

社会参加としての社会調査

—社会調査回答バイアスが与える変数間の関連性への影響の評価*¹—

高田 洋
(札幌学院大学)

【論文要旨】

本研究は、社会調査に応ずることのバイアスが、様々な変数間、特に政治的態度や社会意識に関する変数間の関連性に影響するかどうかを評価する。社会調査に応ずるか否かは、調査対象者の志向や態度との関連が認められる可能性がある。回収されたデータには、調査に応ずるか否かと関連して、特定のバイアスが生じているかもしれない。そのバイアスを評価する。ただし、回収されなかった対象者についての当該変数への回答については知ることができないので、直接的にその偏りを見積もることはできない。そこで本研究では、欠票情報についてのデータに基づき、そこで知ることができる情報により回収の有無を回帰させ欠票ハザード比を推定させる。そこで得られた推定値を回収されたデータにおけるコントロール変数として分析モデルに加え、その結果を、加える前のモデルと比較し、どのように関連性が変動したのかを分析することにより、間接的にバイアスを評価する。結果として、変数間の関係に影響を及ぼす非回収のバイアスは、社会参加はしないが異議申し立てをしにくい消極的忌避と、社会参加はしていないが因習的な志向により調査を拒否する積極的忌避の2つから生ずることを見出した。

キーワード：非回収バイアス、社会参加、社会調査

1. 社会調査に応ずること

本研究においては、「社会調査に応ずること」の意味を考えてみたい。近年の回収率の低下を課題とした欠票分析の研究においては、非回収の対象者の属性的特性や個人情報保護法との関連あるいは調査法上の工夫などが議論されてきた（篠木 2010; 吉川徹 2010; 小野寺・片山他 2010; 松岡亮二・前田忠彦 2015）。そこに本稿では、社会調査という行為が人々にどのような意味を持ってとらえられているかを問題に加えたい。というのも、調査に繰り返し携わる経験を通して、その非回答の傾向に関して、社会調査に対する誤解や偏見があるように感じるからである。その中には、学術や科学に対する懐疑とか研究機関への不信のようなものも含まれる。そして、このことは、社会の状況や変動とも関わり合いがある。学術の権威や信頼の失墜、自由な言論についての社会的雰囲気、流行、政治的風潮、景気、治安などが反映されている。

¹ 本研究は、JSPS 科学研究費特別推進研究事業（課題番号 25000001）「少子高齢化からみる階層構造の変容と格差生成メカニズムに関する総合的研究」に伴う成果の一つであり、本データ使用にあたっては 2015 年 SSM 調査データ管理委員会の許可を得た。また分析に関しては、JSPS 科学研究費助成事業（課題番号 15H03424）「パラデータを活用した訪問調査法の精度管理と不能バイアス補正に関する研究」に伴う成果の一つである。

研究として行う社会調査は、任意性に基づく回答でなくてはならない。そのため、自発的協力が根本である。したがって社会調査を回答するには回答者の協力の意志がある。これはある種の社会参加といえる。そのような社会参加を表象するものが、社会調査に応えることであるとすれば、そのような意志に関わりがある変数や分析について、回収されたデータにおいて偏りがある可能性がある。社会調査に回答することは、社会参加する意志とのかかわりを持っていると考えられるので、(1) 学術的研究に対する態度や偏見、(2) 政治的意向、(3) 公的なものに対する態度や意識が含まれるだろう。これは社会意識といわれる変数に含まれているものである。

社会参加としての側面が社会調査にあるとすれば、回収されたデータは、学術的な社会調査研究への公益性を認め、貢献しようとする協力的な意志が反映されている。とすれば、その逆の回答しないという傾向に対して、次のような二つの仮説を考えることができる。一つは、積極的な忌避ともいうべき、科学研究への偏った見方が生み出す拒否や無理解に基づく無回答である。研究に応ずることはむしろ都合のよい解釈をされるから協力してはいけないというようなことや、大学などは信用できないから協力してはいけないというような回答忌避態度である。このような態度は、権威主義な態度や、合理性に関わる知識との関わりがあるだろう。

