Journal of Choice Modelling 48 (2023) 100431

Contents lists available at ScienceDirect

Journal of Choice Modelling

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

Theory-driven or data-driven? Modelling ride-sourcing mode choices using integrated choice and latent variable model and multi-task learning deep neural networks

Yicong Liu [*], Patrick Loa, Kaili Wang, Khandker Nurul Habib

Department of Civil & Mineral Engineering, University of Toronto, Toronto, ON, Canada

ARTICLE INFO

Keywords:
Ride-sourcing
Machine learning
Deep neural networks
Multi-tasking learning
Integrated choice and latent variable model
COVID-19

ABSTRACT

Ride-sourcing services have had a disruptive impact on urban mobility. However, the perceived risk of contracting the COVID-19 virus while using these services has negatively affected people's willingness to travel by this mode. Therefore, it is essential to understand the factors influencing ride-sourcing usage during and after the pandemic. This study utilized data collected through stated preference experiments to model mode choice decisions during and after the pandemic. The study applied both theory-driven integrated choice and latent variable (ICLV) models and data-driven multi-task learning (MTL) deep neural network framework. The study found that the MTL models achieved the highest prediction accuracies. Additionally, econometric information was derived from both ICLV and MTL models. The marginal effects of level-of-service (LOS) variables were largely agreed between the ICLV and MTL models. However, only the latent variables from the ICLV models presented meaningful behavioural interpretations. The study found that individuals who believed there was greater risk associated with ride-sourcing during the pandemic were less likely to use these services. The ICLV model interpretations also indicate that the perceived safety of using ride-sourcing services is higher during the post-pandemic period compared to during the pandemic period. This finding provides reassurance regarding the recovery and growth of ride-sourcing usage in the post-pandemic era.

1. Introduction

Ride-sourcing service has disrupted urban mobility, primarily by affecting travel patterns and inducing travel demand (Clewlow and Mishra, 2017; Erhardt et al., 2019). The service can improve the accessibility of urban dwellers without private vehicle access (Brown and Williams, 2021). Ride-sourcing usage has been impacted by the pandemic, as evidenced by the significant declines in ridership and the suspension of shared ride-sourcing due to safety concerns (Bogage, 2020; Lee, 2020). The pandemic and the resulting public health policies have produced short-term disruptions in travel behaviour that have the potential to have long-term impacts on travel mode choices (Khamis, 2021). Prior studies have found that the pandemic has reduced the willingness to use shared modes, such as public transit and ride-sourcing, primarily due to the attitudes and perceived risk associated with these modes (De Vos, 2020; Loa and Habib, 2021; Mashrur et al., 2022). This shift has imposed challenges for ride-sourcing services (Kong et al., 2020; Loa et al.,

* Corresponding author.
E-mail addresses: nora.liu@mail.utoronto.ca (Y. Liu), patrick.loa@mail.utoronto.ca (P. Loa), jackkaili.wang@mail.utoronto.ca (K. Wang), khandker.nurulhabib@utoronto.ca (K.N. Habib).

https://doi.org/10.1016/j.jocm.2023.100431
Received 29 August 2022; Received in revised form 20 June 2023; Accepted 5 July 2023
Available online 20 July 2023
1755-5345/

2021d; Tirachini, 2020). Given the substantial impact of ride-sourcing on travel behaviour, it is essential to understand the pandemic's short- and long-term impacts on ride-sourcing use.

This study examines the factors influencing the decision to use ride-sourcing services during and after the pandemic using stated-preference (SP) data. The study applies both theory-driven (discrete choice) and data-driven (machine learning) modelling techniques, and evaluates the conformity and complementarity of the modelling results (Van Cranenburgh et al., 2022). For the data-driven approach, multi-task learning (MTL) deep neural networks (DNN) models are developed to capture the intra-personal correlation between mode choices during and after the pandemic (Wang et al., 2020a). The MTL framework is a variant of the multi-layer perceptron (MLP) structure. The MTL can jointly predict multiple related tasks or outcomes. It leverages the shared information between the multiple related tasks to improve model performance. In the case of this study, the intra-personal correlation between a respondent's during-pandemic and post-pandemic mode choices might exist; therefore, the MTL framework is utilized to jointly model potentially correlated choices.

Previous studies have confirmed that perceived risks provoked by the pandemic were determinants in travellers' choice of using shared mobility (Mashrur et al., 2022; Rahimi et al., 2021). Thus, modelling latent variables is important for examining the influence of the pandemic on travel mode choice. To this end, theory-driven integrated choice and latent variable (ICLV) models are also estimated. ICLV models considered attitudes, opinions and perception as latent variables in the systematic utility function (Ben-Akiva et al., 2002; Kamargianni et al., 2015). The model systematically captures unobserved latent variables through the multiple indicators multiple causes (MIMIC) model (Ben-Akiva et al., 2002; Kamargianni et al., 2015). The ICLV model ensures straightforward econometric interpretability of the effect of latent variables in the choice-making process.

