

集団意思決定を考慮した世帯単位の 交通調査回答行動分析

細谷 謙太¹・川野 倫輝²・渡邊 萌³・佐藤 嘉洋⁴・円山 琢也⁵

¹正会員 熊本県庁 (〒862-8570 熊本県熊本市中央区水前寺 6-18-1) (前熊本大学工学部社会環境工学科)

²学生会員 熊本大学大学院自然科学研究科社会環境工学専攻 (〒860-8555 熊本県熊本市中央区黒髪 2-39-1)
E-mail:178d8811@st.kumamoto-u.ac.jp

³学生会員 熊本大学大学院自然科学研究科社会環境工学専攻 (〒860-8555 熊本県熊本市中央区黒髪 2-39-1)
E-mail:174d8830@st.kumamoto-u.ac.jp

⁴学生会員 熊本大学大学院自然科学研究科社会環境工学専攻 (〒860-8555 熊本県熊本市中央区黒髪 2-39-1)
E-mail:yo-sato@kumamoto-u.ac.jp

⁵正会員 熊本大学准教授 くまもと水循環・減災研究教育センター
(〒860-8555 熊本県熊本市中央区黒髪 2-39-1)
E-mail:takumaru@kumamoto-u.ac.jp

社会調査等の回収率低下の問題に対して、調査未回答者の影響を精査することは重要である。既存研究では、熊本PT調査データにより調査回答選択行動モデルを構築した例があるが、個人単位の分析にとどまっており、世帯単位の回答行動を的確に分析できていなかった。本研究は、熊本PT調査の回答有無と回答方法(紙面とWeb)について、集団意思決定を考慮した世帯単位での分析を行う。個人単位のモデルと比較して世帯単位のモデルは、精度向上が確認され、世帯構成員の性年齢などの影響を的確に表現できている。また、このモデルは世帯内の誰が代表して回答したのかを推定することも可能である。この特性を利用し、世帯内の別の構成員が代理回答したと推定される個人について、そのトリップ数の報告値が過少となる代理回答バイアスの分析例も提示した。

Key Words : *Participation choice model, Proxy response, Group-based choice model, Person trip survey*

1. はじめに

(1) 研究の背景と目的

交通計画等の公共政策の立案の根拠データとしてパーソントリップ調査(以下 PT 調査)を代表する社会調査が実施される。その際に高い回収率と回答データの質の確保は非常に重要である。そのために、どのような世帯が未回答であるのか、またどういった回答方法を選択しているのかを分析する必要がある。

従来、PT 調査の手法は訪問留置回収型であったが人件費の削減や調査員の負担軽減を目的として、郵送・web 併用型に変化してきている。その変化に伴い有効回収率が大幅に低下しており¹⁾、近年では目標回収率が25%程度に設定されることが多い。また、調査方法が郵送・web 併用型に変化したことで回収できるデータに偏りが生じている可能性があり、分析する際に配慮する必要がある。

2012 年秋に第 4 回熊本都市圏パーソントリップ調査(以下 熊本 PT 調査)が実施された。平原・円山²⁾はこの熊本 PT 調査のデータを用い、個人単位による PT 調査の回答有無、紙・web の回答方式の選択行動モデルを構築し、分析している。しかし、PT 調査等の社会調査への回答は個人単位ではなく、世帯での回答行動であり、調査回答の意思決定を的確にモデル化出来ていなかった。

また、世帯単位での回答になることで、子どもや高齢者の回答は世帯内の他者が代理回答をしている可能性が考えられる。特に web 調査回答においては代理回答の可能性が高いことが考えられ、すべての行動を記録できていないおそれがある。これらの代理回答バイアスを分析できる手法が求められる。

以上の背景を踏まえ、本研究では熊本PT調査において世帯に着目し、回答選択行動分析を行う。具体的には以下を目的とする。

1) 世帯単位でのPT調査回答有無、調査回答方法選択

を分析できる新たなモデルの構築

- 2) 調査回答行動や調査回答方法選択行動に影響を与える世帯要素の分析
- 3) 構築したモデルを用いた代理回答バイアスの把握

以上の目的を達成することで、調査回答選択行動モデルの精度向上、世帯属性別に調査方法を変化させる手段の検討、有効回収率を増加させるための調査手法の検討、代理回答による影響を少なくする調査手法の検討につながると思われる。

具体的に、まず PT 調査回答有無、調査回答方法の各段階において、基礎分析を行う。それらから調査回答選択行動に影響を与える説明変数を検討し、ロジットモデル推定を行う。モデル推定は世帯単位でのグループ型選択行動モデルを用いることで、世帯構成員一人ひとりの選択確率を算出する。その結果をもとに世帯の中で誰が回答者となりえるのかの分析も可能である。これらの分析から代理回答の可能性についても考察していく。

(2) 既存研究のレビューと本研究の位置づけ

小嶋・久保田²⁾は、地区交通計画に関する意識調査において意見を言わない人(サイレント層)を分析するため、各種調査への未回答者に関する既存研究を整理し、サイレント層の生成と特徴について明らかにしている。山本・森川³⁾は、買い物頻度モデルの構築のため調査参加選択モデルを推定している。世帯人数に関して参加確率に大きな影響を及ぼしていることを指摘している。また、単身世帯の調査参加率が低いことも明らかにしている。張ら⁴⁾は、世帯の車種選択を対象に集団離散選択モデルの比較分析を実施している。世帯意思決定に際する世帯構成員の相対的な影響力を考慮し、世帯とその構成員を明確に区別した多様な説明変数が利用されている。張ら⁵⁾は従来の交通行動モデルにおける個人単位の分析の限界を指摘し、集団意思決定理論と時間配分理論を融合させ、多項線形効用関数に基づき世帯時間配分モデルを開発した。

名取ら⁶⁾は、PT調査における回答誤差とその発生原因を分析するため、PT調査とアクティビティ・ダイアリー調査を同時に行い比較分析を実施している。PT調査におけるトリップの抜け落ちは約20%であり、帰宅に関するトリップが抜け落ちやすいことを指摘している。

林⁷⁾は、郵送調査について現況を整理し、課題点を指摘している。また、代理記入に関する評価も実施しており、既存調査から郵送調査の代理記入率は、個人面接調査や訪問留め置き調査と同様に5-10%程度であると見込んでいる。

本研究は、熊本PT調査で得られたデータを用いて、世帯単位での調査回答行動モデルを構築する。研究の独自性・新規性として、1) 世帯単位での回答行動分析を構

築している点、2) 構築したモデルを用いて代理回答バイアスの分析を実施している点、が挙げられる。

(3) 本研究の構成

本研究では、第2章において、熊本PT調査データの概要をまとめた上で使用するロジットモデルの構造を説明する。第3章では、まず基礎分析を行いモデルに適用する説明変数の検討を行う。その後モデルを使用することで調査回答有無に関する選択行動を明らかにする。第4章では調査回答方法について第3章と同様に基礎分析後モデルを適用する。第5章では構築したモデルを用いて代理回答される確率を算出し、代理回答による回答バイアスを明らかにする。最後に、第6章で研究の成果、課題についてまとめる。

2. 熊本PT調査と調査回答行動選択モデルの構造

本章では、熊本 PT 調査の分析に利用したデータを詳述し、本研究で想定する調査回答選択行動モデルの構造を示す。

(1) 熊本 PT 調査データの概要

熊本 PT 調査の対象地域は熊本都市圏(熊本市とその周辺の5市6町1村)である。近年の PT 調査は回収率が低下傾向にあるが、熊本 PT 調査では効果的な広報活動により 35.1%の高い回収率を記録している⁸⁾。表-1 に本研究で利用するデータのサンプルサイズの分布を示す。