もう一つは、消極的な忌避ともいうべき、自主規制に基づく無回答である。このような研究に応ずる資格がないと自ら規定してしまう場合や、反体制的意見とみなされてしまうから回答したくないという態度である。これは社会参加や政治についての無力感や、異議申し立ての自由と関わりがある。現に投票率が低下しているが、その背景には政治参加に対する無力感がある。このような態度は、現在の体制的な意向に反する意見を持つほど回答したくないという傾向を促すだろう。このような二つの仮説が考えられる。

本研究では、社会調査に応ずることのバイアスが、変数間の関連性にどのような影響するかどうかを評価する。しかし、回収されなかった対象者の質問項目の選択については知ることができない。そこで、欠票情報についてのデータに基づき、それで知ることができる情報（年齢、地域、住居形態など）で回収の有無を推定し、その傾向を回収されたデータのコントロール変数として利用することにより間接的に評価し、それによって偏りを推測する。

2. 分析

2.1 欠票ハザードの推定

データは、2015年SSM調査第3次配布データ（2017年2月27日版バージョン070）を用いた。また、欠票傾向を推測するために「2015年SSM調査欠票データ」を利用する。欠票データとは、性別、年齢、地点のサンプリングの情報と、調査訪問時に得られる住居形態の情報（一戸建て、共同住宅、オートロック共同住宅、その他の4カテゴリ）が含まれ、回収

の有無が記載されているものである。サンプル数は設計された 16,100 から無効抽出 495 を除いた 15,605 になる。

表 1 回収の有無のプロビットモデル

独立変数	係数		s. e.	t
女性ダミー	.091 ***		.019	4.70
地方(参照カテゴリ:関東)				
北海道	-.022		.051	-.42
東北	.166 ***		.042	3.97
中部	.158 ***		.030	5.23
近畿	-.030		.029	-1.03
中国	.221 ***		.044	4.97
四国	.137 **		.060	2.26
九州	.132 ***		.035	3.76
都市規模(参照カテゴリ:政令指定都市)				
20 万以上	.038		.028	1.35
10-20 万	.127 ***		.031	4.11
10 万未満	.195 ***		.031	6.37
町村	.167 ***		.041	4.13
年齢(参照カテゴリ:70 代)				
20 代	-.391 ***		.035	-11.10
30 代	-.348 ***		.030	-11.54
40 代	-.381 ***		.029	-13.30
50 代	-.326 ***		.030	-10.77
60 代	-.254 ***		.029	-8.82
一戸建てダミー	.150 ***		.023	6.64
<i>Chi square</i>	482.051 ***			
<i>d. f.</i>	17			
<i>n</i>	15605			
<i>McFadden Pseudo R²</i>	.022			

注)有意水準: + p<.10, * p<.05, ** p<.01, *** p<.001

これらの欠票情報に基づき、回収の有無（有=1、無=0 のダミー変数）を従属変数とし、性別、年齢、地域、住居形態を独立変数としたプロビットモデルを推測する。ややモデルのフィットはよくないが、その結果は表 1 のとおりである。女性の方が回収しやすい。関東に比して、東北、中部、中国、四国、九州が回収しやすく、政令指定都市に比して、都市規模が 20 万未満だと回収しやすい。つまり、都市が回収しにくい。年齢は概ね若いほど回収しにくく、一戸建ては回収しやすい。この分析結果から回収の有無の推定値を求め、そのハザード比を求める。つまり、プロビットモデル $\Phi = \beta'x$ から推定値を求め、そのハザード比 $\lambda = \hat{\Phi}/\hat{\Phi}$ を求める。このハザード比をコントロール変数として用いることにより間接的にバイアスを評価する。これは、サンプル選択のトゥービットモデルにおけるヘックマンの 2 段階推定法 (Berk 1983; Breen 1996; Greene 1999; Heckman 1979; Maddala 1992; Winship and Mare 1992) の 1 段階目の推定値をコントロール変数として用いるということである。ヘックマンの 2 段階目の指定は、このハザード比をコントロール変数として加えることにより、元の線型回帰のサンプル選択バイアスを補正するが、これと同じように、さまざまな線型関係にこのハザー

ド比のコントロールをくわえたとき、どのように元の関係が変化するかを見積もる。このハザード比を欠票ハザードとする。

2.2 社会参加の分布と前提

次に、回収されたデータの分析に移る。まず、社会参加について確認しておこう。調査票では、「国政選挙や自治体選挙の際の投票」「政治活動や選挙運動の支援」「市民運動への参加」「ボランティア活動への参加」「自治会・町内会活動への参加」の5つの社会参加項目がある。図1にその分布をまとめる。

