

「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

松岡 亮二¹・前田 忠彦²

(受付 2014 年 10 月 8 日；改訂 2015 年 10 月 13 日；採択 10 月 13 日)

要 旨

「日本人の国民性第 13 次(2013 年)全国調査」に加え、「調査員事後アンケート」および国勢調査のデータを用い、個人と地点の 2 水準を考慮した多項ロジット・マルチレベル分析によって個人・地点・調査員の特性と回収状況(有効回答, 接触不能, 本人拒否, 他者拒否)の関連を実証的に検討した. 先行研究が示す若年層と男性の低回収傾向に加え, 以下の 4 点が明らかになった. 参照カテゴリである有効回答と比較したとき, (1)対象者の住居形態が一戸建であることは, 本人拒否と他者拒否の増加と関連している, (2)地点間に回収状況の差異があり, 中でも, 接触不能に比較的大きな地点間差異がある, (3)国勢調査に基づく町丁字水準の変数が, 地点間の回収状況の差異と関連している, (4)調査員の経験年数が長いと接触不能, 本人拒否, 他者拒否が減る傾向がある. 調査主体が制御可能なのは調査員のみであるので, 今後も調査員の特性と回収状況に関する実証的知見を蓄積し, 回収率の向上に寄与する方策を模索することが求められる.

キーワード：日本人の国民性調査, 欠票分析, 社会調査, 調査員特性, 国勢調査, マルチレベル分析.

1. はじめに

1.1 低下する社会調査の回収率

都市部における拒否や接触不能の増加(坂元, 2001; 玉野, 2003)により社会調査・世論調査の回収率は 1970 年代から長期的に低下傾向(Inaba, 2007; Synodinos and Yamada, 2013)にあり, 2005 年を境に特に低下している(保田 他, 2008; 篠木, 2010). 個人情報関連の犯罪の増加や 2005 年に施行された個人情報の保護に関する法律(いわゆる個人情報保護法)により, かつて回収率の高かった女性や高齢者が調査を忌避し低下傾向に拍車がかかったと考えられる(Inaba, 2007). この傾向は日本における代表的な継続社会調査の一つである「日本人の国民性調査」においても例外ではなく, 回収率は第 7 次(1983 年)調査時に 74%であったものが, 第 9 次(1993 年)調査に 69%となり, 以降 5 年毎に 64%, 56%, 52%と低下した. そして, 本論文の素材である「日本人の国民性第 13 次(2013 年)全国調査」では 50%と, 1993 年に比べて 20 ポイント近く低下している. これは, 主に調査拒否と一時的な不在の増加が原因となっている.

未回収がランダムに起きているわけではない(Groves and Couper, 1998 など)以上, どのような特性を持つ対象者で特に有効票を回収できないかは, 調査不能バイアスの大きさを評価する

¹ 早稲田大学 高等研究所：〒 169-8050 東京都新宿区西早稲田 1-6-1

² 統計数理研究所：〒 190-8562 東京都立川市緑町 10-3

基礎情報として実証的に把握する必要がある。そこで本稿は、「日本人の国民性第 13 次全国調査」の個票データを用い、欠票と関連する個人特性を明らかにする。また、調査員対象に行われた「第 13 次全国調査 調査員事後アンケート」から調査員の特性、2010 年度の国勢調査から町丁目単位の地点の都市度を示す変数を作成し、調査員・地点の特性と回収状況の関連も検証する。なお、同調査における調査不能理由の分布については、中村 他(2015)を参照されたい。

1.2 欠票分析の必要性

個人特性、それに個人が居住している近隣(neighborhood)が調査回答と関連していることが海外(Bethlehem et al., 2011; Groves, 2006 など)だけではなく日本(埴淵 他, 2012)でも明らかにされてきた。中でも、個人水準の特性である対象者の性別と年齢との関連は繰り返し指摘されてきた。具体的には、若年層の男性で未回収が多く(Synodinos and Yamada, 2000; 前田, 2005; 保田, 2008)、未回収を接触不能と拒否に区別すると接触不能によるものが主(田辺, 2003; 保田, 2008)であり、家族と同居していると思われる一戸建住宅であると特に不在であることが多い(三輪, 2008)。また、住居形態が一戸建であると(おそらく)家族による調査拒否が増加し(三輪, 2008)、集合住宅であると接触不能になる傾向がある(三輪, 2008; 埴淵 他, 2012)。接触者のうち拒否する対象者は男性に多く、特にその割合が高いのは男性の 30 代から 50 代前半である(田辺, 2003)。全国規模の調査データをマルチレベルモデルによって分析した近年の研究(Matsuoka and Maeda, 2015)においても、若年層の低回収率傾向や男性の不在傾向は確認されている。なお、接触者のみに限定した分析だと本人による拒否に性差はなく、家族による拒否は男性対象者のほうが多い傾向にある(Matsuoka and Maeda, 2015)。

近隣(neighborhood)水準については、過去 20 年、大都市では一貫して回収率が低く(窪田, 2008)、性別や年齢といった個人特性を統制しても都市化の程度と調査の回収には関連がある(前田, 2008; 三輪, 2008; 埴淵 他, 2012; 山内, 2012)。この傾向は調査員属性を統制しても接触不能と拒否の両方(保田, 2008)で確認されている。都市部に居住する個人が調査拒否する傾向がある一方で、接触不能については住居形態を統制すると都市部の影響は消失するという報告(三輪, 2008)もある。

埴淵 他(2011)に基づき、7 段階の都市規模と 12 カテゴリーの都市特性分類を用い、全国規模の調査である日本版総合的社会調査(Japanese General Social Surveys, JGSS; 大阪商業大学 JGSS 研究センターによる)に対する回収状況との関連をマルチレベル分析によって検討した近年の研究(埴淵 他, 2012)によると、都市規模や個人特性を統制しても、都市分類は接触・接触不能、それに有効回答・拒否とそれぞれ関連していた。また、統計数理研究所による他の全国規模の調査データにマルチレベル分析を適用した研究(Matsuoka and Maeda, 2015)は、国勢調査に基づく市町村の都市度を代理する指標が回収率と関係していることを示した。特に、認知犯罪件数と大学卒者の割合が高い地点で有効回答の割合が下がり、女性である、あるいは大きい住居に住んでいる対象者は認知犯罪件数が多い地点であるとより本人による調査拒否をする傾向がある(Matsuoka and Maeda, 2015)。

