

# 拡張型ウィップル指数を利用した交通調査 における代理回答バイアスの分析

吉川 駿汰・円山 琢也

<sup>1</sup>学生会員 熊本大学大学院自然科学教育部土木建築学専攻 (〒860-8555 熊本市中央区黒髪2-39-1)

<sup>2</sup>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 熊本市中央区黒髪2-39-1)

E-mail: takumaru@kumamoto-u.ac.jp (Corresponding Author)

自記式の交通調査では、トリップの出発・到着時刻を 5,10 分等の時刻に丸めて回答しやすい傾向にあり、この丸め傾向を利用した交通調査の回答品質の評価指標が提案されている。本研究は、この指標を調査回答者の情報を収集している米国の世帯交通調査に適用して、属性別、自己・代理回答別の時刻丸め傾向について分析を行った。分析の結果、代理回答された場合、自己回答でトリップ時刻を報告するよりも時刻の回答が不正確になる傾向にあること等が実証的に示された。また、代理回答者自身のトリップ時刻の不正確さが被代理回答者の不正確さやトリップ数に影響があることも明らかにされた。

**Key Words :** proxy response, rounding error, national household travel survey, survey bias

## 1. はじめに

### (1) 背景及び目的

社会調査における回答の質の高さは、社会課題解決のための政策分析に大きな影響を与えるため重要である。交通計画等に利用されるパーソントリップ調査(以下PT調査)においても調査の質の重要性は認識されている。

しかしながら、自記式のPT調査はトリップの回答漏れや回答値の不正確さなどの課題を有している。本研究の分析対象であるトリップ時刻の丸め誤差もその一つで、これらの誤差やバイアスを精査し、その補正法を検討・開発することは調査の質を確保するうえで重要である。

時刻の丸め誤差に関してRietveld<sup>1)</sup>は交通実態調査において、多くの回答者は出発時刻と到着時刻を本来の時刻ではなく、5、15、30の倍数に丸めて報告する傾向を示している。また、佐藤・円山<sup>2)</sup>は人口統計学で用いられるウィップル指数を応用した指標を提案し、世界の都市での調査の質を分析し、地域間の差などを考察している。しかし、トリップ属性や個人属性別には詳しく分析されおらず、代理回答が丸め誤差に与える影響などさらに詳細な研究が求められる。

本研究では、米国のPT調査データを用いて、時刻の丸め誤差の分析を行う。本調査データは個人属性、トリップ属性に加えて、調査に対して対象者本人が回答したか、他の人が代理で回答したのかのデータを把握してい

る。自己・代理回答間の丸め傾向の違いは筆者らの先行研究<sup>3)</sup>でも分析を行っており、代理回答で報告されたトリップは特に30の倍数に丸められやすいことが明らかになっている。しかし、この研究では個人属性別では分析を行っていないため、さらに詳細な回答時刻の丸め傾向の違いの分析が必要である。

具体的に本研究では、米国PT調査データを利用して、

- ・属性ごとの丸め誤差や回答時刻の精度の把握
- ・自己・代理回答ごとの丸め回答傾向や回答時刻の精度の把握
- ・代理回答者と被代理回答者間の回答時刻精度における因果関係の把握

以上を目的とする。

これらの分析により、どういった属性の人が回答時刻の精度が低い傾向にあるのかを明らかにでき、それに対応した丸め誤差を小さくする手法の開発につながりうると考えられる。

### (2) 本研究の構成

本研究の構成は以下の通りである。まず第2章既往研究について、続いて第3章で米国PT調査の概要と使用データを、第4章で評価手法を説明する。第5章では、個人・トリップ属性ごとの回答時刻の精度の評価分析を、第6章で重回帰分析の結果を示し、最後に第7章で本研究をまとめ、今後の展望について整理する。

## 2. 既往研究のレビュー

代理回答によって報告される情報の正確性に関して分析を行っている研究は複数存在している。Maruyama *et al.*<sup>4)</sup>, Maruyama and Uehara<sup>5)</sup>は代理回答確率の推測式を導出することにより、代理回答情報を収集していない熊本PT調査で得られたサンプルに対して自己・代理回答クラスに分類を行い、代理回答クラスは自己回答クラスと比較してトリップの出発・到着時刻を5,10,15,30の倍数に丸めて回答しやすい傾向にあること、代理回答クラスに分類された人の平均トリップ数は自己回答クラスの平均トリップ数と比較して小さくなる傾向にあることを示した。

Maruyama *et al.*<sup>4)</sup>, Maruyama and Uehara<sup>5)</sup>でレビューされている以外の、交通調査関連の代理回答に関する研究として次がある。まず、Beck *et al.*<sup>6)</sup>は同世帯の二人に対して、購入する自動車の条件の選択に関する表明選好調査において、もう一人が回答するであろうと予測した代理回答の正確性を分析した。その結果、代理回答者は代理回答者自身の選好と同じものを被代理回答者の回答として報告しやすいこと、また代理回答者自身と被代理回答者の選好が著しく異なると把握しているとき、代理回答者の報告する情報が不正確である割合が高くなることが明らかにしている。また、Kamruzzaman and Hine<sup>7)</sup>は小・中学生の通学トリップに対して、代理回答により報告されたトリップ情報の有効性を検討した研究を行い、小学生に関しては親が注意深く監視しているため、代理回答による情報が正確であったが、中学生になると報告する情報の精度が落ちることを示した。

