

Web 調査の偏りの補正：
行動経済学における調査研究への適用

星野崇宏

RCSS

文部科学大臣認定 共同利用・共同研究拠点
関西大学ソシオネットワーク戦略研究機構
関西大学ソシオネットワーク戦略研究センター
(文部科学省私立大学学術フロンティア推進拠点)

Research Center of Socionetwork Strategies,
“Academic Frontier” Project for Private Universities, 2003-2009
Supported by Ministry of Education, Culture, Sports, Science and Technology

The Research Institute for Socionetwork Strategies,

Joint Usage / Research Center, MEXT, Japan

Kansai University

Suita, Osaka, 564-8680 Japan

URL: <http://www.rcss.kansai-u.ac.jp>

<http://www.kansai-u.ac.jp/riss/index.html>

e-mail: rcss@ml.kandai.jp

tel: 06-6368-1228

fax. 06-6330-3304

Web 調査の偏りの補正: 行動経済学における調査研究への適用

星野崇宏¹²

要約

本研究では、これまでに得られた Web 調査の補正に関する知見を元に、大阪大学 21 世紀 COE プロジェクトにおいて実施された「くらしの好みと満足度についてのアンケート」を利用した Web 調査の補正の実証研究を行った。特に、2007 年 2 月に実施されたアンケート調査のうち、「時間割引率」「仕事に対する報酬への危険回避傾向」「労働と報酬」に関連する項目に関して、2009 年 3 月に関西大学 RCSS センターによって実施された Web 調査を補正することで、大阪大学による住民基本台帳からの層化二段無作為抽出に基づく訪問留置調査の結果を予測することが可能かどうかを調べた。結果として、訪問留置調査に対する Web 調査の二乗誤差を補正によって半減することができた。また、補正の効果は調査モードの影響が考えられる「労働と報酬」に関する項目を除き良好であり、Web 調査で二乗誤差が高い項目ほど補正の効果があることがわかった。

また、最後に経済学研究での Web 調査の補正を行う際の留意点について議論した。

Keywords: 選択バイアス、時間割引率、危険回避度、傾向スコア、共変量調整、調査モード

¹ 名古屋大学 大学院経済学研究科 准教授

² 関西大学ソシオネットワーク戦略研究センター 研究員

Email: bayesian@jasmine.ocn.ne.jp

Adjustment of selection bias in Web survey data - application to a survey of behavioral economics –

Takahiro HOSHINO³⁴

Abstract

Based on previous studies on adjustment of selection bias in Web survey, we applied propensity score adjustment method to a Web survey with purposive sampling to predict results of a visit survey with two stage stratified random-sampling conducted by the Osaka University 21th century COE program, "Behavioral Macro-dynamics Based on Surveys and Experiments". The target variables in adjustments are variables for "discount rates", "risk aversion" and "labor and wage".

The sum of squared errors of the estimated rates in the Web survey reduces about by half after using propensity score adjustment method. We also found that the adjustment method does not work well for the variables for "labor and wage" in which the effects of survey mode might be large.

We discussed what should be noted in applying covariate adjustment methods to Web survey of economic research.

Keywords: Selection Bias, Discount rate, Risk aversion, Propensity score, Covariate Adjustment, Effects of Survey mode

³ Associate Professor, Department of Economics and Business Administration, Nagoya University

⁴ Research Fellow, Research Center of Socionetwork Strategies, Kansai University
Email: bayesian@jasmine.ocn.ne.jp

1. 問題意識

Kahneman や Tversky の一連の研究(Tversky and Kahneman,1971; Kahneman and Tversky,1982)以降、人々が必ずしも規範的な意思決定理論から導かれる合理的な意思決定を行っていないことが様々な実験や調査から示されている(竹村,2009)。特に人々の「時間割引率」や「危険回避傾向」などが金融政策や経済成長(Krugman,1991)などを考える上で非常に重要な要因であり、経済合理性からの乖離の程度を積極的に定量化し理論化をしようとする行動経済学への研究関心の高まりがある。

このような潮流から、近年では経済学においても実験研究や調査研究が盛んにおこなわれるようになってきている。

さて、これまで日本における経済学での調査研究の多くは、世論調査や社会調査と同様の方法論によって実施されることが多かった。具体的には住民基本台帳や選挙人名簿から無作為抽出によって抽出した対象者に対する訪問調査・訪問留置調査または郵送調査（以降従来型の調査方法と呼ぶ）を実施するというものである。この調査方式は住民基本台帳と選挙人名簿という明確な抽出台帳が全成人に対して用意されている日本においては当然のものともみなされてきたが、アメリカやフランスでは一般の調査に利用可能な同様の抽出台帳が存在しないことから、エリアサンプリングや電話調査での RDD(Random Digit Dialing)法などが利用されている。これまでの日本での世論調査・社会調査においては住民基本台帳及び選挙人名簿からの無作為抽出が可能であったことから、従来型の調査方法と異なる方法によって実施された調査に対する不信感が根強い。

しかし、このような従来型の調査方法は抽出や調査実施に時間がかかること、また調査費用がかかるといった欠点を有している。加えて近年では個人情報保護の観点から、住民基本台帳の閲覧を実質的に制限する地方自治体が増えたこと、さらには抽出をいくら無作為に行っても、調査への回答拒否などの未回収率が上昇していることも問題となっている。

近年では個人の意識や思考が多様化し、かつ情報化の進行から非常に短期間に個人の考え方が変化するという現象が生じており、また 2008 年のリーマンショック以降の景気後退局面など、これまでなかったような社会経済環境の急変も生じえる。従って標本の代表性を担保しつつ非常に短い実施期間で費用を抑えて実施できる調査方法が必要とされてきている。このようなニーズに対して、民間の市場調査では近年の IT 技術の進歩とインターネットの普及によって Web 調査が広く用いられるようになっており、経済学での調査研究(鶴飼・村上,2008)や、政府関係の調査(例えば内閣府経済諮問委員会「日本 21 世紀ビジョン調査」「公共料金分野における情報公開の推進調査」など)においても利用されるようになってきた。

但し、従来型の厳密な確率的標本抽出に基づく従来型の調査と Web 調査の結果の比較からは、後者は標本の代表性に欠け、バイアスを持つ可能性が高い。例えば、従来型の厳密な確率的標本抽出に基づく調査（訪問調査や郵送調査法）と RDD 電話調査や Web 調査の

結果の比較から、後者は標本の代表性に欠け、バイアスを持つことが示されている(Couper, 2000 ; 星野,2007 ; 前田・土屋,2001 ; 山岡・林,1999).

