Travel Behaviour and Society 33 (2023) 100632

Contents lists available at ScienceDirect

Travel Behaviour and Society

journal homepage: www.elsevier.com/locate/tbs

Empowering revealed preference survey with a supplementary stated preference survey: demonstration of willingness-to-pay estimation within a mode choice case

Maliheh Tabasi [a], Alireza Raei [a], Tim Hillel [b], Rico Krueger [c], Taha Hossein Rashidi [a,*]

[a] Research Center for Integrated Transport Innovation, School of Civil and Environmental Engineering, The University of New South Wales, Sydney 2052, Australia
[b] Department of Civil, Environmental and Geomatic Engineering, University College London (UCL), London, UK
[c] Department of Technology, Management and Economics, Technical University of Denmark (DTU), Denmark

ARTICLE INFO

Keywords:
Mode choice model
joint SP-RP dataset
Nonlinear utility
Multinomial Logit
Willingness to pay

ABSTRACT

Mode choice models play a pivotal role in transport demand modelling and help transport planners, engineers and researchers with policy and infrastructure investment evaluation. Recent mode choice studies primarily use revealed preference (RP) data to reflect individuals' true behaviour. However, this may not be the best practice, given the lack of information in RP data. This study uses a nonlinear utility specification for a multinomial logit mode choice model development using high-quality travel data collected by a GPS-based smartphone application complemented by stated preference (SP) data. The model results highlight the impact of sociodemographic variables on mode choice behaviour and individuals' willingness-to-pay (WTP) when the model is jointly developed compared to stand-alone SP and RP models. The main message of this study is that in addition to collecting RP, which is a reliable and unbiased source of data, collecting complementary SP data is beneficial as it provides information that is not otherwise available in RP data. This may include a proper variation in the public transport cost variable as demonstrated in this study. Moreover, to better understand the travellers' behaviour regarding the trade-off between time and cost a mixed multinomial logit (MMNL) model in the willingness to pay space is developed on the SP data. capturing the unobserved heterogeneity within the estimated WTPs, the MMNL model outputs reveal a higher variation in WTP of car in-vehicle travel time compared to bus in-vehicle travel time.

1. Introduction

Estimating mode choice models are an inseparable part of a transport demand modelling project. They are of great importance to researchers and policymakers to gain insights into individuals' travel behaviour and their willingness to pay for various travel attributes. This information allows authorities to understand travellers' choices more accurately by assisting them in identifying effective transport policies.

Many mode choice models have been developed so far. Most of them studied the effect of sociodemographic characteristics on individuals' travel behaviour for different trip purposes. Examples include commuters (Aziz et al., 2018; Bhat, 2000), non-workers (Manoj and Verma, 2015), women in developing countries (Arman et al., 2018), bicycling obstacles for women (Abasahl et al., 2018), and school-goers mode choice (Assi et al., 2018; Badri, 2013; Müller et al., 2008). Research has also explored influential factors on active mode choice behaviour (Aziz et al., 2018) when competing with other travel modes (Gurumurthy and Kockelman, 2020).

Among mode choice studies, a considerable number of them have focused on the willingness-to-pay (WTP) estimation and model parameter estimation. As an abstract concept, the WTP is a critical input in transport investments and policy evaluation. Being an abstract concept that does not exist in the real world and is subject to extensive personal perceptions, it allows researchers to quantify travel demand and human behaviour (Small, 2012). Moreover, the WTP measures individuals' willingness to pay to reduce their travel time. Therefore, it allows travel behaviour modellers to have a realistic understanding of the trade-off between time and cost of travellers. There is heterogeneity in individuals' WTP (Brownstone and Small, 2005) and it can depend on various factors, including contextual variables, sociodemographic

* Corresponding author.
E-mail address: rashidi@unsw.edu.au (T. Hossein Rashidi).

https://doi.org/10.1016/j.tbs.2023.100632
Received 4 December 2022; Received in revised form 21 June 2023; Accepted 25 June 2023
Available online 5 July 2023
2214-367X/

characteristics, attitudinal factors, and modal attributes (Hossan et al., 2016; Small et al., 2005). It is argued to vary over different geographical areas. Therefore, it is important to estimate it for each area based on the data collected from that specific area.

In the transport literature, there have been mode choice models developed based on SP data (Catalano et al., 2008; Patterson et al., 2005); however, a large part of the recent literature is focused on using RP data for mode choice modelling purposes (Abasahl et al., 2018; Arman et al., 2018; Habib and Weiss, 2014; Malokin et al., 2019; Manoj and Verma, 2015; Ortúzar and Willumsen, , 2011; Parady et al., 2021) which could be due to their advantages in capturing actual travel behaviour of participants. In contrast, SP data are collected through hypothetical scenarios and may not accurately reflect real-life behaviour. Nonetheless, SP data have value in allowing researchers to explore scenarios and alternatives that may not be feasible in real life. RP data typically provide self-reported information, such as travel costs, but may suffer from limitations such as a lack of information on non-selected options. Methods such as the Google Maps directions application programming interface (API) (Hillel et al., 2018) can extract information on these alternatives, but this process can be complex, time-consuming, and prone to errors. Therefore, it is beneficial to use both RP and SP datasets in mode choice modelling to obtain the advantages of each dataset and mitigate the limitations of the other.

Through the development of mode choice models using SP and RP datasets and the analysis of travel behaviour and the value of time among various socio-demographic groups, this study conducts an empirical investigation to emphasise the importance of incorporating SP data in the design phase when planning to collect RP data, as it leads to improved model estimates.

2. Literature review

Two of the first studies, which used the joint SP-RP data instead of RP or SP data to develop a mode choice model, were conducted by Bradley and Daly (1991) and Hensher & Bradley, (1993). They defined generic time and cost parameters, and the only mode-specific parameters were the utility alternative-specific constants associated with SP and RP data in the context of an MNL model. In a vehicle fuel preference study, Brownstone et al., (2000) also highlighted the importance of combining SP and RP datasets within the mixed logit context. Brownstone et al. (2000) developed an intercity travel mode choice model based on joint SP-RP data with psychometric attributes introduced in SP data and entering the choice model as latent variables.

Several recent studies have also used mixed SP-RP data. Developing MNL mode choice models on mixed SP-RP data, Anta et al., (2016) have explored the impact of weather conditions and traffic congestion on travellers' mode choices with a focus on motorised vehicles. They found that rail-based transport is more resilient to weather conditions than buses and cars, and car mode attained the most negative utility to traffic congestion and weather conditions. In a recent study, Lizana et al., (2021) developed several joint mode-departure time choice models with different specifications (MNL, nested logit (NL), and mixed MNL (MMNL), each with three different time intervals including 15, 30, and 60 min) using SP-RP travel data from Santiago workers to analyse the impact of congestion pricing on workers' travel behaviour. Their model's output shows that the value of travel time varies as the time unit of the analysis is altered. Cherchi & Ortúzar (2006) extensively discussed the approaches regarding considering alternative-specific constants for newly introduced options in SP experiments and the related formulation and interpretation for demand prediction using joint SP-RP data. In general, a combination of both RP and SP data can assist the researcher in incorporating the advantages of both types of data into modelling estimation (Ahern and Tapley, 2008; Ben-Akiva and Morikawa, 1990), therefore, may lead to more robust models in terms of reflecting individuals' travel behaviour (Guzman et al., 2021).

This research's RP data are travel diaries of great spatiotemporal precision as they are collected using GPS-based smartphone applications and capture travellers' real choices. Hence, they let us develop models with more accuracy regarding travel choice predictions and policy evaluation, as the model's quality depends on data quality. Moreover, GPS-based smartphone application data can be used to overcome the problem of underreporting trips which happens more frequently in conventional travel survey data (Bricka and Bhat, 2006). Further discussion regarding the benefits and challenges of app-based data, considered an emerging form of data in transport research, is provided by Gadziński (2018) and Harrison et al. (2020). However, it is suggested that relying merely on RP data may not be the best practice.

As transport models and their associated outcomes are context-based and cannot readily transfer from one spatial or temporal point to another (Ortúzar and Willumsen, , 2011), our study provides another benchmark for what to consider as the Sydney residents' WTP for travel time reduction during inner-city travels while emphasising on the impact of SP data on parameter estimation. This detailed information allows policymakers to assess inner-city transport regulations or investments in a more disaggregate approach, making the analysis more rigorous.