第3章の熊本 PT 調査の回答有無に関する分析では、スマホ・アプリを利用したプローブ・パーソン型調査(以下、スマホ調査)への調査依頼者 13,279 世帯(37,232 人)のデータを使用している。スマホ調査の対象者は、PT 調査の第2ロット、第3ロットの依頼対象者のうち熊本市内に住む 20-40 代を含む世帯であり、ランダムに 5,000 世帯ずつ依頼された。第23ロットに加えて、予備ロットとして、年齢制限なく熊本市内居住の 3,279 世帯に依頼が実施され、合計 13,279 世帯にスマホ調査が行われた。13,279 世帯(37,232 人)のうち、熊本 PT 調査に回答した世帯は 3,752 世帯(11,438 人)、未回答が 9,527 世帯(25,794 人)となっている。回答率は 28.3%となっており、熊本 PT 調査全体の回収率 35.1%と比較すると少ないのは、20-40 代の若い世代がスマホ調査の主な対象であったためである。

また、第4章の熊本 PT 調査回答時における回答方法選択では、熊本 PT 調査で回収された 43,520 世帯(97,109 人)のデータを使用している。熊本 PT 調査では、世帯ごとに紙調査票での回答(以下、紙調査)か web 方式調査のうちどちらか一方の方式で回答する調査であった。紙調

表-1 分析対象の熊本PT調査データの概要

分析対象	第3章 回答有無		第4章 回答方法選択	
分析データ	スマホ調査依頼 13,279世帯 (37,232人)		第4回熊本PT調査回収データ 43,520人 (97,109人)	
	PT回答 3,752世帯 (11,438人)	PT未回答 9,527世帯 (25,794人)	web調査 3,906世帯 (9,787人)	紙調査 39,614世帯 (87,332人)
使用可能説明変数	世帯構成員の個人属性 (性別, 年齢) + 世帯属性 (世帯人数, 居住地(Cゾーン))		世帯構成員の個人属性 (性別, 年齢, 職業, トリップ数等) + 世帯属性 (居住地, 世帯人数, 世帯分類)	

査で回答した世帯は 39,614 世帯(87,322 人), web 調査が 3,906 世帯(9,787 人)となっている。

以上のデータから, 各段階 (PT 調査回答有無, 調査回答方法選択) において, 調査回答世帯の属性を基礎分析し, 調査回答に影響を与えそうな要素を検討する。そして, 各段階で世帯単位のロジットモデルを適用する。

なお今回の分析では, 分析対象別に利用できるデータが異なってくることに注意が必要である。第3章で分析する回答有無については, 住民基本台帳から得られる世帯構成員の「性別」, 「年齢」(5 歳階級別), 「世帯人数」, 「居住地」(熊本 PT の C ゾーン別)のみ利用できる。厳重な秘匿処理をすることで, 調査に参加していない方の基礎的な属性データも分析対象としている。第4章で分析する回答方法選択では, 上述のデータに加え, PT 調査から得られる「職業」等が利用できるデータに加わる。

(2) 調査回答行動選択モデルの構造

世帯単位の PT 調査回答行動を把握するために以下のような理論からモデルを構築した。

世帯 i に J_i 人の世帯構成員がいるとする。このうち一人が世帯を代表して回答すると仮定する。世帯 i が調査に回答する場合, j 番目の構成員が回答する確率を $P_{i,j}$ とする。また, 無回答の場合を $j = 0$ とし, その確率を $P_{i,0}$ とする。無回答の効用関数 $V_{i,0}$, および世帯構成員 j が回答を行う効用関数 $V_{i,j}$ を以下のように定義する。

$$V_{i,0} = \sum_k \alpha_k x_{k,i}, \quad V_{i,j} = \sum_l \beta_l y_{l,j} \quad (1)$$

式(1)において $x_{k,i}$ は世帯属性に関する説明変数, $y_{l,j}$ は個人属性に関する説明変数, α_k, β_l はそのパラメータを示す。 $P_{i,0}$ および $P_{i,j}$ に多項ロジットモデルを適用すると, 選択確率は以下の式で表すことが出来る。

$$P_{i,0} = \frac{\exp V_{i,0}}{\exp V_{i,0} + \sum_{j=1}^{J_i} \exp V_{i,j}} \quad (2)$$

$$= \frac{\exp \sum_k \alpha_k x_{k,i}}{\exp \sum_k \alpha_k x_{k,i} + \sum_{j=1}^{J_i} \exp \sum_l \beta_l y_{l,j}}$$

$$P_{i,j} = \frac{\exp V_{i,j}}{\exp V_{i,0} + \sum_{j=1}^{J_i} \exp V_{i,j}} \quad (3)$$

$$= \frac{\exp \sum_l \beta_l y_{l,j}}{\exp \sum_k \alpha_k x_{k,i} + \sum_{j=1}^{J_i} \exp \sum_l \beta_l y_{l,j}}$$

本調査では, 世帯 i の中で何番目の構成員が回答したのかは不明である。しかし本モデルより世帯が回答する確率は $\sum_j P_{i,j}$ で与えられることを利用する。このとき, 実現した状態を示す同時確率 L^* と対数尤度 L はそれぞれ以下で表され, この式を利用してパラメータの最尤推定を行う。

$$L^* = \prod_{i \in N} P_{i,0}^{1-\delta_i} \left(\sum_{j=1}^{J_i} P_{i,j} \right)^{\delta_i} \quad (4)$$

$$L = \ln L^* = \sum_{i \in N} \left\{ (1 - \delta_i) \ln P_{i,0} + \delta_i \ln \sum_{j=1}^{J_i} P_{i,j} \right\} \quad (5)$$

$$\delta_i = \begin{cases} 1 & : \text{世帯}i\text{が回答の場合} \\ 0 & : \text{世帯}i\text{が無回答の場合} \end{cases} \quad (6)$$

以上のモデルを各段階に応じて適切な説明変数を考慮し, 適用していく。なお, このモデルは張ら⁹⁾が提案した集団意思決定モデルにおけるMax-Maxモデルの1種と等価である。この詳細は, 佐藤, 円山⁹⁾を参照されたい。

3. 回答有無に関する分析

(1) PT調査回答有無に関する基礎属性分析

a) 性年齢別

PT 調査の回答有無においては世帯主が大きな影響を

及ぼしていることが考えられる。世帯主については調査データから確認することが出来る。そこで調査対象世帯データから、世帯主の性年齢別の調査への回答率を図-1に示す。まずは、年齢別に回答率を見ていくと 60-69 歳が最も回答率が高くなっている。一方、10-20 歳代の回答率の低さが目立つ。

次に、性別に着目すると 10-20, 30, 40 歳代では主に女性の方が回答率が高くなっているが、50 歳代を境に男性の方が回答率が高くなっている。

円山ら¹⁴⁾は選挙投票率と PT 調査回答率の比較を行い、どちらも 65 歳以下では男性より女性の方が高く、65 歳以上では逆の傾向になることを示した。今回の分析では世帯単位での分析であり、本項では世帯主の年齢に着目しているため、傾向が入れ替わる年齢に差は生じているが同様の傾向が出ており、興味深い。

b) 回答者の世帯人数別

本項では、調査対象世帯データから世帯人数別の調査への回答率を分析する。その分析結果を図-2に示す。世帯人数 3 人の世帯の回答率が最も高く、そこを頂点に山型のグラフになっている。また、単身世帯の回答率が他と比較して低いことが分かる。要因としては日頃から忙しい人が多く、代理回答をする人もいない、ゆっくりと回答している時間がない等が考えられる。