2. 目的

2.1 本稿のねらい

先行研究(Groves and Couper, 1998 など)が示すように欠票はランダムに発生しているわけではない。どのような個人、地点、調査員の特性と欠票が関連しているかという点は、調査不能バイアス評価の基礎情報とするためにも明らかにする必要があるが、日本の社会調査においてこれまでに十分に実証的な検討がされているとは言いがたい。全国規模の調査データを用い、

マルチレベルモデルを適用した近年の研究でも、都市規模と回収状況の関連は確認されているが、埴淵 他(2012)が用いた分類はカテゴリ数が多いため結果の解釈が難しく、Matsuoka and Maeda(2015)は町丁字等のレベル(一部は市町村レベル)の地点変数を用いてその欠票への寄与の分析を行ったものだが、埴淵 他(2012)などが分析に用いた調査員特性が含まれていない。そこで本稿は、先行研究に基づき、日本における典型的な全国規模の社会調査である「日本人の国民性第13次(2013年)全国調査」に加え、同調査の「調査員事後アンケート」、それに国勢調査の町丁字単位のデータを用い、個人、地点、それに調査員の特性と、回収状況(有効回答、接触不能、本人拒否、他者拒否)分類の定義は3.2節にて詳述)との関連について、個人と地点の二水準を考慮したマルチレベルモデルによって分析を行う。

2.2 研究課題

本稿の研究課題(Research Question)は、“個人、調査地点、それに調査員の特性と、非回収状況(接触不能、本人拒否、他者拒否)に関連があるか”である。先行研究に基づき、個人、地点、調査員それぞれの特性が、どの種類の非回収状況と関連しているのかについて、以下の仮説を立てた。

仮説1(対象者の属性の寄与)。

- 仮説1-1. 男性より女性のほうが、年齢が低いより高いほうが、より本人拒否が起りやすい。
- 仮説1-2. 男性であることと年齢が低いことは、接触不能と他者拒否に対する説明要因となる。
- 仮説1-3. 居住する住居形態が一戸建であることが、他者拒否に対する説明要因となる。

若年層の男性は未回収が多く(Synodinos and Yamada, 2000; 前田, 2005; 保田, 2008)、主に接触不能によるもの(田辺, 2003; 保田, 2008)と考えられる。一方、若年層と比べて高齢層、それに男性と比べて女性は、より拒否になりやすい(稲葉, 2010)と想定される。また、一戸建は、同居している家族に接触の際に拒否されていると思われる(三輪, 2008)ので、拒否の中でも〈他者拒否〉の増加と関連する。

仮説2(地点特性変数の寄与)。都市度が高い地点であるほど、接触不能、本人拒否、他者拒否が増加する。

先行研究(前田, 2008; 三輪, 2008; 埴淵 他, 2012; 山内, 2012)が示す通り都市度は有効回答の減少と関連すると考えられる。具体的には、都市度を代理的に示す地点特性として人口密度、第一次産業従事者割合、それに大学卒者割合が、非回収状況と関連する。三輪(2008)は、市郡規模は接触不能については関連がないが接触者の中での拒否とは関連すると論じたが、より近年の研究(埴淵 他, 2012; Matsuoka and Maeda, 2015)は都市の種類・特性と接触・拒否の双方で関連すると報告している。よって、町丁字単位の都市度は、接触不能と2種類の拒否(本人・他者)と関係している一すなわち、地点の都市度が高いほど、接触不能と拒否は増加する—と想定される。

仮説3(調査員特性の寄与)。調査員経験が長いほど、接触不能、本人拒否、他者拒否が減少する。

調査員による回収率の差異に関する先行研究の結果は一致していない。田辺(2003)は同じ地点を担当した調査員による拒否率の差から、調査員による差異を指摘している。一方、保田(2008)は、ロジスティック回帰モデルを利用した解析の結果から、調査員の経験年数は回収状況と関連しないと論じている。埴淵 他(2012)によるマルチレベル・ロジスティック回帰分析の結果では、使用データの年次によって結果が異なり明快な傾向を導き出すことはできていない。調査

員経験年数による地点間の非回収状況に差異が存在し、経験年数が長いほうが接触不能と拒否を抑制し、結果的に有効回答の多さに繋がっているとの想定の当否を検討することとする。

3. 研究手法

3.1 データ

研究課題を実証的に検討するため、「第 13 次全国調査」の計画標本全体についてのデータを用いる。標本設計については中村 他(2015)に詳述されている。要約すれば、標本設計は層化二段無作為抽出であり、層毎の二段抽出の際の第一次抽出単位は、国勢調査の(小地域集計の単位として用いられる)町丁字等、第二次抽出単位が個人であり、各抽出地点内では住民基本台帳から等間隔抽出により個人が選ばれている。これに加えて、地点特性としては調査員対象に行われた「調査員事後アンケート」、それに 2010 年度の国勢調査の小地域集計(町丁字等別集計)データ(政府統計の総合窓口(e-Stat, 2015a))を用いた。また、町丁字等別の面積を「地図で見る統計(統計 GIS)；政府統計の総合窓口(e-Stat, 2015b)」から入手し、町丁字単位の地点の都市度を示す「人口密度」変数を作成した。

なお、「第 13 次(2013 年)全国調査」では、接触不能や回答拒否の場合であっても、調査員が抽出された 6400 人のサンプル(計画標本)すべての住居特性を現地で記録している。大規模な社会調査を用いた欠票分析(たとえば埴淵 他, 2012)でも、年齢と性別しかサンプルの個人特性として含まれないことがあるので、調査員による訪問時に記録される補助的情報(Kreuter and Olson, 2013)が先行研究にない知見をもたらす可能性がある。その具体例を本稿でも示すことになるが、たとえば住居の特徴が欠票となるか否かについての説明力を持ちうるからである。