Some previously estimated WTP for different trip characteristics in the Sydney Great Area are summarised in Table 1. The ranges of reported WTP for both car and transit trips are not sufficiently small (even after correcting for inflation). For instance, in the research by Hensher, (2006), the WTP for car travel time saving is estimated to be 11.66AUD/hour, while in another study two years later, it is reported to be 24.88AUD/hour (Hensher, 2008). Estimating WTP by various methods, data sources, and studies can be beneficial in assisting with the selection of a reliable WTP value (or a range) for policy appraisal because a minor change in WTP value may have a considerable impact on the outcome of the analysis. Furthermore, different components of travel time in public transport may have different values to travellers, which are not considered in these studies except for Hensher & Rose, (2007). On top of that, all the studies in this table either used SP or RP data. However, we will present detailed WTP estimates for different modes and different segments of travel time for public transport users in Sydney using the joint SP-RP dataset and compare our results to these previously estimated WTP.

The primary contribution of this paper is to offer empirical evidence

Table 1
The willingness-to-pay reported for the Sydney Great Area.

Study	Trip characteristic	Willingness-to-pay (AUD/hour)	Dataset type
Economics & Nut Farm, (2004)	rail in-vehicle time	8.76	SP
Hensher, (2006)	car travel	11.66	SP
Hensher & Rose, (2007)	transit in-vehicle time	17.69	SP
	car travel	20.53	
	transit egress time	6	
	transit access time (average)	20.08	
Hensher, (2008)	car commuter trip	24.88	SP
	car non-commuter trip	19.88	
Dixit et al., (2014)	carsharing trips	12.15	RP
Legaspi and Forum, (2015)	bus travel	7.73	
	LRT travel	20.67	SP
	rail travel	13.49	
	ferry	15.01	
Merkert & Beck, (2017)	regional plane trips	126	SP
	regional car trips	37.5	
Douglas et al., (2019)	car travel	14.63	SP
	public transport travel	11.32	
Henn et al., (2011)	rail commuters	16.76	SP

that incorporating a complementary Stated Preference (SP) dataset can enhance the parameter estimation of mode choice models, even when using high-quality Revealed Preference (RP) data. Furthermore, the paper investigates the estimation and comparison of travel time savings for various socio-demographic groups and trip characteristics within the city of Sydney. To accomplish this, the study employs a nonlinear utility specification for the Multinomial Logit (MNL) model on both separate SP and RP datasets, as well as a joint SP-RP dataset.

The paper is structured as follows. Section 3 introduces the SP and RP datasets. Section 4 explains the nonlinear utility specification and the discrete choice modelling framework. Section 5 presents the empirical findings through two subsections, one for the estimated reference WTP parameters of travel attributes and another for the sociodemographic parameters estimated in the developed models. Section 6 presents a comprehensive discussion of the obtained results, accompanied by a comparison to existing literature. Finally, the paper concludes in the last section.

3. Stated and revealed preference data

We use the data collected by a travel study from September through December 2018. The travel study includes two parts: Part one called the “recruit survey”, collected general demographic information as well as stated mode choice preferences and introduced prospective participants to the rMove™ smartphone application (developed by Resource Systems Group (RSG)). The recruitment is conducted by Qualtrics, a market research company that invites participants via email based on their online consumer panel, to ensure that the sample is representative of the population. Using Qualtrics platform, 1772 respondents were asked to select the alternative they prefer the most among three available mode alternatives, each having different values for relating attributes in five different scenarios. We posit that a respondent chooses the alternative with the highest random utility, as is standard practice in discrete choice analysis (Ben-Akiva et al., 2019). The attributes associated with each alternative were distance, access, egress, transfer time, in-vehicle travel time, and total travel cost. The available modes for each scenario were randomly chosen among car, bicycle, walking, train, ferry, and bus modes (Rail was not introduced in SP options). It is important to note that ferries have a significant role in Sydney’s public transport network. It is interesting to know that 778,000 trips are made by ferry within the public transport network of Sydney during July 2022 (NSW Government, 2022).

While it would be valuable to investigate individuals’ preferences for different types of trips (such as work-related versus leisure-related), doing so would require additional context and framing that could potentially burden respondents. Therefore, we elected to use a simple stated preference task design in our survey to minimize the cognitive effort required of participants. This approach allowed us to collect data on individuals’ preferences for various trip attributes without overwhelming respondents with unnecessary complexity. Fig. 1 demonstrates one instance of the SP mode choice survey questions.

Revealed preference data were collected through part two, also called the “travel diary”, which required participants to record their travel for one week using the rMove™ smartphone application. This smartphone application also was used to collect travel data in the UK and develop mode choice models to estimate the value of travel time saving across different modes (Tsoleridis et al., 2022). The rMove™ application collects travel data by using participants’ smartphone GPS. Whenever the smartphone’s GPS detects any movement, thus a potential trip, it starts to track the respondent’s trajectory. At the end of each day, respondents were asked to specify the mode of transport they used for each recorded route. This enabled us to collect high-quality RP travel diary data in this section of the study. Without GPS data, we would not have been able to show trip routes to participants and ask them to provide information about the mode they used for that trip. In addition, by knowing the routes taken by the respondents, we were able to extract

Suppose you are making a **trip** and you can reach your destination via the following **three travel options**:

	Option A	**Option B**	**Option C**
Main mode of transport	train / light rail	car (as driver)	bicycle
Distance [kilometres]	-	-	8.8
Access, transfer and egress time [minutes]	21	4	-
In-vehicle travel time [minutes]	15	18	-
No. of transfers	0	-	-
Travel cost [AU$]	5.2	3.7	0

Which option do you prefer?

Option A - train / light rail

Option B - car (as driver)

Option C - bicycle

Fig. 1. An example of SP mode choice survey questions.

travel time and cost for both chosen and non-chosen alternatives. We used Google APL to extract non-chosen alternative attributes, particularly travel cost, which is not easily perceived or self-reported. For example, for a car trip made by a user, we calculated the trip cost including fuel and potential tolls based on the route and extracted alternative public transport modes with their associated time and costs for that route.

Our RP dataset consists of complete data from 447 participants who fulfilled the study’s base-level completion criteria, which represents approximately 25% of the total number of recruited survey participants. The criteria require participants to specify the mode of transport used for all their daily trips on at least one day during their participation. The final project dataset includes six distinct survey data tables. These tables contain all user-input survey variables, passively collected GPS and location data, specific survey metadata, and derived variables to support data analysis.

A summary of sociodemographic characteristics of data used in the modelling procedure as well as the distribution of each variable according to the Sydney Greater Area 2016 census data (ABS, 2021) is provided in Table 2. For modelling purposes, some numeric variables were categorised into different groups. After removing invalid data and data pre-processing, the dataset consists of 1,732 individuals and their responses to the five-mode choice SP tasks and sociodemographic attributes. The RP trip data is provided for 328 out of those 1,732 individuals. The SP dataset contains approximately the same percentage of female (56%) and male (44%) participants which are aligned with census distribution. Around 28% of participants in SP data collection belong to the age category of 18–25, 31% belong to 26–35, 9% belong to 36–45, 19% belong to 46–55, 9% belong to 56–65, and only 4% of participants in this study were aged 65 or more. The age distribution of the SP sample is representative of the census data except for having fewer participants aged over 65 and more participants aged 18–35 by around 10%. Most participants (26%) have a bachelor’s degree, while 3% completed a doctoral degree. The education distribution in the SP data is largely representative of the Sydney population, except for the senior secondary school education level, which is underrepresented by approximately 24%. The AUD1,200–1,599 per-week income category in SP data has the most frequency amongst the participants. Moreover, the

Table 2
Participants' sociodemographic attributes of joint data.

