Does response lag affect travelers’ stated preference? Evidence from a real-time stated adaptation survey

Keishi Fujiwara
Hiroshima University

Varun Varghese
Hiroshima University

Makoto Chikaraishi (chikaraishim@hiroshima-u.ac.jp)
Hiroshima University

Takuya Maruyama
Kumamoto University

Akimasa Fujiwara
Hiroshima University

Research Article

Keywords: stated adaptation, real-time survey, response lag, hypothetical bias

Posted Date: May 25th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2970403/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Transportation on November 7th, 2023. See the published version at https://doi.org/10.1007/s11116-023-10435-8.
Does response lag affect travelers’ stated preference? Evidence from a real-time stated adaptation survey

Keishi Fujiwara
Hiroshima University
1-5-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, 739-8529, Japan
Email: fujiwara-keishi@hiroshima-u.ac.jp

Varun Varghese
Hiroshima University
1-5-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, 739-8529, Japan
E-mail: varunv@hiroshima-u.ac.jp
ORCiD: 0000-0001-6524-7192

Makoto Chikaraishi, corresponding author
Hiroshima University
1-5-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, 739-8529, Japan
E-mail: chikaraishim@hiroshima-u.ac.jp
ORCiD: 0000-0002-8053-9473

Takuya Maruyama
Kumamoto University
2-39-1, Kurokami, Chuo, Kumamoto, Kumamoto, 860-8555, Japan
E-mail: takumaru@kumamoto-u.ac.jp
ORCiD: 0000-0001-8660-1728

Akimasa Fujiwara
Hiroshima University
1-5-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, 739-8529, Japan
E-mail: afujiw@hiroshima-u.ac.jp
ORCiD: 0000-0002-3354-0631
Does response lag affect travelers’ stated preference?
Evidence from a real-time stated adaptation survey

ABSTRACT

Stated preference (SP) surveys typically ask respondents to make a choice under a hypothetical situation. However, the choice context is often unrealistic, leading to errors and biases in the response. To overcome this, revealed preference (RP) data has been used to create a more realistic choice context for generating SP questions. For instance, in stated adaptation (SA) surveys, users are asked to answer SP questions based on a specific RP context they actually experienced. One challenge in SA surveys is that it is difficult for the respondents to precisely recall the RP context, especially when there is a longer time gap between RP behavior and SP response (response lag). However, no empirical studies have been conducted to test how elicited preferences vary in response to changes in the response lag. This study empirically examines the impact of response lag on SP responses using real-time SA survey data collected from Kumamoto and Hiroshima, Japan. To accomplish this, we developed a survey tool that enables respondents to answer SP questions in real time, i.e., immediately after their RP behavior. The empirical results confirmed that systematic bias increases with an increase in response lag. Additionally, the results showed that the greater the response lag, the more respondents tended to focus on the SP attributes rather than the RP attributes. These findings indicate that the timing of responses is an important survey design parameter when conducting a SA survey.

Keywords: stated adaptation, real-time survey, response lag, hypothetical bias
1. INTRODUCTION

The stated preference (SP) survey can be a useful tool as it helps to capture preferences towards new services or goods that are yet to be introduced. It also allows for measuring travel demand and policy effects without conducting social experiments, which are usually expensive and thus difficult to conduct frequently. However, SP survey data is known to suffer from various biases, including affirmation bias, unconstrained bias, justification bias, and policy response bias. These biases arise essentially because the choice results do not affect the actual life of the respondents (1-4). We often call these biases caused by SP surveys hypothetical biases. Conventional SP surveys ask respondents to answer questions under hypothetical scenarios. Thus, when the respondents answer the survey, they may not be able to consider various contextual factors, i.e., the factors which represent the reality of the travel scenario. This may cause differences between their actual behavior and their stated preferences.

In order to alleviate these biases in SP data, several methods have been proposed, many of which incorporate information on actual behavior taken by respondents, i.e., revealed preference (RP) information (5-10). One such method is the stated adaptation (SA) survey (8-10), where the respondents are asked to answer SP questions within their real decision context (i.e., RP context). The advantage of this approach is that various contextual factors, such as scheduling constraints experienced by the respondents, are taken into account, thereby reducing hypothetical biases (9). With the advancement of the web- and smartphone-based surveys, it has become easier to incorporate RP information while asking SP questions. However, this would be valid only when the respondents could properly recall their memory.

One of the major factors influencing the accuracy of memory recall is the time elapsed after engaging in RP behavior. As the time-gap duration between the RP behavior and SP response (referred to as "response lag" in this study) increases, respondents may find it difficult to accurately recall the RP context. While there have been numerous studies examining the impact of incorporating RP data on SP responses, as reviewed in the following section, to the best of the authors’ knowledge, no study has explored how the response lag influences SP answers.

