者年収、(17)世帯年収、の17項目とする。これらは、SSM プロジェクトで伝統的に重要視されてきた、階層測定、社会移動、地位達成、同類婚、教育機会、階層意識、社会経済的地位、といったテーマの中核的変数であって、1995年から2015年までのSSM調査ではすべて面接調査票に含められてもいる。

⁸ 初職の従業上の地位を対象としなかったのは、初職を聞く際の質問紙構成が2015年SSM調査では変更され、非該当の扱いが変わったゆえである。結果、初職：従業上の地位の無回答の分布は大きく変わり、それ以前との比較が困難になった。

表 4 各年調査における項目無回答率

	1995SSM	2005SSM	2015SSM
現職: 従業上の地位	0.00% (5357)	0.09% (5742)	0.13% (7817)
現職: 仕事内容	0.54% (3914)	0.14% (4139)	0.08% (5120)
初職: 仕事内容	1.49% (5165)	0.36% (5566)	0.12% (7608)
最終学歴	0.17% (5357)	0.14% (5739)	0.01% (7812)
父学歴	18.29% (5357)	20.48% (5742)	16.78% (7817)
父主職: 従業上の地位	3.64% (5357)	5.43% (5742)	5.60% (7817)
父主職: 仕事内容	9.60% (5353)	9.09% (5555)	9.18% (7561)
母学歴	15.12% (5357)	19.91% (5742)	14.66% (7817)
婚姻状況	0.07% (5357)	0.26% (5742)	0.03% (7817)
配偶者学歴	1.53% (4444)	1.23% (4315)	0.43% (5609)
15歳時キョウダイ人数	0.11% (5357)	0.28% (5742)	0.27% (7817)
住居	0.15% (5357)	0.44% (5742)	0.18% (7817)
生活満足感	0.65% (5357)	0.52% (5742)	0.18% (7817)
主観的地位	3.92% (2704)	3.41% (5742)	1.85% (7817)
個人年収	6.91% (5357)	13.41% (5742)	6.58% (7817)
配偶者年収	13.61% (4263)	25.71% (4314)	17.42% (5609)
世帯年収	21.32% (5357)	36.36% (5742)	28.55% (7817)

注: かつこ内は項目無回答率算定の基数.

上の表 4 より理解されるように、項目ごとに無回答率*⁹は大きく異なる。2015 年 SSM 調

⁹ 項目無回答を広義にとらえれば非該当も含まれるが、ここでは答えられるはずなのに回答が得られなかった、無回答・拒否・不明・未記入を、項目無回答として狭義に扱っている。なお、まれに職業コーディング等で起きる「分類不能」も、ここに含めている。

査における項目無回答率を降順に列挙すると、世帯年収（28.6%）、配偶者年収（17.4%）、父学歴（16.8%）、母学歴（14.7%）、父主職：仕事内容（9.2%）、個人年収（6.6%）などが並ぶ。これらが上位6位までを占めるのは、実際、これら3回のSSM調査に共通してみられることである。学歴や収入は、プライバシーに大いに触れる項目であって、「答えたくない」人が多数出るのは納得できよう。だが、同じ収入にかかわる項目の中でも、個人収入と配偶者、世帯収入とではだいぶ無回答率が異なることにも注意が必要である。収入を答えたくないなら、個人収入も世帯収入もすべてが無回答になるはずだが、そうはなっていない。つまり、収入を答えたくない人ももちろん一定程度いるが、それ以上に、そもそも配偶者や世帯全体の収入はわからないから答えられない人が多そうだということではないだろうか。この点、学歴や仕事内容も同じことが言え、本人のそれらならばわかるから答えるが、親の過去の仕事や学歴などはわからないゆえに答えられないのだろう。

それから、ここでひとつ気づく重要な傾向は、2015年SSM調査データでは、10年前の2005年調査データよりも、全体的に項目無回答率が下がっていることである。父主職だけは例外となるけれども、上で挙げた無回答率の高い項目のうちの5つについては、3～8ポイント程度無回答率が低下して、概ね20年前の調査の水準に近いところまで下がっている。また、無回答率の低い項目においても、主観的地位や、配偶者学歴、初職仕事内容など、いくつか無回答率の減少トレンドが観察されるものもある。このたびの2015年SSM調査データは、比較的、項目無回答率が低めに抑制されたデータだということができる。

5.2 項目無回答の属性別分布

続いて、項目無回答率が属性別にどのような値を示すかを記述する。結果は、以下の図5および図6の通りである。無回答率のあまりにも低い項目については、属性間での無回答率に散らばりが無いのはほぼ自明であるので、ここでの分析から割愛した。