Variable name		Frequency in SP sample	Percentage in SP sample	Frequency in RP sample	Percentage in RP sample	Percentage in census 2016 data
Gender	Male	765	44.17	116	35.37	49.33
	Female	967	55.83	212	64.63	50.67
Age category	18–25	477	27.54	57	17.38	14.46
	26–35	544	31.41	91	27.74	21.15
	36–45	161	9.30	81	24.70	18.58
	46–55	330	19.05	48	14.63	16.38
	56–65	150	8.66	36	10.98	13.32
	65-more	70	4.04	15	4.57	16.10
Education level	1 (Secondary school)	192	11.09	24	7.32	7.96
	2 (Senior secondary school)	175	10.10	28	8.54	34.17
	3 (Certificate I, II)	68	3.93	13	3.96	0.07
	4 (Certificate III, IV)	180	10.39	45	13.72	13.24
	5 (Diploma)	180	10.39	31	9.45	5.85
	6 (Associate degree)	72	4.16	11	3.35	4.86
	7 (Bachelor's degree)	452	26.10	102	31.10	22.95
	8 (Graduate Diploma)	99	5.72	32	9.76	2.11
	9 (Master's degree)	249	14.38	36	10.98	7.72
	10[a](Doctoral degree)	48	2.77	6	1.83	1.07
Income interval	\$199[b] per week or less	120	6.93	8	2.44	4.94
	\$200 - \$299 per week	63	3.64	4	1.22	7.79
	\$300 - \$399 per week	68	3.93	2	0.61	8.98
	\$400 - \$599 per week	103	5.95	19	5.79	13.88
	\$600 - \$799 per week	158	9.12	20	6.10	11.14
	\$800 - \$1,199 per week	210	12.12	35	10.67	18.37
	\$1,200 - \$1,599 per week	233	13.45	60	18.29	12.41
	\$1,600 - \$1,999 per week	212	12.24	41	12.50	8.77
	\$2,000 - \$2,499 per week	184	10.62	42	12.80	4.52
	\$2,500 - \$2,999 per week	156	9.01	44	13.41	3.34
	\$3,000 - \$3,999 per week	119	6.87	34	10.37	4.09
	\$4,000 per week or more	106	6.12	19	5.79	1.75

[a]17 people (0.9 % of the SP sample) reported "other" as the level of their education, which is not considered a variable in modelling.
[b]Incomes are stated in Australian dollars (AUD)

income distribution aligns with the census data.

Table 2 presents information on the distribution of participants in the RP data across different sociodemographic variables. The results show that the gender distribution among RP participants is slightly skewed towards females. The age distribution in the RP data is more similar to the census data compared to the SP data, however, similar to the SP data, individuals aged over 65 are also underrepresented in the RP data. Moreover, there is the same issue of underrepresentation of individuals with a senior secondary education level in both the RP and SP data. In terms of income distribution, individuals belonging to the AUD300 to AUD399 per week category is slightly underrepresented in the RP data.

The modal share of the trips collected via the smartphone GPS-based application is shown in Fig. 2. The numbers in this figure represent the total number of trips (and the corresponding percentage) made by each mode of transport within the RP data. As can be seen from Fig. 2, car is the most popular means of transport, accounting for 73.3% of recorded trips. Walking is the next most popular mode, followed by transit, including rail and bus. The bicycle was the least popular method of travel in this survey's statistics, with only 64 journeys.

It is worth mentioning that this research hypothesises that SP and RP data are complementary. In other words, we aimed to use SP data to capture the general mode choice preferences of individuals, which could help fill gaps in RP data when estimating models. For example, if respondents did not make enough trips by public transport during data

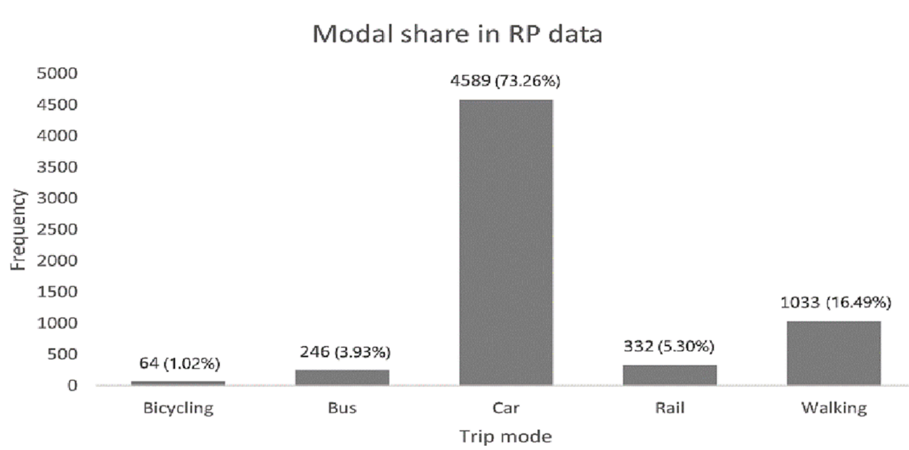

Fig. 2. Modal share in rp data.

collection, their preference could still be captured by their responses to the SP tasks.

4. Multinomial logit mode choice modelling framework

Several multinomial logit and mixed multinomial logit models have been developed during this study. A fundamental premise in discrete choice modelling is that all decision-makers are assumed as rational individuals who seek to maximise the utility obtained from their choice (Train, 2002). In this study, the utility specification is such that in addition to alternative attributes, it can take account of sociodemographic variables with a nonlinear structure. The nonlinear specification is selected since it can highlight the true error terms distribution and better model the travel mode choice behaviour in our data (Cherchi and Ortúzar, 2002). The utility obtained by individual n from selecting alternative j in choice situation t in willingness-to-pay space, in an SP or RP experiment, is defined as equation (1).

$$U_{ntj} = -\beta_{cj}\left(C_{ntj} - \beta\_WTP_{ref,j}(e^{\theta_j Y_n})X_{ntj}\right) + k_j + \varepsilon_{ntj}, \tag{1}$$

where β_{cj} is the cost parameter associated with alternative j, C_{ntj} is the cost of alternative j in choice situation t for individual n, X_{ntj} denotes observed alternative-specific variables of alternative j in choice situation t for individual n except for cost, and $\beta\_WTP_{ref,j}$ is the parameter vector associated with X_{ntj} in willingness-to-pay space. In this equation, Y_n is the vector of the case-specific variables and θ_j is its associated vector of coefficients, e is Euler's number approximately equal to 2.72, k_j is the alternative-specific constant of alternative j, ε_{ntj} is a random variable capturing the unobserved component of utility, and U_{ntj} is the total utility that individual n received from selecting alternative j in choice situation t. For alternatives without monetary cost, like walking, equation (1) simplifies to equation (2) where $\beta_{ref,j}$ is the parameter vector associated with X_{ntj} in preference space.

$$U_{ntj} = \beta_{ref,j}(e^{\theta_j Y_n})X_{ntj} + k_j + \varepsilon_{ntj} \tag{2}$$

Specifying utility like equation (1) allows alternative-specific parameters to be estimated directly in the willingness to pay space. That is, the distribution of the WTP for travel time and the number of transfers can be estimated directly from this specification in MMNL models. This approach can lead to more reasonable WTP estimates although it may cause some inconvenience for the model to fit the data (Train and Weeks, 2005). Although for simple MNL, the WTP estimation in WTP space would not directly add extra information as it tends to be identical to its estimation in preference space, nonetheless, it allows direct estimation of the significance of the WTP parameters, which obviates the need to post-estimate the standard errors.

Assuming each unobserved utility term is independently, identically distributed type I Extreme Value, the Logit choice probabilities can be computed using equation (3).

$$L_{ntj}(\beta) = e^{U_{ntj}} / \sum_j e^{U_{ntj}} \tag{3}$$

Where L_{ntj} is the Logit choice probability which in the MNL context equals the probability of alternative j being chosen by individual n in the choice situation t ($P_{ntj} = L_{ntj}(\beta)$), and U_{ntj} is the utility that individual n obtains from selecting alternative j in the choice situation t which is a function of utility parameters β. On the other hand, within the MMNL model, the choice probability is the integral of $L_{ntj}(\beta)$ over all possible values of β as in equation (4).

$$P_{ntj} = \int L_{ntj}(\beta) f(\beta) d\beta \tag{4}$$

Where $f(\beta)$ is the distribution of parameters vector β.