このような Web 調査のバイアスは統計学においては欠測データの問題として定式化することができ、Heckman(1976)の提起した選択バイアスの問題と類似の構造を有する(星野,2009)ため、標本抽出に関連する共変量を調整することで、バイアスのある程度補正することが可能になると考えられる。実際、星野・前田(2006)、星野(2007)、星野・森本(2007)、Schonlau, van Soest, Kapteyn and Couper(2009)は複数の共変量を利用して算出した傾向スコア(Rosenbaum and Rubin, 1983)による共変量調整によってバイアスの補正が可能となることを示している。但し、星野・前田(2006)や Schonlau ら(2009)では社会調査、星野(2007)では市場調査での補正が行われているが、経済学において関心の高い調査項目についての補正は未だ試行されていなかった。

本研究では、これまでに得られた Web 調査の補正に関する知見を元に、大阪大学 21 世紀 COE プロジェクト(研究代表者:筒井義郎)において実施された「くらしの好みと満足度についてのアンケート」を利用した Web 調査の補正の実証研究を行う。特に、2007 年 2 月に実施されたアンケート調査のうち、「時間割引率」「仕事に対する報酬への危険回避傾向」「労働と報酬」に関連する項目に関して、2009 年 3 月に関西大学 RCSS センターによって実施された Web 調査を補正することで、大阪大学による従来型の調査結果を予測することが可能かどうかを調べる。

2. Web 調査の補正とは？

本研究では Web 調査の結果を補正するための共変量調整法として傾向スコアによる IPW 推定量(Hoshino, Kurata and Shigemasu,2006; Hoshino, 2008)と二重にロバストな推定法を利用する。

従来型調査の手続きが厳密に行われているならば、Web 調査と従来型調査の結果が異なることは Web 調査の結果にバイアスがあることを示す。ここで Web 調査のバイアスを考える際に、欠測データの枠組みを用いることで問題の構造を明確化することができる。以降では「インターネット調査に回答した集団」を“Web 調査の標本”，そして従来型調査は「母集団から無作為抽出された対象者集団」であるはずなので、これを“無作為抽出の標本”と便宜的に名付ける。ここで、ある調査対象者が Web 調査の標本に含まれる場合には $z = 1$ 、無作為抽出の標本に含まれる場合には $z = 0$ となるインディケータ変数を z とする。

また 1 人の調査協力者に対して、本来は 2 つの回答値が得られると考える。具体的には「Web 調査で得られる回答」を y_{web} ，「従来型の調査で得られる回答」を y_{real} とし、本来同一対象者について 2 つの値が得られると考える。このように考えると、「インターネット調査に回答した集団」 $z = 1$ では y_{web} が、そして「無作為抽出の標本」 $z = 0$ では y_{real} が観測され、他方は観測されないということになる。

このように考えれば、Web 調査の結果と従来型の調査の違いは

[1] 標本の違い(両者の抽出方法の違い)

$z=1$ か $z=0$ かで対象者の様々な属性が異なるため、単なる標本変動ではなく、両調査では本来推定しているものが異なる可能性が高い。それぞれの標本平均 \bar{y}_{web} 、 \bar{y}_{real} は $E(y_{web} | z=1)$ 及び $E(y_{real} | z=0)$ の不偏推定量であるが、両者の抽出方法の違いから $E(y_{web} | z=1) \neq E(y_{real} | z=0)$ であれば、たとえサンプルサイズを多くしても二つの標本平均は一致しない。

[2] 回答方法の違い

一般には $y_{web} \neq y_{real}$ であり、これを「調査モードによる違い」と呼ぶ。具体的にはインターネット調査でのラジオボタンのクリックによる回答と、訪問留置調査での筆記による回答では異なる可能性がある。

の2つに起因すると考えることができる。但し、世論調査などで対象となる意見項目ではなく、所有や経験などの事実に基づく項目であれば、あまりこの調査モードの効果は存在しないと考えて良い。また、この調査モードの効果は同一の調査対象者にどちらの調査に対しても回答してもらうことによってのみ調べることが可能であるが、今回利用した調査ではそのような対象者は存在しない。

そこで、本研究では調査モードの効果は存在しない($y \equiv y_{web} = y_{real}$)とすると

$$y = zy_{web} + (1-z)y_{real}$$

と表現することもできる。但しこの場合でも上記[1]の問題は残り、

$$E(y | z=1) \neq E(y | z=0)$$

である。

さて、ここでWeb 調査から従来型調査の結果を推定するという事は、母数

$$E(y_{web} | z=0)$$

つまり「もし無作為抽出標本がWeb 調査に回答した場合の回答の期待値」を推定する、という問題になる。期待値以外に関心がある場合も分布 $p(y_{web} | z=0)$ を推定することが目的となる(星野,2005)。但し、すでに述べたように $z=0$ の場合の y_{web} は観測されない欠測値であり、何も追加的な仮定を置かなければこれを推定することはできない。

そこで、2つの調査の抽出方法の違いによって生じる結果の差異を説明するような様々な共変量 x を利用する共変量調整が可能となる条件、つまり

$$p(y, z | x) = p(y | x)p(z | x)$$

を仮定する。これは「共変量 x を所与とすると項目への回答 y とインディケータ z が独立になる」という条件であり、統計的因果推論では「強く無視できる割り当て条件」(Strongly ignorable treatment assignment; Rosenbaum and Rubin, 1983)と呼ばれるものに相当する。

この仮定が成立するとき、 $e \equiv p(z=1|x)$ とすると、以下の IPW(Inverse Probability Weighting)推定量

$$\hat{E}(y|z=0) = \frac{\sum_{i=1}^N \frac{z_i(1-e_i)}{e_i} y_i}{\sum_{i=1}^N \frac{z_i(1-e_i)}{e_i}}$$

は $E(y|z=0) = (y_{web} | z=0)$ の不偏推定量となる(星野,2009)。但し i は調査協力者 i の値を示し、 N は従来型調査と Web 調査を合わせたサンプルサイズである。実際には e_i (傾向スコア(propensity score)と呼ばれる) は未知なので、インディケータ z を共変量 x によって説明する回帰分析(ロジスティック回帰やプロビット回帰)を行って得られた推定値で代入するが、その場合には $\hat{E}(y|z=0)$ は一致推定量になる。これが傾向スコアを用いた Web 調査のバイアスの補正のための推定量である。

IPW 推定量以外にも、回答が連続変数である場合には、Hoshino(2007)による二重にロバストな推定法を用いて得られる「訪問留置調査に回答した標本における Web 調査の回答の平均」の推定量

$$\frac{1}{N_0} \sum_{i=1}^N \left[\frac{z_i(1-e(x_i, \hat{\alpha}))y_i}{e(x_i, \hat{\alpha})} + \left(1 - \frac{z_i}{e(x_i, \hat{\alpha})} \right) g(x_i, \hat{\beta}_1) \right]$$