代理回答に着目した研究はほかにも多く存在するが、本研究では自己・代理回答別のトリップ開始時刻の丸め傾向の違いについて分析を行っている。交通調査における測定誤差は度々指摘されている問題であり、代理回答による影響も多々あるため、代理回答と丸め誤差の関係についての実態の把握を行うことは非常に重要であり、これを行うことが本研究の新規性および独自性である。

## 3. 調査概要と使用データ

本研究では2017年に米国で実施されたPT調査データ(National Household Travel Survey, NHTS)を用いる。調査の概要は表-1に示す。調査対象世帯は全米50州を対象にランダムに抽出される。

調査手順は事前に抽出された世帯に依頼書を郵送し、参加意向を確認した後、交通行動情報を報告する2段階の調査方式となっている。交通行動調査は指定された日程の一日のトリップを報告するものであり、報告終了後調査参加の謝礼(\$20)が郵送される。

交通行動調査の報告はWeb、電話もしくは両方の3種類あり、調査対象者は世帯ごとに報告方法を任意で選択することが出来る。調査で得られる情報としては調査対象者の性別、年齢、就業状態等の個人属性やトリップ情報、公共交通機関に対しての意識などがある。

本研究ではマスターデータのうち非外出者を除くサンプルを分析対象とする。サンプルサイズは自己回答をした人が159,013人、代理回答をされた人は60,181人、代理回答をした人が42,360人である。なお、本調査は5歳以上の世帯構成員を調査対象としているが、調査において16歳未満の世帯員に関してはその世帯の成人が必ず代理回答をするようになっている。その他、調査の詳細は別途<sup>4)</sup>に報告されている。なお、本研究では、調査サンプル内での代理回答バイアスの精査を目的としているため、拡大係数は利用していない。ここで関連する用語を以下のように定義する。

- 自己回答者 (self respondents): 自分の交通行動等を自分で記入した人。この回答データを単に自己回答と呼ぶ。
- 被代理回答者 (proxy respondents): 自分の交通行動等を他者に記入してもらった人。代理回答された人。この回答データを単に代理回答と呼ぶ。
- 代理回答者 (proxy completers): 他者の交通行動等を代理で記入した人。

表-1 2017年米国NHTS概要

調査期間	2016年3月～2017年5月
調査方法	Web調査, 電話調査 もしくは両方
調査対象地	米国全土
調査対象世帯数	129,696世帯
調査対象人数	264,234人
分析対象人数	219,194人
回答総トリップ数	923,572トリップ

出典)資料<sup>4)</sup>

## 4. 評価手法

### (1) ウィップル指数の概要

ここでは、人口統計学で用いられているウィップル指数について説明する。ウィップル指数は年齢を調査した際、自分の正確な年齢を知らない人が、年齢の下一桁を0もしくは5の自分の年齢に近いと思われる数字で報告してしまうことで発生するエイジヒーピングを測る指標である。ウィップル指数での評価の対象は23歳から62歳までの範囲であり、式(1)、(2)のように定義されている。

$$W_5 = \frac{P_{25} + P_{30} + \dots + P_{60}}{\frac{1}{5}(P_{23} + P_{24} + \dots + P_{62})} \times 100 \quad (1)$$

$$W_{10} = \frac{P_{30} + P_{40} + P_{50} + P_{60}}{\frac{1}{10}(P_{23} + P_{24} + \dots + P_{62})} \times 100 \quad (2)$$

式(1), (2)はそれぞれ5年, 10年間隔での年齢層での不正確さを評価するウィップル指数であり,  $P_n$ は $n$ 歳人口を表している. ウィップル指数は一般的に 100 から 500 の範囲で表され, 数値の小さいほうが正確であると判断される. 式から得られた数値に対しての基準例を表-2に示す.

表-2 ウィップル指数の基準

ウィップル指数	データの質
<105	非常に正確
105-110	比較的正確
110-125	許容
125-175	不良
>175	非常に不良

出典)資料<sup>4)</sup>

次に佐藤・円山<sup>3)</sup>によって定式化されたRDT (Rounding Departure Time) 指数について説明する. RDT指数はウィップル指数の概念を応用し, 報告されたトリップの出発時刻の不正確さを評価する指標である. 出発時刻を $m$ 分単位で回答するときの指数を $RDT_m$ とすると, 5, 10, 15, 30, 60 分単位でのRDT指数はそれぞれ式(3)~(7)のように表される.

$$RDT_5 = \frac{N_{00} + N_5 + \dots + N_{55}}{\frac{1}{5}(N_{00} + N_{01} + \dots + N_{59})} \times 100 \quad (3)$$

$$RDT_{10} = \frac{N_{00} + N_{10} + \dots + N_{50}}{\frac{1}{10}(N_{00} + N_{01} + \dots + N_{59})} \times 100 \quad (4)$$

$$RDT_{15} = \frac{N_{00} + N_{15} + \dots + N_{45}}{\frac{1}{15}(N_{00} + N_{01} + \dots + N_{59})} \times 100 \quad (5)$$

$$RDT_{30} = \frac{N_{00} + N_{30}}{\frac{1}{30}(N_{00} + N_{01} + \dots + N_{59})} \times 100 \quad (6)$$

$$RDT_{60} = \frac{N_{00}}{\frac{1}{60}(N_{00} + N_{01} + \dots + N_{59})} \times 100 \quad (7)$$

$N_m$ は出発時刻が $m$ 分と報告されたトリップ数である. ウィップル指数は個人単位での評価に対して, RDT指数はトリップ単位での評価指標となっている.