A maximum likelihood estimator is used for parameter estimation. The likelihood function of the joint SP-RP estimation comprises of multiplication of two components: the SP component and the RP component as presented in equation (5). The SP or RP components of the joint likelihood function will be used for separate estimation of SP or RP data.

$$L(\beta_{cj}, \beta_{ref,j}, \beta_{WTP\,ref,j}, \theta_j, k_j) = \prod_{n=1}^{N_{SP}} \prod_{t=1}^{T_{SP}} \prod_{j \in J_{nt,SP}} (P_{ntj})^{y_{ntj}} \times \prod_{n=1}^{N_{RP}} \prod_{t=1}^{T_{n,RP}} \prod_{j \in J_{nt,RP}} (P_{ntj})^{y_{ntj}} \tag{5}$$

In equation (5), y_{ntj} equals one if individual n selected alternative j in choice situation t, and zero otherwise. L denotes the likelihood function, N_{SP} is the total number of participants in the SP experiment and N_{RP} is the total number of participants in RP data collection, T_{SP} is the total number of tasks in the SP experiment, and $T_{RP,n}$ is the total number of tracked trips for individual n. $T_{RP,n}$ is different for each individual as they made a different number of trips during the data collection period while T_{SP} is constant for all individuals and equal to five. As individuals may face a different set of alternatives in each choice situation t, $J_{nt,SP}$ defines the choice set of the individual n in situation t in the SP experiment context, while $J_{nt,RP}$ is the available alternatives to individual n in the choice situation t in the RP data. The estimated parameters of the model, namely $\beta_{cj}, \beta_{ref,j}, \beta_{WTP\,ref,j}, \theta_j, k_j$ are the values that maximise the natural logarithm of the likelihood function in equation (5). In MMNL, the parameters of the distribution $f(\beta)$ are estimated. The readers are referred to for an extensive discussion regarding the derivation of probability and likelihood functions (Train, 2002).

In the joint estimation, a scale parameter, namely σ^2, is also introduced to capture the difference between SP and RP data regarding the variance of the unobserved component of utility (Hensher & Bradley, 1993). The scale parameter is multiplied by the utility of SP data; Therefore, it can be defined as equation (6). The interpretation of the scale parameter will further be investigated in section 5 where we analyse the joint estimation outputs.

$$\sigma^2 = \frac{Var(\varepsilon_{ntj,RP})}{Var(\varepsilon_{ntj,SP})} \tag{6}$$

5. Empirical findings

The research outputs are presented in this section. An extensive manual specification search is performed to find the models that best fit the data. The estimated parameters of the final models are provided in this section. These models are developed in the willingness-to-pay space; therefore, the value of travel time savings and the impact of sociodemographic attributes on them are estimated explicitly. The time unit is considered minute, and the costs are in Australian dollars (AUD). The Biogeme package runs the maximum log-likelihood optimisation to estimate parameters (Bierlaire, 2020). The impact of sociodemographic variables is considered in the utility specification by multiplying an exponential term by the associated reference alternative-specific attributes. The index in the exponential term is a linear combination of sociodemographic variables. This exponential term serves as an adjustment factor to the estimated reference parameters. The positive or negative estimated parameters for the sociodemographic attributes indicate that those specific sociodemographic characteristics can strengthen or weaken the influence of the associated alternative attribute on the utility, respectively. The models' outputs are discussed in the following subsections.

6. Reference parameters of travel attributes

Table 3 presents the estimated reference parameters and their corresponding t-statistics in parenthese for the MNL models applied to SP, RP, and combined SP-RP data, as well as the MMNL model on SP data while assuming a lognormal distribution of costs and a normal distribution of willingness-to-pays. It is important to note that some parameters are specific to a particular dataset, for instance, parameters related to rail are RP-specific because rail mode is only available in RP data and

Table 3
Estimated reference parameters of MNL models on SP, RP, and joint SP-RP data and the MMNL model on SP data where costs are assumed to be lognormally distributed and WTP parameters are assumed to be normally distributed.

Variable name	MNL on SP	MNL on RP	MNL on joint SP-RP	MMNL on SP (mean)	MMNL on SP (standard deviation)
Bus fare	0.116 (-6.72)	NS*	0.0471 (-3.38)	−2.44 (-8.98)	1.07(5.94)
Rail fare	–	NS	NS	–	–
Travel cost by car in SP data	0.0345 (-3.54)	–	0.0262 (-5.31)	−3.52 (-20.8)	2.16(11.6)
Travel cost by car in RP data	–	0.0688 (-8.38)	0.0529 (-6.58)	–	–
Ferry fare (SP-specific)	0.176 (-4.93)	–	0.143 (-3.63)	−2.13 (-3.56)	1.16(4.97)
Train fare (SP-specific)	0.106 (-6.4)	–	0.0815 (-4.48)	−2.96 (-9.64)	1.85(9.34)
WTP for in-vehicle travel time in car	−0.248 (-2.89)	−0.334 (-4.34)	−0.287 (-4.92)	−0.301 (-7.95)	1.57(18.9)
WTP for in-vehicle travel time in bus	−0.0893 (-4.02)	−0.0242 (-8.22)	−0.292 (-2.77)	−0.174 (-2.78)	0.135 (2.83)
WTP for in-vehicle travel time in train (SP-specific)	−0.083 (-4.14)	–	−0.103 (-3.93)	−0.0614 (-2.33)	−0.0348 (-2.24)
WTP for in-vehicle travel time in ferry (SP-specific)	−0.0635 (-2.76)	–	−0.0784 (-2.67)	−0.151 (-1.18)	−0.117 (-1.42)
Rail travel time (RP-specific)	–	−0.0264 (-8.22)	−0.016 (-4.03)	–	–
WTP for access-egress travel time in bus	−0.0796 (-2.52)	−0.0247 (-3.62)	−0.172 (-2.21)	−0.0774 (-1.8)	0.0605 (2.59)
WTP for access-egress travel time in train (SP-specific)	−0.0281 (-1.25)	–	−0.0173 (-0.823)	−0.0684 (-2.2)	0.222 (4.46)
WTP for access-egress travel time in ferry (SP-specific)	−0.0803 (-1.64)	–	−0.0618 (-1.67)	−0.383 (-1.26)	0.471 (1.69)
WTP for access-egress travel time in rail (RP-specific)	–	−0.0281 (-3.91)	−0.0157 (-2.07)	–	–
WTP for number of transfers in bus	−0.49 (-1.86)	−0.289 (-2.65)	−1.57 (-1.89)	−1.03 (-1.4)	1.6(1.29)
WTP for number of transfers in train (SP-specific)	−0.397 (-1.7)	–	−0.562 (-2.02)	−2.76 (-2.43)	3.26(3.64)
WTP for number of transfers in ferry (SP-specific)	−0.361 (-1.3)	–	−0.5 (-1.49)	−2.69 (-1.21)	4.99(1.76)
WTP for number of transfers in rail (RP-specific)	–	−0.195 (-2.45)	−0.175 (-2.56)	–	–
Bicycling travel time	−0.0327 (-12.9)	−0.0256 (-5.2)	−0.0226 (-6.63)	−0.0352 (-10.6)	–
Walking travel time	−0.00929 (-10.2)	−0.00614 (−11)	−0.00532 (-8.53)	−0.00833 (-8.46)	–
Alternative-specific constant of alternative 1 (SP-specific)	0.153 (5.47)	–	0.131 (3.58)	0.212 (6.09)	–
Alternative-specific constant of alternative 2 (SP-specific)	0.169 (6.07)	–	0.145 (3.73)	0.241 (6.93)	–
Alternative-specific constant of bicycle (RP-specific)	–	−4.39 (-25.4)	−4.3 (-31.8)	–	–
Alternative-specific constant of bus (RP-specific)	–	−2.46 (-19.5)	−2.75 (–23.2)	–	–
Alternative-specific constant of rail (RP-specific)	–	−2.17 (-19.3)	−2.29 (-19.3)	–	–
Alternative-specific constant of walking (RP-specific)	–	−1.78 (–33.7)	−1.68 (–32.6)	–	–
The square root of the scale parameter (σ)	–	–	1.16 (4.72)	–	–

* Not significant

is not included in the SP tasks. On the other hand, ferry and train are only introduced in the SP dataset. The 'NS' value in the table stands for Not Significant and indicates that our data could not reveal a significant relationship between that variable and the utility of the associated alternative.

Furthermore, the positive and intuitive sign of the travel cost coefficients (see the specification of Equation (1) suggests that the utility for individuals decreases as the travel cost increases for each mode. The willingness to pay (WTP) for reducing in-vehicle travel time (IVTT), access-egress travel time (ATE), and the number of transfers in public transport are estimated to be negative. This implies an inverse relationship between travel time and travel cost while maintaining the mode's utility for individuals constant. The estimated parameters for walking and bicycling travel time are not WTP as no cost attributes were associated with using a bicycle or walking. These coefficients are negative, indicating that longer walking or bicycling times have a negative impact on the utility individuals gain from these modes of transport.