を用いることで、傾向スコアの推定が誤っている場合においてもより頑健な推定が可能である(星野,2009)。本研究の目的変数は 2 値および多値であるが、このような離散変数の場合であっても、まず目的変数を 2 値化に、回帰関数 $g(x_i, \hat{\beta}_1)$ にロジスティック回帰分析モデル

$$\frac{1}{1 + \exp(-(\gamma + \hat{\beta}_1^t x_i))}$$

を利用することで適用することができる。

さて、傾向スコアを用いた解析を行う場合には、傾向スコアを算出するために複数の共変量を同時に解析に投入する。従って、すべての共変量において無回答や回答拒否などの欠測がない対象者のみを用いて解析を行うと、サンプルサイズが非常に小さくなってしまふ。この問題に対する対処法はいくつか考えられるが、共変量がカテゴリカル変数であり、両方のデータに欠測が存在する場合については、欠測そのものを 1 つのカテゴリとみなして解析を行えばよい。一方、訪問留置調査では書き忘れや回答拒否などによる欠測があるが、Web 調査では質問場面の遷移において欠測を認めない調査形式が利用されているために欠測が生じない、という場合には欠測を新しいカテゴリとしても Web 調査にはその

カテゴリーに分類される調査対象者がいないために傾向スコアの算出ができなくなる。

このような場合には訪問留置調査における欠測値を単一代入法や多重代入法で補完する必要がある。

3. 利用したデータについて

今回 Web 調査の補正の対象として利用した大阪大学による調査の対象者は 20 歳以上 75 歳までの全国成人である。そこで関西大学による Web 調査のうち、同じ年齢幅のデータを解析対象とした。大阪大学による訪問留置調査での回収標本は 3112 人であり、回収率は 85.03%であった。また、関西大学による Web 調査は 1798 人であった。また、Web 調査は 2009 年 3 月に実施されたのに対して、訪問留置調査は 2007 年 2 月から 3 月に実施されていることから、Web 調査の質問文では 2007 年 2 月時点での回答を記載するように指示している。

本研究では「時間割引率」「仕事に対する報酬への危険回避傾向」「労働と報酬」に関連する項目を目的項目とした。具体的には「時間割引率」は「A 時点で金額 X を受け取るか、それ以降の B 時点で X に利息が加わった部分を受け取るか」の 2 枝選択をさせる項目によって測定されている。ここで時点 A,B,X の組み合わせは (2 日、9 日、1 万円)、(90 日、97 日、1 万円) (1 ヶ月後、13 ヶ月後、1 万円) (1 ヶ月後、13 ヶ月後、100 万円) である。同様に支払わなくてはならない場合についても (1 ヶ月後、13 ヶ月後、100 万円) の組み合わせについて選択させる。また各項目で利子を 8 種類設定しており、合計して $5 \times 8 = 40$ の回答を求めている。

さらに、「危険回避度」を測定する項目として、仕事に対する報酬の支払い方法として「月収が半々の確率で現在の月収の 2 倍になるか、現在の月収の 30%減になる仕事」と「現在の月収の 5%増しに確定している仕事」のどちらかを選ばせる項目、および「月収が半々の確率で現在の月収の 60%増しになるか、現在の月収の 10%減になる仕事」と「現在の月収の 5%増しに確定している仕事」のどちらかを選ばせる項目を利用する。

また他にも、2 年間収入がないと仮定して、金利の違いによって「1 年目により多く消費する」か「2 年目により多く消費する」かを選ばせる項目 (金利が 6 種類あるため計 6 項目) や、10 年間の生活レベルの分配に関する項目、50%の確率で当選し 2000 円もらえる 200 円の宝くじを買うか (および保有して売却するか) どうかについての項目、1%の確率で当選し 10 万円もらえる 200 円の宝くじを買うか (および保有して売却するか) どうかについての項目、「50%の確率で 2000 円の損失が発生するとし、200 円の保険料で全額損失補償がされるとする場合に、保険料を支払うか」という項目、「50%の確率で 2000 円もらえるか 1000 円支払わなくてはならないくじを受け取るか」という項目があり、これらも危険回避度の高さを測定するものである。

加えて、労働と報酬に関する項目として、自由に労働時間を変更できるという仮定のも

とで、「世帯所得と同額の賞金が生存中毎年支払われる宝くじに当たった場合に、労働時間を増やすかどうか」という項目、「勤め先が賃金を今年のみ半減させるとした場合に、労働時間を増やすかどうか」という項目、「勤め先が賃金を永久に2倍にした場合に、労働時間を増やすかどうか」という項目、「勤め先が賃金を今年限り2倍にした場合に、労働時間を増やすかどうか」という項目を利用した。

まとめると、「時間割引率」40項目、「危険回避度」14項目、「労働と賃金」4項目の計58項目を調整の目的変数とした。

今回のWeb調査はRCSSセンターによって実施されたものであり、筆者は項目設計を行ったものではない。結果として阪大21世紀COEの調査と複数の項目で選択肢が異なっていたため、カテゴリーを合併するなどの処理を行い、変数が共通になるようにした。

またWeb調査において多くの項目で回答項目が必須の設定となっていたが、訪問留置調査では収入や学歴などの項目で無回答の比率が高かった。2節で論じたように、共変量として利用する項目がカテゴリカル変数の場合であれば、無回答に対応するカテゴリーを作成して共変量調整に利用すればよいが、Web調査で無回答に対応するカテゴリーが無い、または非常に頻度が少ない場合には共変量として利用することはできない。そこで本研究では共変量として利用する項目については、無回答に単一代入を行った。具体的には各変数のモードを無回答による欠測値に代入した。

4. 解析結果

4. 1 Web調査の偏りについて

これまで指摘されてきたように(Couper, 2000; 大隅, 2002; Taylor, 2000)、従来型の調査とWeb調査の結果の乖離はきわめて大きくなることが多い。その理由は大きく分けて標本抽出法の違いと調査モードの違いによる。実際、本研究で利用した2つの調査については、大阪大学が実施した訪問留置調査での標本抽出法は層化2段無作為抽出(10地域4都市規模別による層化)であり、母集団は2007年2月時点で住民基本台帳に記載されている成人である。一方、Web調査での標本抽出法は調査実施を委託した調査会社が有する調査協力者パネルからの割当法であり、母集団を訪問留置調査と同じと考えてして議論を行う場合には有意抽出であると言える(星野, 2009)。

本研究で利用した調査での目的項目は世論調査のような意見を聞くようものではなく、所有や収入などの事実や割引率などの経済学的な構成概念を測定するものであるため、調査モードの影響は小さいと考えてよい。