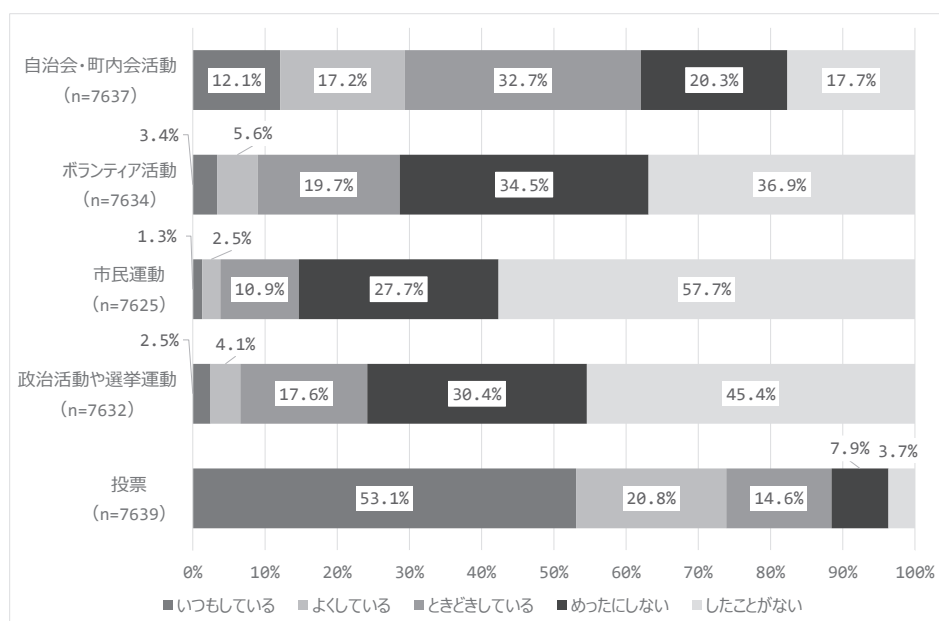


図1 社会参加の分布

投票態度は高い関心を示している。近年の低い投票率に反して、回収票における投票態度は高い。低い投票率は、政治的な無関心と政治的な満足の相反する双方を示すともいえるが、社会参加の指標として考えれば、得られたデータは高い参加を示す。また、別の項目の「あなたは、政治にどの程度関心がありますか」（有効回答数 7,604）という質問については、「非常に関心がある」が 10.0%、「ある程度関心がある」が 52.2%、「あまり関心がない」が 31.3%、「全然関心がない」が 6.5%であり、ある程度高い政治関心を示しているといえる。

投票と比べて、他の4つの指標の社会参加の程度は低いですが、それでも一般的なそれらの活動の頻度に比べれば高い参加率を示しているといえるだろう。この4つの指標はのちの分析において、一つの合成指標「社会活動態度」として用いる。主成分得点として構成する。寄与率は51.2%で、「政治活動や選挙運動の支援」の主成分負荷量は.624で共通性は.386、同じく「市民運動への参加」は.680と.462、「ボランティア活動への参加」.739と.546、「自治会・

町内会活動への参加」は.808と.653である。一つの主成分を構成するととらえてよいだろう。本研究では、以上の分布を確認したうえで、社会参加する人は社会調査も応えやすいということを仮定する。この仮定そのものも重要な課題であり、厳密に確認するためにはもっと慎重な分析が必要ではあるが、本研究ではこのことを分析の仮定とする。

そのうえで次のような分析枠組みを考える。図2はそれを表したものである。横軸に社会参加を、縦軸に任意の社会意識変数の関心の度合いを示す。右側の第一象限と第四象限の部分はこの仮定により回収できていると考える。つまり右側についてはデータがあると仮定される。このとき、回収されたデータについて線型関係を推測させたときに、例えば点線のように正の関係の推定が得られたとしよう。その時、欠票ハザードコントロール変数を加えたときに、その関係が実線のように傾きが下がるように変化した場合、第二象限のデータが得られた場合の変化としてとらえる。つまり、実際は第二象限のデータがあったのに、回収されなかったために、点線の評価は過大評価されていたと間接的に見積もるのである。このように考え、どの部分のサンプルが欠けていたのかを推定する。関係の在り方としては、正・不・無の3通りと過大・過小の2通りを組み合わせた6通りのパターンがあり、その評価を行う²。