Van Cranenburgh et al. (2022) raised the question on the possibility of integration of machine learning (ML) and discrete choice paradigm. They recognized that improving the explainability of ML models could lead to greater application of ML techniques in choice modelling. Recently, several studies delivered early wins on deriving econometric information from ML models (Wang et al., 2020b; Zhao et al., 2020). However, their inference was limited to socio-economic and level-of-service (LOS) variables. This study extends the experiment by extracting econometric information of attitudinal variables from ML models. This exercise sheds light on the applicability of state-of-art ML models to infer the effects of latent variables on choice outcomes.

In a nutshell, this study aims to contribute to the literature from the following perspectives. Firstly, this study provides empirical evidence on factors influencing ride-sourcing use and how their influence varies based on the pandemic context. This information can contribute to understanding the nature of ride-sourcing use, its subsequent impacts on the transportation system and the environment, and the applicability of the findings of pre-pandemic ride-sourcing studies, in the post-pandemic period. Secondly, this study demonstrates the applicability of the novel MTL-DNN framework on a stated-preference dataset with intra-personal correlation. This paper contributes by showcasing how to exploit the strengths of the MTL model with specific SP experimental designs. The SP design asked respondents to make mode choice decisions for the same trip in two contexts: 1) if the trip was made during the pandemic; 2) if the trip were to be made after the pandemic. The design artificially induces intra-personal correlation between choice scenarios. Thirdly, this study contributes to collective efforts of integrating the data-driven and theory-driven paradigms. The study evaluates the conformity and complementarity of the modelling results from the two modelling paradigms, examining the feasibility of merging the two paradigms.

The remainder of the paper is organized as follows: in Section 2, a review of relevant studies is presented. Section 3 describes the data used in this study. Section 4 presents the methodology. Section 5 presents modelling results. Finally, Section 6 concludes the study.

2. Literature review

The shift in modal preferences that resulted from the pandemic has been attributed primarily to its impacts on attitudes and perceptions of risk (De Vos, 2020). Specifically, there is evidence that attitudes towards private vehicles and active modes have become more positive or remained unchanged. In contrast, public transit and ride-sourcing attitudes have become more negative (De Haas et al., 2020; Ozbilen et al., 2021; Rahimi et al., 2021; Shamshiripour et al., 2020). It has resulted in a decline in ride-sourcing trips and reduced ride-sourcing frequency (City of Toronto, 2021; Loa et al., 2021b; Matson et al., 2021). Additionally, the belief that the risk associated with travel has increased due to the pandemic has been found to influence the frequency of ride-sourcing use during the pandemic (Loa et al., 2021b).

However, there is also evidence that certain individuals have increased ride-sourcing usage due to the pandemic. For example, Loa et al. (2021b) reported that 8% of respondents increased their ride-sourcing use, with the most common reasons being the desire to avoid crowded areas and their reluctance to use transit. A similar shift was reported by Costa et al. (2022), who found that ride-sourcing use has declined among car users and increased among transit users. The shift from public transit to ride-sourcing persisting post-pandemic could result in an increase in both vehicle-kilometers travelled and emissions, due to the increase in the number of ride-sourcing trips and deadheading among drivers. Notably, there is evidence that ride-sourcing use has recovered more strongly than transit ridership, which could stem from a continued apprehension toward using shared modes (Dzisi et al., 2021). The shift in modal preference and policies restricting ride-sourcing services during the pandemic may have significant impacts on the transportation system and the environment. Through Monto-Carlo Markov Chain simulations, Yang et al. (2022) estimated that the suspension of ride-sourcing services could result in a 60% reduction in vehicle emissions, although it would reduce travel efficiency at the same time.