また、本研究で利用したWeb調査データと訪問留置調査は2年ほど実施時期が異なることは重大な影響を与えうるが、これについても今回解析の対象とした調査項目のうち、調

整目的項目が「時間割引率」「危険回避度」「労働と賃金」といった成人においては比較的安定と考えられる項目である。また共変量として利用する項目についても、たとえば収入は2006年での収入を聞くなど、2年前の調査対象者の状態を聞いていることや、こちらもある程度時間の経過によらず安定している。これらのことを考えると、2年の時間のズレはあまり問題にならないと考えられる。また、本研究の目的は時間割引率などの経済学的変数に対してこれまで開発された共変量調整法が有効であるかどうかを調べることであり、たとえ2年間で目的変数の母集団平均が異なっているとしても、2年前の調査を予測できるかどうかを調べることには意味があると考えられる。

これに関連した先行研究として、星野・前田（2006）は日本版総合的社会調査(Japanese General Social Survey:JGSS)の2001・2002・2003年版調査の一部項目について2004年にWeb調査を実施し、共変量調整による補正後の推定値を算出している。その結果、各年のJGSS調査での推定値に対するWeb調査での推定値の二乗誤差を基準とすると、調整後の推定値の二乗誤差は5~6割程度縮減されており、実施時期が3年ほど異なっている場合においても調整が有効に作用することが示されている。本研究においても、Web調査への共変量調整を行うことで、2年前の訪問留置調査の予測を行うことが可能であると考えられる。

ここで、今回のWeb調査の偏りがどの程度のものであるかを示す。本研究で補正の目的項目とした58項目について、各選択肢への訪問留置調査の選択比率の推定値を横軸に、Web調査での選択比率の推定値を縦軸にプロットしたのが図1である。図1にある直線は「訪問留置調査の結果=Web調査の結果」を示しており、多くの項目でこの直線からの著しい乖離が見られる。但し、2つの標本の比率の推定値の差異が標本変動の結果である可能性があることから、2つの調査での比率が等しいとした場合の95%信頼区間を計算する必要がある。この信頼区間を構成するために、まず訪問留置調査の母集団比率を p_r とするときの漸近信頼区間が

$$\left(p_r - 1.96 \times \sqrt{\frac{\hat{p}_w \times (1 - \hat{p}_w)}{N_w}}, p_r + 1.96 \times \sqrt{\frac{\hat{p}_w \times (1 - \hat{p}_w)}{N_w}} \right)$$

と計算できることを利用する。ここで \hat{p}_w, N_w をそれぞれWeb調査での比率推定値とサンプルサイズとする。

ここで実際には真値 p_r は得られないため、その推定値 \hat{p}_r で代用する必要があるが、この際に訪問留置調査とWeb調査は独立な標本であることから、 \hat{p}_r を推定したことを考慮した信頼区間は

$$\left(\hat{p}_r - 1.96 \times \sqrt{\frac{\hat{p}_r \times (1 - \hat{p}_r)}{N_r} + \frac{\hat{p}_w \times (1 - \hat{p}_w)}{N_w}}, \hat{p}_r + 1.96 \times \sqrt{\frac{\hat{p}_r \times (1 - \hat{p}_r)}{N_r} + \frac{\hat{p}_w \times (1 - \hat{p}_w)}{N_w}} \right)$$

となる。但し N_r は訪問留置調査のサンプルサイズであり、ここでは訪問留置調査と Web 調査の結果が等しい場合の信頼区間を考えるので、 $\hat{p}_r = \hat{p}_w$ とすればよい。

図 1 は訪問留置調査において様々な比率推定値が得られた場合の 95%信頼区間もプロットしている。図 1 から明らかなように、訪問留置調査と Web 調査の結果の差異は標本変動によるものではないことから、今回取り上げたような経済学的な変数を調査する際にも 2 つの調査の抽出方法等が異なることによる系統的なバイアスが発生しえることが明白である。特に危険回避度についての回答で大きな差異があるようである。

ところで、訪問留置調査と Web 調査の差異の大きさの程度はどの程度であろうか？性別や年代の違いなどよりも大きいであろうか？この問いに答えるために、訪問留置調査で得られた回答に限定して、男性での選択比率を横軸に、女性での選択確率を縦軸にプロットしたのが図 2 である。図 2 からは時間割引率について大きな性差が存在することが分かる。

調整の目的項目すべてを考えると、訪問留置調査と Web 調査の比率推定値の差の二乗和は 6456.94 であり、男性と女性での比率推定値の差の二乗和 7800.42 と大きな差異はない。