また, 回答者一人当たりの出発時刻の不正確さを評価するwRDT (weighted Rounding Departure Time) 指数も同じく

佐藤・円山によって定式化されている. wRDT指数は回答者一人当たりのトリップ数の影響を除くために, 個人ごとのトリップ数を逆数とした重みづけを行っており, 式(8)~(12)のように表される.

$$wRDT_{05} = \frac{\sum_i \frac{N_{i,00}}{n_i} + \sum_i \frac{N_{i,05}}{n_i} + \dots + \sum_i \frac{N_{i,55}}{n_i}}{\frac{1}{5}(\sum_i \frac{N_{i,00}}{n_i} + \sum_i \frac{N_{i,01}}{n_i} + \dots + \sum_i \frac{N_{i,59}}{n_i})} \times 100 \quad (8)$$

$$wRDT_{10} = \frac{\sum_i \frac{N_{i,00}}{n_i} + \sum_i \frac{N_{i,10}}{n_i} + \dots + \sum_i \frac{N_{i,50}}{n_i}}{\frac{1}{10}(\sum_i \frac{N_{i,00}}{n_i} + \sum_i \frac{N_{i,01}}{n_i} + \dots + \sum_i \frac{N_{i,59}}{n_i})} \times 100 \quad (9)$$

$$wRDT_{15} = \frac{\sum_i \frac{N_{i,00}}{n_i} + \sum_i \frac{N_{i,15}}{n_i} + \sum_i \frac{N_{i,30}}{n_i} + \sum_i \frac{N_{i,45}}{n_i}}{\frac{1}{15}(\sum_i \frac{N_{i,00}}{n_i} + \sum_i \frac{N_{i,01}}{n_i} + \dots + \sum_i \frac{N_{i,59}}{n_i})} \times 100 \quad (10)$$

$$wRDT_{30} = \frac{\sum_i \frac{N_{i,00}}{n_i} + \sum_i \frac{N_{i,30}}{n_i}}{\frac{1}{30}(\sum_i \frac{N_{i,00}}{n_i} + \sum_i \frac{N_{i,01}}{n_i} + \dots + \sum_i \frac{N_{i,59}}{n_i})} \times 100 \quad (11)$$

$$wRDT_{60} = \frac{\sum_i \frac{N_{i,00}}{n_i}}{\frac{1}{60}(\sum_i \frac{N_{i,00}}{n_i} + \sum_i \frac{N_{i,01}}{n_i} + \dots + \sum_i \frac{N_{i,59}}{n_i})} \times 100 \quad (12)$$

ここで,  $N_{i,m}$ は個人 $i$ が $m$ 分と回答したトリップ数,  $n_i$ は個人 $i$ が報告した総トリップ数である.

なお, ウィップル指数は報告された年齢分布が特定の年齢に集中せず, 一様な分布であるときに正確と評価される. 一方でRDT指数およびwRDT指数は報告されたトリップの出発時刻が特定の「分」に集中していないときに正確であるとなる. すなわち, RDT指数, wRDT指数はウィップル指数と同じくトリップの真の出発時刻が一樣に分布していると仮定を置いていることに留意されたい.

ここで、 $wRDT_m/m$  ( $m = 5, 10, 15, 30, 60$ ) という指標は、 $m$ 分単位に丸める割合の平均値となる。この値とその標準偏差を表-3に示す。この表から、全トリップの開始時刻の内、84.5%が5分単位に丸められ、23.6%が60分単位に丸められていることがわかる。また、丸められる割合のばらつきは15分単位が最も大きいことが読み取れる。本稿では、データのばらつきを最も表現できている $wRDT_{15}$  に絞って分析を続ける。

表-3 各丸め時刻の回答割合と標準偏差

$m$	$wRDT_m/m$	標準偏差
5	0.845	0.262
10	0.579	0.325
15	0.587	0.339
30	0.423	0.335
60	0.236	0.277

## 5. 個人・トリップ属性別の $wRDT$ ・ $RDT$ 指数

### (1) 個人属性別 $wRDT$ 指数

本章では個人属性に着目し、どういった人が時刻を丸めやすいのか、また丸められやすいのかを分析する。

本研究では、15分単位での $RDT$ 指数および $wRDT$ 指数に着目して分析を行っている。なお、5、10、30、60分単位での $RDT$ 指数及び $wRDT$ 指数については付録に示す。ここで、以降の図の凡例に関して以下のように定義する。

- ・自己回答: 自己回答者自身の $wRDT_{15}$
- ・被代理回答: 代理回答者により報告された被代理回答者の $wRDT_{15}$
- ・代理回答: 代理回答者が報告した代理回答者自身の $wRDT_{15}$