In this study, we quantitatively analyze the influence of response lag on preference elicitation results using a real-time SA survey in Japan. To accomplish this, we developed a survey tool that can reflect the RP context and allow respondents to answer SP questions in real time (11). The real-time SA survey asked respondents to state their behavioral adaptations to a hypothetically introduced congestion pricing scheme, where the alternatives include: (1) keep the present trip under a road pricing scheme; (2) cancel the trip; (3) change the time of day; (4) change the destination; (5) change the travel mode; and (6) change the route. It is worth noting that the 1st and 6th options may not significantly impact their activity-travel schedule after the trip, while the other options could affect travel decisions for the remainder of the day. For example, “cancel trip” and “change the destination” options could affect the destination of the next trip due to the change in the origin; “change the travel mode” from car to public transit may make it difficult to reach some of the destinations where public transit services are not available; and “change the time of the day” could potentially force them to cancel some of the trips they had planned to do after the current one. In summary, some of the behavioral adaptations would require travelers to make considerable changes to their schedule, i.e., have a higher cost of schedule adjustment, while others would not. It can be expected that the respondents can recall such schedule adjustment costs when answering the SP question immediately after making the trip, while they may not be able to do so when
answering the SP question at a later time. To investigate these impacts of response lag, this study
formulates the following three hypotheses:

H1. The greater the response lag, the greater the systematic bias in the choice result, i.e., the
respondents tend to choose the options with relatively higher schedule adjustment costs.

H2. The greater the response lag, the more difficult it becomes for the respondent to reconstruct
the RP context in answering SP questions, i.e., the response lag acts as a moderator variable
that reduces the impacts of the RP attributes on the choice result.

H3. The greater the response lag, the more hypothetical (and the less real) the choice context
becomes, resulting in a more dominant influence of the SP attributes on the choice result, i.e.,
the response lag acts as a moderator variable that increases the impacts of the SP attributes on
the choice result.

For the first hypothesis, we test the possibility of systematic bias in SP responses
attributed to the response lag, i.e., whether or not options with relatively higher schedule
adjustment costs tend to be chosen with the increase in response lag.

Regarding the second and third hypotheses, we test the hypothesis that the relative
impacts of RP (or SP) attributes on the choice gets weaker (or stronger) as the responses lag
increases (or decreases), i.e., by evaluating the impact of the response lag as a moderator variable.
Essentially, we investigate whether longer response lags lead respondents to rely predominantly
on SP attributes while disregarding the effects of the RP context.

The structure of this research is as follows. Section 2 provides a literature review, while
Section 3 describes data utilized in the study. The results of the basic analysis are presented in
Section 4. Section 5 describes the model used in the estimation. Section 6 presents the estimation
results using data from the real-time SA survey, and Section 7 concludes the paper by discussing
key findings, contributions, and directions for future research.
2. LITERATURE REVIEW

2.1 How to improve the reliability of SP surveys

There is a rich body of literature which demonstrates the usefulness of SP surveys for obtaining users’ preferences in hypothetical situations (1, 12-13). However, when respondents answer in the SP survey, various biases such as affirmation bias may occur due to the experimental nature of the method, so it is challenging to obtain accurate users’ preferences.

In order to deal with the problem associated with SP surveys, a number of methods have been proposed such as sending a message to users about a possibility of bias in the survey when they answer (Cheap Talk (14)) or requesting them to attach a confidence level (e.g., 0-10 points (14)) to their choices. Methods that combine SP and RP surveys have also been widely used to reduce the above-mentioned biases. In the following sections, we review the methods that combine SP and RP surveys.

2.2 Bias adjustment method for SP survey using RP information

The combined RP-SP model was proposed to enhance the advantages and reduce the disadvantages of both RP and SP data, allowing to control biases potentially contained in SP data (1). Additionally, an RP-SP simultaneous model with correlated error terms has been proposed (15). However, these models do not directly consider how RP contextual factors affect the SP choice, and thus the RP-SP model can be seen as a compromise method that is useful when RP and SP data have already been independently collected.

2.3 SP survey design method combined with RP survey

In order to make the choice task more realistic, a number of SP survey designs that utilize RP information have been proposed. One such method is the pivoting SP survey method (5), which creates SP attribute levels based on the user’s information obtained in the RP survey. It adjusts the attribute levels by increasing or decreasing them based on a reference obtained from the RP survey. However, in the pivoting SP survey, RP information is only utilized to make the levels of SP attributes realistic, and the survey is not designed to encourage respondents to consider the RP context while answering SP questions. For instance, in a real-life context, some people may decide to pay the congestion charge if they have to attend an important meeting and cannot afford to be delayed. Such RP contextual information would be very important to estimate how many people would change their behavior after the introduction of congestion charge, but the pivoting SP survey cannot take this aspect into account. It has been reported that ignoring these contextual factors can lead to bias, for example, overestimating the preference for new travel modes (16, 17).

Another method of combining RP information with SP surveys is the SA survey (8-10). In an SA survey, questions are designed to reflect the user’s actual context and experience, allowing for SP responses that consider the RP context. The SP-off-RP method (6) also aims to obtain SP responses considering the RP context, similar to the SA survey. However, in SA surveys, respondents may find it difficult to precisely recall the RP context with the increase in time after their response, i.e., an increase in response lag. As the response lag increases, the SA survey may not be able to gather SP responses that adequately account for the influence of the RP context. Therefore, we have developed a survey tool that can reflect the RP context in real time, while
measuring the response lag (11). The real-time stated adaptation (SA) survey allows respondents to answer SA questions in real time, utilizing RP information. This approach facilitates easier recall of the RP context and is expected to yield relatively reliable SP answers. However, to the best of our knowledge, there are no existing studies on real-time SA surveys in the transportation field except ours. In summary, the effectiveness of real-time responses has not been confirmed yet, and this study attempts to address this gap in research.