In the SP data, since there are three alternatives, two alternative-specific constants are defined for the first and second alternatives, with this coefficient being normalized to zero for the third option. As modes are randomly displayed to individuals and the first alternative could be associated with different modes in different tasks, we could not define mode-specific constants for the SP data. However, in the RP data, four alternative-specific constants are defined for rail, bus, walking, and biking, while this constant is normalized to zero for cars due to specification considerations. Lastly, the scale parameter is estimated in the joint estimation to capture the difference in the variance of the error term in the utility specification across the SP and RP datasets.

According to the estimated WTPs for in-vehicle travel time (IVTT) in Table 3, individuals tend to pay higher costs to reduce their in-vehicle travel time in a car compared to public transport. Among public transport options, the ferry appears to have a slightly less value of IVTT. This could be due to the individuals' enjoyment while riding on a ferry. At a 95% confidence interval, the WTP parameter for access-egress time (ATE) is significantly from zero for the bus and the negative sign implies the adverse effect of ATE on the utility of the bus. However, ATE parameters are significant and intuitive for train and ferry within a 90% and 80% confidence interval. The WTP for the number of transfers in public transport has an intuitive sign for all public transport modes with being statistically significant at a 95% confidence interval for the bus and train and 80% for the ferry. Moreover, the alternative-specific constants for the first and second alternatives in SP data have positive signs and are significantly different from zero, which suggests in the SP experiment, individuals have some positive attitudes towards selecting the first and second alternatives as their preferred mode for travel. This could be because individuals tend to choose the options introduced earlier (Hess et al., 2010).

We could not estimate WTP for public transport travel time and the number of transfers in RP data as the corresponding cost parameters are statistically insignificant which means our data could not reveal any relationship between public transport fare and their utilities. This could happen because of the relatively small variation in transit cost in RP data (Louviere et al., 2000). However, the travel time and the number of transfer parameters are estimated in preference space which are negative and statistically different from zero. As can be seen in Table 3, the alternative-specific constants are all negative and significantly different from zero. This means individuals have a negative attitude towards public transport and active modes (walking and bicycling) compared to car. In other words, without concerning alternative attributes, car is the preferable mode among individuals who participated in this study.

As Table 3 shows SP data enable us to estimate a cost coefficient for bus and therefore estimate parameters in WTP space in the joint estimation. The scale parameter (σ^2), which is defined in the joint SP-RP estimation to capture the variation between the random error terms variances across SP and RP datasets, is estimated to be 1.3456 and is statistically significant. This finding indicates that the variance of unobserved factors in the utility of SP data is 34.56% higher compared to RP data. This could happen as the unobserved influential factors on utility may differ across different datasets, and it's essential to consider it while conducting joint estimation.

To capture the unobserved heterogeneity in WTP, MMNL models with normal, lognormal, and non-random cost parameters are estimated while the distributions of WTPs are assumed to be normal on SP data. The model with a lognormally distributed cost parameter best fits the data. According to Table 3, the estimated mean and standard deviation of cost parameters for all modes are statistically significant at a 95% confidence level, with the cost parameter of the car having the highest variance among the participants. Moreover, all estimated WTP means and variances are statistically significant at a 95% confidence level except for the means and standard deviations of the WTP of ferry in-vehicle travel time and WTP of the number of transfers of the bus (which are significant at about 80% confidence level). This result confirms the existence of unobserved heterogeneity in the WTP across individuals. The variance of WTP of travelling in car is estimated to be 1.57, while for the in-vehicle and access-egress travel time in bus, the variances are 0.135 and 0.0605, respectively. The results suggest the variance of WTP for car travel time is higher than bus travel time. In addition, WTP for bus in-vehicle travel time is more widely spread around its mean than WTP for bus access-egress time. The number of ferry transfers is also valued in a wide range by individuals as it has a statistically significant estimated standard deviation. It is worth mentioning that we faced some empirical issues in the convergence of the MMNL models on RP and joint data, therefore, those models are not presented in this paper.

6.1. Sociodemographic parameters

Table 4 shows the estimated coefficients and their corresponding t-statistics in parenthese for the sociodemographic variables that are presented in the index of the exponential term in the utility function. All parameters are significantly different from zero at a 95% confidence interval. It is worth mentioning that these parameters are not considered to be random in the MMNL estimation. Some abbreviations are used in this table as follows: B-IVTT-WTP denotes that the estimated sociodemographic parameter modifies the reference coefficient of the willingness to pay to reduce in-vehicle travel time, B-ATE-WTP means the same concept for access-egress time and B_NO_TRANSFERS indicates that sociodemographic is modifying the reference parameter for willingness to pay to reduce the number of transfers in public transport. We will discuss each model's estimates in the following subsections. Lastly, B_TIME indicates the corresponding sociodemographic is adjusting the reference parameter of walking or biking travel time.

According to the estimated parameters in Table 4 for SP data, individuals aged 26 to 35 value their in-vehicle travel time more than other age groups. This might happen as individuals in this age interval may be at their maximum productivity regarding their job and income which then leads to greater WTP. Moreover, the results suggest that individuals with education code 4 corresponding to Certificates III and IV are more inclined to pay to reduce access-egress time. On top of that, decision-makers aged between 36 and 45 years and those with an educational background of codes 5 and 7 (corresponding to a diploma and bachelor's degree) are more willing to pay to reduce the number of transfers in public transport indicated by the positive sign of their corresponding coefficients in Table 4.

Gender, income, age, and education seem to be influential on the impact of walking and bicycling time on the utility individuals obtained from these modes in SP data. According to Table 3, walking and bicycling time tend to influence the utility of these modes more negatively for women than men (Abasahl et al., 2018). In addition, individuals with a higher income level tend to experience a more negative impact on their overall utility when it comes to walking and bicycling time. Moreover, the travel utility obtained by younger individuals (aged between 18 and 25) is less adversely affected by the increased walking and bicycling travel time compared to other age groups. In contrast, the elderly, who are aged 56 or older, experience more reduction in utility due to increased walking or biking time than other individuals. This is consistent with previous findings (Handy et al., 2006; Kim and Ulfarsson,

Table 4
Estimated coefficients and their corresponding t-statistics in parenthese for the sociodemographic variables that are presented in the index of the exponential term in the utility function.

Variable	MNL on SP	MNL on RP	MNL on joint SP-RP	MMNL on SP
B-IVTT-WTP for individuals with education level 4	NS*	−1.46 (-2.98)	−0.882 (-2.12)	NS
B-IVTT-WTP for individuals with education level 9	NS	0.486 (4.17)	0.95(5.18)	NS
B-IVTT-WTP for individuals aged between 26 and 35	0.855 (5.41)	NS	0.54(4.78)	NS
B_ATE_WTP for individuals with education level 4	0.995 (2.93)	1.1 (5.71)	1.18(5.76)	1.13(5.71)
B_ATE_WTP for individuals aged more than 65	NS	1.18 (3.3)	1.78(3.93)	NS
B_ATE_WTP for women	NS	0.436 (2.52)	0.841 (2.15)	NS
B_NO_TRANSFERS for individuals aged between 36 and 45	1.03(3.04)	0.688 (1.9)	0.947 (3.63)	NS
B_NO_TRANSFERS for individuals with education level 5	1.25(2.82)	NS	0.907 (2.49)	NS
B_NO_TRANSFERS for individuals with education level 7	0.828 (1.99)	NS	0.552 (1.71)	NS
B_TIME for women	0.334 (7.35)	NS	0.446 (7.95)	0.435 (6.71)
B_TIME for the amount of income in AUD	0.0000673 (3.87)	NS	0.000103 (5.37)	0.0000663 (2.77)
B_TIME for individuals aged between 18 and 25	−0.213 (-3.61)	NS	−0.161 (-2.77)	−0.113 (-1.58)
B_TIME for individuals aged between 26 and 35	NS	0.324 (3.63)	0.362 (3.55)	NS
B_TIME for individuals aged between 56 and 65	0.221 (2.82)	−0.509 (-5.2)	NS	0.595 (5.97)
B_TIME for individuals older than 65	0.961 (8.84)	NS	1.12(9.34)	1.46(10.7)
B_TIME for individuals with education level 1	0.195 (2.88)	NS	0.288 (3.89)	NS
B_TIME for individuals with education level 2	NS	−0.465 (-2.94)	−0.628 (-3.02)	NS
B_TIME for individuals with education level 3	0.378 (3.59)	NS	0.533 (4.63)	0.421 (3.35)
B_TIME for individuals with education level 7	−0.151 (-2.79)	NS	−0.126 (-2.16)	−0.309 (-4.08)
B_TIME for individuals with education level 9	−0.196 (-2.94)	NS	−0.177 (-2.44)	−0.309 (-4.08)
B_TIME for individuals with education level 10	NS	0.412 (1.87)	0.281 (1.12)	NS

* Not significant

2004) and is intuitive as, generally, walking and bicycling are active travel modes that require engaging in some physical activity and therefore, with the increase in age, more extended walking and biking time seems more difficult to people, and more negatively influence associated utility. The results also indicate that higher education levels keep individuals' utility less negatively influenced by walking and bicycling time maybe due to being more educated regarding the adverse environmental impacts of non-active transport modes (Villena-Sanchez et al., 2022).