また、性別はそれぞれ自己回答者、被代理回答者、代理回答者自身の性別を表している。

図-1は性年齢別自己・代理回答別の $wRDT_{15}$ を示している。男女ともに被代理回答者の $wRDT_{15}$ が高いという結果になった。これより、被代理回答者のトリップのほうが15分単位に丸めて回答されやすい傾向にあることがわかる。これは代理回答者が被代理回答者の実際の出発時刻を把握できておらず、憶測で回答されてしまうためだと考える。また代理回答された人の $wRDT_{15}$ は自己回答ほど男女間に大きな差はなかった。

図-2は就労状態別の $wRDT_{15}$ を示している。代理回答されたときの平均 $wRDT_{15}$ は「14歳以下の学生」が最も小さいという結果となった。これは主に親が子供のトリップに同伴するケースが多く、時刻を把握できているた

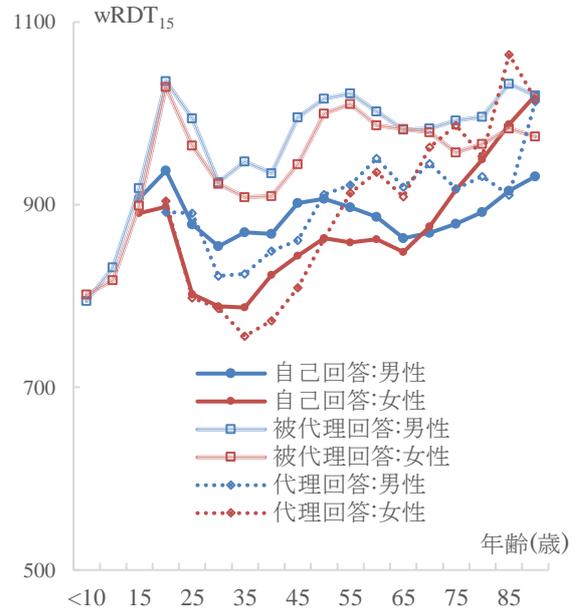


図-1 性年齢別回答種別 $wRDT_{15}$

注)本調査では15歳以下については、20歳以上の世帯構成員が代理回答すると定められている。よって、15歳未満の自己回答と20歳未満の代理回答は除外している。

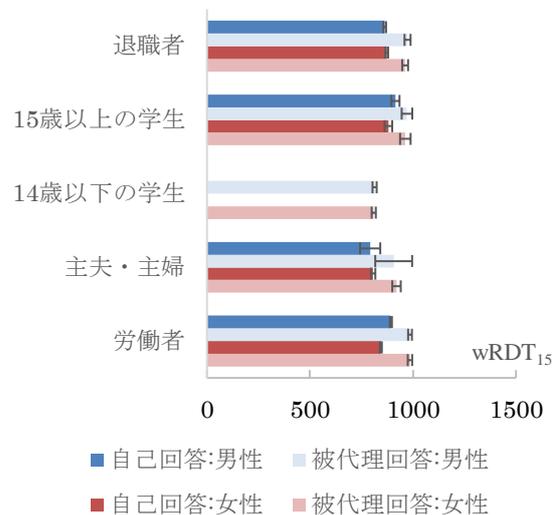


図-2 就労状態別 $wRDT_{15}$

注)図のひげは95%信頼区間を示す

めであると考え。一方、「労働者」や「15歳以上の学生」を代理回答した場合は丸めやすい傾向にあることが分かる。これは日々の通勤・通学のトリップはある程度決まった時間におこなわれるため、時刻を丸めて回答してしまっているのではないかと考える。

図-3はサンプルの最終学歴別のwRDT<sub>15</sub>を示している。「高卒未満」のみ、代理回答された場合のほうが自己回答した場合よりも回答時刻の精度が高くなる傾向になっている。また、自己回答者は教育水準が上がるにつれて正確な時刻を回答しやすい傾向にあることがわかるが、代理回答をされる場合でも同様の傾向が出ている。これは調査対象者の教育レベルが高水準であるとき、その世帯自体の教育レベルが高いと推測されるため、このような結果になったと考える。

図-4、5はそれぞれ代理回答者が男性・女性の場合の被代理回答者の性別wRDT<sub>15</sub>を示している。被代理回答者の年齢が20代から50代において、代理回答者が異性の場合、正確な時刻を回答しやすいことが分かる。この年代では代理回答は多くの場合が夫婦間で行われており、自分のパートナーが代理回答するときには、正確な時刻を回答しやすいことがわかる。一方で15歳以下では比較的wRDT<sub>15</sub>が小さいことがわかる。これは小さな子供のトリップには代理回答者である親が同伴するケースが