2.4 Studies on the effect of response time on choice decisions

Several studies have examined the impact of response time on choices in SP surveys. For instance, research on paper-based surveys has shown that when respondents were provided with self-reported reflection time, their preference for new goods decreased (18). Meanwhile, in web-based SP surveys, it was observed that an increase in response time increased the preference for new goods (19). Moreover, in SP surveys, a longer response time was observed to decrease the variance of error terms and the variance of randomly distributed taste parameters in a mixed logit model (20-21). However, another study found contrary evidence and observed that response time increased error variance (22).

It is important to note that the response time examined in the aforementioned studies differs from the response lag addressed in this study. Specifically, two distinct mechanisms may influence the choice results based on the timing gap between RP behavior and SP response: 1) memory effects (also called “response lag” effects in this study), where a larger timing gap makes it more challenging for the respondents to accurately recall the RP context, and 2) time to think effects, where more time spent contemplating the choice task may lead to more precise SP answers. While the latter effects cannot be disregarded for complex choice tasks (such as selecting different environmental policy options affecting individual and collective welfare), we assume that the former effects dominate in the context of behavioral adaptations to a hypothetical congestion pricing scheme. This is because the survey primarily focuses on individual preferences and in principle, do not directly involve others' preferences. Additionally, the time-to-think effects are expected to manifest within a relatively short time gap (e.g., whether the respondent takes an additional 5 minutes to think), while the memory effects are expected to arise with longer time gaps (e.g., whether the respondent answers SP questions today or the following day). This study mainly focuses on relatively longer time gaps that would allow us to consider the influence of time gap as memory or response lag effects.

3. DATA COLLECTION METHOD

A real-time SA survey was conducted in Kumamoto and Hiroshima metropolitan areas during January and February 2020. The survey included 150 people selected from each city who regularly pass through or visit the city center by car. Each respondent was requested to install a mobile phone application on their device, which recorded their RP behavioral histories, including travel times, for all trips made over the subsequent two weeks. The respondents were asked to fill in an RP survey before each trip to record their travel mode and purpose. During each trip, the respondent received a push notification to answer the SA question immediately, if he or she met the following three conditions: a) using a car, b) passing through the congestion charging area, departed from the area, or arrived at the area (Figure 2), and c) traveling during a specific time of day (from 6:00
to 19:00). This push notification allows respondents to answer the question in real-time. In this study, the term "response lag" refers to the time difference between pressing the button after completing the trip and answering the SP survey question.

In the SP survey, participants were presented with a hypothetical scenario involving the payment of a congestion charge to enter and move around the city center. They were then asked to answer questions on behavioral changes. The SP survey displayed various combinations of attributes to the respondents on their smartphone screens, as shown in Figure 3. The survey included four SP attributes: 1) travel time reduction, 2) basic pricing level, 3) pricing level during off-peak hours, and 4) start time of off-peak hours. Based on the above-mentioned attributes, the travel time reduction and the congestion charge were calculated using formulas in Table 1. Two different plans were utilized to compare user preferences for different pricing schemes. In plan 1, the congestion charge was set randomly, while in plan 2, the congestion charge was calculated based on travel time in the congestion charging area (see Table 1). The respondents were divided into two user groups, and both the congestion pricing schemes were implemented for each group during different time periods of the survey, as shown in Figure 4. The survey included six alternative options for behavioral change: 1) keep the present trip under congestion pricing scheme, 2) cancel the trip, 3) change the time of day, 4) change the destination, 5) change the travel mode, and 6) change the route.

Figure 3 shows an example of an SP choice scenario presented to the user in Hiroshima (translated from Japanese to English in the figure) after completing their trip. The question is presented to them via a push notification once their trip is completed. In this particular example, the congestion charge was set at 500 JPY, travel time reduction would be 6 min (30 min of total travel time), and they made the trip between 09:00-16:00. The respondents were then asked to choose from the six alternatives mentioned above.
FIGURE 2 The metropolitan areas of Hiroshima (left) and Kumamoto (right)

<table>
<thead>
<tr>
<th></th>
<th>Plan 1</th>
<th>Plan 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Travel time reduction</td>
<td>$z \times x$ (minute)</td>
<td>$z \times x$ (minute)</td>
</tr>
<tr>
<td>1. [6:00 – $t$]</td>
<td>$y$ (JPY)</td>
<td>$z/30 \times y$ (JPY)</td>
</tr>
<tr>
<td>2. [t – 16:00]</td>
<td>$y \times m$ (JPY)</td>
<td>$z/30 \times y \times m$ (JPY)</td>
</tr>
<tr>
<td>3. [16:00 – 19:00]</td>
<td>$y$ (JPY)</td>
<td>$z/30 \times y$ (JPY)</td>
</tr>
<tr>
<td>4. others</td>
<td>$0$ (JPY)</td>
<td>$0$ (JPY)</td>
</tr>
</tbody>
</table>

$z$: Travel time in the pricing area
$y$: SP attribute representing the basic pricing level. {50, 100, 250, 500, 1000}
$m$: SP attribute representing the price level during off-peak. {0.3, 0.4, 0.5}
$t$: SP attribute representing start time of off-peak. {9:00, 10:00, 11:00}
$x$: SP attribute representing travel time reduction. {0.1, 0.2, 0.3, 0.4}
FIGURE 3 Screen of SP survey (translated from Japanese to English)
4. THE RESULTS OF DESCRIPTIVE ANALYSIS