3.2 使用変数

表 1(後掲)に示したように、被説明変数を有効回答と 3 種類の非回収状況(接触不能, 本人拒否, 他者拒否)に分類した。以下、本データの被説明変数の 4 つの状態については「 $\langle \quad \rangle$ 」で括弧で示す。また、以下の記述で「#」記号を伴う数字は日本人の国民性調査で共通に用いる項目の整理番号を表す。 \langle 接触不能 \rangle は移転, 長期不在, 一時不在を意味し, 死亡(0.2%), 該当者なし(0.8%), 尋ね当たらず(住所不明)(1.1%), 病気(2.4%), 老すい(0.8%), その他(.0%)は分析から除外した。これらの欠損値の合計数は 346 ケースで, サンプル全体(6400 ケース)の 5.4%にあたる。 \langle 本人拒否 \rangle は本人に調査現地で面会した際の拒否, \langle 他者拒否 \rangle は本人に調査現地で面会した上で拒否された場合以外の拒否であり, 通常は家族によるものであるが, 事前の電話連絡によって拒否された場合など, 調査員の本人接触状況が確認できていないものも含んでいる。電話連絡による拒否を, \langle 他者拒否 \rangle にまとめる積極的な根拠はないが, 本人との調査現地での接触が可能でその上で本人が拒否したケースを明確に区分することを目的としてこの方針とした。なお, 日本の国民性調査において \langle 本人拒否 \rangle と \langle 他者拒否 \rangle を分けたのは「第 13 次(2013 年)調査」が最初である。

個人水準の説明変数は, 表 1 に示したように 8 項目からダミーコーディングにより 12 変数を作成した。性別(#1.1)は女性を 1(X_1)とし, 男性を 0(参照カテゴリ)とした。年齢(#1.2)は 20 代を参照カテゴリとする 5 つのダミー変数($X_2 \sim X_6$)で表されている。これらは対象者名簿からの情報であり, また対象者個人の特性と言える。その他の変数は対象者の住居に関する調査員の観察記録に基づいており, 対象者の住居に関する特性である。一戸建てか否か, オートロックの有無, 敷地内駐車スペースの有無, 住居の様子: 新しいか否か, 住居の様子: 広い・大きいか否か, 住居の様子: 表札の有無, に関する 2 値変数($X_7 \sim X_{12}$)となっている。以上により参照カテゴリを除き, 個人水準の説明変数が 12 個用意されることになる。同じく表 1 に地点水

準として示されているのは調査員の経験年数を意味する2つのダミー変数(Z_1, Z_2)で、参照カテゴリは調査員経験が6年未満である。

後掲の表2には地点水準における連続変数の内容を示してある。これらは2010年度の国勢調査による小地域集計に基づいた地点情報で、市町村ではなく町丁字単位のデータを用いている。実際には最終抽出単位である個人の系統抽出の手続きに際して、一つの町丁字に収まらずに隣接の地域に移動する場合があり、このときは同一地点内に複数の町丁字が含まれることになるが、地理的に近接した町丁字等の性格が似ていることを踏まえて、本論文では抽出起点となった町丁字のデータを地点特性として用いている。なお、Matsuoka and Maeda(2015)は国勢調査以外の公的統計(たとえば「統計でみる市区町村のすがた」による)に基づく市区町村レベルでの集計値も利用しているが、本稿では町丁字等のより細かい地域区分での地域特性の効果を検討するために、国勢調査情報のみを使用している。これらの連続変数は都市度を代理的に示し、地点水準で標準化(平均0, 標準偏差1)された。人口密度 Z_3 と大学卒者割合 Z_5 が高いほど都市度が高く、第一次産業従事者割合 Z_4 が高いほど都市度が低いと考えられる。事実、3つの地域水準の変数の相関関係は、人口密度が第一次産業従事者割合と負の関係(-0.436)、大学卒者割合とは正の関係(0.536)にあり、第一次産業従事者割合と大学卒者割合も負の関係(-0.465)である。なお、分析結果の解釈を容易にするため、標準化した変数を分析に用いた。

3.3 分析モデル

個人、地点、調査員の特性と非回収状況に関連があるか検討するため、多項ロジット・マルチレベル分析を行った。まずは、級内相関係数(intra-class correlation coefficient)を確認するため、回収状況を示す変数を被説明変数とし、ヌルモデル(説明変数を含まないランダム切片モデル)による分析を行った。その後、レベル1(個人水準)に説明変数を加えたモデル、レベル2(地点水準)に地点特性に関わる変数と調査員変数を追加したモデルを作成した。地点によって個人特性の効果の大小に差異があるか検証するため、ランダム傾きの有無を確認し、クロスレベルや同水準の交互作用項など探索した上で最終モデルを確定した。被説明変数は〈有効回答〉を参照カテゴリとし、〈接触不能(移転, 長期不在, 一時不在)〉, 〈本人拒否〉, 〈他者拒否〉と比較するランダム切片モデルを用いた。なお、分析はすべてHLM7.0(Raudenbush et al., 2011)によって行った。この分析の代わりに、〈接触不能〉に加えて本人以外の拒否も除外したケースでの分析を行うことも意味があるが、ここではこの二つの先行研究との相互参照が容易となることを優先した。

レベル1(個人水準)モデル: 3つのカテゴリ($k = 1, 2, 3$)について共通

検討したランダム切片モデルにおいて、個人水準の m 番目の説明変数についての地点 j 個人 i の測定値を X_{mij} と表記する($m = 1, 2, \dots, 12$)。個人水準のモデルでは、第 j 地点の第 i 番目の対象者が、地点毎の回帰係数ベクトル β_j (後述)が与えられたもとで被説明変数(Y)の第 k カテゴリに正反応(該当)する確率を ϕ_{kij} として、参照カテゴリの正反応確率に対する比の対数が次のように回帰式により表現される:

$$\log[\phi_{kij}/\phi_{0ij}] = \beta_{0j(k)} + \sum_{m=1}^{12} \beta_{mj(k)} X_{mij} \quad (k = 1, 2, 3)$$