Using ML approaches to model mode choice decisions has been a rising area of research in recent years. Compared to discrete choice models, ML models are shown to have higher prediction accuracies (Golshani et al., 2018; Lee et al., 2018; Ma and Zhang, 2020). The ML approaches that have been applied for mode choice modelling include tree-based algorithms (Ma and Zhang, 2020; Moons et al., 2007), support vector machines (Moons et al., 2007), and various DNN models (Golshani et al., 2018; Lee et al., 2018; Ma and Zhang, 2020). In particular, DNN models have been increasingly adopted in transportation studies due to their excellent performance and flexible structures. For example, Ma and Zhang (2020) developed a DNN model with entity embeddings to predict travel mode choices, achieving an overall accuracy of 88.3% with four travel modes: driving, walking, cycling, and public transit. The structure of a DNN model can be tailored specifically for the given datasets and objectives of a study. The MTL structure utilized in this study is a special form of DNN designed to jointly model multiple related tasks. It is a popular DNN structure widely applied in image classification (Bensaoud and Kalita, 2022; Kuang et al., 2017) and natural language processing (Chen et al., 2021). However, MTL models are rarely utilized in transportation and choice modelling. The only such study is Wang et al. (2020a), who used an MTL DNN framework to jointly estimate travel mode choices from revealed preference (RP) and SP data. They argued that the MTL DNN framework is theoretically appealing in jointly analyzing RP and SP mode choice decisions. The advantages of this framework include its automatic feature learning ability and flexible modelling structure. It can be applied to any related tasks (i.e., RP and SP mode choices), and flexibles constraints may also be applied to capture the similarities and differences between the related tasks (Wang et al., 2020a). They also discovered that the performance of the MTL DNN model is comparable to that of a nested logit (NL) model (Wang et al., 2020a).

In terms of model interpretability, there have been some recent developments in interpreting the effects of explanatory variables from DNN models. Particularly for choice analysis, Wang et al. (2020b) argued that DNN models can be interpreted and analyzed for economic information such as choice probabilities, market share, social welfare, and elasticities. They proposed both function-based interpretation and gradient-based interpretation for extracting economic information from DNN models and found such information to be reasonable and more flexible than those of discrete choice models. However, Wang et al. (2020b) also pointed out that DNN models are highly sensitive to hyperparameters and initialization, and their probability functions can be locally irregular. These can potentially limit the interpretation of DNN models and need to be addressed. Zhao et al. (2020) also investigated the behavioural outputs and variables' importance of a neural network (NN) model for travel mode choice. They compared the marginal effects and elasticities of the variables computed from the NN model and a multinomial logit (MNL) model. They found that derived marginal effects and elasticities have the same signs but rather different magnitudes. In addition to economic and behavioural information, the relative importance of the explanatory variables on model outputs is also examined by studies using DNN to model travel mode choices. Several methods have been proposed to compute variable importance for DNN models, such as absolute values of weights (Zhang et al., 2020; Zhao et al., 2020) and Garson's algorithm (Golshani et al., 2018). As an extension of the work stated above, this paper compares the economic information extracted from ICLV models with MTL models. None of the existing studies summarized above contained attitudinal factors. This paper can service as a reference for future development of inferring latent variables in machine learning models.

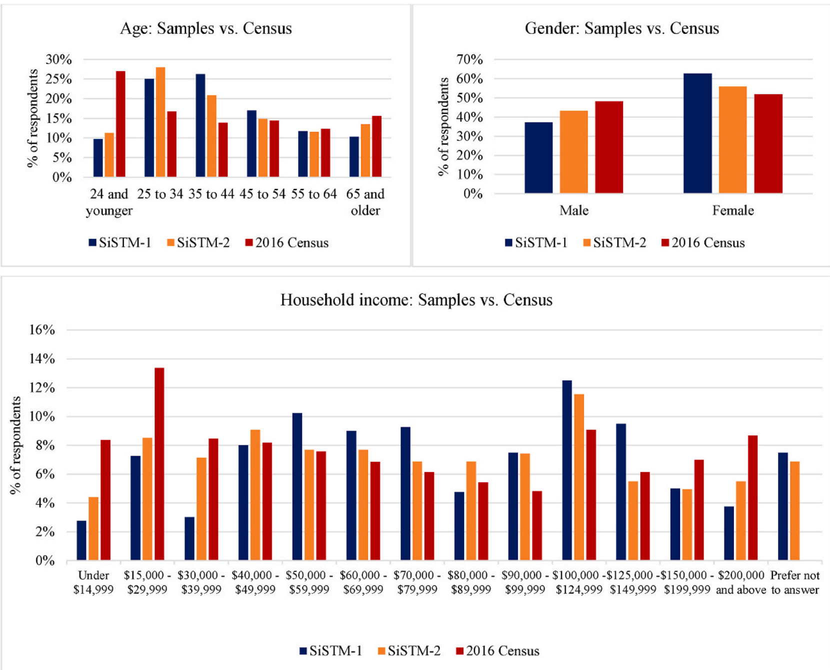

Fig. 1. Comparison of key socio-economic attributes – samples vs. census.