世帯人数は多いほど回答者となりえる対象者が増え回答率が増加されることが予想されたが、そのような分析結果は得られなかった。

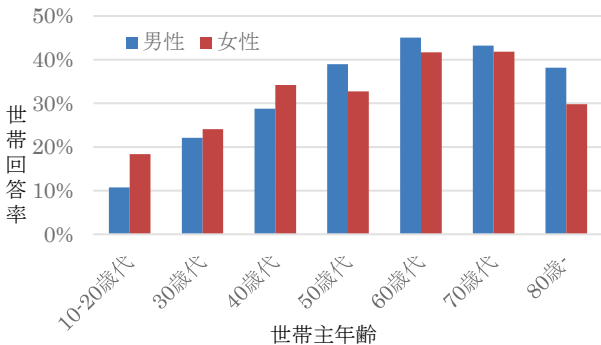


図-1 世帯主性年齢別の PT 調査世帯回答率

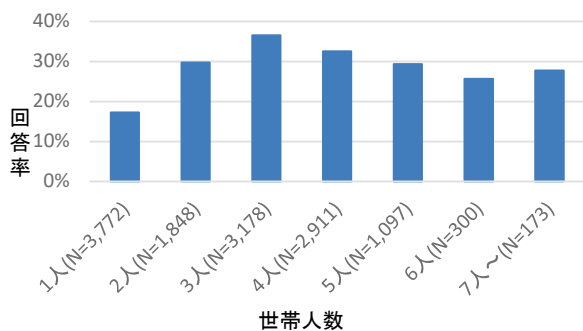


図-2 世帯人数別の PT 調査世帯回答率

(2) PT調査回答有無選択モデルの推定結果

第 2 章で述べたモデルを適用し PT 調査回答有無について推定を行った。今回の分析では、 $x_{k,i}$ に単身世帯ダミー、世帯人数、居住地Cゾーンの用途地域割合を設定し、 $y_{l,j}$ には年齢、性別の個人属性を設定した。 α_k はPT調査に回答があるとパラメータが低く、 β_l は回答があるとパラメータが高くなる。

表-2, 3に推定結果を示す。PT調査回答率には、性年齢、世帯人数などが影響を及ぼしている。表-2において用途地域の推定値に着目すると、第1,2種低層住居専用地域、第1,2種中高層住居専用地域では有意に符号が負であり、それらの地域に住む人はPT調査回答の傾向がある。具体的には、第1,2種低層住居専用地域に比べ第1,2種中高層住居専用地域のパラメータが大きくなっており、これらは第1,2種低層住居専用地域が一戸建てが多く、PT調査をはじめとする社会調査に回答しやすいことを示している。また、表-3のように性年齢をダミー

表-2 世帯単位のPT調査回答有無選択モデル推定結果 (1)

説明変数	パラメータ	t値	
世帯属性(α_k)			
定数項	1.67	15.23	***
単身世帯	0.11	1.49	
世帯人数	0.25	12.88	***
第1,2種低層住居専用地域	-0.60	-2.84	***
第1,2種中高層住居専用地域	-0.25	-3.46	***
近隣商業・商業地域	-0.08	-0.83	
準工業・工業地域	-0.22	-1.51	
個人属性(β_l)			
男性20歳ダミー	-0.68	-5.14	***
男性:年齢(20歳代を除く)	0.01	4.99	***
女性:年齢	0.02	14.35	***
サンプルサイズ		13,279	
ρ^2		0.184	
修正済み ρ^2		0.183	

*:10%有意, **:5%有意, ***:1%有意
 α_k :回答ありだとパラメータ低い, β_l :回答ありだとパラメータ高い

表-3 世帯単位のPT調査回答有無選択モデル推定結果 (2)

説明変数	パラメータ	t値	
世帯属性(α_k)			
定数項	2.62	9.57	***
単身世帯	-0.04	-0.47	
世帯人数	0.22	10.98	***
第1,2種低層住居専用地域	-0.56	-2.62	***
第1,2種中高層住居専用地域	-0.23	-3.12	***
近隣商業・商業地域	-0.06	-0.61	
準工業工業地域	-0.22	-1.52	
個人属性(β_l)			
20歳代ダミー	0.09	0.32	
30歳代ダミー	0.79	2.84	***
40歳代ダミー	1.17	4.19	***
50歳代ダミー	1.67	5.30	***
60歳代ダミー	2.17	7.37	***
70歳以上ダミー	2.24	7.38	***
20歳代ダミー	1.18	4.50	***
30歳代ダミー	1.82	7.03	***
40歳代ダミー	2.21	8.14	***
50歳代ダミー	2.24	8.26	***
60歳代ダミー	2.17	7.55	***
70歳以上ダミー	1.46	4.53	***
サンプルサイズ		13,279	
ρ^2		0.192	
修正済み ρ^2		0.190	

*:10%有意, **:5%有意, ***:1%有意
 α_k :回答ありだとパラメータ低い, β_l :回答ありだとパラメータ高い

として説明変数に導入したところ、男性20歳代のみ値が有意にならず、その他の性年齢のダミーが有意になっていることがわかる。このことから男性20歳代のみ傾向が異なることが考えられる。そこで表-2の個人属性では、男性20歳代のみをダミーにしている。実際にパラメータが負を示し、 β 値も有意であることから男性の20歳代は未回答の傾向にある。要因としては男性の20歳代は単身世帯が多いことが考えられる。年齢に着目すると、男女とも年齢が高くなるほど回答しやすく、高齢者がいる世帯ほど回答する傾向が読み取れる。

表-3は説明変数の組み入れ方を変更している。 $x_{k,i}$ には表-2と同様の説明変数を設定し、 $y_{l,j}$ には性年齢をダミー変数として設定した。世帯属性については表-2と同様の傾向を示している。図-3は性年齢のパラメータをグラフで示したものであるが、性年齢の傾向においては若い世代では女性の回答率が高いが、60歳を境に男性の回答率が高くなってきている。これは世帯主に着目した基礎集計と同様の結果になっており、モデルを用いることで世帯主以外の世帯構成員の性年齢の影響を表現することが出来た。

なお、表-4は既存研究²⁾において個人単位で推定されたモデルである。修正済み尤度比を比較すると、0.119から0.184、0.190と大きくなっており、世帯単位の分析法によるモデル精度向上が確認できた。

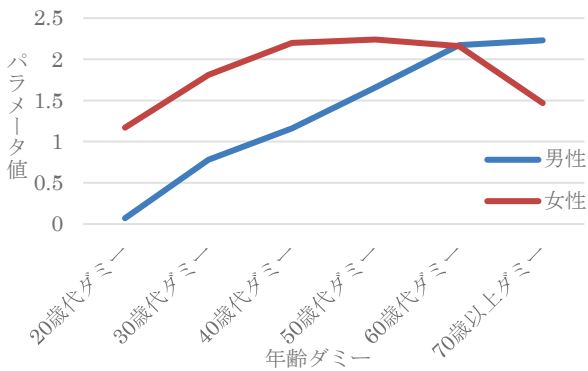


図-3 PT調査回答有無選択モデルにおける性年齢別ダミーのパラメータ推定結果

表-4 個人単位でのPT調査回答有無選択モデル推定結果²⁾

説明変数	推定値	t値	
性別(男:1)	-0.099	4.37	***
年齢	0.011	19.81	***
居住地ダミー(東区)	0.069	2.39	**
世帯人数	0.042	5.27	***
第1, 2種低層住居専用地域	0.568	4.70	***
第1, 2種中高層住居専用地域	0.191	4.26	***
準住居地域	-0.559	-1.14	
近隣商業・商業地域	-0.020	-0.31	
準工業・工業地域	-0.021	-0.25	
定数項(PT調査参加)	-1.542	-27.50	***
サンプルサイズ		37,232	
ρ^2		0.120	
修正済み ρ^2		0.119	
t値:1.96以上5%有意**, 2.58以上1%有意***			

4. 回答方法選択に関する分析

本章では、熊本 PT 調査回答世帯の回答方法選択と世帯属性の関係を分析する。熊本 PT 調査回収サンプル 43,520 世帯のうち、世帯全員の年齢が不明であった 2 世帯、世帯データが子どものみの 97 世帯を除いた 43,421 世帯で分析を行っている。