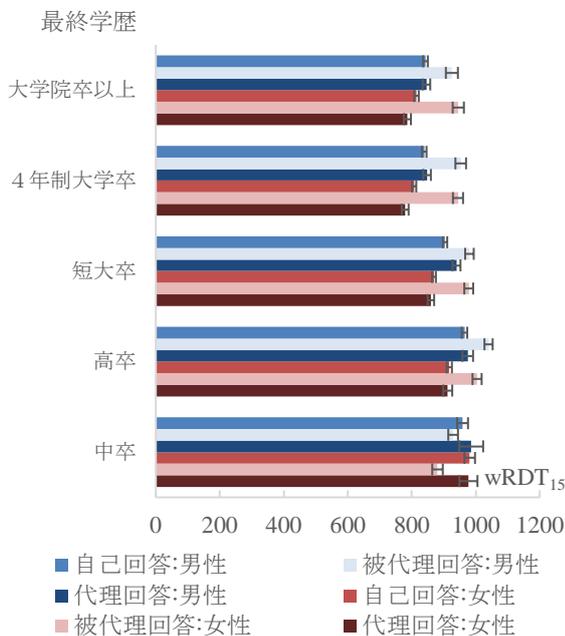


図-3 最終学歴別wRDT<sub>15</sub>

注)図のひげは95%信頼区間を示す

多いため、詳細な時刻を把握できていたためと推測する。

表-4は代理回答者との関係性別のwRDT<sub>15</sub>を示している。特に母親または父親に代理回答をされた場合、wRDT<sub>15</sub>が小さく、正確な時刻が回答されやすい傾向が出ている。一方で自身の子供に代理回答をされる場合にはwRDT<sub>15</sub>が高く、回答された時刻の精度が低くなりやすいことが分かった。

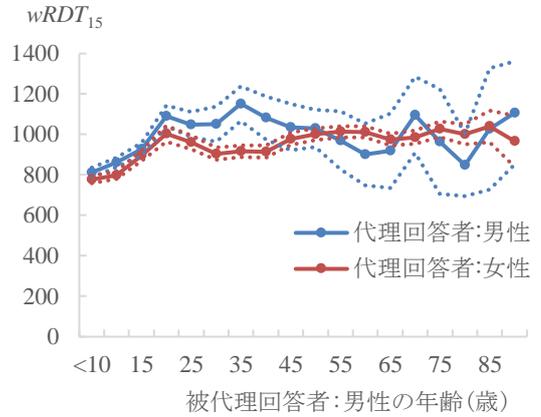


図-4 男性が被代理回答者の場合のwRDT<sub>15</sub>

注)図の点線は95%信頼区間を示す

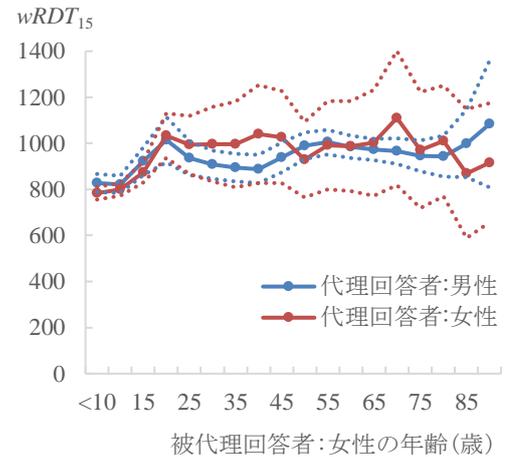


図-5 女性が被代理回答者の場合のwRDT<sub>15</sub>

注)図の点線は95%信頼区間を示す

表-4 代理関係者との関係性別のwRDT<sub>15</sub>

被代理回答者の性別		妻	夫	娘	息子	母親	父親
男性	平均wRDT <sub>15</sub>	957.84	—	999.56	943.26	832.40	881.59
	標準偏差	487.08	—	468.31	484.06	524.45	527.37
	サンプルサイズ	10423	—	160	134	8318	5145
女性	平均wRDT <sub>15</sub>	—	954.01	980.33	941.96	823.79	861.23
	標準偏差	—	501.38	510.63	519.88	520.06	521.45
	サンプルサイズ	—	11690	421	212	7869	4609

(2) トリップ属性別RDT指数

図-6, 7, 8 はそれぞれ HBW(Home-Based Work), HBD(Home-Based Discretionary), NHB(NonHome-Based)トリップの男女別自己・代理回答別RDT<sub>15</sub>を示している。以下にそれぞれのトリップ分類の定義を示す。

- ・ HBW: 出発・到着地点がそれぞれ自宅かつ勤務先のトリップ
- ・ HBD: NHTs データの内, Home-based Shop, Home-based Social / recreational, Home-based other を組み合わせたトリップ。いずれも出発・到着地点がそれぞれ自宅かつ任意の目的地トリップ。
- ・ NHB: 出発・到着地点のいずれも自宅が含まれないトリップ

いずれのトリップにおいても被代理回答者のRDT<sub>15</sub>は自己回答者のRDT<sub>15</sub>よりも大きい傾向にあり、トリップ目的に関わらず、代理回答されたトリップは自己回答されたトリップよりも時刻を丸めやすい傾向にあることが分かる。これらのトリップのうち、最も平均RDT<sub>15</sub>が大きいのは HBW トリップであった。特に自宅から仕事先へ向かうトリップは毎日のように行うトリップであり、出発時刻が固定されていることが考えられるため、時刻

を丸めて報告しやすいと推測される。また、HBW トリップが最も自己・代理回答間で平均RDT<sub>15</sub>に大きな差が見られた。代理回答者も自己回答者と同じくある程度固定された時刻を想像して回答がしやすかったためこのような結果になったと推測される。一方で最もRDT<sub>15</sub>が低かったのは NHB トリップであった。NHB トリップは出