3. Data description

3.1. Survey design

The data used to develop the models were collected as part of the *study into the use of shared travel modes* (SiSTM). The goals of SiSTM were to understand the impacts of the pandemic on the use of ride-sourcing in the Greater Toronto Area (GTA) and the influence of attitudes on ride-sourcing use. Two cycles of the SiSTM survey were conducted – the first ("SiSTM-1") in July 2020 and the second ('SiSTM-2″) in July 2021. The SiSTM-1 and SiSTM-2 surveys asked respondents to provide information on personal and household attributes, their pre-pandemic travel behaviour, and the impacts of the pandemic on their ride-sourcing usage. Respondents were also asked to answer a series of attitudinal questions about the during- and post-pandemic periods. Additionally, respondents were asked to complete a series of SP experiments on travel mode choices for commute and non-commute trips during the pandemic and post-pandemic periods.

Both surveys were conducted using a web-based interface, with a random sample of market research panel members living in the GTA invited to complete the survey. Non-monetary compensation was provided to participants upon completion of the survey by the market research company. After the data were cleaned and invalid responses were removed, 920 and 806 completed responses remained for SiSTM-1 and SiSTM-2, respectively. For more information on the SiSTM-1 and SiSTM-2 surveys, see Loa et al. (2021a) and Loa et al. (2021c), respectively. This study uses the responses from residents of Toronto due to the higher levels of ride-sourcing and transit use compared to the surrounding municipalities; there were 400 and 364 responses from Toronto residents in SiSTM-1 and SiSTM-2, respectively.

3.2. Sample description

The distributions of key socio-economic attributes in the two samples are compared to the 2016 Canadian Census in Fig. 1. Respondents under the age of 24 and over 65 were under-represented in the sample, while those between the ages of 25 and 44 were over-represented. This is likely due to the administration of the survey to the members of a market research panel and the use of a web-based survey interface. Similarly, individuals from households earning less than $40,000 and over $150,000 annually were under-represented, while those from households earning between $40,000 to $150,000 were over-represented. As shown in Fig. 2, most respondents possessed a driver's license and had access to a private vehicle, while slightly less than half owned a transit pass at the time of the survey.

SiSTM-1 and SiSTM-2 included attitudinal questions pertaining to both the during-pandemic and post-pandemic periods. As outlined in Fig. 3, most respondents believe that the pandemic increased the risk associated with travel and with using shared mobility services. However, compared to SiSTM-1 respondents, SiSTM-2 respondents were less likely to believe that the level of risk at the time of the survey was greater than that of the pre-pandemic period. Moreover, SiSTM-2 respondents were more likely to indicate that they would feel safe using taxis and ride-sourcing during the pandemic. Similarly, as shown in Fig. 4, SiSTM-1 respondents were more likely to report apprehension about spending time outside their homes and increasing their reliance on online orders. Conversely, a greater percentage of SiSTM-2 respondents indicated that they would like to return to their normal routine despite COVID-19 remaining a public health threat.

SiSTM-2 also included several new attitudinal questions to help understand how attitudes and perceptions of risk changed during the pandemic. As summarized in Fig. 5, the use of ride-sourcing and taxi is influenced by the severity of the public health measures that are currently in place. Additionally, roughly half of the respondents indicated their concern regarding COVID-19 increases when public health measures become more restrictive. Besides, the responses to the attitudinal questions highlight how the change in the level of concern about COVID-19 differs among members of the sample. Specifically, while 43% of respondents indicated they were less concerned about COVID-19 than earlier in the pandemic, 37% indicated more concern.

SiSTM-1 and SiSTM-2 respondents were also asked about their attitudes toward travel in the post-pandemic period, and the results are illustrated in Fig. 6. Overall, SiSTM-2 respondents were less likely to believe that there would be more risk associated with using ride-sourcing post-pandemic and less willing to spend time outside their homes. Additionally, SiSTM-2 respondents were more likely to

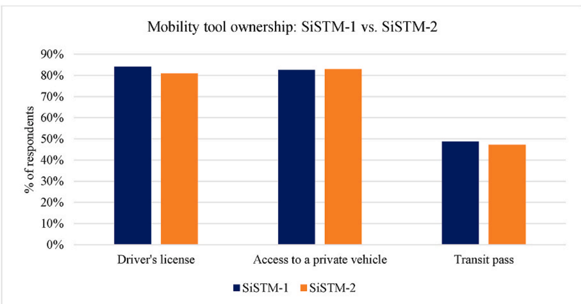

Fig. 2. Mobility tool ownership – SiSTM-1 vs. SiSTM-2.