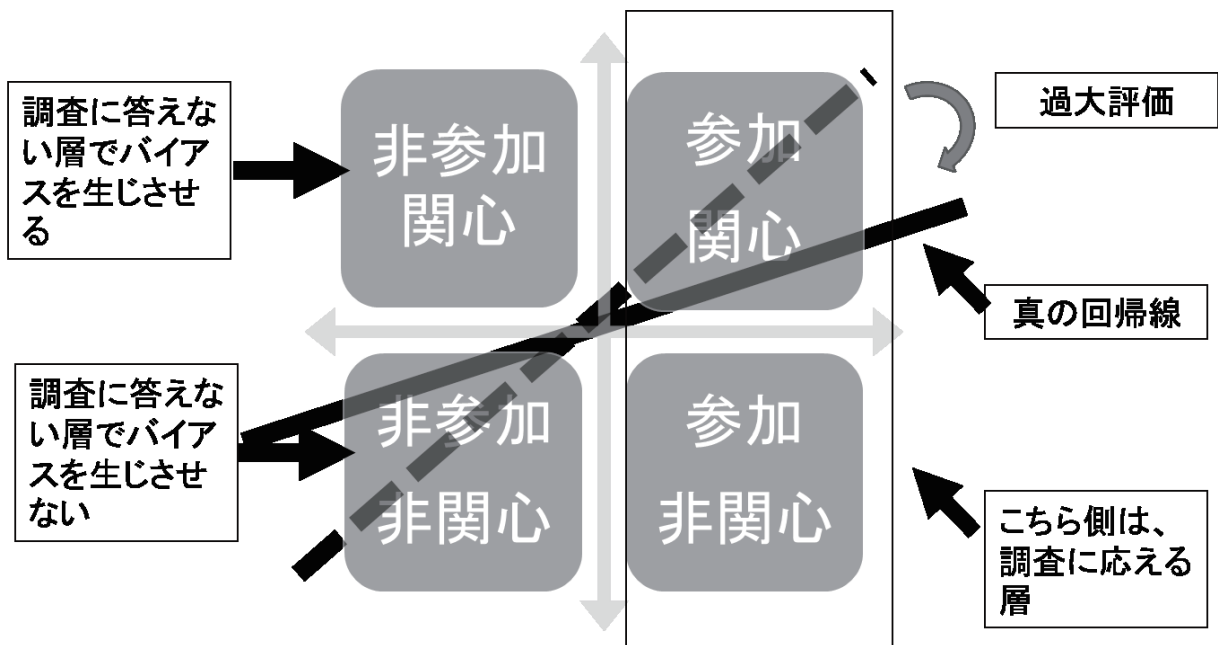


図2 分析枠組み

² なお、欠票には拒否以外の理由（不在など）によるものも含まれるので、欠票のすべてが社会調査に応えない意図を反映しているわけではないが、欠票理由の大多数は拒否によって占められているので（三輪・前田 2018）、ここでは欠票ハザードをそのまま指標として用いる。

2.3 社会意識変数

社会調査に答えることを社会参加の一つとしてみなすことを前提し、回収と非回収のバイアスを評価する。その場合、その関係の変化は社会意識変数に表れる。つまり、社会参加と関わりがあるような、積極的忌避と消極的忌避の仮説の変数を考えることになる。ここでは、その変数として次の8つの変数を扱う。