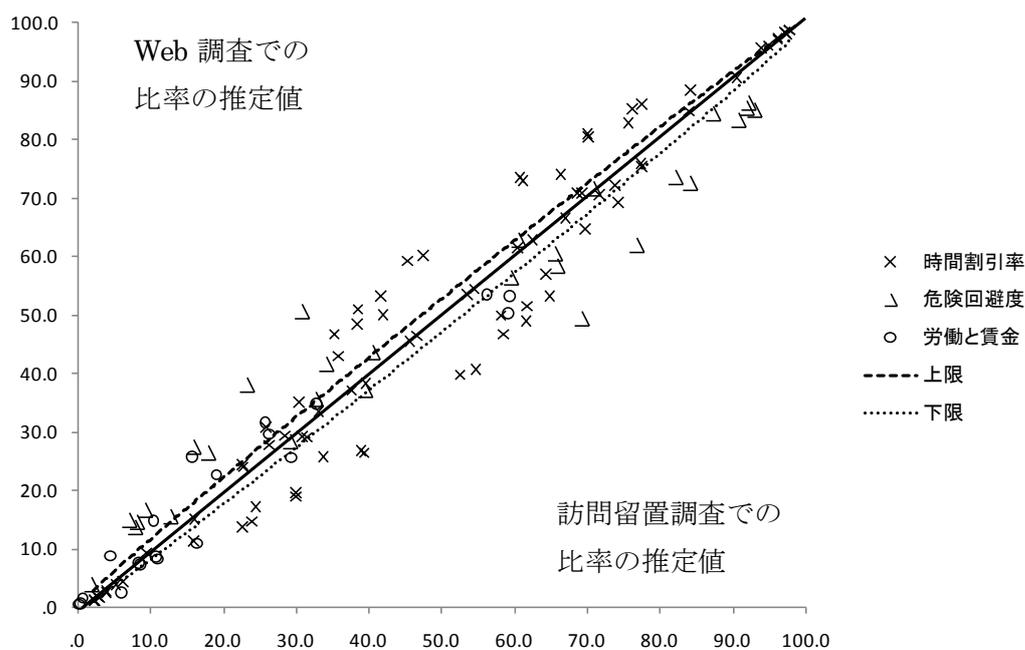


図 1：訪問留置調査と Web 調査の差異

*注意：直線は訪問留置調査と Web 調査の結果が等しい場合にこの直線上にプロットされることを示す。

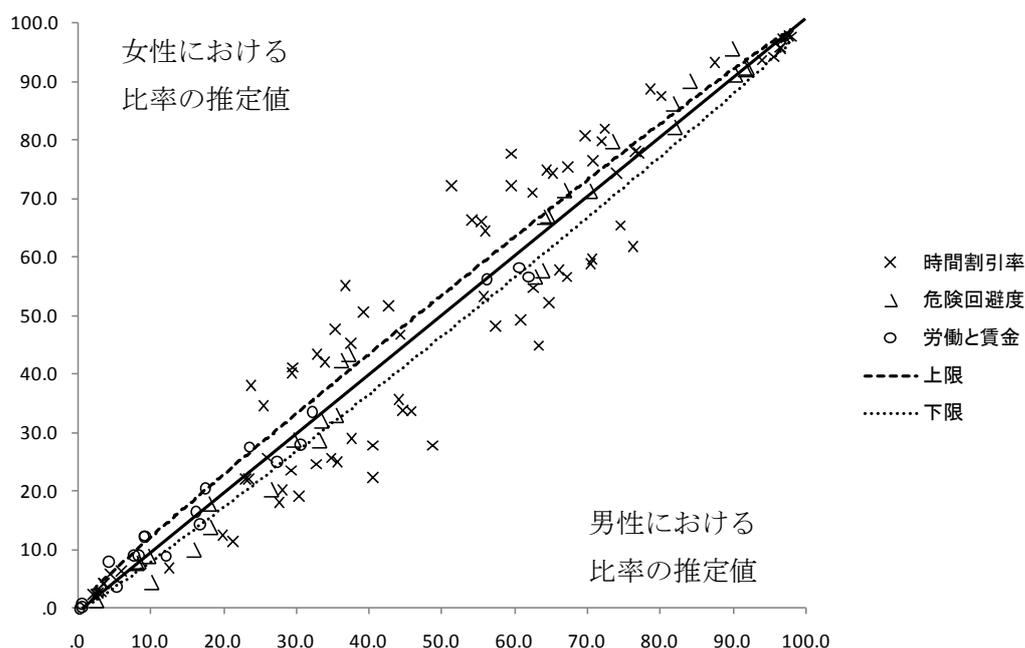


図 2：男女による差異

*注意：直線は男性と女性の結果が等しい場合にこの直線上にプロットされることを示す。

つまり、時間割引率や危険回避度といった、進化的に考えれば性別による差異が当然生じるであろう項目に関してであっても、Web 調査の回答のバイアスの程度は性別程度に大きいということである。性別による差異が生じにくいような項目についてであれば、Web 調査のバイアスは性別での結果の差異よりも大きくなることが予想できる。

このような結果からも Web 調査で得られたデータをそのまま利用して解析することには大きな問題があると言ってよいだろう。

4. 2 共変量の選択について

本研究では星野・前田(2006)によって提案された調整方法を利用した。これは、まず訪問留置調査とweb調査で共変量の候補となる項目をピックアップし、以下の2段階で共変量を絞り込んだ⁵。

[1] まず両調査間で差異のある項目のみ選定する。

今回は2群での差の検定として、共変量の候補項目が連続変数の場合にはt検定を、2値変数の場合にはロジスティック回帰分析を、順序尺度の場合には順序ロジスティック

⁵ 但し web 調査では性別年代についての割当法が実施されていることから、年代層などは2群間で有意ではないが、あえて共変量として加えている。他にも整合性の観点からこの2段階に完全には従っていないわけではない。

ク回帰分析を、名義尺度の場合にはクロス表に対する χ^2 乗検定を行い、有意な差がある項目のみを残した。結果として共変量調整の候補である69項目のうち45項目が候補として残った。

- [2] 上記の項目を説明変数として、目的変数を説明する回帰分析を行い、偏回帰係数が有意でなかったものを除外した。実際には目的変数は58項目であることから、6項目以上で有意になった項目を共変量として残した。

結果として利用することとした共変量は「性別」「年代層」「都道府県」「周りの人の生活水準」「配偶者の有無」「15歳のころの兄弟姉妹の人数(4項目)」「最終学歴」「職業」「勤務形態」「一年間の労働時間」「子供の有無」「2006年のボーナスを含めた個人年収」「同居家族形態」「世帯人数」「現在仕事を探しているか」「父親の最終学歴」「母親の最終学歴」「扶養か被扶養か」「住居形態」「生活水準」「15歳のころの生活水準」「中学校でのクラス委員の経験有無」「高等学校でのクラス委員の経験有無」「中学校での成績(3項目)」である。

これら29項目の共変量を用いて群別のインディケータ z を説明するロジスティック回帰分析を行った。ロジスティック回帰分析の正判別率は82.3%、Cox & Snell R^2 乗は0.389、Nagelkerke R^2 乗は0.532と、利用した29項目の共変量から群別のインディケータを十分よく説明できていることがわかる。この分析で得られた「web調査群に所属する」ことについての予測確率が傾向スコアである。但し、web調査対象者と分類される予測確率を傾向スコアとするが、ここで傾向スコアが0.01以下の場合にはすべて0.01とした。これは、傾向スコアを用いたIPW推定量では傾向スコアの逆数による重み付けを行っていることから、非常に小さな傾向スコアの値を持つ対象者が存在することによって推定が不安定になることを避けるためである。

4. 3 補正結果について

補正が作用したかどうかを調べるため。まず時間割引率に関する項目の訪問留置調査、web調査、及び補正後の推定値の比較を表1に記載した。また、「危険回避傾向」「労働と報酬」についても同様に表2に記載した。

本研究で補正の目的項目とした58項目について、各選択肢への訪問留置調査の選択比率の推定値を横軸に、補正後のWeb調査での選択比率の推定値を縦軸にプロットしたのが図3である。ここで、調整後の推定値の分散は通常の比率の推定値と異なり、サンプルサイズと比率の真値だけではなくデータそのものに依存する。従って信頼区間を図2のように計算することはできない。そこでCopas and Eguchi(2005)に従い、推定値のロバストな分散を計算した結果を図3に記載した。図から明らかなように、補正を行ってもまだ有意な差のある項目は多いが、補正前に比べて補正後の方が差が減少している項目が多いこともわ

かる。実際、訪問留置調査と Web 調査の比率推定値の差の二乗和は 6456.94 であったが、訪問留置調査と補正後の Web 調査の比率推定値の差の二乗和は 3284.14 であり、前者に比べて後者では 50.86%に縮減しており、補正の効果は十分あったと言える。