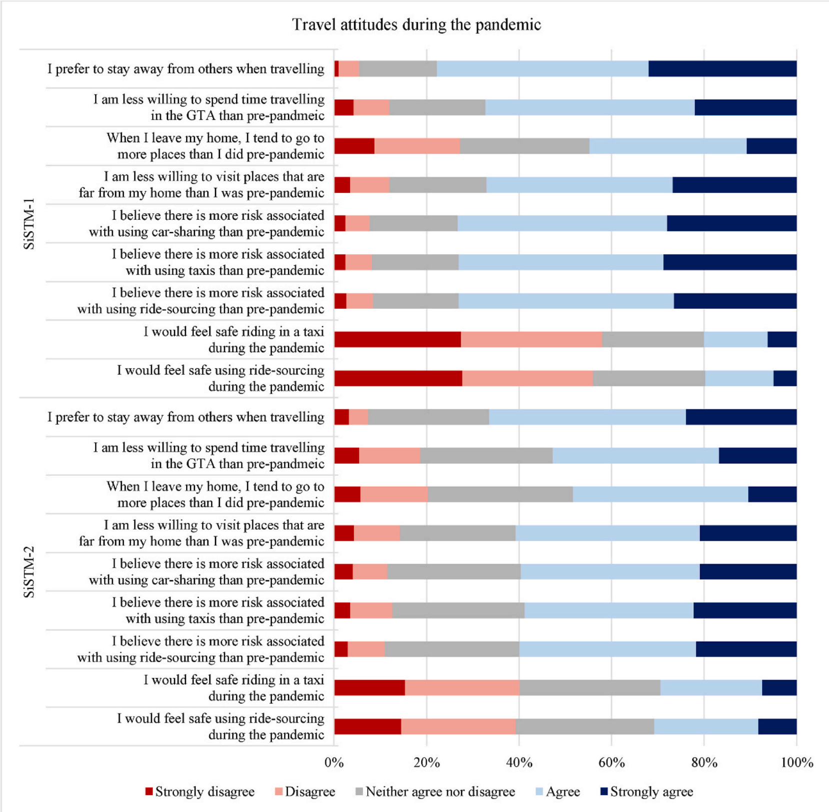

Fig. 3. Comparison of during-pandemic travel attitudes.

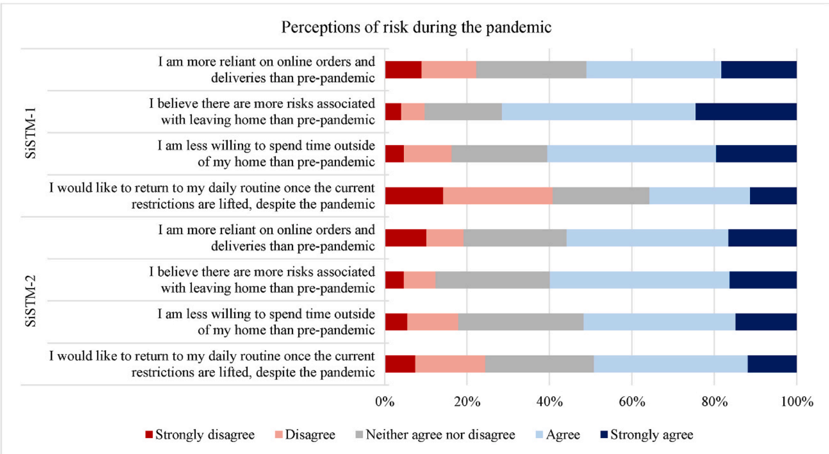

Fig. 4. Comparison of during-pandemic perceptions of risk.

believe that daily life would return to normal post-pandemic; however, roughly one-quarter of respondents believe that post-pandemic life will differ from pre-pandemic life.

3.3. Experimental design

Survey respondents were also asked to complete a series of SP experiments concerning travel mode choices at different pandemic stages. The experiments included eight modes – auto-drive (i.e., drive yourself), auto-passenger (i.e., driven by someone you know), public transit, exclusive ride-sourcing, shared ride-sourcing, taxi, bicycle, and walking. The two driving modes were only available to

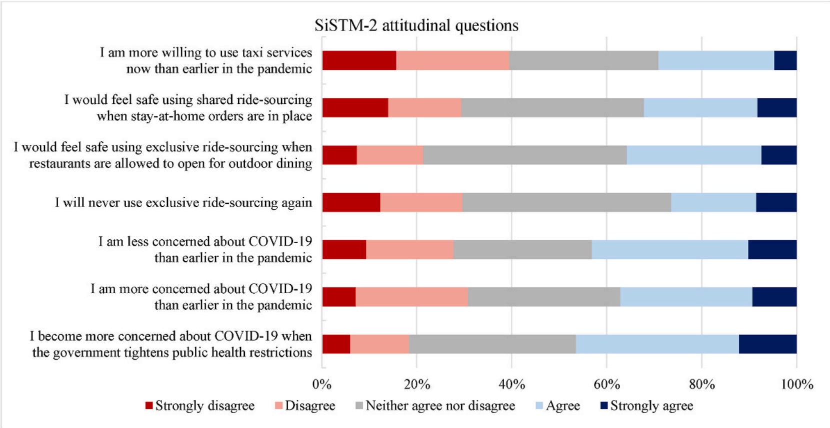

Fig. 5. Summary of attitudinal questions only included in SiSTM-2.