ただし、左辺の対数内の確率は

$\phi_{1ij} = \text{Probability}(Y = 1|\beta_j)$: 〈接触不能〉の確率

$\phi_{2ij} = \text{Probability}(Y = 2|\beta_j)$: 〈本人拒否〉の確率

$\phi_{3ij} = \text{Probability}(Y = 3|\beta_j)$: 〈他者拒否〉の確率

$\phi_{0ij} = \text{Probability}(Y = 0|\beta_j) = 1 - \phi_{1ij} - \phi_{2ij} - \phi_{3ij}$: 〈有効回答〉の確率(参照カテゴリ)

であり, $\beta_{0j(k)}$ は第 k カテゴリに対する回帰式における切片項, $\beta_{mj(k)}$ は m 番目の説明変数に対する傾きである. なお 3 本の各回帰式の地点毎の回帰係数をまとめて, $\beta_{j(k)} = [\beta_{0j(k)} \beta_{1j(k)} \cdots \beta_{12j(k)}]'$ ($k = 1, 2, 3$) とすると, 前述の回帰係数ベクトルは $\beta_j = [\beta'_{j(1)} \beta'_{j(2)} \beta'_{j(3)}]'$ と更にまとめて示したものである ($'$ は転置記号).

レベル 2 (地点水準) モデル: 3 つのカテゴリ ($k = 1, 2, 3$) について共通

地点水準のモデルでは, 傾きを固定効果として, 切片のみにランダム効果を仮定し, 5 つの地点水準の説明変数 $Z_1 \sim Z_5$ を導入する (t 番目の地点水準説明変数の地点 j の値を Z_{tj} とする; $t = 1, 2, \dots, 5$).

ランダム切片:

$$\beta_{0j(k)} = \gamma_{00(k)} + \sum_{t=1}^5 \gamma_{0t(k)} Z_{tj} + u_{0j(k)} \quad (k = 1, 2, 3)$$

ただし, $\gamma_{0t(k)}$ は t 番目の説明変数に対する回帰係数, $u_{0j(k)}$ は残差項である.

固定傾き:

$$\beta_{mj(k)} = \gamma_{0m(k)} \quad (k = 1, 2, 3; m = 1, 2, \dots, 12)$$

4. 分析結果

個人水準と地点水準のカテゴリカルな変数の頻度 (表 1) と地点の連続変数の記述統計量 (表 2) を確認した.

地点間差異の程度を明らかにするため, 説明変数を含まないランダム切片モデルであるヌルモデルによる多項ロジット・マルチレベル分析によって級内相関係数を求めた (ロジットによる級内相関については, たとえば, Snijders and Bosker, 2012, pp. 304–305). その結果, 〈接触不能〉は 0.123, 〈本人拒否〉は 0.078, 〈他者拒否〉は 0.074 であった. これらは全分散のうち地点間分散の寄与分であり, 残りは個人間分散であることを意味している.

地点と個人水準双方に説明変数を追加した最終モデルの分散成分を基に級内相関係数を算出した結果, 地点間の分散は〈接触不能〉は 0.104, 〈本人拒否〉は 0.055, 〈他者拒否〉は 0.071 となった. ヌルモデルと比較すると, 減少幅は大きくないが, 説明変数の追加によって減少したことが確認できる. 表 3 は参照カテゴリである〈有効回答〉と比較した, 〈接触不能〉, 〈本人拒否〉, 〈他者拒否〉それぞれの結果を示している. 地点水準変数と個人水準変数が非回収状況のそれぞれのカテゴリと有意に関連している. 以下 3 つのカテゴリ毎に主な結果を述べる. 以下の記述では, 2 値のロジスティック回帰分析のオッズ比に対応させ, 参照カテゴリに対する確率比 $[\phi_{kij}/\phi_{3ij}]$ の変化によって係数の効果が評価される (単に確率比が \sim 倍などと記す). 多項ロジットモデルは関心下のカテゴリを参照カテゴリと比較した場合の寄与要因をモデル化したものであり, 各説明変数の効果を表す確率比も常にこの比較に対する寄与度である (当該カテゴリ単独で考えた割合の大小に対する直接の評価ではない) 点に留意しながら, 以下の結果を解釈する.

〈接触不能〉

〈有効回答〉と比較して〈接触不能〉を予測するモデルでは, 3 つの地点特性変数が有意となった. 人口密度と大学卒者割合は正の係数, 第一次産業従事者割合は負の係数となっていることから, このモデルでは都市度が高いほど〈接触不能〉となっていることがわかる. 人口密度が平均よりも 1 標準偏差高いと, 〈有効回答〉と比べた〈接触不能〉の確率比は 1.341 倍 ($p < 0.001$) である (以下では, 係数である確率比が有意な場合についてだけ述べる. p 値については表 3 を参照のこと). 同様に, 説明変数の平均とそれよりも 1 標準偏差高い場合との比較で評価した場合, 第一次産業従事者割合が高いと〈接触不能〉の確率比は 0.859 倍, 大学卒者割合が高いと 1.130 倍

表1. 個人水準および地点水準のカテゴリ変数の分布.

	度数	%		度数	%
個人水準(6054人)変数					
被説明変数					
有効回答(0:参照)	3170	52.4	本人拒否(2)	1054	17.4
接触不能(1)	989	16.3	他者拒否(3)	841	13.9
説明変数					
個人特性					
性別			オートロックの有無		
女性(X ₁)	3047	50.3	オートロック(X ₈)	644	10.6
男性(参照)	3007	49.7	その他(参照)	5410	89.4
年齢			敷地内駐車スペース		
20代(参照)	823	13.6	有り(X ₉)	3632	60.0
30代(X ₂)	1084	17.9	その他(参照)	2422	40.0
40代(X ₃)	1065	17.6	住居の様子(新しさ)		
50代(X ₄)	1008	16.7	新しい(X ₁₀)	1542	25.5
60代(X ₅)	1146	18.9	その他(参照)	4512	74.5
70歳以上(X ₆)	928	15.3	住居の様子(広さ・大きさ)		
			広い・大きい(X ₁₁)	1437	23.7
			その他(参照)	4617	76.3
住居特性			住居の様子(表札)		
形態			有り(X ₁₂)	4052	66.9
一戸建(X ₇)	4041	66.7	その他(参照)	2002	33.1
その他(参照)	2013	33.3			
地点水準(400地点)説明変数					
調査員総経験年数					
6年未満(参照)	53	13.3	10年以上(Z ₂)	285	71.3
6年以上10年未満(Z ₁)	62	15.5			

表2. 地点水準(町丁字等单位)における地点特性を示す連続変数の記述統計.