積極的な忌避にかかわるのは、偏見や誤解、凝り固まった意識や柔軟な知識に基づく変数である。調査票にある質問項目から「権威主義」「婚姻出産自由」「性別役割分業」「文化資本」に関する社会意識変数を取り上げる。権威主義は、「権威のある人々には常に敬意をはらわなければならない」「以前からなされたやり方を守ることが、最上の結果を生む」「この複雑な世の中で何をなすべきかを知るいちばんよい方法は、指導者や専門家にたよることである」の合成尺度である。主成分スコアを用い、寄与率は57.0%、負荷量は順に、.715、.787、.762、共通性は.511、.619、.580である。結婚・恋愛や出産についての伝統的な尺度として、「同性どうしが、愛しあってもよい」と「結婚しても、必ずしも子どもを持つ必要はない」を取り上げる。寄与率は75.2%、負荷量は.872と.862、共通性は.760と.743である。婚姻出産自由とする。性別役割分業主義は、「男性は外で働き、女性は家庭を守るべきである」「男の子と女の子は違った育て方をすべきである」「家事や育児には、男性よりも女性がむいている」から成る。58.4%の寄与率、主成分負荷は.759、.711、.820、共通性は、.576、.505、.672である。知識にかかわることとして、文化資本を取り上げる。「クラシック音楽のコンサートへ行く」「美術館や博物館に行く」「図書館に行く」「小説や歴史などの本を読む」で構成し、49.8%の寄与率、.400、.628、.797、.896の負荷量、.160、.395、.636、.802の共通性である。クラシックコンサートのみやや尺度信頼性が低い、この4つの変数で構成する。

次に、消極的な忌避の変数である。これは、異議申し立ての自主規制にかかわる。現政策に対する反対意見や、口を噤む行為にかかわる変数である。「平等志向」「多元主義」「人的投資志向」「信頼」に関する社会意識変数を取り上げる。新自由主義に反する格差是正や平等化にかかわる質問項目は、「チャンスが平等にあたえられるなら、競争で貧富の差がついてもしかたがない」「競争の自由をまもるよりも、格差をなくしていくことの方が大切だ」「今後、日本で格差が広がってもかまわない」「富む者と貧しい者とのあいだの所得の格差を小さくすべきだ」から成る。寄与率は52.3%、主成分負荷量は-.625、.752、-.764、.742、であり、1番目と3番目は負の係数、共通性は、それぞれ、.391、.566、.584、.551である。平等志向とする。多元主義は、「違った考えかたをもった人がたくさんいる方が社会にとって望ましい」という単独質問で構成する。子どもへの人的資本的投資志向として、「子どもにはできるだけ高い教育を受けさせるのがよい」「子どもには、学校教育のほかに家庭教師をつけたり、塾に通わせた方がよい」「子どもにはできるだけ多くの財産を残してやるのがよい」の3つからなる自由主義的な教育態度を分析する。55.0%の寄与率、.793、.794、.625の負荷

量、.629、.631、.390 の共通性である。人的投資志向とする。政府ではなく共同体的な連帯志向の概念として、「たいていの人は信用できる」「近所の人たちはお互い仲良くやっている」「困っているとき、近所の人たちは手助けしてくれる」からなる信頼の尺度を利用する。62.9%の寄与率、.670、.854、.841 の負荷量、.449、.730、.708 の共通性である。

2.3 欠損の評価

これらの社会意識変数間の相関をとり、欠票ハザードを統制変数とした時の偏相関をとる。その差がどのように変化したかを見ることにより、どこかのデータが欠損していたかを間接的に推測する。

表2は、投票態度と、社会活動態度（政治参加、市民運動、ボランティア、自治会・町内会の合成尺度）と、8つの社会意識変数との相関係数、ハザード比を統制変数とした時の偏相関係数、その相関係数と偏相関係数の差を示している。相関係数は、正の相関を濃い網掛け、負の相関を薄い淡い掛け、5%水準で有意でない無相関を網掛けなしで示している。差については、±.01以上の差を網掛けし、正を濃い網掛け、負を淡い網掛けで示している。

6通りのパターンがあるので順に検討していこう。まず第一は、正の関連性が過小評価されていた組み合わせである。これは、投票態度と文化資本、社会活動態度と文化資本、投票態度と人的投資志向の3つの組み合わせである。統制後の傾きが上昇するパターンで、社会活動態度や投票態度が少ない傾向のある人で文化資本や人的投資志向が低い人が回収されていなかったという間接的な評価となる。文化的な志向が低い人が回収していないパターンで積極的忌避にあたる。