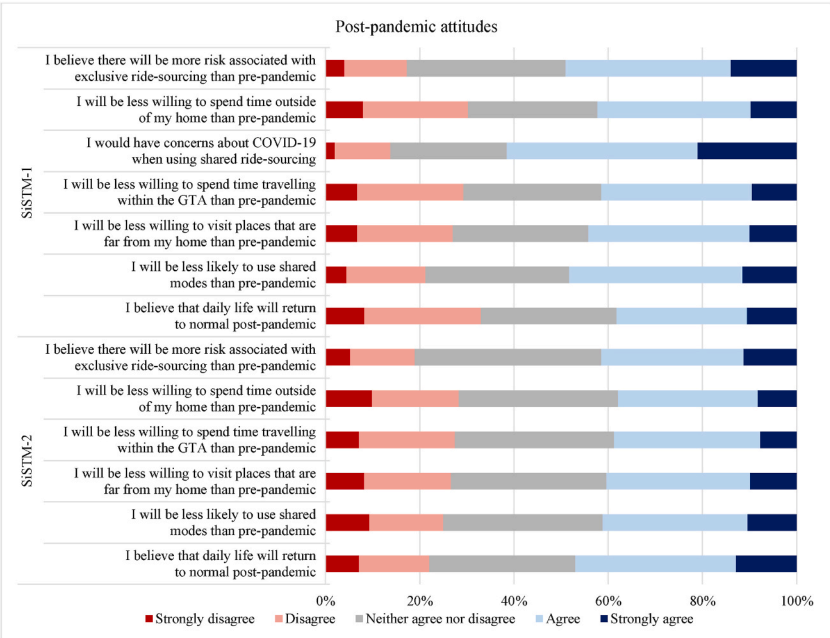

Fig. 6. Comparison of post-pandemic attitudes.

You have been fully vaccinated against COVID-19, but mass vaccination has not been achieved

Attributes	Drive yourself	Driven by someone you know	Public transit	Exclusive ride-sourcing	Shared ride-sourcing	Taxi	Bicycle	Walking
Travel time (mins)	46.4	46.4	56.6	34.3	34.3	34.3	36.5	44.4
Travel cost ($)	4.21	6.32	3.25	34.79	27.84	43.12	-	-
Parking cost ($)	5.00	-	-	-	-	-	-	-
Waiting time (mins)	-	-	8	8	8	2	-	-
Walking time (mins)	-	-	1	-	-	-	-	-
Number of other passengers	-	-	-	-	1	-	-	-
Level of crowding	-	-	Highly crowded (Physical distancing cannot be maintained)	-	-	-	-	-
All passengers and operators are required to wear masks	-	-	-	-	-	Yes	-	-
Vehicles are disinfected at the end of each day	-	-	-	Yes	Yes	Yes	-	-
There is a physical barrier between The driver and the passengers	-	-	-	Yes	-	Yes	-	-

Fig. 7. An example of the SP experiments used in the SiSTM surveys.

those who indicated they had access to a private vehicle; all other modes were always available. The alternatives and attributes included in the SP experiments are demonstrated in Fig. 7. The SP experiments were designed based on a hypothetical commuting and non-commuting trip, with the most recent iteration of the regional household travel survey being used to determine the reference trip distances. Specifically, the average trip distances were used for motorized modes (i.e., private vehicle, taxi, and ride-sourcing) and public transit, while the 95th percentile values were used for active modes.

The reference distance values were then used to define the reference travel time and cost values for each mode and each trip purpose, based on a set of assumed travel speeds and per-distance costs. The levels of the travel time and cost attributes were determined by modifying the reference values using information from previous stated preference studies regarding mode choices. The values of the waiting time attribute for public transit were determined based on the service standards of the local transit agencies. In contrast, the values of the walking time attribute were based on established standards for the maximum access distance for transit (Vuchic, 2005). Besides, the values for the taxi and ride-sourcing modes were based on the values used by Weiss et al. (2019). In addition to the attributes presented in Fig. 7, two contextual variables were included in the SP experiments of SiSTM-2 – the number of vaccine doses received by the respondent and whether mass vaccination has been achieved.