	数	最小値	最大値	平均	標準偏差	歪度	尖度
標準化前							
人口密度(Z ₃)	400	15.400	35359.800	7361.035	7101.583	1.030	0.401
第一次産業従事者割合(Z ₄)	400	0.000	0.530	0.0406	0.08217	3.180	11.285
大学卒者割合(Z ₅)	400	0.010	0.510	0.1616	0.09021	0.885	0.683
標準化後							
人口密度(Z ₃)	400	-1.034	3.943	0	1	1.030	6.76
第一次産業従事者割合(Z ₄)	400	-0.494	5.986	0	1	3.180	11.285
大学卒者割合(Z ₅)	400	-1.678	3.827	0	1	0.885	0.683

であった。調査員経験年数も、〈有効回答〉と比較して〈接触不能〉となるかに関係がある。ただし、調査員経験年数が6年未満の調査員と比べて6年以上10年未満であると確率比は0.675倍であるが、10年以上だと0.796倍で有意ではなかった。

表 3. 多項ロジット・マルチレベル分析の結果.

固定効果	接触不能			本人拒否			他者拒否		
	係数	標準 誤差	確率 比	係数	標準 誤差	確率 比	係数	標準 誤差	確率 比
地点水準(400地点)変数									
切片	-1.372 ***	0.053	0.254	-1.131 ***	0.045	0.323	-1.423 ***	0.049	0.241
地点特性									
人口密度	0.294 ***	0.068	1.341	0.260 ***	0.054	1.297	0.116 *	0.052	1.123
第一次産業従事者割合	-0.152 *	0.065	0.859	-0.073	0.053	0.930	-0.098	0.063	0.907
大学卒者割合	0.122 *	0.058	1.130	0.063	0.050	1.065	0.124 *	0.055	1.132
調査員経験年数(参照カテゴリ:6年未満)									
6年以上10年未満	-0.393 †	0.206	0.675	-0.183	0.177	0.833	-0.569 **	0.185	0.566
10年以上	-0.228	0.177	0.796	-0.303 *	0.146	0.739	-0.396 **	0.150	0.673
個人水準(6054人)変数									
個人特性									
女性	-0.459 ***	0.080	0.632	0.098	0.076	1.103	-0.687 ***	0.081	0.503
年齢(参照:20代)									
30代	-0.365 **	0.128	0.695	-0.004	0.162	0.996	-0.683 ***	0.137	0.505
40代	-0.683 ***	0.147	0.505	0.272 †	0.145	1.312	-0.694 ***	0.134	0.500
50代	-1.033 ***	0.148	0.356	0.311 *	0.146	1.365	-1.034 ***	0.137	0.355
60代	-1.451 ***	0.157	0.234	0.450 **	0.147	1.568	-1.383 ***	0.140	0.251
70歳以上	-1.883 ***	0.175	0.152	0.025	0.151	1.026	-2.009 ***	0.190	0.134
住居特性									
形態:一戸建	-0.129	0.160	0.879	0.266 †	0.159	1.305	0.529 **	0.196	1.696
オートロック:有り	0.350 †	0.184	1.419	0.028	0.180	1.029	0.240	0.207	1.271
駐車スペース:有り	-0.183	0.152	0.833	-0.250 †	0.138	0.779	-0.026	0.170	0.974
住居:新しい	-0.321 **	0.106	0.726	-0.168 †	0.092	0.845	-0.365 **	0.114	0.694
住居:広い・大きい	-0.213 †	0.121	0.808	-0.352 ***	0.106	0.703	0.052	0.108	1.053
住居:表札有り	-0.816 ***	0.122	0.442	-0.321 **	0.117	0.725	-0.203	0.126	0.816
変量効果									
切片	標準偏差	分散成分		標準偏差	分散成分		標準偏差	分散成分	
切片	0.617 ***	0.380		0.437 ***	0.191		0.500 ***	0.250	

† = $p < .10$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$, 参照カテゴリ = 有効回答

個人水準の特性も〈接触不能〉と関連を示している。他の変数を統制したとき、女性の確率比は 0.632 倍、20 代と比べて上の年代であればあるほど〈接触不能〉となる確率比が下がる傾向にある。事実、30 代の 0.695 から 70 歳以上の 0.152 まで単調に確率比は下がっている。また、住居形態(一戸建)と敷地内駐車スペースが調査員の目で明らかに存在する場合は、〈接触不能〉とは有意に関連していない。一方、それぞれ他の変数を統制しても〈接触不能〉と関連する住居の特徴もある。それぞれの確率比は、オートロックがある場合 1.419 倍、‘新しい’は 0.726 倍、‘広い・大きい’は 0.808 倍、そして‘表札有り’は 0.442 倍という傾向を示している。

〈本人拒否〉

表 3 の中央の列は〈有効回答〉と比べ〈本人拒否〉となる地点・個人水準の特徴を示している。都市度を代理的に示す人口密度が高いと〈本人拒否〉が上昇する傾向にある(確率比は 1.297 倍)。調査員経験年数も 6 年未満と比べ、10 年以上であると確率比は 0.739 倍で、調査員経験が長いことが〈本人拒否〉を下げる説明要因となっている。個人水準の結果によると、性別、それに年齢は 20 代と 30 代には有意な差はない。一方、20 代と比べて 40 代、50 代、60 代と〈本人拒否〉