Table 2 presents the distribution of SP responses, which can be categorized into unchanged and changed options. In one-third of SP scenarios, respondents chose “unchanged”, indicating their intention to continue the current trip and pay the congestion charge. Meanwhile, in the remaining two-thirds of SP scenarios, respondents chose “changed”, which included five different behavioral changes. These results indicate that the congestion pricing scheme has a non-negligible impact on travel behavior. Among the travelers who opted to change their behavior, a relatively small number of people chose “cancel the trip” (2nd alternative) and “change the destination” (4th alternative). This could be attributed to the challenges involved in cancelling or altering plans for commuting and business trips, as they would require significant adjustments to schedules. Conversely, respondents who chose “change the route” (6th alternative) were the largest, which implies that travelers tend to keep the original departure time, destination, and travel mode, but change their route to avoid the congestion charge. Figure 5 shows the relationship between SP choices and trip purpose. Respondents with duty purpose (business trips) showed the highest percentage (42.4%) for the 1st alternative, i.e., “no behavior change and pay the congestion charge”. Meanwhile, respondents traveling for commute purpose have the lowest percentage of selecting the 1st alternative (30.8%). This suggests that for business travelers, it is more difficult to change their travel behavior patterns (routes and travel modes) and they prefer to pay the congestion charge instead. On the other hand, commuters avoid the charge by changing their travel behavior more often as compared to other travel purposes. Figure 6 shows the relationship between SP choices and the levels of congestion charge. It was observed that as the congestion charge increased, the percentage of users choosing to pay the charge and continuing with their current behavior (i.e., “no behavior change and pay the congestion charge”) decreased. Additionally, when the congestion charge level was less than 100 JPY, the percentage of people who chose to pay the charge was extremely high (50.6%). Figure 7 demonstrates the relationship between SP choices and response lag. It was observed that a longer response lag corresponded to a higher percentage of respondents choosing the 1st alternative (“no behavior change and pay the congestion charge”). This suggests that a longer response lag may diminish the impact of the congestion charge, indicating the presence of potential systematic bias in stated preferences.
TABLE 2 Distribution of SP responses

<table>
<thead>
<tr>
<th>Alternative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unchanged</td>
<td>635 (34.4%)</td>
</tr>
<tr>
<td>1 Keep the present trip under congestion pricing scheme</td>
<td></td>
</tr>
<tr>
<td>Changed</td>
<td>1211 (65.6%)</td>
</tr>
<tr>
<td>2 Cancel the trip</td>
<td>43 (2.3%)</td>
</tr>
<tr>
<td>3 Change the time of day</td>
<td>150 (8.1%)</td>
</tr>
<tr>
<td>4 Change the destination</td>
<td>51 (2.8%)</td>
</tr>
<tr>
<td>5 Change the travel mode</td>
<td>144 (7.8%)</td>
</tr>
<tr>
<td>6 Change the route</td>
<td>823 (44.6%)</td>
</tr>
<tr>
<td>Total</td>
<td>1846 (100%)</td>
</tr>
</tbody>
</table>

FIGURE 5 Relationship between choices and trip purpose [sample size]
5. MODELING FRAMEWORK

To test the three hypotheses mentioned in Section 1, we utilize the data from the SA survey to model behavioral changes resulting from the introduction of a congestion pricing scheme. Although the survey presented six adaptation options to the respondents, for the purpose of modeling the behavioral change, we merged them into three alternatives based on the cost of scheduling adjustment: (1) “no behavioral change”, which is used as a base alternative, (2) “change the route”, and (3) “other behavior change”, which includes options 2-5 (cancel the trip, change the time of day, change the destination and change the travel mode). The 3rd alternative entails a higher schedule adjustment cost, while the 2nd alternative (change the route) incurs a relatively lower schedule adjustment cost. Since the real-time SA survey involves multiple trips made by the
same respondents, resulting in multiple SP answers from each respondent. This would cause correlations across answers from the same respondent. To account for this correlation, we employ a panel mixed logit model (MXL model). We compare the estimation results of the MXL model with that of a multinominal logit model (MNL model).

5.1 The MNL model with response lag effects

In the formulation of the MNL model with response lag effects, the random utility function is defined as:

\[ U_{ikj} = V_{ikj} + \varepsilon_{ikj}, \quad (1) \]

where \( V_{ikj} \) represents the deterministic part of the utility function for individual \( i \)'s \( k \)-th choice of alternative \( j \), and \( \varepsilon_{ikj} \) is an error term following the standard Gumbel distribution. It is worth noting that in the SA survey, the error term (unobserved factor) would contain unobserved RP contextual information. Furthermore, the effect of such RP elements is expected to decrease as the response lag increases due to the memory gap. To account for this, we formulate the systematic utility function \( V_{ikj} \) as follows:

\[ V_{ikj} = \theta_{ik}^{SP} \beta_{j}^{SP} x_{ikj}^{SP} + \theta_{ik}^{RP} \beta_{j}^{RP} x_{ikj}^{RP} + \beta_{j}^{At} x_{ikj}^{At} + \gamma_j \ln(L_{ik}) + \beta_j^{const} \quad (2) \]