(1) 熊本PT調査回答方法に関する基礎分析

a) 世帯主の性年齢別

回答方法選択時においてもPT調査回答有無と同様に、世帯主の意向が影響していると考えられる。そのためまず、世帯主の性年齢別の回答方法の分析を行う。分析した結果を図-4に示す。

年齢に着目すると、性別問わず 20 歳代が最大になっており、年齢が高くなるほどその割合は低下している。若い世代は日常的に web に接する機会がその他の世帯と比較すると多いことが要因として考えられる。80-90 歳代で web の割合が増加しているのは世帯主以外の世代が代理回答していることが考えられる。性別では、女性に比べ男性の方が web で回答している。男性の方が web で回答している要因として、男性の方がインターネット利用者が多いということが考えられる。総務省¹²⁾によると女性に比べ男性の方が約 7%、インターネット利用者が多い。

b) 世帯主の職業別

本項では、PT 調査回答世帯のデータから、世帯主の職業ごとの回答方法割合をみていく。分析結果のグラフを図-5に示す。まず、紙調査の割合が比較的高い職業は、農林漁業・畜産・育林、鉱業・採石、砂利採取業、飲食サービス業が挙げられる。これらの職業の web 調査の選択割合が低い要因としては、業務のなかで web に接する機会が少ないため、紙調査を選択している可能性がある。一方で、web 調査の割合が高い職業は、情報通信業、学術研究・専門・技術、サービス業、公務員等が挙げられる。これらの職業では、比較的業務中に web に接する機

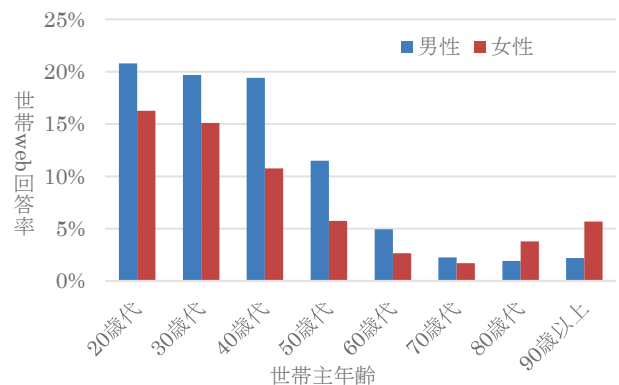


図-4 世帯主性年齢別のweb調査回答率

会が多いと考えられる。

c) 世帯人数別

図-6 では世帯人数別の回答世帯を分析している。スマホ調査依頼者のみ分析対象としていた既存研究²⁾では単身世帯の web 調査回答率が低かった。しかし、本研究で分析対象を熊本 PT 調査回答全世帯にすると web 回答率は高くなり、異なる結果となった。単身世帯では代理回答が出来ず、時間がないことが考えられるため時間、場所を問わずに回答がしやすい web 調査が増えたのではないかと考えられる。また、世帯人数が 4 人の世帯が最も web 調査回答率が高い。これは本節 a) 項で明らかになった年齢が若い人が世帯構成員に多いからだと考えられる。

d) 世帯属性別

対象世帯を65歳以上の高齢者のみの世帯と75歳以上の後期高齢者のみ世帯、子どものいる世帯、単身世帯、15-65歳の現役のみ世帯に分類し回答方法の傾向を把握する。最もweb調査回答率が高い世帯は子どものいる世

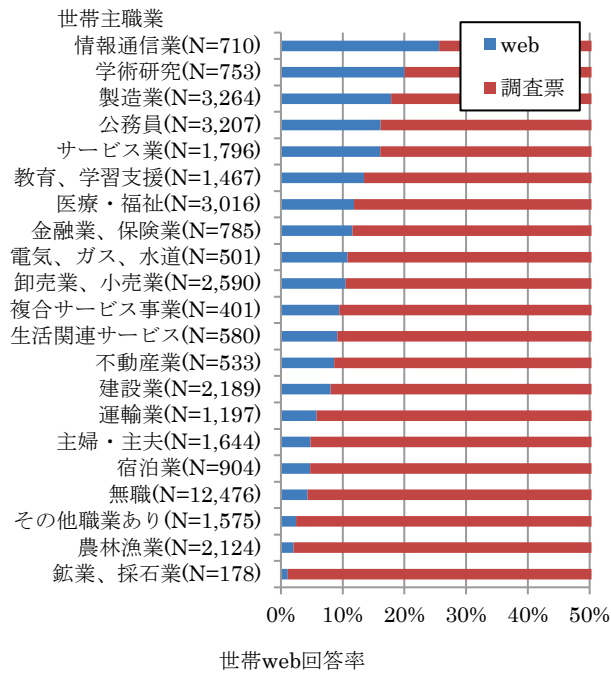


図-5 世帯主職業別のweb調査回答率

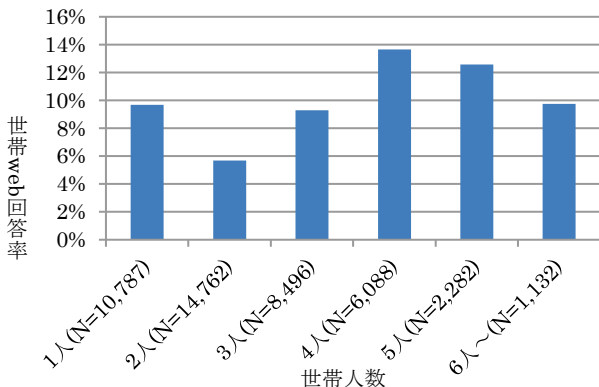


図-6 世帯人数別の web 調査回答率

帯 (16.4%) である。これは、比較的若い世帯が多く日常的にwebを使っていることが要因の一つとして考えられる。これに現役のみ世帯(11.6%)と単身世帯(9.8%)が続いている。一方、高齢者のみ世帯 (2.0%)、後期高齢者のみ世帯 (2.0%)についてはほとんどが紙調査での回答となっており、予想された結果になっている。

(2) 回答方法選択モデルの推定結果

本節a)e)項の分析結果をもとに説明変数の候補を検討し、推定を行った結果を表-5に示す。表-5の推定結果よりt値については、いずれの説明変数のt値も高く統計的に有意であると言える。ただし、サンプルサイズが大きいに留意は必要である。

世帯人数の符号が正であることから、世帯人数が多い世帯ほど紙調査を、世帯人数が少ない世帯ほど web 調査を選択する傾向がある。また世帯属性のダミーでは高齢者のみ世帯、子どものいる世帯は t 値が有意になっており、高齢者のみ世帯は紙調査、子どものいる世帯は web 調査での回答をしやすいことが推定できている。現役世帯のみ t 値が低く、他と比較し傾向が異なることが考えられる。

表-5 世帯単位のPT調査回答方法選択モデル推定結果

説明変数	パラメータ	t値
紙回答(α_k)		
定数項	-2.92	-4.19 ***
世帯人数	0.45	24.08 ***
世帯平均年齢	0.04	10.76 ***
高齢者のみ世帯ダミー	0.42	3.48 ***
子どものいる世帯ダミー	-0.31	-2.99 ***
現役世帯ダミー	-0.15	-1.72 *
市区町村ダミー	0.41	2.72 ***
web回答(β_l)		
20歳代ダミー	-3.01	-4.43 ***
30歳代ダミー	-2.50	-3.71 ***
40歳代ダミー	-2.22	-3.30 ***
50歳代ダミー	-2.52	-3.73 ***
60歳代ダミー	-3.11	-4.55 ***
70歳以上ダミー	-3.39	-4.75 ***
20歳代ダミー	-3.55	-5.22 ***
30歳代ダミー	-3.23	-4.78 ***
40歳代ダミー	-3.07	-4.54 ***
50歳代ダミー	-3.58	-5.25 ***
60歳代ダミー	-3.80	-5.52 ***
70歳以上ダミー	-2.73	-3.87 ***
サービス業ダミー	0.56	6.94 ***
学術研究ダミー	0.74	6.86 ***
情報通信業ダミー	0.95	9.32 ***
無職ダミー	0.16	2.16 **
主婦・主夫ダミー	0.27	2.76 ***
鉱業・採石業ダミー	-2.04	-2.07 **
農林漁業ダミー	-1.56	-5.65 ***
その他(職業あり)ダミー	-1.65	-6.58 ***
運輸業ダミー	-1.14	-5.28 ***
複合サービス業ダミー	-0.52	-2.22 **
宿泊業ダミー	-1.08	-5.35 ***
サンプルサイズ		43,421
ρ^2		0.611
修正済み ρ^2		0.610