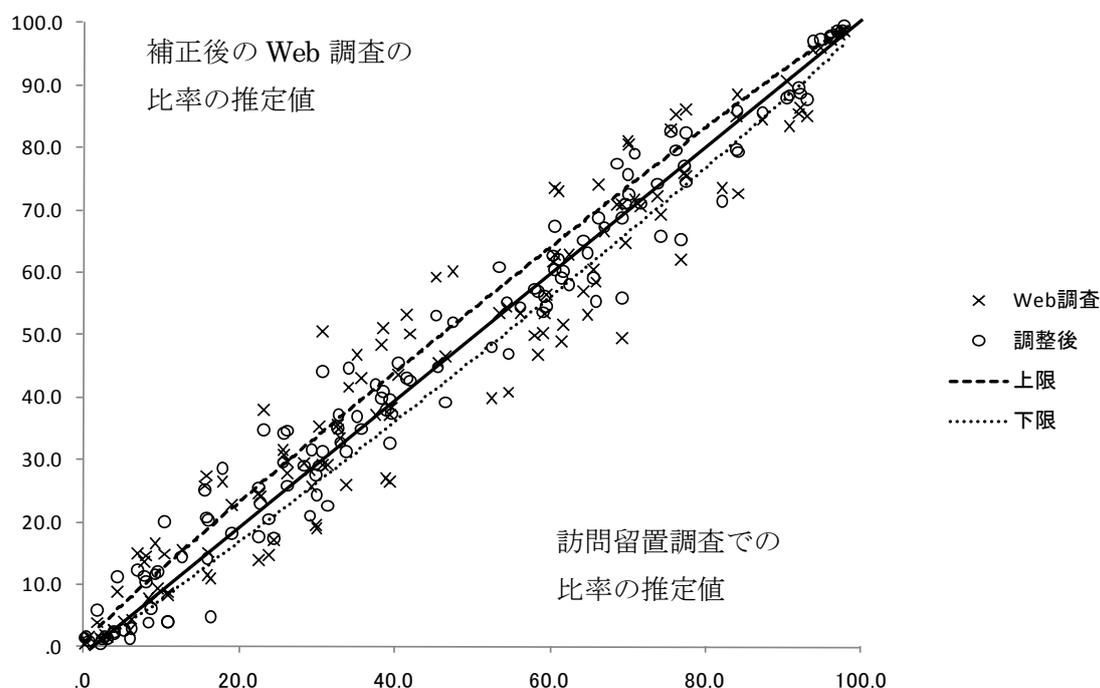


図 3：訪問留置調査と Web 調査のバイアス補正後の推定値の比較

さて、今回の解析で特に補正が有効であったのは「時間割引率」の項目群についてである。この項目での Web 調査の二乗誤差は 3826.775 であり、一方補正後の二乗誤差は 1142.959 と、補正によって二乗誤差が 29.867%に縮減した。また、「危険回避度」の項目群についても Web 調査の二乗誤差は 2236.352 であり、一方補正後の二乗誤差は 1484.357 と、補正によって二乗誤差が 66.374%に縮減した。一方、「労働と報酬」に関する項目では Web 調査の二乗誤差は 393.8149 であり、そもそも二乗誤差は小さい。この項目群についての補正後の二乗誤差は 656.8283 と、補正によって二乗誤差が 66.786%増加しており、調整が有効に作用していないことがわかる。

このような結果になって理由として、以下の 2 点の要素が考えられる。

- [1] 「労働と報酬」については”わからない”という選択肢が存在している
- [2] 二乗誤差が小さい項目では調整が有効に作用しない。

表 2 から明白であるが、“わからない”に対する選択比率において共変量調整による推定値は訪問留置調査における比率から大きく乖離しており、調査モードの影響などの可能性が示唆される。調査モードの影響が存在する場合には、たとえ共変量調整によって「無作

為抽出標本が Web 調査に回答した際の比率の推定値」が正しく推定されたとしても、それが「無作為抽出標本が訪問留置調査に回答した差異の比率」とは大きく異なれば、Web 調査の単純比率よりも二乗誤差がかえって広がる可能性がある。

一方、2 点目にある二乗誤差が小さい項目において調整が有効に作用しない、というのはどうしてだろうか？

このことを考えるために、横軸に各項目の Web 調査の（訪問調査に対する）二乗誤差を、縦軸の補正後の二乗誤差をプロットしたものを図 4 に記載した。図からは、「Web 調査の二乗誤差が小さい項目では両者の二乗誤差にはあまり相関がないこと」「Web 調査の二乗誤差が大きい項目では補正後の二乗誤差が小さくなる」という傾向があることがわかる。この現象は星野・前田(2006)でも指摘されているが、 θ を母数とし、 $\hat{\theta}$ をその推定量とすると、平均二乗誤差は

$$E(\theta - \hat{\theta})^2 = E(\theta - E(\hat{\theta}))^2 + E(\hat{\theta} - E(\hat{\theta}))^2$$

のようにバイアスの二乗の項と推定量の分散に分解できることに起因すると考えられる。すなわち、そもそもバイアスのない部分については標本変動に伴う推定量の分散が二乗誤差の主要素であるために、Web 調査での二乗誤差が小さい部分では補正後の二乗誤差も一定程度存在する⁶。一方、Web 調査での二乗誤差が大きい部分については、Web 調査の結果ではバイアスにくらべて共変量調整によって得られる推定量のバイアスが小さくなるために、このような現象が生じると考えられる。

5. 議論

本研究で行った解析結果からは、これまで行われてきた Web 調査の補正に関する多くの先行研究と同様に、経済学に関する変数を測定する Web 調査においても一定の補正が可能であることが分かった。特に補正のターゲットとしていた大阪大学の訪問留置調査と、関西大学の Web 調査では調査時期が 2 年も異なっているが、それにも関わらず補正が一定程度成功している。両調査の結果の差異は「抽出方法の違い」「調査モードの違い」「調査時期の違い」によって成立していると考えられるが、そのうち「抽出方法の違い」を補正するだけでも、トータルでみた両調査の二乗誤差のうち半分程度を縮減できているということは注目するに値すると考えられる。

但し、今回の解析結果からは、「労働と報酬」に関する項目については補正が作用しなかったが、この原因は前節で考察した 2 つの要因によると考えられる。特に、調査モードの影響は共変量調整法によって除去することはできないため、web 調査を実施する際には、なるべく調整のターゲットとする調査と回答方法が異ならないように注意すべきである。

⁶ 実際には Web 調査での推定量と共変量調整による推定量は相関を有するために、推定量の分散の大きさについても関連はあるはずである。

本研究で利用した Web 調査は RCSS センターによって実施されたものであり、大阪大学の調査と複数の項目で選択肢が異なっており、カテゴリーを合併するなどの処理を行わなければ変数が共通にならないなど、もともと補正することが難しい調査データであった。加えて、訪問留置調査と Web 調査で回答を必須にするかどうかの設定が異なっており、訪問留置調査では無回答による欠測が非常に多かった。このようなことは、共変量項目の利用における欠測の処理を難しくするだけではなく、調査モードの影響、つまり同一調査対象者が両調査において異なる回答を行う可能性を与えるものである。