where \( x_{ikj}^{SP} \) and \( x_{ikj}^{RP} \) are vectors of explanatory variables obtained from the SP and RP surveys, respectively. \( L_{ik} \) is the response lag in minutes for individual \( i \) during the choice context \( k \), and \( \beta_{j}^{SP} \) and \( \beta_{j}^{RP} \) are vectors of coefficient parameters for the SP and RP variables, respectively. \( x_{ikj}^{At} \) is a vector of explanatory variables representing individual attributes, while \( \beta_{j}^{At} \) is a vector of coefficient parameters for the same. \( \beta_j^{const} \) is a constant term for each alternative, and \( \gamma_j \) is the parameter coefficient for the response lag. In this research, we introduce the scale parameters \( \theta_{ik}^{SP} \) and \( \theta_{ik}^{RP} \) to quantify the influence of the response lag on the effects of RP and SP attributes. The scale parameters \( \theta_{ik}^{SP} \) and \( \theta_{ik}^{RP} \) are expressed as follows:

\[ \theta_{ik}^{SP} = f_B(\alpha^{SP} \times \ln(L_{ik})) \quad (3) \]
\[ \theta_{ik}^{RP} = f_B(\alpha^{RP} \times \ln(L_{ik})) \quad (4) \]

where, \( \alpha^{SP} \) and \( \alpha^{RP} \) are unknown parameters to be estimated, they capture the effect of response lag on the scale parameters \( \theta_{ik}^{SP} \) and \( \theta_{ik}^{RP} \). When determining the scale parameters, the form of the utility function is modified to increase the expected utility. Function \( f_B \) is expressed by the following equation:

\[ f_B(x) = \begin{cases} 
\exp x, & (\lambda = 1) \\
(1 + (1 + \lambda)x)^{1-\lambda}, & (otherwise) 
\end{cases} \quad (5) \]

The above Box-Cox type of \( f_B \) is introduced to reflect a potential non-linear impact of response lag on the scale parameters. In this research, parameters are estimated by manually
varying $\lambda$, and the parameter value of $\lambda$ which produces the highest likelihood is retained. Overall, by directly including the response lag as an explanatory variable and introducing scale parameters for SP and RP attributes as functions of the response lag, we aim to confirm the presence of systematic bias due to the response lag (hypothesis 1) and quantify the influence of the response lag on the effects of RP and SP attributes on the choice results (hypotheses 2 and 3).

5.2 MXL model

In this research, we incorporate the unobserved inter-individual heterogeneity using a MXL model. The MXL model accounts for this heterogeneity through an additional error term $\eta_{ij}$. The random utility function is defined as:

$$U_{ikj} = V_{ikj} + \eta_{ij} + \epsilon_{ikj} \tag{6}$$

where $\eta_{ij}$ is an error term that varies across alternatives and individuals, following a normal distribution with a mean of 0 and standard deviation of $\sigma_j$. The MXL model uses the same idiosyncratic error term $\epsilon_{ikj}$ and equations for the systematic utility $V_{ikj}$ as the MNL model.

6. ESTIMATION RESULTS

The explanatory variables representing RP attributes ($x_{ikj}^{RP}$) and individual attributes ($x_{ikj}^{At}$) are shown in Table 3. Since the primary focus of this study was to examine the influence of response lag on SP choices and its moderation on the effects of RP and SP variables, only two dummy variables representing individual attributes were included in the models, specifically to test the impact of age and income (one dummy variable for each attribute). Meanwhile, SP attributes ($x_{ikj}^{SP}$) consist of the congestion charge (JPY) and time saving (min) variables, as explained in Section 3. To estimate the models, we initially tested the potential non-linear impact of response lag by varying the value of $\lambda$ in eq. (5). Through this analysis, we determined that the final likelihood is maximized when $\lambda = 1$.

The estimation results using the value of $\lambda = 1$ are shown in Table 4 for the MNL model and Table 5 for the MXL model. Comparing the two models, the MXL model demonstrates improved performance with higher log final likelihood and adjusted $\rho^2$ values, indicating an improved model performance. This also suggests that the effect of unobserved heterogeneity among individuals is likely to be present.

The estimation results reveal two major findings. Firstly, as shown in Table 4 and Table 5, the parameter representing the response lag ($\gamma_f$) is significant with a negative sign for alternatives 2-5 (cancel the trip, change the time of day, change the destination, and change the travel mode), as well as alternative 6 (change the route) in both the MNL and MXL models. This implies that the respondents are more likely to choose alternative 1 (no behavioral change) as response lag increases. This tendency suggests that the response lag introduces systematic bias in the SP data and diminishes the impact of the congestion pricing scheme on the respondent’s choices. Secondly, the results of both models show that the estimated parameters $\alpha^{SP}$ and $\alpha^{RP}$ are positive and negative, respectively. Furthermore, $\alpha^{RP}$ is statistically significant for both models (at the 10 percent significance level for MXL). This indicates that when the response lag is large,
the SP attributes have a larger impact on the respondents’ choices, while the RP attributes have smaller impacts. These estimation results support the three hypotheses presented in Section 1 regarding the impact of response lags. However, it should be noted that the parameter $\alpha^{SP}$ was not statistically significant.

To visualize the effect of the response lag on the contribution of SP attributes, RP attributes, individual attributes, and the error term towards the total variance of utility differences, we employed the variance decomposition (23) and represented it graphically. Figure 8 illustrates the results for MNL, where the vertical axis shows the proportion of variance attributed to SP attributes, RP attributes, individual attributes (At), and the error term, with the sum set to 100. The results for MXL, which also include the error term for individual heterogeneity ($\eta_{ij}$) is presented in Figure 9. It is worth noting that the error term (white noise) may include unobserved RP contextual factors, and thus we anticipate that its contribution will decrease as the response lag increases. The horizontal axis of both figures represents the response lag, indicating the time elapsed between presenting the SP survey questions to the users and their actual response (1[minute], 1[hour], 3[hour], 1[day], 3[day], 10[day]).