In the RP data, education level seems to be effective on individuals with an education level of 4 (Certificate III, IV) are less sensitive to in-vehicle travel time, while those with an education level of 9 (Master's degree) are more likely to be adversely influenced by the increase in in-vehicle travel time rather than others. Considering higher education as an income indicator, the results suggest individuals with higher income value their in-vehicle travel time more than others.

7. Discussion

As a result of our study on the effect of including SP data in the estimation of mode choice models for RP data, we were able to extract the WTP of Sydney residents for various modes of travel directly in the WTP space. It is important to note that WTPs are used in this study as a measure of individuals' travel behaviour concerning their mode choice. Table 5 summarises the estimated reference WTPs from MNL models applied to SP, RP, and combined SP-RP data and their 90% confidence intervals in brackets.

According to Table 5, the reference WTP for car travel time is estimated to be 14.88 and 20.04 AUD/hour based on SP and RP data, respectively, while it equals 17.22 AUD/hour in joint SP-RP estimation. The car WTP that is estimated based on SP data is almost identical to the WTP estimated by Douglas et al., (2019) and close to the WTP estimated by Hensher et al., (2006). The estimated WTP for car travel time using RP data is aligned with previous literature (Hensher, 2008; Hensher and Rose, 2007) although they have used SP data for their studies. According to Table 5, the willingness to pay to reduce car travel time is lower in SP data compared to RP data which is consistent with the findings of previous studies (Isacsson, 2007; Wardman, 2001). This difference can be explained by the existence of hypothetical bias in SP data as we deviate from objective evidence by conducting experiments (Hensher, 2010). The ratio of the estimated WTP parameters for car travel time obtained through SP data compared to RP data is 0.743, which is in proximity to the value of 0.9 calculated by Li et al., (2020). A well-designed SP experiment can result in a WTP to RP ratio that is closer to one, indicating that the experiment is better at capturing individuals' true behaviour.

The information presented in Table 5 demonstrates that a joint SP-RP estimation approach allowed us to estimate the WTP for bus in-vehicle travel time, which we could not accomplish by relying solely on RP

Table 5
The estimated WTP for travel time and number of transfers in MNL models applied to SP, RP, and combined SP-RP data and their 90% confidence intervals.

Willingness-to-pay (AUD /hour)	SP	RP	Joint SP-RP
In-vehicle travel time (IVTT)			
Car	−14.88 [-6.42, −23.36]	−20.04 [-10.98, −29.10]	−17.22 [-10.38, −24.06]
Bus	−5.36 [-3.17, −7.55]	NS*	−19.02 [-5.17, −29.87]
Rail	NA	NS	–
Train	−4.98 [-3.00, −6.96]	–	−6.18 [-3.09, −9.27]
Ferry	−3.81 [-1.54, −6.08]	–	−4.70 [-1.26, −8.15]
Access transfer egress travel time (ATE)			
Bus	−4.78 [-1.66, −7.89]	NS	−10.32 [-1.19, −19.45]
Rail	NA	NS	–
Train	−1.69 [0.53, −3.91]	–	−1.04 [1.44, −3.52]
Ferry	−4.82 [0.01, −9.64]	–	−3.71 [0.65, −8.07]
Number of transfers			
Bus	−0.49 [-0.06, −0.92]	NS	−1.57 [0.06, −3.20]
Rail	NA	NS	–
Train	−0.40 [-0.01, −0.78]	–	−0.56 [-0.02, −1.11]
Ferry	−0.36 [-0.32, −0.40]	–	−0.50 [0.16, −1.16]

* Statistically significant coefficient could not be estimated using the provided data

data. Therefore, incorporating SP data's additional information and variation into the mode choice modelling on RP data is recommended to travel behaviour modellers. The estimated willingness-to-pay (WTP) for in-vehicle travel time on the bus, as obtained through joint analysis, is close to the WTP for car travel, which is consistent with the findings of Hensher and Rose, (2007). However, in contrast to their study, the WTP for the bus is slightly higher than the car in our study. This may be due to the fact that they estimated the WTP for transit in general, rather than for the bus specifically. Moreover, Table 5 shows that WTP for bus in-vehicle travel time is higher compared to train and ferry unlike Legaspi and Forum, (2015) finding. Our results also indicate that WTP for reducing access-egress travel time associated with public transport is lower than the WTP for reducing in-vehicle travel time in public transport. It is important to note that, no jointly estimated coefficient is available for rail WTP since it was not included as an alternative in the stated preference tasks. As a result, no additional information was available to assist with the parameter estimation for rail WTP using the RP data.

Based on the values in Table 5, an increase of one transfer for bus travel may result in a negative impact on the utility that individuals obtain from bus travel, equivalent to an increase in bus fare of 1.57 AUD. The estimated WTPs for the number of transfers in public transport are generally higher in joint estimation compared to estimates on SP data. Based on these estimations, individuals exhibit a greater willingness to pay to reduce the number of transfers when travelling by bus compared to train or ferry.

Some other exciting outputs were found in this research. This study's outputs reveal that individuals' WTP and travel behaviour are influenced by their sociodemographic attributes and trip characteristics like trip mode (Börjesson and Eliasson, 2014). More specifically, the impact of age, gender, income, education, and trip mode on the WTP and travel behaviour is empirically investigated during this research and extensively presented in the empirical findings section. For instance, results indicate that the elderly and individuals with higher income are more sensitive to walking or biking travel time and they prefer shorter walking or biking travel time compared to others (Kim and Ulfarsson, 2004; Villena-Sanchez et al., 2022). These are important findings to understand the travel behaviour of individuals with different sociodemographic characteristics.

The results from the developed MMNL model on SP data in Table 3 also reveal that the variance of WTP may be higher among car users compared to bus riders, which emphasises the necessity of considering the unobserved heterogeneity among car users' WTP when evaluating relative transport policies or investment alternatives. In addition, the estimated means and standard deviations for WTP of different components of bus travel time reveal that bus in-vehicle travel time is valued higher by riders and is subject to higher variance across the participants than the associated access-egress. Generally, the means of WTP estimated from the MMNL model are higher than the corresponding estimates from the MNL model in this study which aligns with previous research (Hensher, 2001). For instance, considering SP data, the WTP for bus in-vehicle travel time is estimated to be normally distributed with a mean of 10.44AUD/hour and a standard deviation of 8.1 from the MMNL model. At the same time, it is calculated to be 5.358AUD/hour from the MNL model.

Defining a scale parameter (σ^2), the difference between SP and RP data in terms of the variance of error term of utility specification was investigated. According to Table 3, the scale parameter was estimated to be 1.3456 and statistically different from zero, confirming that the variance of the random term of utility is 34.56% larger in SP data compared to RP data. Estimating the scale parameter during joint estimation can avoid misinterpretation of the estimated coefficients based on two separate data sources (Train, 2002).

8. Conclusion

Using a mixed dataset of RP and SP data, we empirically investigated the effect of accommodating SP data in the estimation of mode choice models on RP data and compared the estimates across three MNL mode choice models on SP data, RP data, and joint SP-RP data. This study utilises high-quality travel data from a GPS-based smartphone application called rMove™. GPS-based travel diary data helps researchers overcome the underreporting of trips by participants (Bricka and Bhat, 2006). As an outcome of our investigation, the WTP of Sydney residents for several travel modes is extracted directly in willingness-to-pay space using those MNL models while accounting for the heterogeneity resulting from the sociodemographic attributes. Moreover, the unobserved heterogeneity in WTP is captured by developing the MMNL model on SP data in the willingness-to-pay space, allowing the direct estimation of the significance level of the WTP.