は増加する。特に60代の確率比は1.568倍と高く、70歳以上は20代と違いがない。住居の特徴では、「一戸建」の係数は正であるが、「駐車スペース」、「新しい」、「広い・大きい」、「表札有り」は〈本人拒否〉と負の関連にある。

〈他者拒否〉

〈他者拒否〉について〈有効回答〉と対比したモデルは、地点特性のうち人口密度と大学卒者割合が有意となった。また、調査員経験年数を示す2つのダミー変数は、双方共に有意に〈他者拒否〉と関係している。調査員経験が6年以上10年未満であると確率比は0.566倍、10年以上は0.673倍と、それぞれ〈有効回答〉と比べた場合の〈他者拒否〉に対する減少が確認できる。

個人水準の変数も〈他者拒否〉と関連している。男性と比べて女性であると〈他者拒否〉の確率は低い(0.503倍)。また、20代と比べると、年齢が上がるほど〈他者拒否〉の割合は単調に下がっていく。30代では0.505であるが70歳以上であると0.134倍となる。「一戸建」であると1.696で〈他者拒否〉傾向を上昇させ、住居が「新しい」ことはその逆の効果(0.694倍)となっている。

5. 考察

5.1 本研究の方法面の限界と課題

本研究には方法面でいくつかの限界があるので、その点についてははじめにまとめておく。

第一に、4.1節で述べた通り級内相関係数は0.074~0.123と大きくなかった。本研究では全国規模の社会調査のサンプルサイズとしても小さいとはいえない個人水準6400(分析に用いたサイズは6054)、地点水準400を基本とした条件でも、有意な説明要因を探す際には効果の検出をやや難しくしていたと考えられる。

第二に、説明変数を両レベルに追加した最終モデルとヌルモデルの級内相関係数に大きな差異がなかったことから、地点水準の説明変数の選択についての課題も残る。即ち、より説明力の高い説明変数が本研究で用いた変数以外に見いだされる可能性は否定できない。地点水準の変数のうち町丁字単位の地域の性格に関する変数は、国勢調査資料の中で候補となる多数の要因がある中から、先行研究に基づいて予備的な解析を行い、都市度と関連すると思われる有力な変数を探索した上でモデルを構築した。こうした要因をより網羅的に検討するような変数選択の方法を検討することは、統計的な手法の開発の上でも重要と思われる。より洗練された統計的方法による分析は今後の課題としたい。

このような限界を抱えている中でも、本研究は4章での分析により、非回収状況に対する異なる地点変数・調査員変数の寄与について一定の成果を上げたと考えるので、それらの点と、結果が調査実務に与える示唆について、次の5.2節で論じる。

5.2 分析結果の総合討論

本稿の分析結果は3つの仮説を部分的に支持している。仮説1-1については、30代と70歳以上を除いて年齢が高いことが〈本人拒否〉の説明要因であることが確認されたが、性別による差異はなかった。一方、仮説1-2は全面的に支持され、男性と年齢が低い者ほど〈有効回収〉と比較して〈接触不能〉と〈他者拒否〉につながりやすい傾向が確認された。事実、〈他者拒否〉の年齢による影響については、10歳毎のダミー変数を用いた確認では、年齢の上昇とともに単調に低下している。また、仮説1-3の通り、一戸建は〈他者拒否〉の増加と関連していた。

これらの結果は先行研究とほぼ一致する。主に接触不能による(田辺, 2003; 保田, 2008)若年層の男性の未回収の多さ(Synodinos and Yamada, 2000; 前田, 2005; 保田, 2008)、それに高齢層の拒否(稲葉, 2010)傾向が「日本人の国民性第13次(2013年)全国調査」によっても確認されたことになる。〈接触不能〉と〈他者拒否〉が男性に多いのは、性別による在宅率の違いが反映してい

ると考えられる。また、〈一戸建〉が〈他者拒否〉の増加と関連していることは、三輪(2008)が推論したように同居している家族からの拒否が増えているからと解釈できる。

個人水準の変数としては、対象者の住居の様子も複数の不能カテゴリに対する説明要因となることが示された。これは従来から指摘されている「一戸建か否か」という変数以外にも、住居の広さや新しさのような社会経済的な地位との関連が予想される変数と調査不能との間の関連を示唆している。この結果は調査員が現地で取得する補助情報(Kreuter and Olson, 2013)の有用性の例示となるものである。

地点間のばらつきについては、5.1節で述べたように大きい値ではないが3種類の非回収状況の級内相関から、一定の地点間異質性が確認された。中でも〈接触不能〉は2種類の拒否と比べると比較的地点間の異質性が大きい。調査対象者が不在であったり、独居であるがゆえに家族などによる〈他者拒否〉に繋がらなかったりする対象者が特に都市部に多いことが、比較的大きな地点間のばらつきの理由と考えられる。これらの地点間差異は、部分的に国勢調査による町丁字単位の地点特性の変数によって説明された。カテゴリによって有意水準は異なるが、有意でなかった変数の結果も含め、都市度が高い地点であるほど〈有効回答〉に比べて〈接触不能〉、〈本人拒否〉、〈他者拒否〉が増加する傾向があった。都市度と有効回答の関連(前田, 2008; 三輪, 2008; 埴淵 他, 2012; 山内, 2012)は、町丁字単位で都市度を操作化することで、先行研究よりも厳密に検討した本稿の分析でも確認されたことになる。なお、都市度と〈接触不能〉の有意な関連は三輪(2008)の報告とは異なり、近年の研究(埴淵 他, 2012; Matsuoka and Maeda, 2015)と整合的であった。都市度が高いほど、〈接触不能〉と2種類の拒否(本人・他者)が増加する結果は、おそらく、地点における他者に対する信頼感(地域住民の信頼感のベースライン)の違いが影響していると考えられる。都市度が低い地点では地域に対する信頼感や安心度が高い傾向にあり、一方、都市部では、Matsuoka and Maeda(2015)が論じるように、犯罪に対するリスク意識と他者に対する警戒心が高いことから、非回収の傾向の強さに繋がったと考えられる。