With regard to the proportions of variances, both MNL and MXL models showed that RP attributes accounted for a relatively small proportion compared to SP attributes. As the response lag increased for all alternatives, the proportion of SP attributes increased, while the proportions of RP attributes, individual attributes, and error terms decreased. Moreover, it was observed that the contribution of SP attributes was greater for the alternatives 2-5 (cancel the trip, change the time of day, change the destination, and change the travel mode) as compared to alternative 6 (change the route). These findings suggest that respondents tend to consider RP context more when responding to SP questions with shorter response lags. This highlights the importance of real-time responses in encouraging respondents to consider the RP context effectively. In addition, it was observed that the effect of the idiosyncratic error term representing unobserved inter-individual heterogeneity is higher than the effect of unobserved inter-trip heterogeneity.

Table: Note that the RP attribute may act as a covariate that influences both the response lag and the choice result, which can introduce potential bias in the estimation results (24). To address this issue and ensure the robustness of our findings, we conducted additional estimation considering the impact of RP contextual covariates using propensity scores. The method and estimation results can be found in the Appendix. Overall, the results suggest that although there are some differences in the estimated parameters, the effect of response lag on the choice outcome and the signs of the parameters $\alpha^{SP}$ and $\alpha^{RP}$ are consistent with the results obtained without considering the propensity score. This further supports our hypotheses.

**TABLE 3 Variables for RP data**

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Type of Variable</th>
<th>Variable description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Individual attribute</td>
<td>1: ≤ 35 years old 0: &gt; 35 years old</td>
</tr>
<tr>
<td>Income</td>
<td>Individual attribute</td>
<td>1: ≥ 10 million JPY/year 0: &lt; 10 million JPY/year</td>
</tr>
<tr>
<td>Arrival time to destination</td>
<td>RP attribute</td>
<td>1: between 6 am - 9 am 0: otherwise</td>
</tr>
<tr>
<td>Commuting</td>
<td>RP attribute</td>
<td>1: if trip purpose is commuting 0: otherwise</td>
</tr>
<tr>
<td>Duty</td>
<td>RP attribute</td>
<td>1: if trip purpose is duty 0: otherwise</td>
</tr>
</tbody>
</table>
Table 4 Estimation result of behavioral change model (MNL)

<table>
<thead>
<tr>
<th>Choice</th>
<th>Change the route</th>
<th>Other behavior change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimation</td>
<td>t value</td>
</tr>
<tr>
<td>Constant</td>
<td>$-1.40 \times 10^{-1}$</td>
<td>$-0.834$</td>
</tr>
<tr>
<td>Congestion</td>
<td>$2.73 \times 10^{-3}$</td>
<td>$3.37^{**}$</td>
</tr>
<tr>
<td>Time saving (min)</td>
<td>$-1.94 \times 10^{-2}$</td>
<td>$-2.86^{**}$</td>
</tr>
<tr>
<td>Commuting</td>
<td>1.08</td>
<td>$3.23^{**}$</td>
</tr>
<tr>
<td>Duty</td>
<td>$-2.45 \times 10^{-1}$</td>
<td>$-1.01$</td>
</tr>
<tr>
<td>Pick-up</td>
<td>$-1.37 \times 10^{-1}$</td>
<td>$-0.386$</td>
</tr>
<tr>
<td>Arrival time to destination</td>
<td>$-6.25 \times 10^{-1}$</td>
<td>$-2.14^{*}$</td>
</tr>
<tr>
<td>Age</td>
<td>$4.93 \times 10^{-1}$</td>
<td>$3.92^{**}$</td>
</tr>
<tr>
<td>Income</td>
<td>$-4.96 \times 10^{-1}$</td>
<td>$-2.73^{**}$</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>$-7.46 \times 10^{-2}$</td>
<td>$-2.59^{**}$</td>
</tr>
<tr>
<td>$\alpha_{SP}$</td>
<td>$6.57 \times 10^{-2}$ (0.798)</td>
<td></td>
</tr>
</tbody>
</table>

Initial likelihood: $-2028.038$
Final likelihood: $-1794.357$
Number of observations: 1846
Adjusted $\rho^2$: 0.1152

Note: “Other behavior change” include cancel the trip, change the time of day, change the destination and change the travel mode. Significance levels: ‘***’ 1%, ‘*’ 5%, ‘+’ 10%. “Pay the fee and perform the same action as the current one” was the base alternative.