図5および図6は、地域ブロック、性別、年代の3変数によってブレイクダウンした項目無回答率を、2005年SSM調査データと2015年とで比較したものである。まず図4は、収入にかかわる3項目にかんする分析結果である。地域ブロックについてはいくらか変動はあるものの明確な系統性は見出しがたい。やや2005年データで、個人・配偶者・世帯収入ともに北海道・東北、九州・沖縄において無回答率が低い傾向があるようにも読めるくらいである。性別による差異は明瞭で、世帯年収と配偶者年収の無回答は女性のほうがおしなべて高い。そして年代による収入の無回答パターンは、基本的には緩やかな右上がりの傾向があるのと同時に、世帯年収だけは20代について突出して無回答率が高いことが特徴的である。20代で親と同居の場合、親の収入がよくわからない人が多いということだろう。

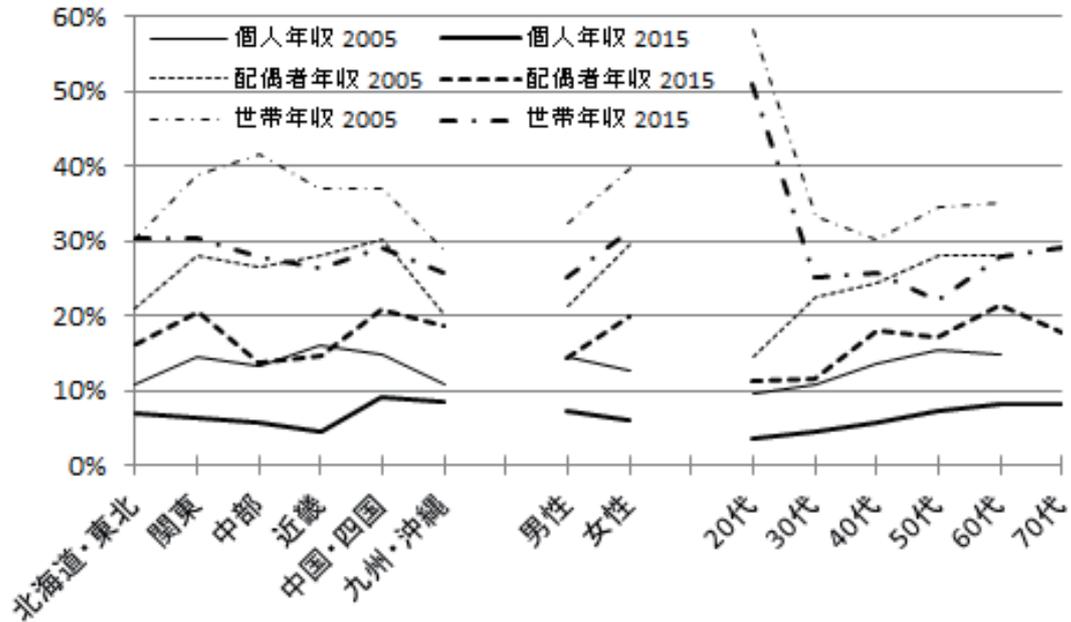


図5 収入3項目の項目無回答率（2005年SSM、2015年SSM）

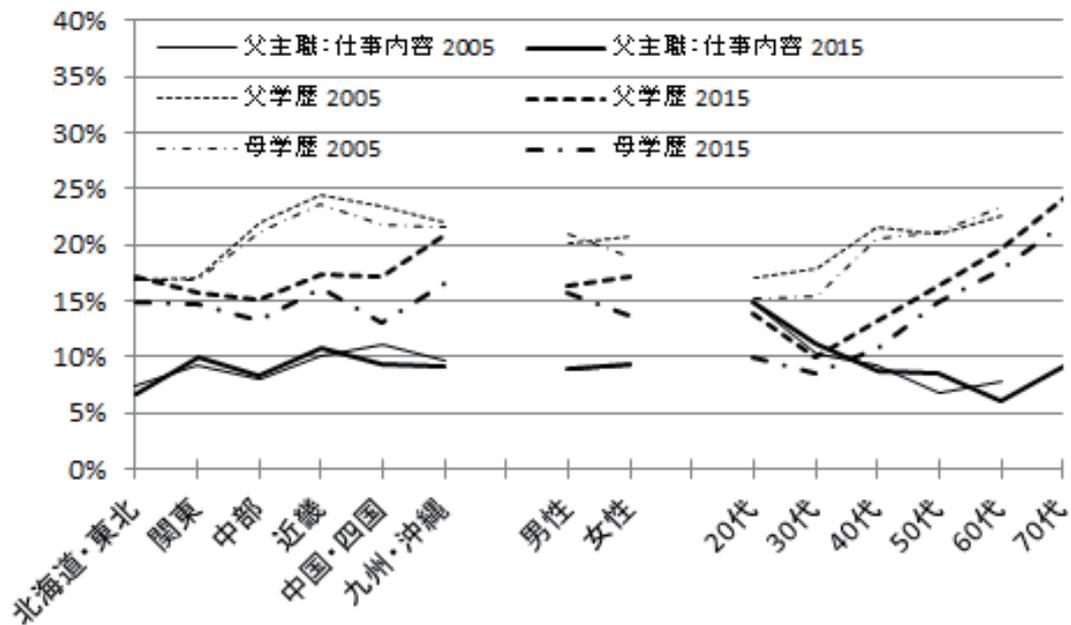


図6 親属性3項目の項目無回答率（2005年SSM、2015年SSM）

図6は、親の属性にかかわる無回答の多い3項目として、父主職：仕事内容、父学歴、母学歴の3つをとりあげて、先ほどと同様に図示したものである。地域差はここでもはっきりとはしていない。2005年データだと親学歴について「西高東低」の傾向がみられたが、2015年データでは地域差は若干平準化された。性別にかんしては、特に違いはみられない。年代

については、興味深いパターンがみられた。父職だと右下がり、すなわち年代が高くなるほど無回答率が低くなる。だが、父母の学歴については反対に右上がり、年代が高くなるほど無回答率が高くなっている。被雇用者が増え職住分離が進んでいったがゆえに、若い層のほうが親の働く姿を見る機会が減少したことが、前者の背景要因になっているのではないかと推察する。後者は、高齢層の親となると旧制学歴諸学校の出身である蓋然性が高く、学歴を回答できなかつたり、思い出せないのかもしれない。

以上のように、性別、年代において、社会的意味のありうる項目無回答パターンが観察された。これらから得られる一般的な含意は、仮に回収されたサンプル全体では性別や年代の分布が歪んでいなかったとしても、使用する項目の無回答パターンによっては、分析に含めた瞬間にデータの偏りが発生するということだ。分析者のおこないたいデータ解析に先立って、項目の無回答率や無回答パターンを把握しておく慎重さが求められるところである。

6. さいごに