The D-efficient design method was applied in the experimental design software Ngene to produce 12 experimental designs for both commuting and non-commuting trips. Each respondent completed a total of 12 choice experiments. Each respondent was presented with three randomly selected experiments for each trip purpose and was asked to select their preferred mode if the trip was made during the pandemic period. Respondents were then presented with the same set of choice experiments and asked to select their preferred mode if the trip was made in the post-pandemic period. This design aimed to capture the difference in respondents' choices under different external contexts. In addition, the design deliberately creates a correlation between the during-pandemic (DP) and post-pandemic (PP) choice scenarios.

As shown in Figs. 8 and 9, SiSTM-2 respondents were more likely to choose the public transit, taxi, and bicycle alternatives and less likely to choose the driving modes. This could stem from changes in attitudes and perceptions of risk over the pandemic and the increases in cycling observed in the first year of the pandemic (Budd and Ison, 2020; Marsden et al., 2021).

4. Methodology

4.1. *Deep neural network (DNN) models*

4.1.1. *Multi-layer perceptron (MLP) model*

DNN models are popular ML models that have been applied in various fields. They are known for their flexible structures and the ability to learn nonlinear relationships between input variables (Golshani et al., 2018). One of the most commonly used structures of DNN models is the fully connected feedforward network, also known as the multi-layer perceptron (MLP) model. An MLP model comprises one input layer, several hidden layers, and one output layer. Each layer contains a set of neurons. The neurons in the input layer represent the explanatory variables, and the classification or regression results are expressed through the neurons in the output layer, depending on whether it is a classification or regression model. In this study, classification models are utilized to predict mode choices; therefore, the output layer generates the estimated probabilities of being classified as each alternative mode. The numbers of hidden layers and their corresponding neurons are hyperparameters of the model that need to be tuned to fit the dataset better. The structure of the MLP model developed in this study is shown in Fig. 10, which contains four hidden layers. There are 200 neurons in each of the first two hidden layers, and 100 neurons in the last two hidden layers. This model structure is consistent with the structure of the MTL model used in this study.

The underlying mechanisms of the DNN models are pattern association and error correction (Hensher and Ton, 2000). The information on pattern association is processed through the neurons in each layer. Each neuron receives data from the connected neurons in the previous layer and then processes them through a nonlinear activation function. Equation (1) provides a general formulation of this process for a neuron u:

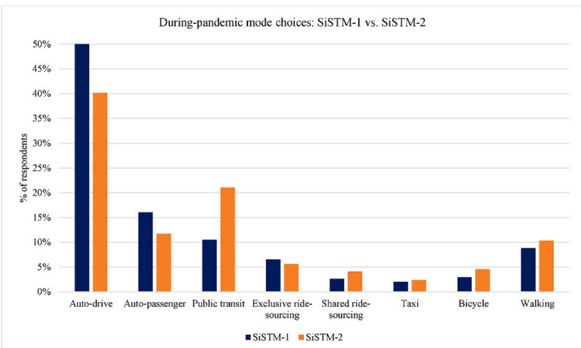

Fig. 8. Comparison of modes chosen in the during-pandemic experiments.

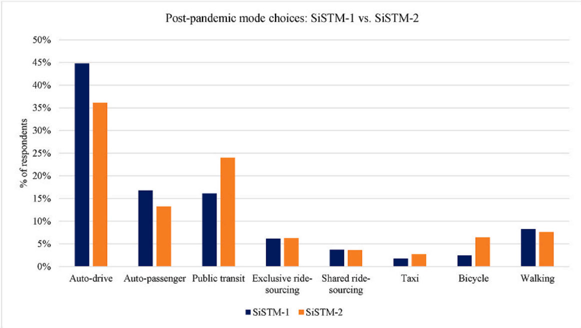

Fig. 9. Comparison of modes chosen in the post-pandemic experiments.

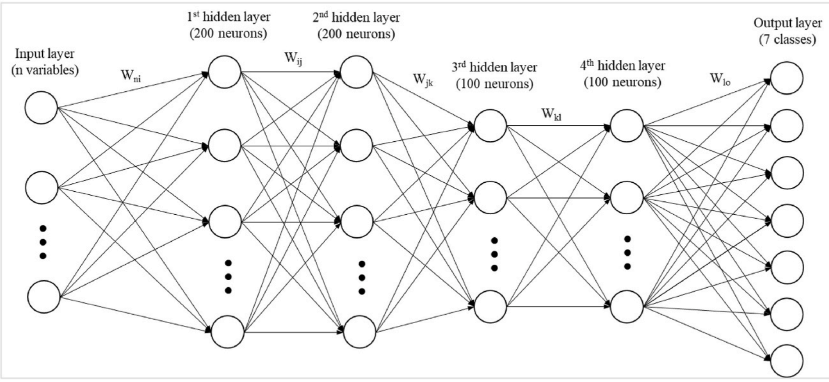

Fig. 10. Structure of MLP model.

$$h_u = \varphi\left(\sum_{m=1}^{M} W_{mu} X_m + \beta_u\right) \tag{1}$$

where,

M = number of neurons in the previous layer,
X_m = input to the current neuron u from the previous layer that contains M neurons,
W_{mu} = weight that connects the previous neuron m to the current neuron u,
β_u = bias term of the current neuron u,
h_u = output of the nonlinear activation function $\varphi(\,)$.