表2 相関係数、偏相関係数、その差

		相関係数							
	権威主義	婚姻出産自由	性別役割分業	文化資本	平等志向	多元主義	人的投資志向	信頼	
投票態度	-.007	-.183	.026	.176	.013	.007	.039	.219	
社会活動態度	.012	-.218	.015	.137	.070	-.071	.006	.302	
		偏相関係数							
	権威主義	婚姻出産自由	性別役割分業	文化資本	平等志向	多元主義	人的投資志向	信頼	
投票態度	-.025	-.130	.020	.207	-.018	.038	.061	.193	
社会活動態度	-.014	-.139	.007	.184	.026	-.027	.038	.269	
		相関係数の差(偏相関-相関)							
	権威主義	婚姻出産自由	性別役割分業	文化資本	平等志向	多元主義	人的投資志向	信頼	
投票態度	-.018	.053	-.006	.031	-.031	.032	.022	-.027	
社会活動態度	-.027	.079	-.009	.047	-.044	.043	.032	-.034	

第二は、正の関連性が過大評価されていたパターンである。投票と信頼、社会活動態度と平等志向、社会活動態度と信頼の3つの組み合わせがあった。過大に傾きが評価されていた場合で、社会参加はしないが、信頼が高く、平等を志向する人が回収されていなかったと評価される。現在の新自由主義的な格差容認をよしとしてないが、社会参加はしない層である。はっきりと異議申し立てをしにくい人々が回収されていない。また、信頼は、行政的政策的な福祉政策の代替となっているという説があるが、つまりは、政府に頼らないということを示しているともいえる。これらは消極的な忌避にあたる。

第三に、負の関連性が過大評価されていたパターンで、これも3つの組み合わせである。投票と婚姻出産自由、社会活動態度と多元主義、社会活動態度と婚姻出産の自由で、いずれも社会参加はしていないが、多元主義を認めない、因習的な婚姻出産にこだわる層が含まれる。これは積極的な忌避にあたる。このようなところも回収されていない。

第四に、負の関連性が過小評価されていたパターンだが、これにあたる組み合わせはなかった。

第五に、無相関が、統制後に正の相関を持つようになった組み合わせである。これは、投票と多元主義、社会活動態度と人的投資志向の2つの組み合わせがあった。この場合は、社会参加していない人で、多元主義的でない人、人的投資的でない人が含まれる。これは積極的な忌避で、そのような人も回収されていないと思われる。

第六に、無相関が統制後に負の相関を持つようになる組み合わせで、投票と平等志向、投票と権威主義、社会活動態度と権威主義が含まれる。これは、消極的な忌避の、社会参加していないが反自由主義的な意見を持つ消極的忌避と、社会参加はしていない権威主義的な人の積極的な忌避の両方の結果である。

最後に、性別役割分業主義については変化があまり見られなかった。この変数と、社会活動態度及び投票との関連性については、欠損に影響されず頑健である。

3. 結論

本研究では、社会調査に答えることを社会参加の一種とみなし、社会参加の意向に関連性を持っているような社会意識変数のバイアスを評価することを試みた。社会参加をしていない人は社会調査に答えない可能性が高く、当然ながら社会参加しないということと関連のある質問項目や変数について、その意見が反映されていない可能性がある。現に近年の投票率の低下に反して、回収されたデータの投票態度は高く、社会参加している人が調査にも回答しているという傾向が読みとれる。得られているデータは社会参加に積極的な人のデータである可能性がある。

ただし、本研究では記述統計に代表されるようなそれぞれの変数の分布についてのバイアスを評価しているわけではない。応えてくれなかった人の変数の分布については知ることが

できないからでもある。そこで、本研究では、変数間の関連性、特に、社会参加と社会意識の関連におけるバイアスを評価することにした。回収されていない欠票がどのように関連性に影響を及ぼすかということの評価するのである。つまり、欠票の影響を考慮していない関連性と、考慮したい場合の関連の違いによって、バイアスを評価することを試みたのである。

社会参加している人は社会調査にも応えるということを仮定すると、当然ながら、回収されていないデータは社会参加していない人ということになる。そうすると、バイアスの評価は、社会参加していない人で、どのような人が、調査に応えていないのかということが分かればよい。線型モデルを仮定すれば、欠票を考慮した場合に、その傾きが変化した方向が分かれば、正確ではないがどの部分が欠けていたかを間接的に評価できるのである。

トウビットモデルのヘックマン推定法は、サンプル選択ハザードを推定し、それをコントロール変数とすることにより、そのバイアスを評価するやり方なので、この方法を応用した。まず、サンプル情報に基づき、回収の有無をプロビットモデルで推定し、ハザード比を求め、次に、それをコントロール変数とすることにより、その傾きの変化をみればどこに欠票があったのかを間接的に知ることができる。今回は、相関と偏相関をつかったが、そのような方法で分析を行った。