最後に、調査員経験の長さとして3つの非回収状況に関連があることがわかった。これらの知見は、調査員による差があるとした田辺(2003)、関連がないとした保田(2008)、それに結果がばらついて傾向が見出せなかった埴淵 他(2012)とも異なる。〈接触不能〉については、経験が6年以上10年未満であると有意で、10年以上では関連が見られなかったことから留意する必要があるが、〈本人拒否〉と〈他者拒否〉については経験が10年以上であると減少傾向にある。本稿の結果は、調査員の経験に基づく交渉術で本人や同居する家族からの拒否を減少させていることを示唆している。

5.3 分析結果からの提言

本研究が調査の実施に与える示唆としては、調査員特性と回収率に関する検討を一層深める必要性が挙げられる。調査員事後アンケートで調査員についてより詳細な特性を調べ、経験年数以外の特性で回収状況と関連している事項を明らかにする必要がある。Groves and Couper(1998)は、対象者の協力に影響を与える調査員の他の態度特性(調査に関わる様々な信念や調査時における行動)なども取り上げているが、それらが影響することの十分な証拠は得ていない。本稿でも事後調査の中に含まれている調査員自身の学歴や調査員を続けている動機等の変数を説明要因に取り込むことも試みたが、有意な結果は得られなかった。

また、調査員特性と回収状況の関連に加えて、調査員と調査対象者の組み合わせによって回収率が異なるのか検討されることが望ましい。保田(2008)と保田 他(2008)は、Groves and Couper(1998)が示した調査対象者と調査員の相互作用によって調査協力がどうかの判断に繋がるといふ枠組みを示した上で、調査に対する信頼などは制御できないが調査員の行動は制御可能であるとしている。もっとも、保田 他(2008)が言及するように、大規模調査では調査会社に委託

することから調査員の制御は難しい。その上、一地点あたりのケース数に現実的な上限があるため相互作用の分析も現実的ではない。地域を限定した調査であったとしても一地点あたりのケース数が多い調査で相互作用を検討することや、多数の調査にわたって調査員を紐付けて調査員特性と回収状況の関連を調べるなどを通して、回収率を向上させる具体的な方策を制御可能な調査員に提示する必要がある。

調査員に過大な負荷をかけずに可能な調査法研究としては、回収状況に関する詳細な情報の収集が考えられる。たとえば、〈有効回答〉となったケースの中でも、当初は対象者本人が難色を示したが調査員の説得によって調査に応じた事例もあるはずである。それは〈他者拒否〉でも同じで、同居人から難色を示されたものの、本人と直接話せるまでと粘り強く交渉することが、おそらく拒否と調査員経験年数の負の関連であると考えられるが、これは推測の域を出ない。調査員経験の長短や他の特性が、どのような対処法・交渉術の差異となり本人と家族による拒否を減少させているのか、今後の検討事項である。また、接触不能の中に、居留守であったり家族が本人は病気だと申告したりするケースも考えられ、これらについては拒否を〈本人拒否〉、もしくは〈他者拒否〉と分類することで、より個人、地点、調査員特性と回収状況の厳密な検証が可能となる。

以上を踏まえた調査の実査に関する示唆としては、まず、都市度が高い地点の担当に、可能な限り6年以上の調査経験を持つ調査員を振り分けることが考えられる。ただし、地方によっては調査会社が選択できるほど調査員候補者が多くとも考えられず、現実的な改善は難しいという事態も起こりうる。調査員経験が長いほうが地点における回収状況が改善されるが、調査員の高齢化もあり、今後の社会調査の実査のことを考慮すると、経験年数が多い調査員による若手の育成が急務といえる。

最後に、回収率の向上法を模索することも重要であるが、調査不能バイアス(星野, 2010)やその補正(Fushiki and Maeda, 2014; 伏木・前田, 2015; 土屋, 2010)の知見を活かし、回収されたデータの分析における工夫も重要な課題となることを指摘しておきたい。

参 考 文 献

- Bethlehem, J., Cobben, F. and Schouten, B. (2011). *Handbook of Nonresponse in Household Surveys*, Wiley, New Jersey.
- Fushiki, T. and Maeda, T. (2014). Nonresponse adjustments for estimates of proportions in the 2010 Survey on Stratification and Social Psychology, *Behaviormetrika*, **41**(1), 99–114, doi: 10.2333/bhmk.41.99.
- 伏木忠義, 前田忠彦 (2015). 調査不能を伴う社会調査における推定: 国民性に関する意識動向調査を題材に, 新潟大学教育学部研究紀要 自然科学編, **7**(2), 63–71.
- Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys, *Public Opinion Quarterly*, **70**(5), 646–675.
- Groves, R. M. and Couper, M. (1998). *Nonresponse in Household Interview Surveys*, John Wiley and Sons, New York.
- 埴淵知哉, 村中亮夫, 花岡和聖, 中谷友樹 (2011). 社会調査における回収率の地域差—JGSS 累積データ 2000–2006 の回収状況データを用いた分析, 日本版総合的社会調査共同研究拠点研究論文集, **11**, 181–192.
- 埴淵知哉, 中谷友樹, 村中亮夫 (2012). 社会調査における回収率の地域差とその規定要因: 個人および地域特性を考慮したマルチレベル分析, 地理学評論, [Series A], **85**(5), 447–467.
- 星野崇宏 (2010). 調査不能がある場合の標本調査におけるセミパラメトリック推定と感度分析: 日本人の国民性調査データへの適用(特集 日本人の国民性調査研究—平成期の20年—), 統計数理, **58**(1),

3-23.