Figure 8 Variance decomposition of MXL model for other behavior change (including cancel the trip, change the time of day, change the destination and change the travel mode) (left) and 6(change the route) (right)
TABLE 5 Estimation result of behavioral change model (MXL)

<table>
<thead>
<tr>
<th>Choice</th>
<th>Change the route</th>
<th>Other behavior change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimation</td>
<td>t value</td>
</tr>
<tr>
<td>Constant</td>
<td>$-5.54 \times 10^{-1}$</td>
<td>$-1.12$</td>
</tr>
<tr>
<td>Congestion</td>
<td>$4.01 \times 10^{-3}$</td>
<td>$3.68^{**}$</td>
</tr>
<tr>
<td>Charge (JPY)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time saving (min)</td>
<td>$-3.39 \times 10^{-2}$</td>
<td>$-2.53^{**}$</td>
</tr>
<tr>
<td>Commuting</td>
<td>$5.46 \times 10^{-1}$</td>
<td>$0.714$</td>
</tr>
<tr>
<td>Duty</td>
<td>$-2.86 \times 10^{-1}$</td>
<td>$-0.433$</td>
</tr>
<tr>
<td>Pick-up</td>
<td>$1.21 \times 10^{-1}$</td>
<td>$0.150$</td>
</tr>
<tr>
<td>Arrival time to destination</td>
<td>$-4.52 \times 10^{-1}$</td>
<td>$-0.850$</td>
</tr>
<tr>
<td>Age</td>
<td>$7.01 \times 10^{-1}$</td>
<td>$1.02$</td>
</tr>
<tr>
<td>Income</td>
<td>$-1.24$</td>
<td>$-2.36^{**}$</td>
</tr>
</tbody>
</table>

$\gamma$ $-7.38 \times 10^{-2}$ $-1.66^+$ $-1.62 \times 10^{-1}$ $-3.17^{**}$

$\sigma$ $2.92$ $9.89^{**}$ $2.29$ $10.1^{**}$

$\alpha^SP$ $8.05 \times 10^{-2}$ $(1.12)$

$\alpha^{RP}$ $-1.02 \times 10^{-1}$ $(-1.72^+)$

Initial likelihood $-2028.038$

Final likelihood $-1328.416$

Number of observations 1846

(Respondents) (150)

Adjusted $\rho^2$ 0.3331

Note: “Other behavior change” include cancel the trip, change the time of day, change the destination and change the travel mode. Significance levels: ‘**’ 1%, ‘*’ 5%, ‘+’ 10%. “Pay the fee and perform the same action as the current one” was the base alternative.

FIGURE 9 Variance decomposition of MXL model for Other behavior change (including cancel the trip, change the time of day, change the destination and change the travel mode) (left) and 6(change the route) (right)
In this research, we evaluated the impact of response lag on choice behavior using data from real-time SA surveys conducted in Kumamoto and Hiroshima metropolitan areas in January and February 2020. The findings of this research support the hypotheses formulated in Section 1. The key findings can be summarized as follows:

1. The discrete choice models show significant negative impact of response lag on choices of other alternatives (including cancel the trip, change the time of day, change the destination and change the travel mode) and route change, relative to paying the congestion charge. These findings support the first hypothesis, which suggests that a longer response lag leads to a larger systematic bias in the choice results.

2. The analysis of variance decomposition provided insights into the changes in the contribution of different attributes with varying response lag. The results support the second hypothesis, indicating that as the response lag increases, the impact of RP attributes on the choice result tends to decrease.

3. Finally, it was observed that as the response lag increases, the impact of SP attributes on the choice result tends to increase, supporting the third hypothesis.

Based on the results of the study, it can be argued that a longer response lag in SA surveys, leads to respondents placing more emphasis on the SP attribute rather than considering the RP context. This introduces biases in the results of their responses and subsequent policy implications. For example, the results show that even with a high congestion pricing scheme, respondents often choose not to change their behavior, when the response lag was large. Such findings may encourage policy-makers to implement pricing schemes, but this study confirms that the systematic bias in responses increases with an increase in response lag. Typically, the decision-making process in congestion pricing schemes is likely to take place well before the start of the trip. This is especially true for regular daily commute trips, as users can prepare for a behavior change more carefully while taking the congestion charge into consideration. However, in certain scenarios the users’ decision-making processes may be similar to that of the SP survey in this study, and real-time responses may be effective. These scenarios include non-daily or non-regular trips, situations when travelers do not have prior information about the congestion charge, dynamic application of the congestion charge, and during the implementation of dynamic incentive systems such as Incentrip (25).

This study has some limitations and a number of remaining tasks. Firstly, we measured systematic bias through the estimated parameter for the response lag variable. Ideally, systematic bias should be measured by comparing actual with predicted choices. However, as that was not possible in this study, we formulated a hypothesis based on estimated model parameters. Secondly, studies in the future should try to estimate the model considering trip chaining, since respondents may answer the SP questions collectively (i.e., after the completion of all trips in a trip chain) and the response lag may be large if the trip chaining is complex.

The results of this research, which demonstrate the influence of response lag on SP responses, hold valuable implications for the development of transportation survey applications and the formulation of reliable transportation policies. Survey applications can be designed to encourage and incentivize lower response lags, prompting users to consider their real contexts when responding to SP questions. This approach would help mitigate the biases associated with SP surveys. To the authors’ knowledge, there are presently no commercial applications capable of integrating real-time SP questions into smartphone-based travel diary surveys. Developing such
an application would be crucial for obtaining more reliable SP data.

**APPENDIX TESTING THE ROBUSTNESS OF THE RESULTS USING PROPENSITY SCORES**

When estimating the model in Section 5, it was assumed that RP attributes only influence the choice result. However, in reality, RP attributes may also affect the response lag as well, leading to differences in response lag depending on trip-related contextual factors. For instance, commuters may find it more challenging to respond in the morning compared to the evening. To address this, the RP attribute is represented as a covariate $X$ that affects the response lag and choice result, the response lag is considered as a treatment variable $Z$, and the choice result is considered as an outcome variable $Y$. Using propensity scores, the study empirically examines the influence the RP attribute on both the response lag and the choice result.