その結果、回収されないことについては積極的忌避と消極的忌避の2種類があることがわかった。積極的忌避は、科学研究についての誤解や偏見によって、回答しないという態度で、権威主義的態度や、文化資本、婚姻出産の自由などにかかわる。社会参加はしない層で、権威主義的で、文化資本があまりなく、婚姻出産の多様性を認めない層が社会調査にも応えていないということが推測された。

もう一つの消極的忌避は、体制に対する異議があるが申し立てしにくい層が自主的に自分の意見を規制している層である。社会参加はしていないが、反自由主義的に平等を志向する人、多元主義的な人、信頼が高い人である。こうした人々が調査に回答していない。これは言論や表現の自由との関わりが強く、統制が強まれば、こうした口を噤む自主規制は広まる。ただし、このような層は自由な雰囲気が強まれば、より積極的に調査に協力する可能性もある。しかし、現在の状況は自主規制が広がっている傾向にあるかもしれない。

社会調査は自発的協力を基調とするので、どうしてもこのような市民性との関わりがあるだろう。市民性の一部としての社会調査というテーマはこれまではあまり明示的には論じられてはいなかったが、科学研究の公共性を理解してもらうためにも、そのような研究は今後重要になってくるように思われる。

[文献]

Berk, Richard A. 1983. "An Introduction to Sample Selection Bias in Sociological Data," *American Sociological Review* 48: 386-398.

- Breen, Richard. 1996. *Regression Models: Censored, Sample Selected or Truncated Data*. SAGE Publications.
- Greene, William H. 1999. *Econometric Analysis, 4th Edition*. 斯波恒正・中妻照雄・浅井学 訳. 2000. 『計量経済分析』エコノミスト社.
- Heckman, James J. 1979. "Sample Selection Bias as a Specification Error," *Econometrica* 47(1): 153-161.
- 吉川徹. 2010. 「拒否増加にいかに対応するか」『社会と調査』5: 16-25.
- Maddala, G. S. 1992. *Introduction to Econometrics, Second Edition*. 和合肇 訳. 1996. 『計量経済分析の方法』シーエーピー出版.
- 松岡亮二・前田忠彦. 2015. 「『日本人の国民性第13次全国調査』の欠票分析：個人・地点・調査員の特性と調査回収状況の関連」『統計数理』63(2): 229-242.
- 三輪哲・前田忠彦. 2018. 「2015年SSM調査による調査不能と項目無回答の基礎分析」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: 13-27.
- 小野寺典子・片山朗・佐藤 嘉倫・前田 忠彦・松田映二・吉川徹・篠木幹子・大谷信介. 2010. 「座談会『回収率を考える』」『社会と調査』5: 26-65.
- 篠木幹子. 2010. 「社会調査の回収率の変化」『社会と調査』5: 5-15.
- Winship, Christopher and Robert D. Mare. 1992. "Models for Sample Selection Bias," *Annual Review of Sociology*, 18: 327-350.

Social Survey as a Social Participation: Evaluating the influence of non-response bias on the relationships between variables ^{*}

**Hirosi TAKADA
(Sapporo Gakuin University)**

Abstract

This study evaluates whether the bias of responding to social surveys affects the relationship between various variables, especially political attitudes and social consciousness. Whether to respond to a social survey or not may be related to the orientation and attitudes of the surveyed person. Hence, a specific bias may arise in the collected survey data related to whether or not survey targets respond to the survey. However, we cannot know the responses to the relevant variables from survey targets whose data has not been collected, so we cannot estimate this bias directly. Therefore, in this research, based on data about the missing information, response or non-response is regressed by the information that can be learned there, so as to estimate the hazard ratio of non-response. The estimated value obtained is then added to the analysis model as a control variable in the collected data. The result of this model is compared with that from the model before adding the control variable, and by analyzing how the relationship changes, the bias is evaluated indirectly. As a result, it was found that the bias affecting the relationship between the variables consists of the following two factors. The first is a passive avoidance in which, although the target does not participate in society, they find it difficult to object, and the second is an active avoidance that rejects the survey due to a conventional orientation.

Keywords: Non-response bias, Social participation, Social survey

* The study was supported by JSPS KAKENHI Grant Number JP25000001.