また、これまでの市場調査や社会調査での先行研究では、収入よりも旅行の頻度などが虚偽や回答拒否が少なく可処分所得や生活水準のレベルを正しく測定しうること、その結果として旅行の頻度等の調査対象者の行動関連項目が補正に有効であることが報告されている(星野,2007)。本研究で利用した両調査ではこのような行動関連項目が調査されておらず、共変量候補という点でもやや不足していた。今後同様の共変量調整を行う場合には上記の観点を考慮した上で Web 調査の設計を行う必要がある。

今後、経済学に関連する項目についての同様の実験的な検討を行っていくことで、より精度の高い補正が可能になると考えられる。

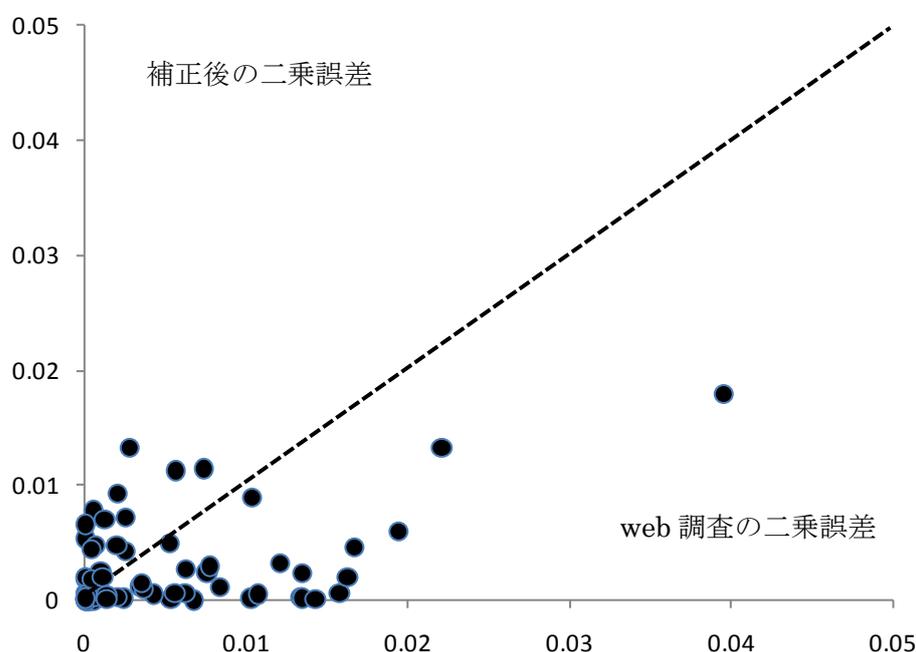


図 4 : Web 調査での推定量と補正後の推定量の二乗誤差の関係
*但し二乗誤差を計算する前に%表示の比率を 100 で除している。

				訪問留置調査	web調査	補正後	
時間 割 引 率	A	B	X	利率(%)			
	2日	9日	1万円	-10	97.3271%	98.0534%	
	受け取り			0	93.8230%	95.6062%	
				10	64.8179%	53.2814%	
				20	61.5436%	48.9989%	
				50	54.6823%	40.7675%	
				100	39.3322%	26.4182%	
				200	29.9933%	19.0211%	
				300	22.5258%	13.8487%	
		90日	97日	1万円	-10	97.0270%	98.3871%
		受け取り			0	94.8381%	96.0512%
				10	61.6863%	51.5573%	
				20	58.3868%	46.7742%	
				50	52.5143%	39.7664%	
				100	38.8589%	26.9188%	
				200	29.9223%	19.5773%	
				300	23.8705%	14.7386%	
		1ヶ月	13ヶ月	1万円	-5	97.6775%	98.7764%
		受け取り			0	96.1253%	97.1635%
				2	69.6531%	64.7942%	
			4	64.2857%	57.0078%		
			6	58.0776%	49.8888%		
			10	33.7151%	25.8621%		
			20	24.4272%	17.1858%		
			40	15.9152%	11.4572%		
	1ヶ月	13ヶ月	100万円	-5	97.8033%	98.7208%	
	受け取り			0	96.1188%	97.3860%	
			0.1	68.5917%	70.9121%		
			0.5	53.4797%	53.5595%		
			1	37.5676%	37.1524%		
			2	30.8030%	29.1991%		
			6	16.0256%	15.1279%		
			10	9.4708%	9.3437%		
	1ヶ月	13ヶ月	100万円	-5	25.8283%	30.8120%	
	支払い			0	45.5683%	45.5506%	
			0.1	60.4329%	61.6240%		
			0.5	66.9712%	66.5184%		
			1	71.5928%	70.5784%		
			2	73.7821%	72.2469%		
			6	77.3342%	75.9733%		
			10	77.4434%	75.4727%		

表1：時間割引率に関する項目の訪問留置調査、Web調査、及び補正後の推定値の比較
 ＊質問形式は「A時点で金額Xを受け取るか、それ以降のB時点でXに利息が加わった部分を受け取るか」の2枝選択であるため、「A時点で金額Xを受け取る」ことを選択する比率を記載している。

謝辞

本研究は、大阪大学 21 世紀 COE プロジェクト「アンケートと実験によるマクロ動学」によって実施された「くらしの好みと満足度についてのアンケート」の結果を利用している。本アンケート調査の作成に寄与された、筒井義郎、大竹文雄、池田新介の各氏に感謝する。関西大学ソシオネットワーク戦略研究センターによって実施されたデータを提供頂いた。センター長である鶴飼康東氏、調査実施にあたった竹村敏彦氏に感謝する。