Since the processing variable $Z$ (response lag) is a continuous variable in this study, the propensity score method for continuous quantities as proposed by Imai (26) is applied. A propensity score function is characterized by $\theta$. Firstly, the estimated value of $\hat{\theta}(X) = X^T \zeta$ is calculated by employing a linear regression model using the covariate $X$, treatment variable $Z$ and an unknown parameter $\zeta$. This estimated value serves as the propensity score. Subsequently, the propensity scores are used to stratify the data, dividing it into different strata based on the estimated propensity scores. The weight, denoted as $W_{kl}$, for a particular response lag $k$ in stratum $l$, is calculated by dividing the number of individuals ($n_{kl}$) with that response lag in the stratum by the total number of individuals ($N_l$) in the stratum. This weight calculation is performed using eq. (A-1).

$$W_{kl} = \frac{n_{kl}}{N_l} \quad (A-1)$$

The estimation process incorporates a weighted likelihood function to account for the propensity score weights. The weight for individual $i$ at response delay $k$ in stratum $l$, denoted as $w_{kli}$, is the reciprocal of $W_{kl}$ divided by the expected value of $W_{kl}$. This weight is used to weigh the contributions of each individual in the likelihood function.

$$LL = \sum_{k=1}^{K} \sum_{j=1}^{J} w_{kli} * d_{ij} * \ln P_i(j) \quad (A-2)$$

where, $P_i(j)$ is the probability of individual $i$ choosing alternative $j$ and $d_{ij}$ is a dummy variable which indicates whether individual $i$ chose alternative $j$ or not. The results of the model estimation using the above equations are shown in Table 6. Meanwhile, the results of variance decomposition are shown in Figure 10. The results show that $\alpha^{RP}$ is negative and statistically significant, indicating a similar trend to the results obtained without considering the propensity score. Similarly, for variance decomposition, the results are similar to that of the models without considering the propensity score. These results confirm that the three hypotheses are supported in this model, as they were in the case without propensity score consideration.
**TABLE 6** Estimation result of behavioral change model using MNL model with propensity score

<table>
<thead>
<tr>
<th>Choice</th>
<th>Change the route</th>
<th>Other behavior change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimation</td>
<td>t value</td>
</tr>
<tr>
<td>Constant</td>
<td>$-7.49 \times 10^{-2}$</td>
<td>$-0.394$</td>
</tr>
<tr>
<td>Congestion charge</td>
<td>$3.70 \times 10^{-3}$</td>
<td>$3.94^{**}$</td>
</tr>
<tr>
<td>Time saving</td>
<td>$-2.74 \times 10^{-2}$</td>
<td>$-3.26^{**}$</td>
</tr>
<tr>
<td>Commuting</td>
<td>$9.28 \times 10^{-1}$</td>
<td>$2.41^{**}$</td>
</tr>
<tr>
<td>Duty</td>
<td>$-1.07 \times 10^{-1}$</td>
<td>$-0.384$</td>
</tr>
<tr>
<td>Pick-up</td>
<td>$-2.15 \times 10^{-1}$</td>
<td>$-0.529$</td>
</tr>
<tr>
<td>Arrival time to destination</td>
<td>$-5.33 \times 10^{-1}$</td>
<td>$-1.52$</td>
</tr>
<tr>
<td>Age</td>
<td>$5.00 \times 10^{-1}$</td>
<td>$3.59^{**}$</td>
</tr>
<tr>
<td>Income</td>
<td>$-4.52 \times 10^{-1}$</td>
<td>$-1.93^{+}$</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>$-8.53 \times 10^{-2}$</td>
<td>$-2.79^{**}$</td>
</tr>
<tr>
<td>$\alpha_{SP}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_{RP}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial likelihood</td>
<td>$-2.028.038$</td>
<td></td>
</tr>
<tr>
<td>Final likelihood</td>
<td>$-1.771.375$</td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>1846</td>
<td></td>
</tr>
<tr>
<td>Adjusted $\rho^2$</td>
<td>0.1157</td>
<td></td>
</tr>
</tbody>
</table>

Note: Choices 2-5 are cancel the trip, change the time of day, change the destination and change the travel mode. Significance levels: ‘**’ 1%, ‘*’ 5%, ‘+’ 10%, “Pay the fee and perform the same action as the current one” was the base alternative.
AUTHOR CONTRIBUTION STATEMENT

The authors confirm contribution to the paper as follows: study conception and design: K. Fujiwara, V. Varghese, M. Chikaraishi, T. Maruyama; analysis and interpretation of results: K. Fujiwara, V. Varghese, M. Chikaraishi, T. Maruyama; draft manuscript preparation: K. Fujiwara, V. Varghese, M. Chikaraishi, T. Maruyama, A. Fujiwara. All authors reviewed the results and approved the final version of the manuscript. The authors do not have any conflicts of interest to declare.

ACKNOWLEDGEMENT

The authors would like to acknowledge that a part of this research was conducted under the research project “Short-term travel demand prediction and comprehensive transport demand management”, supported by the Committee on Advanced Road Technology under the authority of the Ministry of Land, Infrastructure, Transport, and Tourism of Japan.

REFERENCES


25. incenTrip, https://incentrip.org