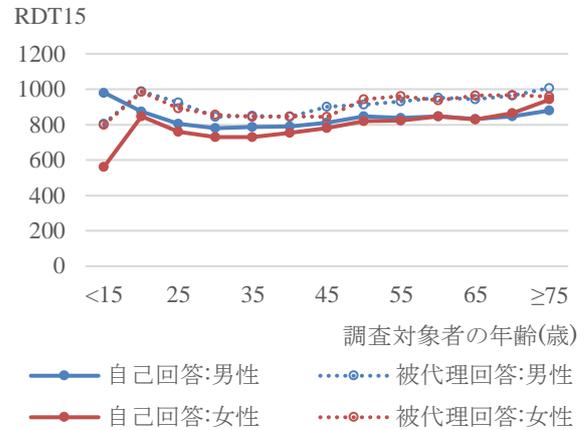


図-7 自己・代理回答別HBDのRDT<sub>15</sub>

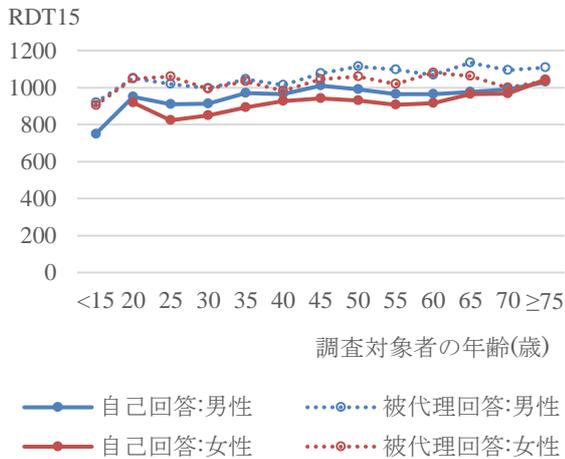


図-6 自己・代理回答別HBWのRDT<sub>15</sub>

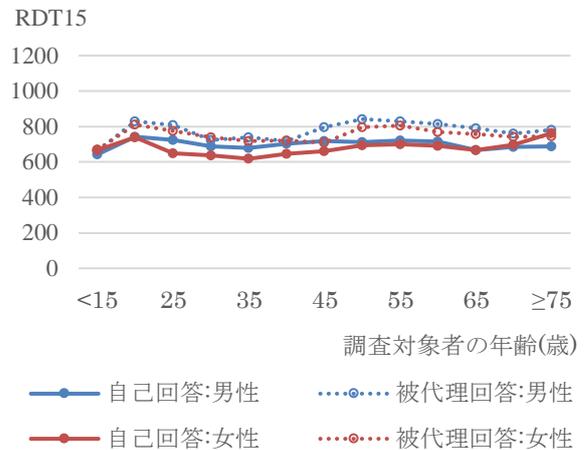


図-8 自己・代理回答別NHBのRDT<sub>15</sub>

表-5 各トリップ分類の自己・代理回答別平均RDT<sub>15</sub>

トリップ分類		自己回答:男性	被代理回答:男性	自己回答:女性	被代理回答:女性
HBW	平均RDT <sub>15</sub>	964.35	1057.98	910.67	1031.53
	標準偏差	718.72	683.85	732.59	695.16
	サンプルサイズ	50232	13591	43588	9881
HBD	平均RDT <sub>15</sub>	836.37	874.77	820.78	873.16
	標準偏差	745.01	739.55	746.65	739.82
	サンプルサイズ	161676	66794	198350	68242
NHB	平均RDT <sub>15</sub>	700.99	745.92	681.25	731.07
	標準偏差	748.40	749.99	746.84	749.76
	サンプルサイズ	108567	31602	139291	31015

表-6 代理回答者同行非同行人別平均 $wRDT_{15}$ 

被代理回答者の性別		同行:同性	同行:異性	非同行人:同性	非同行人:異性
男性	平均 $wRDT_{15}$	814.40	799.16	931.86	889.08
	標準偏差	747.23	748.39	727.62	736.99
	サンプルサイズ	7174	22379	5452	43420
女性	平均 $wRDT_{15}$	768.80	860.98	898.50	875.80
	標準偏差	749.76	741.74	735.15	739.37
	サンプルサイズ	17515	25281	7222	28111

発・到着地点がいずれも自宅ではないため、多くの場合がトリップチェーンの一部である。そのため、調査対象者はより多くのトリップ情報の回答が必要な NHB トリップの回答を控える人がいたのではないかと考えられる。一方で NHB トリップの回答を行う人は、調査全般を正確に回答しており、丸めの少ない時刻を報告しやすかったと推測する。

表-6は代理回答で報告されたトリップの内、代理回答者が被代理回答者のトリップに同行した場合と同行していない場合の平均 $wRDT_{15}$ を代理・被代理回答者が同性か異性が別で示している。代理回答者がトリップに同行すると同行していない場合に比べて、そのトリップの時刻を把握しやすいため、より詳細な時刻を報告しやすいのではないかと考えられる。また、女性の代理回答者が女性の被代理回答者のトリップに同行した場合、 $wRDT_{15}$ がほかの条件よりも比較的小さく、正確な時刻を報告しやすいという結果となった。