*:10%有意, **:5%有意, ***:1%有意
 α_k :紙回答だとパラメータ高い, β_l :web回答だとパラメータ高い
 注)市区町村ダミー(甲佐町, 御船町:1, その他市区町村:0)

図-7は性年齢のパラメータ値を示したものであるが、性別問わず 40 歳代が最も web 調査を選択しやすい傾向になっている。しかし、世帯主の性年齢の基礎分析では若い人ほど web 調査で回答する傾向になっていた。この傾向は基礎分析では明らかにすることが出来なかったことであり、世帯単位の行動モデル推定結果の有用性を示している。また性別では男性の方が 70 歳以上を除いて、web 調査で回答しやすく、これは基礎分析でも同様の結果が得られている。職業ダミーにおいては世帯主職業別の分析において web 回答率が高かった職業は正の値、低かった職業は負の値が推定されている。世帯主職業において web 回答率が低かった主婦・主夫、無職が正の値を推定しているのは世帯主以外の世帯構成員の職業を表現できているためと考える。

なお、表-6は既存研究²⁾における個人単位による調査回答方法選択行動モデルである。尤度比を比較すると個人単位での推定結果の0.543から世帯単位では0.606と大きくなっておりモデル精度の向上が出来ている。

5. 構築したモデルを利用した代理回答の分析

熊本PT調査等の社会調査においては世帯単位での回答が行われている。そこで、ある個人の回答を別の世帯構成員が行う代理回答の影響が考えられる。代理回答バイアスとは、本人の移動を他者が推測して記入することで短トリップや私事トリップが抜け落ちることで生じる回答のバイアスである。

(1) 代理回答バイアスの基礎分析

代理回答されやすい人としては高齢者が挙げられる。図-8は熊本PT調査において高齢者のみの世帯と高齢者の他に同居している人がいる世帯(以下 高齢者のいる世帯)の高齢者のトリップ数を比較したものである。0トリップに顕著な差が生じている。また、高齢者のみの世帯と高齢者のいる世帯とで高齢者の平均トリップ数を比較すると、高齢者のみが2.56トリップであるのに対し、高齢者のいる世帯では2.32と少なくなっている。

表-7はさらに高齢者を紙・web調査、高齢者のみ世帯、

表-7 平均トリップ数比較

	紙調査 (25,071人)		web回答 (945人)		全体
	高齢者のみの世帯 (16,217人)	高齢者のいる世帯 (8,854人)	高齢者のみの世帯 (342人)	高齢者のみの世帯 (603人)	
平均トリップ数	2.59	2.36	1.88	1.42	2.68
標準偏差	1.9	1.98	1.8	1.75	
t値	8.63***		3.79***		

***:1%有意

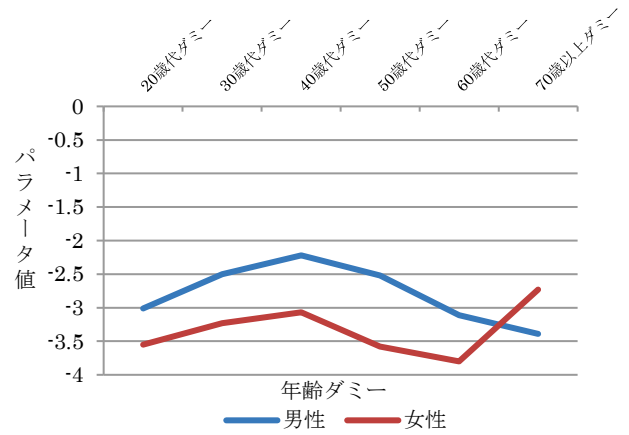


図-7 PT調査回答方法選択モデルにおける性年齢ダミーの推定値

表-6 個人単位の回答方法選択モデル推定結果²⁾

説明変数	推定値	t値	
性別 (男:1)	-0.07	-2.78	***
年齢	0.021	32.76	***
世帯人数	-0.066	-6.638	***
職業ダミー (情報通信)	0.862	10.52	***
居住地ダミー (西区)	0.19	4.56	***
用途地域面積0ダミー	0.491	8.66	***
定数項 (紙面調査選択)	1.487	28.83	***
サンプルサイズ		71,039	
ρ^2		0.543	
修正済み ρ^2		0.543	

t値: 1.96以上5%有意**, 2.58以上1%有意**

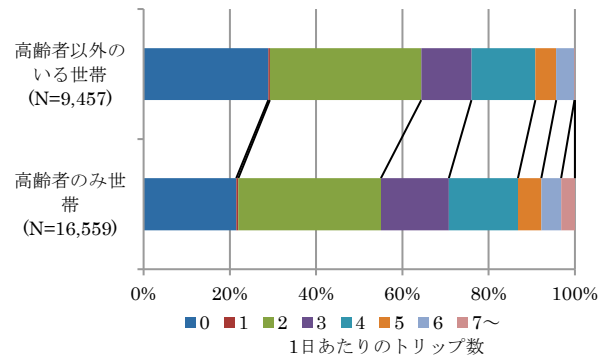


図-8 高齢者トリップ数比較

高齢者のいる世帯に分類し、平均トリップ数を分析した結果である。熊本PT調査の紙調査は世帯表と個人票があり、個人票については1人一枚配布される。よって、紙調査はweb調査と比較し代理回答が少ない、また高齢者のみ世帯は高齢者のいる世帯に比べ代理回答が少ないと考えられる。各分類の平均トリップ数から、代理回答の可能性が高くなるほど平均トリップ数が少なくなっている。このように代理回答による代理回答バイアスの影響を推測できる。

(2) 代理回答される確率の算出法

代理回答を分析するために、第3章(3)節の回答有無におけるロジットモデルの推定結果を用いる。Aを個人jが回答する事象、Bを世帯iが回答する事象とする。また、世帯の回答確率を $\sum_{j=1}^{J_i} P_{i,j}$ 、個人の回答確率を $P_{i,j}$ とすると、代理回答される確率 P_{proxy} は条件付き確率を利用して以下で与えられる。

$$P(B) = \sum_{j=1}^{J_i} P_{i,j}, P(A) = P_{i,j}$$

$$P(\bar{A}) = 1 - P(A) = 1 - P_{i,j} \quad (7)$$

$$P(\bar{A} \cap B) = P(B) - P(A)$$

$$= \sum_{j=1}^{J_i} P_{i,j} - P_{i,j} \quad (8)$$

$$P(\bar{A}|B) = \frac{P(\bar{A} \cap B)}{P(B)}$$

$$= \frac{\sum_{j=1}^{J_i} P_{i,j} - P_{i,j}}{\sum_{j=1}^{J_i} P_{i,j}} \quad (9)$$

$$= P_{proxy}$$

以上の式から代理回答される確率を算出し、代理回答をされる人がどのような属性の人であるかを分析する。

(3) 代理回答される確率の基礎分析

式(9)を用いて代理回答される確率を算出し分析を行う。図-9は年齢別に代理回答される確率を分析した結果である。高齢者、子どもにおいて代理回答される確率が高い傾向が読み取れる。図-10は性別に代理回答される確率を示している。男性が代理回答されやすいことが確認できる。また、図-11はトリップ数別に代理回答される確率を分析した結果である。0-2トリップにおいて代理回答される確率が80-100%の割合が高い。これらは、代理回答によって正しいトリップ数が記録されていない可能性を示唆する。