本稿では、2015年SSM調査の調査不能と項目無回答にかんして、基礎分析をおこなった。その結果を要約すると、明らかとなったことは、以下述べる諸点にまとめられる。第一に、回収率はだいたい5割程度であった。それは正規対象だけに限っても、留置票まで含めても、大きく変わるものではなかった。第二に、欠票理由の分布は、前回2005年SSM日本調査のそれと比べ、概ね変わらなかった。ただし、部分的には、転居の割合が増え、短期不在が減るという変化はみられた。第三に、2005年SSM日本調査に比して、今回の2015年SSM調査ではデータの基本属性の分布にかんする偏りが小さくなった。第四に、人口的属性や住居形態など個人特性は接触不能と関連し、地域特性は拒否と関連していた。そして第五に、10年前よりも、項目無回答率はほとんどの基本質問項目について下がった。

社会調査の困難さが語られるようになって久しいが（玉野 2003）、今回の2015年SSM調査でも、それは変わらずに直面したというのが正直なところである。とはいえ、10年前と対比的に比較すると、回収率についても、項目無回答率についても、やや改善したとみることができる。どう悲観的にみたとしても、今回収集されたデータが量と質の両面において、従来継続されてきたものよりもひどく劣っていると非難を浴びるものでは決してないだろう。

本稿は、調査プロジェクト全体で共有すべき、調査不能と項目無回答にかかわる基本的な情報をまとめるものであった。だがそれに留まるものあり、調査設計や実査における戦略へのさらなる改善の示唆には至らなかった。もしそれを望むなら、調査員の属性や行動記録、当日の気象条件など、詳細なデータまで用いて解析に挑む必要があるだろう。ただ調査環境をなげくばかりではなく、欠票を含むデータの分析結果を証拠として、実践的なヒントを探し出す努力も、着実に積み上げていくべきである。

[文献]

- 三隅一人・三輪哲. 2008. 「2005年SSM日本調査の欠票・回収状況の分析」三輪哲・小林大祐編『2005年SSM調査シリーズ1 2005年SSM日本調査の基礎分析』2005年SSM調査研究会: 17-29.
- 白波瀬佐和子. 2018. 「2015年「社会階層と社会移動に関する全国調査(SSM調査)」実施の概要」保田時男編『2015年SSM調査報告書1 調査方法・概要』2015年SSM調査研究会: 1-12.
- 玉野和志. 2003. 「サーベイ調査の困難と社会学の課題」『社会学評論』53(4): 537-551.
- 保田時男. 2007. 「低下する回収率と回収不能の要因」谷岡一郎ほか編『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.