### (3) 代理回答者の $wRDT$ 指数が被代理回答者の報告内容に与える影響

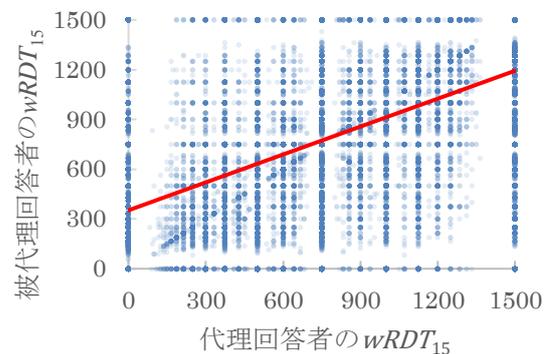
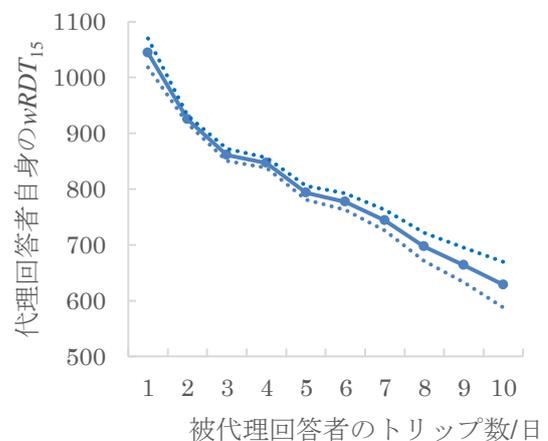
図-8 は代理回答が発生した場合の代理回答者の $wRDT_{15}$ と被代理回答者の $wRDT_{15}$ の相関関係を示した散布図である。なお図-8の赤線は近似直線であり、欠損データを除いた 56,653 サンプルが対象である。2 群間の相関係数は 0.607 と比較強い相関である。すなわち、代理回答者が時刻を丸めて回答しやすいならば、被代理回答者の時刻も丸めて報告しやすく、逆に正確な時刻を報告する代理回答者は被代理回答者の時刻も正確に報告しやすいということである。このことから、代理回答者によって報告された被代理回答者の $wRDT_{15}$ は代理回答者が時刻を丸めて回答しやすいか否かに少なからず影響を受けることがわかる。

図-9 は代理回答者自身の $wRDT_{15}$ と被代理回答者のトリップ数との関係性を示している。ただし、代理回答をされた人のデータのみを用いた分析結果となっており、自己回答を行った人のデータは除いている。

代理回答者が報告する被代理回答者のトリップ数が小さいほど、代理回答者自身の $wRDT_{15}$ は大きくなっていく。

すなわち、代理回答者が時刻を丸めて回答しやすい人ならば、被代理回答者のトリップを少なく回答してしまう傾向にあるということであり、すべてのトリップを報告するのが面倒で数トリップを省いて回答してしまっていると考えられる。

この結果より、代理回答が発生する場合、代理回答者

図-8 代理・被代理回答者の  $wRDT_{15}$  の相関関係図-9 被代理回答者のトリップ数別代理回答者の  $wRDT_{15}$ 

注) 図の点線は95%信頼区間を示す

表-7 代理回答者同行非同性別平均 $wRDT_{15}$ の重回帰分析結果

説明変数	モデル1				モデル2			
	偏回帰 係数	t値	***	標準化偏 回帰係数	偏回帰 係数	t値	***	標準化偏 回帰係数
<b>個人属性</b>								
定数項	949.55	167.9	***		871.67	159.49	***	
男性ダミー	25.07	11.42	***	0.02	8.61	3.93	***	0.01
代理回答ダミー	62.17	21.12	***	0.05	-	-		-
学生ダミー	64.39	11.17	***	0.02	64.12	11.39	***	0.02
労働者ダミー	34.38	14.20	***	0.03	37.09	15.61	***	0.04
ドライバーダミー	24.96	6.71	***	0.02	91.68	26.04	***	0.07
平日ダミー	-53.32	-19.71	***	-0.04	-47.55	-18.00	***	-0.04
<b>世帯属性</b>								
世帯人数	-7.12	-8.01	***	-0.02	-16.30	-19.00	***	-0.04
世帯年収	-8.21	-24.50	***	-0.05	-7.41	-22.56	***	-0.05
都市部ダミー	-36.19	-13.70	***	-0.03	-28.96	-11.23	***	-0.02
<b>代理回答者属性</b>								
代理回答ダミー× 代理回答者の $wRDT_{15}$					30.58	103.33	***	0.27
代理回答ダミー× 男性ダミー(代理回答者)					-125.09	-30.11	***	-0.08
代理回答ダミー× 主婦・主夫ダミー(代理回答者)					-43.55	-4.06	***	-0.01
自由度調整済み決定係数	0.01				0.06			
サンプルサイズ	214,670							

\*: 10%有意, \*\*: 5%有意, \*\*\*: 1%有意

の $wRDT_{15}$ 、つまり代理回答者がどれほど時刻を丸めて回答しやすいかによって、被代理回答者のデータにも影響が出てしまうことが明らかとなった。

## 6. 調査対象者の $wRDT_{15}$ の重回帰分析

本章では重回帰分析を用いて本研究の主題である $wRDT_{15}$ と様々な要因との関係性を明らかにするために2つのモデルを構築し検証を行った。1つ目は代理回答ダミーを、2つ目は代理回答者属性を予測因子に含んだモデルである。本分析ではマスターデータの内、調査対象者および代理回答者の非外出者を除いたサンプルを用いている。なお本分析における重回帰分析は従属変数である調査対象者の $wRDT_{15}$ と説明変数の相関関係を把握することが目的であり、予測を目的としたモデリングではないため、決定係数の大きさは重視しない。