		訪問留置調査	web調査	補正後	
危険回避傾向	報酬の支払方法	2倍もしくは30%減 確実に5%増し	12.6630% 87.3370%	15.5172% 85.5376%	14.4624% 85.5376%
	報酬の支払方法2	60%増しか10%減 確実に5%増し	34.0971% 65.9029%	41.6018% 58.3982%	44.7194% 55.2806%
	金利0%による1年め2年めの消費量の違い	1年めにより多く消費する 2年めにより多く消費する	30.7420% 69.2580%	50.6118% 49.3882%	44.1511% 55.8489%
	金利2%による1年め2年めの消費量の違い	1年めにより多く消費する 2年めにより多く消費する	23.1858% 76.8142%	38.0423% 61.9577%	34.6957% 65.3043%
	金利4%による1年め2年めの消費量の違い	1年めにより多く消費する 2年めにより多く消費する	15.8342% 84.1658%	27.4194% 72.5806%	20.7224% 79.2776%
	金利6%による1年め2年めの消費量の違い	1年めにより多く消費する 2年めにより多く消費する	9.1746% 90.8254%	16.6296% 83.3704%	11.6357% 88.3643%
	金利8%による1年め2年めの消費量の違い	1年めにより多く消費する 2年めにより多く消費する	8.0468% 91.9532%	14.5717% 85.4283%	10.4237% 89.5763%
	金利10%による1年め2年めの消費量の違い	1年めにより多く消費する 2年めにより多く消費する	7.7982% 92.2018%	13.6819% 86.3181%	11.3408% 88.6592%
	10年間の生活レベルへの分配	最初の生活レベルを高くし、低くしていく 毎年同じ生活レベル	1.7775% 65.5695%	3.8932% 60.5117%	5.8854% 59.0688%
		最初の生活レベルを低くし、高くしていく	32.6531%	35.5951%	35.0457%
	2000円当たる宝くじを買うか	買う 買わない	60.5514% 39.4486%	62.8476% 37.1524%	60.4895% 39.5105%
	10万円当たる宝くじを買うか	買う 買わない	59.5207% 40.4793%	56.3960% 43.6040%	54.4836% 45.5164%
	2000円当たる宝くじを売るか	売る 売らない	17.8477% 82.1523%	26.4182% 73.5818%	28.5387% 71.4613%
	1/2で2000円の損失に対して200円の保険料を払うか	支払う 支払わない	70.8748% 29.1252%	71.6352% 28.3648%	79.0115% 20.9885%
	半々で当たり2000円受取、外れ1000円支払のくじをもらう	もらう もらわない	6.9821% 93.0179%	14.9055% 85.0945%	12.2136% 87.7864%
	宝くじに当選したら、仕事量を変えるか	増やす	0.6859%	1.6393%	0.9524%
		変えない	56.1957%	53.5519%	54.3913%
		減らす	26.2003%	29.6903%	34.6210%
		すぐに仕事を辞めて働かない	8.5963%	7.3770%	6.2009%
		わからない	8.3219%	7.7413%	3.8344%
賃金が今年1年半分になったら、仕事量を変えるか	増やす	29.3285%	25.6831%	31.5660%	
	変えない	32.8004%	34.8816%	37.1614%	
	減らす	15.6236%	25.7741%	25.0631%	
	すぐに仕事を辞めて働かない	5.9388%	2.6412%	1.4092%	
	わからない	16.3088%	11.0200%	4.8004%	
賃金が永久に2倍になったら仕事量を変えるか	増やす	19.0476%	22.6919%	18.1368%	
	変えない	59.3407%	53.3370%	56.2334%	
	減らす	10.3896%	14.8498%	20.0211%	
	すぐに仕事を辞めて働かない	0.3330%	0.7786%	1.6070%	
	わからない	10.8891%	8.3426%	4.0017%	
賃金が今年1年2倍になったら、仕事量を変えるか	増やす	25.6784%	31.5907%	29.5347%	
	変えない	59.0668%	50.2781%	53.6294%	
	減らす	4.4011%	8.8432%	11.3024%	
	すぐに仕事を辞めて働かない	0.0993%	0.6118%	1.4452%	
	わからない	10.7545%	8.6763%	4.0883%	

表2：「危険回避傾向」「労働と報酬」に関する項目の訪問留置調査、Web調査、及び補正後の推定値の比較

引用文献

- Copas, J. and Eguchi, S. (2005). "Local model uncertainty and incomplete data bias (with discussion)", *Journal of the Royal Statistical Society, ser. B*, **67**, 459-512.
- Couper, M.P. (2000). "Web surveys: a review of issues and approaches". *Public Opinion Quarterly*, **64**, 464-494.
- 大隅昇 (2002). インターネット調査. 社会調査ハンドブック (林知己夫編) pp.200-240. 朝倉書店.
- Hoshino, T., Kurata, H., and Shigemasu, K. (2006). "A Propensity Score Adjustment for Multiple Group Structural Equation Modeling", *Psychometrika*, **71**, 691-712.
- Hoshino, T. (2007). "Doubly Robust type Estimation for Covariate Adjustment in Latent Variable Modeling", *Psychometrika*, **72**, 535-549.
- Hoshino, T. (2008). "A Bayesian Propensity Score Adjustment for Latent Variable Modeling and MCMC algorithm", *Computational Statistics & Data Analysis*, **52**, 1413-1429.
- 星野崇宏 (2005). "欠測群の周辺分布の母数に対する傾向スコアを用いた重み付き M 推定量の提案と介入効果研究への応用" 行動計量学. 第 32 巻 2 号. 121-132.
- 星野崇宏・前田忠彦 (2006). "傾向スコアを用いた補正法の有意抽出による標本調査への応用と共変量の選択法の提案" 統計数理, 第 54 巻 1 号 191-206.
- 星野崇宏 (2007). "インターネット調査に対する共変量調整法のマーケティングリサーチへの適用と調整効果の再現性の検討" 行動計量学, 第 34 巻 1 号 33-48.
- 星野崇宏・森本栄一 (2007). "インターネット調査の偏りを補正する方法について: 傾向スコアを用いた共変量調整法" 「Web マーケティングの科学—リサーチとネットワーク—」(井上哲浩・日本マーケティングサイエンス学会 編) 第 1 章 27-59 千倉書房
- 星野崇宏 (2009). 「調査観察データの統計科学: 因果推論・選択バイアス・データ融合」 岩波書店
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment Under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.
- Krugman, K. (1991). "History Versus Expectations", *Quarterly Journal of Economics*, **106**, 651-667.
- 前田忠彦・土屋隆裕(2001). 「日本人の国民性 2000 年度吟味調査報告」 統計数理研究所 研究レポート, 87.
- Rosenbaum, P.R., and Rubin, D.B. (1983). "The central role of the propensity score in observational studies for causal effects", *Biometrika*, **70**, 41-55.
- Schonlau, M., van Soest, A., Kapteyn, A., and Couper, M. (2009). "Selection Bias in Web Surveys and the Use of Propensity Scores", *Sociological Methods & Research*, **37**,

219-238.

Taylor, H. (2000). "Does internet research work?" *International Journal of Market Research*, **42**, 51-63.

竹村和久 (2009). 「行動意思決定論」 日本評論社

Tversky, A., & Kahneman, D. (1971). "Belief in the law of small numbers", *Psychological Bulletin*, **76**, 105-110.

鵜飼康東・村上雅俊 (2008). 「国民年金納付者行動 Web アンケート結果の概要と探索的検討」 RCSS Discussion Paper Series.

山岡和枝・林知己夫 (1999). 「電話帳記載・非記載者をめぐる諸問題：首都圏調査から」 行動計量学, **26**, 114-124.