- Inaba, A. (2007). Problems relating to declining response rates to social survey research in Japan: Trends after 2000, *International Journal of Japanese Sociology*, **16**, 10-22.
- 稲葉昭英 (2010). NFRJ(全国家族調査)レポート NFRJ08 のデータ特性：予備標本・回収率・有配偶率, *家族社会学研究*, **22**(2), 226-231, doi: 10.4234/jjoffamilysociology.22.226.
- Kreuter, F. and Olson, K. (2013). Paradata for nonresponse error investigation, *Improving Surveys with Paradata: Analytic Uses of Process Information* (ed. F. Kreuter), John Wiley and Sons, Hoboken, New Jersey.
- 前田忠彦 (2005). 郵送調査法の特徴に関する一研究—面接調査法との比較を中心として—(特集 日本人の国民性調査 50 年), *統計数理*, **53**(1), 57-81.
- 前田忠彦 (2008). SSM2005 国内調査の標本設計と調査精度に関する検討, 2005 年 SSM 日本調査の基礎分析：構造・趨勢・方法, Vol.1 (三輪哲, 小林大祐 編), 1-15, 2005 年 SSM 調査研究会, 仙台.
- Matsuoka, R. and Maeda, T. (2015). Neighborhood and individual factors associated with survey response behavior: A multilevel multinomial regression analysis of a nationwide survey in Japan, *Social Science Japan Journal*, **18**(2), 217-232.
- 三輪 哲 (2008). 働き方とライフスタイルの変化に関する全国調査 2007 における標本特性と欠票についての基礎分析, 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, No.10, 1-17.
- 中村 隆, 土屋隆裕, 前田忠彦 (2015). 国民性の研究第 13 次全国調査—2013 年全国調査—, *統計数理研究所調査研究レポート*, No.116.
- Raudenbush, Stephen W., Bryk, Anthony, Cheong, Yuk Fai, Congdon, Richard T. and du Toit, Mathilda (2011). *HLM 7: Hierarchical Linear and Nonlinear Modeling*, Scientific Software International, Lincolnwood, Illinois.
- 坂元慶行 (2001). 「日本人の国民性調査」：社会調査研究のある最前線(〈特集〉実証の姿：その思惟と展開), *理論と方法*, **16**(1), 75-88.
- 政府統計の総合窓口(e-stat)(2015a). 平成 22 年度国勢調査小地域集計(Web サイト), <http://www.e-stat.go.jp/SG1/estat/NewList.do?tid=000001039448>, 最終閲覧 2015 年 2 月 8 日.
- 政府統計の総合窓口(e-stat)(2015b). 地図で見る統計(統計 GIS)データダウンロード(Web サイト), <http://www.e-stat.go.jp/SG1/estat/toukeiChiri.do?method=init>, 最終閲覧 2015 年 2 月 8 日.
- 篠木幹子 (2010). 社会調査の回収率の変化(特集 回収率を考える), *社会と調査*, **5**, 5-15.
- Snijders, T. A. B. and Bosker, R. J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, 2nd ed., SAGE Publications, London.
- Synodinos, N. E. and Yamada, S. (2000). Response rate trends in Japanese surveys, *International Journal of Public Opinion Research*, **12**(1), 48-72.
- Synodinos, N. E. and Yamada, S. (2013). Japanese public opinion surveys: 20-year trends, *Behaviormetrika*, **40**(2), 101-127.
- 玉野和志 (2003). サーベイ調査の困難と社会学の課題(〈特集〉社会調査：その困難をこえて), *社会学評論*, **53**(4), 537-551.
- 田辺俊介 (2003). 面接調査の欠票理由の検討—面接調査の回収率向上のための一提言, *社会学論考*, **24**, 1-27.
- 土屋隆裕 (2010). 調査への指向性変数を用いた調査不能バイアスの二段補正—「日本人の国民性第 12 次全国調査」への適用—(特集 日本人の国民性調査研究—平成期の 20 年—), *統計数理*, **58**(1), 25-38.
- 崔田知久 (2008). 面接調査の現状と課題(〈特集〉世論調査方法の再検証—「総合調査学」へ向けて—), *行動計量学*, **35**(1), 5-16.
- 山内昌和 (2012). 第 4 回全国家庭動向調査の無回答に関する検討(特集『第 4 回全国家庭動向調査(2008 年)』の個票データを利用した実証研究(その 3)), *人口問題研究*, **68**(1), 70-89.
- 保田時男 (2008). 低下する回収率と回収不能の要因, 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(谷岡一郎, 仁田道夫, 岩井紀子 編), 447-458, 東京大学出版会, 東京.
保田時男, 宍戸邦章, 岩井紀子 (2008). 社会調査とデータ管理の諸方法 (1) 大規模調査の回収率改善のための調査員の行動把握—JGSS における訪問記録の分析から(社会調査とデータ管理の諸方法), 理論と方法, **23**(2), 129-136.

A Study of Survey Non-response Using the 13th Nationwide Survey
of the Japanese National Character: Assessing Sampled Individuals',
Neighborhood' and Survey Takers' Characteristics
Associated with Non-response Behavior

Ryoji Matsuoka¹ and Tadahiko Maeda²

¹Waseda Institute for Advanced Study, Waseda University

²The Institute of Statistical Mathematics

This study assesses nonresponse behaviors of a large-scale social survey. Using data of the 13th nationwide survey of the Japanese National Character, a survey for survey takers, and national population census of 2010, this study investigates whether sampled individuals', neighborhood', and survey takers' characteristics relate to the individuals' survey response behaviors. As sampled individuals (N=6054) are nested in each sampled neighborhood (N=400), a multinomial two-level multilevel modeling approach was applied. Results of the study indicate that, when compared to a reference category (i.e., response), (1) whether a person lives in a detached home is related to refusal by target persons and by others (i.e., other than sampled individuals, likely family members), (2) there are differences between neighborhoods in survey response behaviors, especially between-neighborhood-variation of non-contact cases is relatively large, (3) neighborhood-level variables obtained from national population census of 2010 partly explain differences between neighborhoods in the behaviors, and (4) survey takers' years of experience are associated negatively with non-contact and refusal by target individuals and by others. The study's findings indicate that more experienced survey takers are capable of avoiding non-contact and refusal cases. Since survey administrators cannot force sampled individuals to respond to a survey, but can select and train survey takers, further studies need to examine relationships between survey takers' characteristics and sampled individuals' non-response behaviors to have practical implications that can improve survey response rates.