(4) トリップ数に着目した代理回答バイアスの分析

図-12は代理回答される確率別に平均トリップ数を算出し、年齢別に示したものである。概ねどの年齢でも代理回答される確率が60-80%、80-100%の平均トリップ数

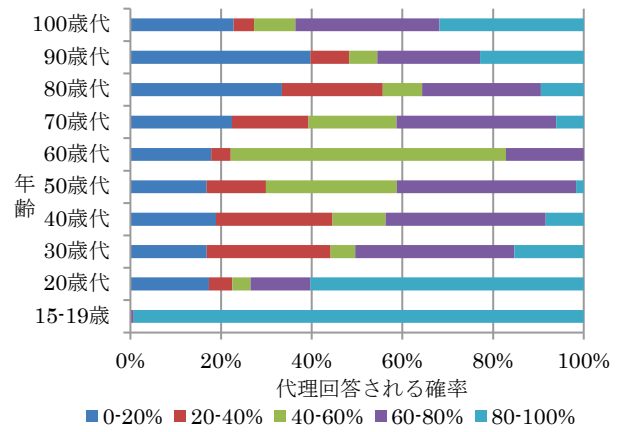


図-9 年齢別代理回答される確率の分布

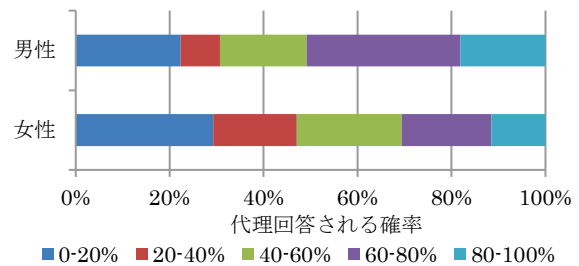


図-10 性別代理回答される確率の分布

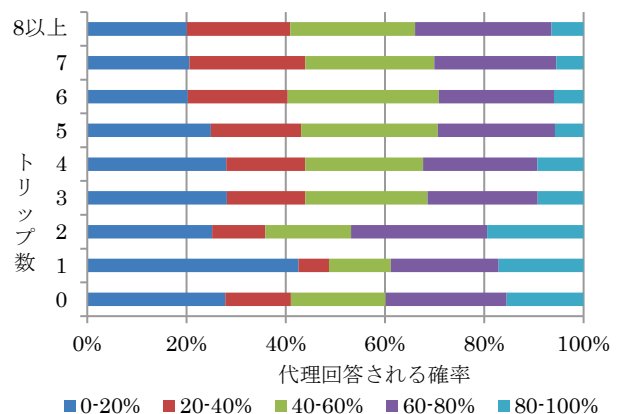


図-11 トリップ数別代理回答される確率の分布

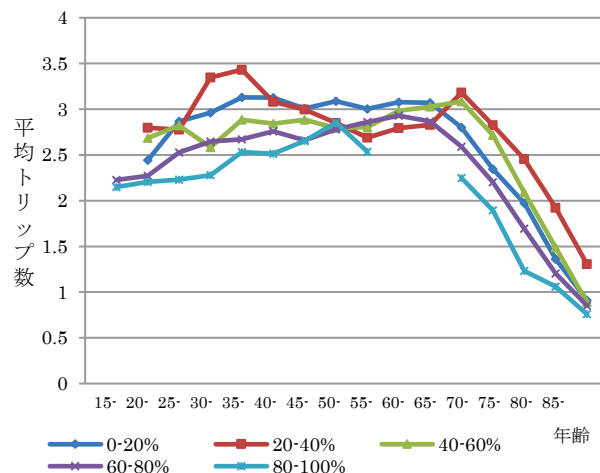


図-12 代理回答される確率別平均トリップ数

は少なくなっていることが分かる。また、代理回答される確率が 0 と 0-20%の人の平均トリップ数が一番大きくなるのが予想されたがそのような結果は得られなかった。これらは、代理回答される確率が 0%の人が大きく影響していることが要因として考えられる。

図-13は代理回答される確率を0-50%、50-100%に分類し、性・年齢別に平均トリップ数を比較した結果である。表-8には図-13から得られた平均トリップ数から検定を行い、差が生じているのかを確認している。検定結果から男性は25-29歳、35-39歳、40-44歳、50-54歳、60-64歳、70歳以上の各年代において有意な差が生じていることが明らかになった。男性はどの年代においても代理回答される確率が0-50%の人が平均トリップ数が多く、0.2トリップほど差が出ている。構築したモデルでは女性の方が調査に回答しやすいという結果が得られていたため、男性は代理回答による影響を受けやすく差が大きくなったことが考えられる。女性は20-49歳の各年代、60-64歳、65-69歳、70-74歳、80-84歳において有意な差が生じている。20-40歳の年代でトリップ数の差が非常に大きくなっている。これは、単身の就業女性と複数人世帯の未就業(主夫)女性とのトリップ数の差が要因の一つと考えられる。

次に、就業状況別に代理回答される確率からトリップ数の代理回答バイアスを分析する。就業者はPT調査の個人表からわかる職業において無職、主婦・主夫、その他(職業なし)を除いたものとする。図-14は男性の就業状況別に代理回答される確率から平均トリップ数を比較した結果である。表-10がその検定結果である。男性就業者に関しては25-29歳、35-54歳の各年代において有意な差が生じており、働き盛りの年代に差が大きいことが分かった。一方、年齢が高くなるほどトリップ数の差は小さくなっている。高齢で就業している人は代理回答の影響を受けにくいと言える。男性未就業では20-24歳、30-

34歳、35-39歳、55-59歳、70歳以上の各年代において有意な差が得られている。若い世帯に差がある一方、就業者と異なり高齢者にも差が生じている。代理回答される確率が低い未就業の高齢者は単身もしくは同年代の女性と2人世帯が多く、代理回答される確率が高い未就業高齢者と比べ、買い物など私事トリップが多いことが考えられる。