表-7に重回帰分析の結果を示す。モデル1では代理回答ダミーが正に有意となっている。先の分析でも示していたように自己回答、代理回答の選択と時刻の丸めとの相関が示された。また就業状態や世帯属性など各変数と関係性がみられた。また標準化偏回帰係数に着目すると、最も従属変数の $wRDT_{15}$ に大きな影響を与えているのは

代理回答ダミーと世帯年収であった。これより様々な因子がある中でも代理回答は時刻の丸めに相当の影響を持つ変数であるといえる。また世帯年収は負に有意になっている。世帯年収の高い世帯の人は仕事等でトリップが多いこと、在宅時間が短いことなどがその要因として考えられる。

一方でモデル2ではモデル1で有意になった変数に加え、代理回答者属性が有意になる結果となった。中でも代理回答者の $wRDT_{15}$ は標準偏回帰係数が0.27と最も大きい結果となり影響力が強いことがわかった。次いで標準偏回帰係数の値が大きいのは男性ダミー(代理回答者)であった。これらの結果より調査対象者の $wRDT_{15}$ に対して強い影響力のある説明変数は代理回答の有無や代理回答者属性であり、報告される時刻の精度に対して自己回答の重要性が大きいと結論することが出来る。

## 7. 結論

本研究では米国PT調査におけるトリップの出発時刻の丸め誤差の実態の把握と代理回答者の時刻の丸めやすさが与える被代理回答者の回答の影響の把握を行った。本研究の成果を以下に示す。

- ・自己回答者と比較して、代理回答者が報告したトリップ出発時刻の  $RDT$  指数,  $wRDT$  指数は大きい, すなわち時刻を丸めて回答しやすい傾向にある.
- ・60代の男性は女性よりも時刻を15分単位で丸めやすい一方で70歳以降になると女性のほうが時刻を丸めて回答しやすい傾向にある.
- ・代理回答者と被代理回答者が同性か異性かでも傾向に違いがあり, 20代から50代において, 代理回答者が異性の場合, 正確な時刻を回答しやすいことが明らかになった.
- ・代理回答者がトリップに同行している場合, 同行していない場合に比べ, 正確な時刻を報告しやすい傾向にある.
- ・代理回答者の  $wRDT_{15}$  は被代理回答者の  $wRDT_{15}$  および報告されるトリップ数に影響を与えていることを示した.
- ・重回帰分析によって調査対象者の  $wRDT_{15}$  と説明変数との相関を示した. また, 特に代理回答者の  $wRDT_{15}$  の影響が強いことも明らかにした.

今後の展開として以下が挙げられる.

- ・本分析で用いた2017年米国PT調査データは電話調査もしくはWeb調査による回答であったが, 調査手法により,  $wRDT$  指数は変化すると考えられる. よって電話調査のみを利用している2009年調査や自宅訪問も行って調査した1983年調査との比較を行い, 調査方法による丸め回答の実態を把握する.
- ・諸外国, 諸都市との比較を行い,  $wRDT$  指数がどのような要因から影響を受けているかを把握する.
- ・傾向スコア分析等の因果推論を用いて自己・代理回答の回答種の要因のみがトリップ数に与える影響を把握する

謝辞: 本研究は, JSPS 科研費 JP18H01561, JP21H01458 の支援を受けた成果の一部です.

#### 参考文献

- 1) Rietveld, P.: Rounding of arrival and departure times in travel surveys: An interpretation in terms of scheduled activities, *Journal of Transportation and Statistics*, Vol. 5, pp. 71–82, 2002.
- 2) 佐藤 嘉洋, 円山 琢也: ウィップル指数を応用した交通調査データ質評価指標の開発と適用, 第63回土木計画学研究発表会(春大会), 2021.
- 3) 吉川 駿汰, 円山 琢也: 交通調査における代理回答バイアスと丸め誤差: 米国PT調査の事例, 第63回土木計画学研究発表会(春大会), 2021.
- 4) Maruyama, T., Hosotani, K., and Kawano, T.: Inferring proxy response in household travel surveys with unknown completer using a group-based choice model, *Transportation*, Vol. 48(1), pp. 283–302, 2021.
- 5) Maruyama, T. and Uehara, K.: Validating inferred proxy response using rounding of departure/arrival times in travel surveys, *Asian Transport Studies*, Vol. 6, 100027, 2020.
- 6) Beck, M.J., Rose, J.M., and Hensher, D.A.: The accuracy of proxy responses in a stated choice setting: A re-examination and some controversial conclusions, *Transportation Research Part A*, Vol. 46, pp. 226–239, 2012.
- 7) Kamruzzaman, M. and Hine, J.: Self-proxy agreement and weekly school travel behaviour in a sectarian divided society, *Journal of Transport Geography*, Vol. 29, pp. 74–85, 2013.
- 8) United Nations Department of Economic and Social Affairs: *United Nations Demographic Yearbook 1973*, UN, New York, 1974.
- 9) Abounassif, L.: Proxy reporting and travel diary completion status and their impacts on reported trips in National Household Travel Surveys, MS thesis, Tennessee Tech University, 2020.

(Received October 1, 2021)

## ANALYZING PROXY-RESPONSE BIAS IN TRAVEL SURVEYS USING EXTENDED WHIPPLE'S INDEX

Shunta YOSHIKAWA and Takuya MARUYAMA