Based on the results of this research, it is highly recommended that travel behaviour modellers not only collect RP data for their research but also allocate some time and effort to collect complementary SP data. This would bring more benefits than its cost; in general, collecting SP data is more cost-beneficial than RP data (Bradley and Daly, 1991). Although SP data seems biased, it can carry information that cannot be gained from RP data. In other words, parameter estimation based on SP data is more potent than RP data due to the sufficient variation in independent variables in SP data (Louviere et al., 2000). This study shed light on an empirical instance of this statement, where the bus cost coefficient was insignificant in estimation using merely RP data. However, it became significant when we used SP data jointly with RP data for the analysis. In SP data, it is possible to define fares with more variation, allowing us to capture the impact of public transport cost on individuals' mode choices. Therefore, incorporating SP data's additional information and variation into the mode choice modelling on RP data is recommended (Hensher and Bradley, 1993). The phenomenon that this practice will lead to the correctness of the parameters or not may be investigated in the future direction of this study. One other potential future direction of this study could be to examine whether the order of presenting SP and RP surveys to participants has an impact on the supplementary effect of SP data on RP data.

It's worth mentioning that in this study, a separate analysis of individuals' preferences based on different trip purposes is not provided to avoid burdening respondents. However, conducting a travel behaviour analysis for each trip purpose would contribute valuable information to the literature. Furthermore, our RP dataset includes data from 447 participants which is relatively smaller compared to the 1772 participants in the SP survey. In the SP tasks, respondents were asked to select their most preferred alternative among three randomly displayed mode alternatives. However, a better survey design approach would be to anchor available alternatives on a specific reference trip, such as the respondents' latest trip. Additionally, designing an SP survey based on the choice set of RP data would be beneficial as having the same choice set in both surveys would facilitate the comparison.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abasahl, F., Kelarestaghi, K.B., Ermagun, A., 2018. Gender gap generators for bicycle mode choice in Baltimore college campuses. Travel Behav. Soc. 11, 78–85. https://doi.org/10.1016/j.tbs.2018.01.002.

Ahern, A.A., Tapley, N., 2008. The use of stated preference techniques to model modal choices on interurban trips in Ireland. Transp. Res. Part A Policy Pract. 42, 15–27. https://doi.org/10.1016/J.TRA.2007.06.005.

Anta, J., Pérez-López, J.B., Martínez-Pardo, A., Novales, M., Orro, A., 2016. Influence of the weather on mode choice in corridors with time-varying congestion: a mixed data study. Transportation (Amst). 43, 337–355. https://doi.org/10.1007/s11116-015-9578-1.

Arman, M.A., Khademi, N., de Lapparent, M., 2018. Women's mode and trip structure choices in daily activity-travel: a developing country perspective. Transp. Plan. Technol. 41, 845–877. https://doi.org/10.1080/03081060.2018.1526931.

Assi, K.J., Nahiduzzaman, K.M., Ratrout, N.T., Aldosary, A.S., 2018. Mode choice behavior of high school goers: Evaluating logistic regression and MLP neural networks. Case Stud. Transp. Policy 6, 225–230. https://doi.org/10.1016/j.cstp.2018.04.006.

Australian Bureau of Statistics. TableBuilder [Internet]. Canberra: ABS; 2021 November 19 [cited 2023 March 22]. Available from: https://www.abs.gov.au/statistics/microdata-tablebuilder/tablebuilder., n.d.

Aziz, H.M.A., Nagle, N.N., Morton, A.M., Hilliard, M.R., White, D.A., Stewart, R.N., 2018. Exploring the impact of walk–bike infrastructure, safety perception, and built-environment on active transportation mode choice: a random parameter model using New York City commuter data. Transportation (Amst). 45, 1207–1229. https://doi.org/10.1007/s11116-017-9760-8.

Badri, M.A., 2013. School travel modes: factors influencing parental choice in Abu Dhabi. Int. J. Educ. Econ. Dev. 4, 203. https://doi.org/10.1504/ijeed.2013.056010.

Ben-Akiva, M., Morikawa, T., 1990. Estimation of switching models from revealed preferences and stated intentions. Transp. Res. Part A Gen. 24, 485–495. https://doi.org/10.1016/0191-2607(90)90037-7.

Ben-Akiva, M., McFadden, D., Train, K., 2019. Foundations of stated preference elicitation: Consumer Behavior and Choice-based Conjoint Analysis. Found. Trends Econom. 10, 1–144. https://doi.org/10.1561/0800000036.

Bhat, C.R., 2000. Incorporating observed and unobserved heterogeneity in urban work travel mode choice modeling. Transp. Sci. 34, 228–238. https://doi.org/10.1287/trsc.34.2.228.12306.

Bierlaire, M., 2020. A short introduction to PandasBiogeme. Tech. Rep. TRANSP-OR 200605. Transp. Mobil. Lab. ENAC, EPFL.

Börjesson, M., Eliasson, J., 2014. Experiences from the Swedish Value of Time study. Transp. Res. Part A Policy Pract. 59, 144–158. https://doi.org/10.1016/j.tra.2013.10.022.

Bradley, M.A.&, Daly, A.J., 1991. Estimation of Logit Choice Models Using Mixed Stated Preference and Revealed Preference Information, in: 6th International Conference on Travel Behaviour Quebec, May 1991. 6th International Conference on Travel Behaviour Quebec, May 1991.

Bricka, S., Bhat, C.R., 2006. Comparative analysis of global positioning system-based and travel survey-based data. Transp. Res. Rec. 1972 (1), 9–20.

Brownstone, D., Small, K.A., 2005. Valuing time and reliability: Assessing the evidence from road pricing demonstrations. Transp. Res. Part A Policy Pract. 39, 279–293. https://doi.org/10.1016/j.tra.2004.11.001.

Brownstone, D., Bunch, D.S., Train, K., 2000. Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. Transp. Res. Part B Methodol. 34, 315–338. https://doi.org/10.1016/S0191-2615(99)00031-4.

Catalano, M., Casto, B.L., Migliore, M., 2008. Car sharing demand estimation and urban transport demand modelling using stated preference techniques. Eur. Transp. Trasp. Eur. 40, 33–50.

Cherchi, E., de Dios Ortúzar, J., 2006. On fitting mode specific constants in the presence of new options in RP/SP models. Transp. Res. Part A Policy Pract. 40, 1–18. https://doi.org/10.1016/j.tra.2005.04.002.

Cherchi, E., Ortúzar, J.D.D., 2002. Mixed RP/SP models incorporating interaction effects: Modelling new suburban train services in Cagliari. Transportation. https://doi.org/10.1023/A:1016307308723.

Ortúzar, J. de D., Willumsen, L.G., 2011. Modelling Transport, Modelling Transport. John wiley & sons. 10.1002/9781119993308.

Dixit, V., Dixit, V., Trieu, J., Jian, S., Li, X., Trieu, J., Jian, S., Li, X., 2014. VALUE OF TRAVEL TIME SAVINGS FOR CARSHARE USERS IN SYDNEY, in: Australian Institute of Traffic Planning and Management (AITPM) National Conference, 2014, Adelaide, South Australia, Australia.

Douglas, N., Thornton, P.B., Thornton, P., 2019. Australasian Transport Research Forum, ATRF 2019 - Proceedings. Australas. Transp. Res. Forum, ATRF 2019 - Proc.

Economics, D., Nut Farm, T., 2004. Value of Travel Time Savings for Sydney Rail Passengers. Henry Labrador Publ, Co.

Gadziński, J., 2018. Perspectives of the use of smartphones in travel behaviour studies: Findings from a literature review and a pilot study. Transp. Res. Part C Emerg. Technol. 88, 74–86. https://doi.org/10.1016/j.trc.2018.01.011.

Gurumurthy, K.M., Kockelman, K.M., 2020. Modeling Americans' autonomous vehicle preferences: A focus on dynamic ride-sharing, privacy & long-distance mode choices. Technol. Forecast. Soc. Change 150, 119792. https://doi.org/10.1016/j.techfore.2019.119792.