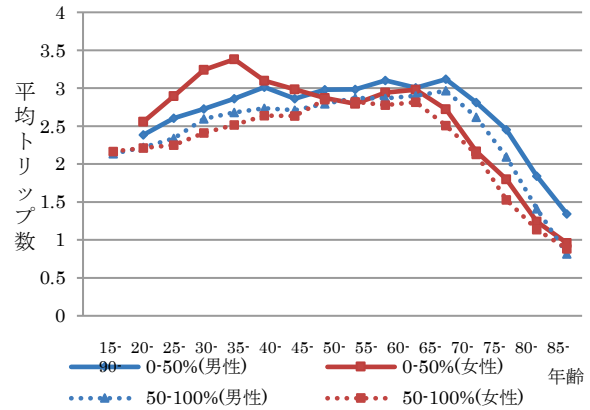


図-13 性別の代理回答される確率別平均トリップ数

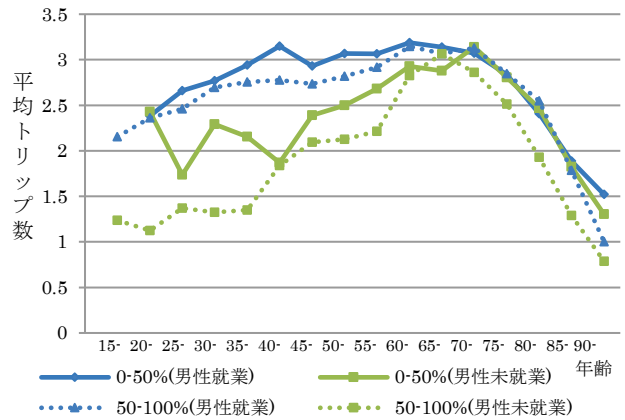


図-14 男性就業状況別代理回答される確率別平均トリップ数

表-9 性別の代理回答される確率別平均トリップ数検定結果

		15-19歳	20-24歳	25-29歳	30-34歳	35-39歳	40-44歳	45-49歳	50-54歳	55-59歳	60-64歳	65-69歳	70-74歳	75-79歳	80-84歳	85-89歳	90歳以上
0-50%(男性)	平均トリップ数		2.38	2.61	2.73	2.86	3.01	2.86	2.98	2.99	3.10	3.01	3.12	2.81	2.45	1.84	1.34
	サンプル数		234	327	287	375	357	319	492	576	833	668	2186	1949	1351	590	134
	標準偏差		1.77	1.67	1.75	1.84	1.79	1.63	1.83	1.95	1.82	1.94	2.07	1.97	1.85	1.68	1.59
50-100%(男性)	平均トリップ数	2.14	2.22	2.34	2.60	2.68	2.74	2.71	2.79	2.87	2.87	2.90	2.97	2.62	2.09	1.41	0.82
	サンプル数	2294	1306	1706	2046	2735	2575	2416	2513	2885	3680	2804	703	489	393	207	71
	標準偏差	1.08	1.38	1.48	1.67	1.71	1.76	1.64	1.77	1.76	1.93	1.98	2.03	2.09	1.90	1.63	1.18
	t値		1.33	2.69	1.18	1.79	2.77	1.58	2.09	1.34	3.33	1.26	1.72	1.87	3.29	3.24	2.69
			***		*	***		**		***		*	*	***	***	***	
0-50%(女性)	平均トリップ数		2.56	2.90	3.24	3.38	3.10	2.98	2.87	2.79	2.94	2.98	2.72	2.17	1.80	1.24	0.96
	サンプル数		396	1176	2018	2775	3015	3004	2818	2833	1559	1113	976	948	873	617	239
	標準偏差		1.64	1.80	1.97	2.05	1.94	1.85	1.70	1.68	1.74	1.88	1.85	1.71	1.72	1.48	1.54
50-100%(女性)	平均トリップ数	2.16	2.21	2.25	2.41	2.52	2.64	2.63	2.85	2.82	2.78	2.82	2.51	2.13	1.53	1.13	0.88
	サンプル数	2224	1610	1403	947	892	583	691	1087	1362	3672	2888	2428	1922	1229	676	320
	標準偏差	1.06	1.19	1.24	1.43	1.58	1.85	1.57	1.75	1.77	1.76	1.84	1.80	1.71	1.55	1.36	1.25
	t値		3.95	10.43	13.00	13.13	5.48	5.10	0.36	-0.42	3.10	2.46	3.10	0.60	3.73	1.35	0.61
			***	***	***	***	***	***			***	**	***		***		

***:1%有意, **:5%有意, *:10%有意

表-10 男性就業状況別の代理回答される確率別平均トリップ数検定結果

		15-19歳	20-24歳	25-29歳	30-34歳	35-39歳	40-44歳	45-49歳	50-54歳	55-59歳	60-64歳	65-69歳	70-74歳	75-79歳	80-84歳	85-89歳	90歳以上
0-50%(男性就業)	平均トリップ数		2.38	2.66	2.77	2.94	3.15	2.93	3.07	3.06	3.19	3.14	3.07	2.84	2.41	1.89	1.52
	サンプル数		227	308	263	336	319	278	416	460	559	335	717	550	359	151	25
	標準偏差		1.78	1.63	1.69	1.80	1.73	1.59	1.78	1.87	1.79	1.95	2.13	1.94	1.85	1.57	1.42
50-100%(男性就業)	平均トリップ数	2.15	2.36	2.46	2.70	2.76	2.78	2.73	2.82	2.92	3.15	3.07	3.13	2.85	2.55	1.78	1.00
	サンプル数	2252	1160	1517	1901	2589	2457	2322	2426	2690	2705	1270	274	155	104	51	10
	標準偏差	1.06	1.30	1.43	1.63	1.68	1.75	1.62	1.75	1.72	1.88	1.94	1.92	2.04	1.88	1.82	1.61
	t値		0.18	2.00	0.65	1.80	3.61	1.97	2.67	1.57	0.50	0.57	-0.39	-0.05	-0.68	0.36	0.89
			**		*	***	**	***									
0-50%(男性未就業)	平均トリップ数	2.43	1.74	2.29	2.15	1.87	2.39	2.50	2.68	2.93	2.88	3.14	2.80	2.47	1.82	1.30	
	サンプル数	7	19	24	39	38	41	76	116	274	333	1469	1399	992	439	109	
	標準偏差	1.40	1.45	2.26	1.98	1.85	1.85	2.03	2.20	1.86	1.92	2.03	1.98	1.84	1.71	1.62	
50-100%(男性未就業)	平均トリップ数	1.24	1.12	1.37	1.32	1.35	1.84	2.10	2.13	2.22	2.83	3.06	2.86	2.51	1.93	1.29	0.79
	サンプル数	42	146	189	145	146	118	94	87	195	975	1534	429	334	289	156	61
	標準偏差	1.44	1.53	1.47	1.61	1.55	1.69	1.97	2.04	2.06	2.05	2.01	2.09	2.11	1.89	1.54	1.09
	t値		2.40	1.05	2.01	2.35	0.09	0.83	1.17	1.85	0.77	-1.60	2.44	2.30	4.27	3.62	2.47
		**		**	**				*			**	**	***	***	**	

***:1%有意, **:5%有意, *:10%有意

表-11 女性就業状況別の代理回答される確率別平均トリップ数検定結果

		15-19歳	20-24歳	25-29歳	30-34歳	35-39歳	40-44歳	45-49歳	50-54歳	55-59歳	60-64歳	65-69歳	70-74歳	75-79歳	80-84歳	85-89歳	90歳以上
0-50%(女性就業)	平均トリップ数		2.59	3.04	3.35	3.46	3.17	3.07	3.01	2.89	3.07	2.98	2.63	2.23	2.05	1.55	1.49
	サンプル数		331	853	1294	1837	2087	2198	2045	1768	882	441	348	290	242	121	39
	標準偏差		1.55	1.69	1.80	1.86	1.79	1.65	1.61	1.51	1.60	1.78	1.91	1.79	1.68	1.71	2.02
50-100%(女性就業)	平均トリップ数	2.18	2.30	2.31	2.51	2.59	2.70	2.76	2.98	2.96	2.86	2.88	2.54	2.19	1.70	1.35	0.87
	サンプル数	2187	1461	1221	800	724	471	543	769	869	1439	776	527	369	187	71	23
	標準偏差	1.0	1.1	1.1	1.3	1.4	1.5	1.4	1.6	1.6	1.6	1.8	1.8	1.7	1.7	1.4	1.2
	t値		3.13	10.94	12.17	12.99	5.89	4.42	0.45	-1.14	3.01	0.91	0.68	0.30	2.16	0.89	1.53
			***	***	***	***	***	***		***				*			
0-50%(女性未就業)	平均トリップ数	2.42	2.52	3.05	3.22	2.94	2.75	2.50	2.64	2.78	2.98	2.78	2.14	1.71	1.17	0.86	
	サンプル数	65	323	724	938	928	806	773	1065	677	672	628	658	631	496	200	
	標準偏差	2.06	2.01	2.24	2.36	2.24	2.30	1.89	1.91	1.89	1.94	1.81	1.67	1.73	1.41	1.41	
50-100%(女性未就業)	平均トリップ数	1.05	1.30	1.81	1.84	2.19	2.35	2.18	2.53	2.57	2.73	2.79	2.50	2.11	1.50	1.11	0.89
	サンプル数	37	149	182	147	168	112	148	318	493	2233	2112	1901	1553	1042	605	297
	標準偏差	1.69	1.41	1.83	1.77	2.28	2.90	1.94	1.96	2.06	1.84	1.85	1.81	1.71	1.53	1.36	1.25
	t値		3.97	4.06	7.16	5.39	2.09	3.20	-0.24	0.66	0.66	2.18	3.33	0.34	2.49	0.67	-0.25
			***	***	***	***	**	***			**	***		**			

***:1%有意, **:5%有意, *:10%有意

図-15 は女性の就業状況別に代理回答される確率から平均トリップ数を比較した結果である。表-11 がその検定結果である。女性の就業者では 20-49 歳の各年代、60-64 歳、80-84 歳で有意な差が出ている。また、未就業者では 20-49 歳の各年代、65-69 歳、70-74 歳、80-84 歳で有意な差が生じており、女性においては差が生じている年代に就業状況別で同じような傾向がある。女性の若い世帯は就業、未就業に関わらず平均トリップ数の差が大きく、代理回答による影響を大きく受けている属性であるといえる。

以上の分析から代理回答バイアスの一部を実証的に示すことができた。

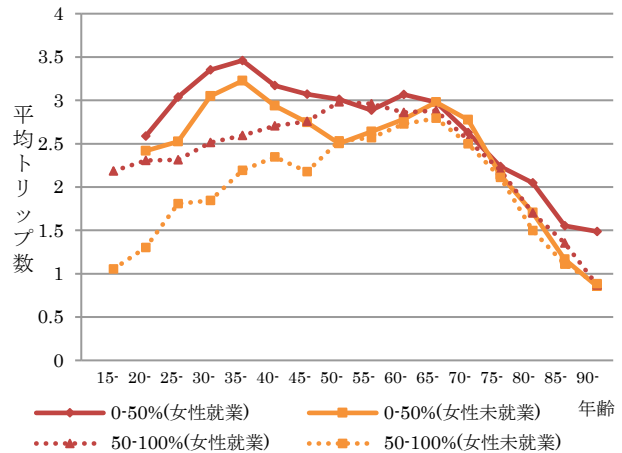


図-15 女性就業状況別代理回答される確率別平均トリップ

6. おわりに

本研究では、熊本 PT 調査でデータを用いて、調査回答選択行動に影響を与える要素を明らかにし、回答行動選択モデルを構築した。さらに構築したモデルを用いて代理回答バイアスの分析を行った。

本研究の成果を以下に示す。

- (1) 世帯単位での分析を可能にする新たな集団意思決定型のロジットモデルを構築した。
- (2) 熊本 PT 調査における回答有無選択行動について、集団意思決定モデルを適用することで、高齢者が

いる世帯ほど PT 調査に参加しやすいこと等を明らかにした。

- (3) 熊本 PT 調査における回答方法選択行動について、集団意思決定モデルを適用することで、世帯人数が多い世帯ほど紙調査を、世帯人数が少ない世帯ほど web 調査を選択することを明らかにした。また、高齢者のみの世帯は紙調査を選ぶ傾向にあり、子供のいる世帯では web 調査を選ぶ傾向にあることを示した。
- (4) 構築したモデルを用いて、世帯単位の選択確率や

世帯構成員一人一人の選択確率を算出することで、代理回答バイアスを分析する手法を提案した。

- (5) 代理回答される確率別に属性別でトリップ数の分析することで代理回答バイアスを実証的に明らかにした。さらに代理回答バイアスを受けやすい属性についても示した。

今後の課題について以下に示す。

- (1) 世帯の中で回答者となりうる人を考慮した説明変数の検討、また集団選択モデルにおいて構成員間の相互作用を考慮した集団離散選択モデルの適用を検討する必要がある。
- (2) サンプルサイズが大きいことでクロスバリデーションの適用が可能である。モデルの有効性を示唆するためにも検討の価値がある。
- (3) 代理回答バイアスを分析する回答項目としてトリップ数以外の回答項目(トリップ時間やトリップ目的)に着目した分析を行う。
- (4) 代理回答の影響によるデータサンプルの偏りを補正する手法の開発を行うことで代理回答バイアスを少なくすることができる。
- (5) 本研究における代理回答とは推測にとどまっている。世帯票の回答者の記入を求める等により、現実の代理回答者のデータを収集することも可能である。それらを利用した分析も今後求められる。

謝辞：本研究は、JSPS 科研費 JP18H01561の助成を受けた成果の一部です。

参考文献

- 1) 高橋勝美, 平見憲司, 森尾淳, 西野仁: 我が国のパーソントリップ調査の無回答状況とその要因に関する考察, 第 39 回土木計画学研究発表会・講演集,#283,2009.
- 2) 円山琢也, 平原雄太郎: スマホ調査を実施した PT 調査における調査参加選択行動モデルの構築, 第 49

回土木計画学研究発表会, 2014.6.

- 3) 小嶋文, 久保田尚: 調査主題への関りから見たサイレント層の生成と特徴~地区交通問題に関する調査を対象として~, 土木学会論文 D, Vol.64, No.3, pp.367-379, 2008.
- 4) 山本俊行, 森川高行: 地域間競争を考慮した買い物物品でモデルの構築~大規模小売店舗の中心市街地への出店時の買い物行動変化の分析の適用~, 都市計画論文集, Vol.48, No.3, pp459-464, 2013.
- 5) 張峻屹, 桑野将司, 藤原章正: 集団離散選択モデルの比較分析~世帯の車種選択を例に~, 土木計画学研究・論文集, No.23, no.2, 2006.
- 6) 張峻屹, A.Borgers, H.Timmermans: 集団効用関数に基づく世帯時間配分モデルの開発及び実証的分析, 土木計画学研究・論文集, Vol.19, No.3, 2002.
- 7) 名取義和, 谷下雅義, 鹿島茂: パーソントリップ調査における回答誤差とその発生要因, 土木計画学研究・論文集, No.17, 2000.
- 8) 林英夫: 郵送調査法の再評価と今後の課題, 行動計量学, 第 37 巻, 第 2 号, pp.127-145, 2010.
- 9) 佐藤嘉洋, 円山琢也: 集団意思決定モデルを用いた益城町仮設住宅居住者の郵送調査回答行動分析, 第 57 回土木計画学研究発表会, 2018.
- 10) 円山琢也, 宮原進, 三ヶ尻裕司, 佐藤嘉洋: 郵送型 PT 調査における動的な予備調査票の投入手法の提案, 交通工学論文集, 第 1 巻, 第 2 号, pp76-81, 2015.2.
- 11) 円山琢也, 照屋尚大, 日高陸生: 選挙投票率と PT 調査参加率の比較分析, 土木学会論文集 D3, Vol.72, No.5, pp.955-962, 2016.
- 12) 総務省ホームページ: 平成 28 年度通信利用動向調査の結果, 男女別インターネットの利用状況の推移 <http://www.soumu.go.jp/johotsusintokei/statistics/statistics05.htm>

(2018. 4. 27 受付)

Group Decision Modelling Approach to Analyze Response Behavior of Household Travel Survey

Kenta HOSOTANI, Tomoki KAWANO, Hajime WATANABE, Yoshihiro SATO
and Takuya MARUYAMA

It is important to carefully examine effects of the increase in non-respondents to social surveys (i.e., a declining response rate). This study analyzes response behavior, including participation choice and response method choice (paper-based or web-based), for the Kumamoto Person Trip (PT) survey. Existing studies have employed individual-based analyses, but this study involved a household-based analysis using a group decision model. We found that the household-based model has greater accuracy than the individual-based model and describe the influence of household members. We also examined the proxy-response bias caused by completion of a travel survey by a household member other than the target one.