Guzman, L.A., Arellana, J., Cantillo-García, V., Ortúzar, J.d.D., Le Pira, M., 2021. Revisiting the Benefits of Combining Data of a Different Nature: Strategic Forecasting of New Mode Alternatives. J. Adv. Transp. 2021, 1–15.

Habib, K.N., Weiss, A., 2014. Evolution of latent modal captivity and mode choice patterns for commuting trips: A longitudinal analysis using repeated cross-sectional datasets. Transp. Res. Part A Policy Pract. 66, 39–51. https://doi.org/10.1016/J.TRA.2014.04.013.

Handy, S., Cao, X., Mokhtarian, P.L., 2006. Self-selection in the relationship between the built environment and walking: Empirical evidence from Northern California. J. Am. Plan. Assoc. 72, 55–74. https://doi.org/10.1080/01944360608976724.

Harrison, G., Grant-Muller, S.M., Hodgson, F.C., 2020. New and emerging data forms in transportation planning and policy: Opportunities and challenges for "Track and Trace" data. Transp. Res. Part C Emerg. Technol. 117, 102672.

Henn, L., Douglas, N., Sloan, K., 2011. Surveying Sydney rail commuters' willingness to change travel time. Conf. Pap. Deliv. 34th Australas. Transp. Res. Forum Proc. held 28-30 Sept. 2011 Adelaide, Aust.

Hensher, D.A., 2001. The valuation of commuter travel time savings for car drivers: Evaluating alternative model specifications. Transportation (Amst). 28, 101–118. https://doi.org/10.1023/A:1010302117979.

Hensher, D.A., 2006. The signs of the times: Imposing a globally signed condition on willingness to pay distributions. Transportation (Amst). 33, 205–222. https://doi.org/10.1007/s11116-005-2302-9.

Hensher, D.A., 2010. Hypothetical bias, choice experiments and willingness to pay. Transp. Res. Part B Methodol. 44, 735–752. https://doi.org/10.1016/j.trb.2009.12.012.

Hensher, D.A., Bradley, M., 1993. Using stated response choice data to enrich revealed preference discrete choice models. Marketing Letters 4 (2), 139–151.

Hensher, D.A., Rose, J.M., 2007. Development of commuter and non-commuter mode choice models for the assessment of new public transport infrastructure projects: A case study. Transp. Res. Part A Policy Pract. 41, 428–443. https://doi.org/10.1016/j.tra.2006.09.006.

Hensher, D.A., Greene, W.H., Rose, J.M., 2006. Deriving willingness-to-pay estimates of travel-time savings from individual-based parameters. Environ. Plan. A 38, 2365–2376. https://doi.org/10.1068/a37395.

Hensher, D.A., 2008. Valuation of Travel Time Savings: A Synthesis of Evidence from Five studies undertaken for Thiess and Consortia members 2004-2007. Sydney NSW 2006 Australia.

Hess, S., Rose, J.M., Polak, J., 2010. Non-trading, lexicographic and inconsistent behaviour in stated choice data. Transp. Res. Part D Transp. Environ. 15, 405–417. https://doi.org/10.1016/j.trd.2010.04.008.

Hillel, T., Elshafie, M.Z.E.B., Jin, Y., 2018. Recreating passenger mode choice-sets for transport simulation: A case study of London, UK. Proc. Inst. Civ. Eng. - Smart Infrastruct. Constr. 171, 29–42. https://doi.org/10.1680/jsmic.17.00018.

Hossan, M.S., Asgari, H., Jin, X., 2016. Investigating preference heterogeneity in Value of Time (VOT) and Value of Reliability (VOR) estimation for managed lanes. Transp. Res. Part A Policy Pract. 94, 638–649. https://doi.org/10.1016/J.TRA.2016.10.022.

Isacsson, G., 2007. The trade off between time and money: Is there a difference between real and hypothetical choices?, Working papers in transport economics NV - 2007: 3 . Statens väg- och transportforskningsinstitut, Transport economics Stockholm, Society, environment and transport. Swedish National Road and Transport Research Institute.

Kim, S., Ulfarsson, G.F., 2004. Travel mode choice of the elderly: Effects of personal, household, neighborhood, and trip characteristics. Transp. Res. Rec. 1894 (1), 117–126.

Legaspi, J., Forum, N.D.-P.A.T.R., 2015, undefined, 2015. Value of travel time revisited–NSW experiment. transport.nsw.gov.au.

Li, Z., Hensher, D.A., Ho, C., 2020. An empirical investigation of values of travel time savings from stated preference data and revealed preference data. Transp. Lett. 12, 166–171. https://doi.org/10.1080/19427867.2018.1546806.

Lizana, P., Ortúzar, J.d.D., Arellana, J., Rizzi, L.I., 2021. Forecasting with a joint mode/time-of-day choice model based on combined RP and SC data. Transp. Res. Part A Policy Pract. 150, 302–316.

Louviere, J.J., Hensher, D.A., Swait, J.D., Adamowicz, W., 2000. Stated Choice Methods Analysis and Applications (with a contribution by. Cambridge University Press.

Malokin, A., Circella, G., Mokhtarian, P.L., 2019. How do activities conducted while commuting influence mode choice? Using revealed preference models to inform public transportation advantage and autonomous vehicle scenarios. Transp. Res. Part A Policy Pract. 124, 82–114. https://doi.org/10.1016/j.tra.2018.12.015.

Manoj, M., Verma, A., 2015. Activity-travel behaviour of non-workers belonging to different income group households in Bangalore. India. J. Transp. Geogr. 49, 99–109. https://doi.org/10.1016/j.jtrangeo.2015.10.017.

Merkert, R., Beck, M., 2017. Value of travel time savings and willingness to pay for regional aviation. Transp. Res. Part A Policy Pract. 96, 29–42. https://doi.org/10.1016/j.tra.2016.11.022.

Müller, S., Tscharaktschiew, S., Haase, K., 2008. Travel-to-school mode choice modelling and patterns of school choice in urban areas. J. Transp. Geogr. 16, 342–357. https://doi.org/10.1016/j.jtrangeo.2007.12.004.

NSW Government, T. for N., 2022. Public Transport Patronage - Top Level Chart [WWW Document]. URL https://www.transport.nsw.gov.au/data-and-research/passenger-travel/public-transport-patronage/public-transport-patronage-top-level.

Parady, G., Ory, D., Walker, J., 2021. The overreliance on statistical goodness-of-fit and under-reliance on model validation in discrete choice models: A review of validation practices in the transportation academic literature. J. Choice Model. 38, 100257 https://doi.org/10.1016/J.JOCM.2020.100257.

Patterson, Z., Ewing, G., Haider, M., 2005. Gender-based analysis of work trip mode choice of commuters in suburban Montreal, Canada, with stated preference data. Transp. Res. Rec. 1924 (1), 85–93.

Resource Systems Group (RSG), n.d. rMove [WWW Document]. URL https://rmove.rsginc.com/what-is-rmove/.

Small, K.A., 2012. Valuation of travel time. Valuation of travel time. Econ. Transp. 1 (1-2), 2–14.

Small, K.A., Winston, C., Yan, J., 2005. Uncovering the distribution of motorists' preferences for travel time and reliability. Econometrica 73, 1367–1382. https://doi.org/10.1111/j.1468-0262.2005.00619.x.

Train, K., 2002. Discrete Choice Methods with Simulation [WWW Document]. Cambridge Univ, Press https://eml.berkeley.edu/books/choice2.html (accessed 1.12.22.

Train, K., Weeks, M., 2005. Discrete Choice Models in Preference Space and Willingness-to-Pay Space. Appl. Simul. Methods Environ. Resour. Econ. 1–16 https://doi.org/10.1007/1-4020-3684-1_1.

Tsoleridis, P., Choudhury, C.F., Hess, S., 2022. Deriving transport appraisal values from emerging revealed preference data. Transp. Res. Part A Policy Pract. 165, 225–245. https://doi.org/10.1016/j.tra.2022.08.016.

Villena-Sanchez, J., Boschmann, E.E., Avila-Forcada, S., 2022. Daily travel behaviors and transport mode choice of older adults in Mexico City. J. Transp. Geogr. 104, 103445 https://doi.org/10.1016/j.jtrangeo.2022.103445.

Wardman, M., 2001. A review of British evidence on time and service quality valuations. Transp. Res. Part E Logist. Transp. Rev. 37, 107–128. https://doi.org/10.1016/S1366-5545(00)00012-0.