The outputs of the current layer h_i then serve as the inputs to the subsequent layer. Several types of nonlinear activation functions exist, such as sigmoid, hyperbolic tangent, step, and rectified linear unit (ReLU) functions (Lee et al., 2018). This study utilizes the ReLU function as the nonlinear activation function in all hidden layers of the models. The formulation of the ReLU function is shown in Equation (2).

$$ReLU(Z_i) = \begin{cases} Z_i & if\ Z_i \geq 0 \\ 0 & if\ \ Z_i < 0 \end{cases} \tag{2}$$

The ReLU activation function addresses the vanishing gradient problem and its detrimental impact on estimating weights (Nair and Hinton, 2019). The gradient descent method adjusts the weights to minimize the errors between predictions and targets. Such errors are computed through a loss function. The cross-entropy loss function is utilized in this study, whose formulation is presented in Equation (3). It is a common loss function used for classification models. The softmax function shown in Equation (4) is the activation function of the output layer, which is used to estimate the probability of choosing each choice alternative. The alternative with the highest probability is then determined to be the predicted choice of the observation.

$$L_{CE} = -\frac{1}{M}\sum_{m}^{M}\sum_{k=1}^{K} t_k \log P_k \tag{3}$$

$$P_k = softmax(Z) = \frac{exp(z_i)}{\sum_{k=1}^{K} exp(z_k)} \tag{4}$$

where,

M = number of observations in the dataset,
K = number of choice alternatives to be classified,
t_k = a binary indicating whether alternative k is the correct prediction,
P_k = the predicted probability of choice being alternative k,
Z = the input to the softmax function located at the output layer.

4.1.2. Multi-task learning (MTL) model

The multi-task learning (MTL) framework is a DNN framework that leverages the shared information between multiple learning tasks to improve the prediction accuracy for each task (Zhang and Yang, 2018). This framework is often applied when multiple related tasks need to be identified from the same datasets. In this study, the related tasks are the survey respondents' mode choices for the during- and post-pandemic SP scenarios. Since the SP scenarios are the same for the two periods, the explanatory variables remain constant across the two related tasks. Based on this characteristic of the dataset, the structure of the MTL model is developed, as shown in Fig. 11. The MTL model includes four hidden layers, similar to the MLP model. However, the difference is that there are shared and task-specific layers in the MTL model. The first two hidden layers are shared layers that are fully connected and trained by data from both tasks. Each of the shared layers has 200 neurons. Following the shared layers are task-specific layers with 100 neurons. Two task-specific layers are fully connected for each task and can only be trained by data from that specific task. The shared layers are to capture the similarities between the two tasks, and the task-specific layers are to learn the unique characteristics of each task (Caruana, 2004; Wang et al., 2020a). The MTL model adopts the same activation functions as the MLP model. The cross-entropy loss is calculated for each task, and the model is trained by minimizing the average cross-entropy loss of the two tasks.

4.1.3. Model training and regularization

Other than the number of hidden layers and neurons, the hyperparameters of the DNN models also include the optimization algorithms, learning rate, batch size, training iterations, and regularization methods. The hyperparameters were tuned using the random search method and evaluated with a 5-fold cross-validation process (Bergstra and Bengio, 2012; Zhang et al., 2020).

Different optimization algorithms were tested for the models, including stochastic gradient descent, the Adagrad, and the Adam algorithms. The Adam optimization algorithm is chosen for both MLP and MTL models to minimize the average cross-entropy loss. It extends the stochastic gradient descent (SGD) optimization algorithm. Unlike the classical SGD, which employs a constant learning rate to update all parameters, the Adam optimizer generates adaptive learning rates for different parameters by estimating the first and second moments of gradients (Kingma and Ba, 2015). Its effectiveness has been exemplified in many deep learning applications (Kingma and Ba, 2015).

The learning rate determines the proportion of parameters to be updated along the gradient direction (Ma and Zhang, 2020). The search space for the learning rate was between 1.00E-05 and 0.01, and a learning rate of 5.00E-05 was chosen because it yielded the

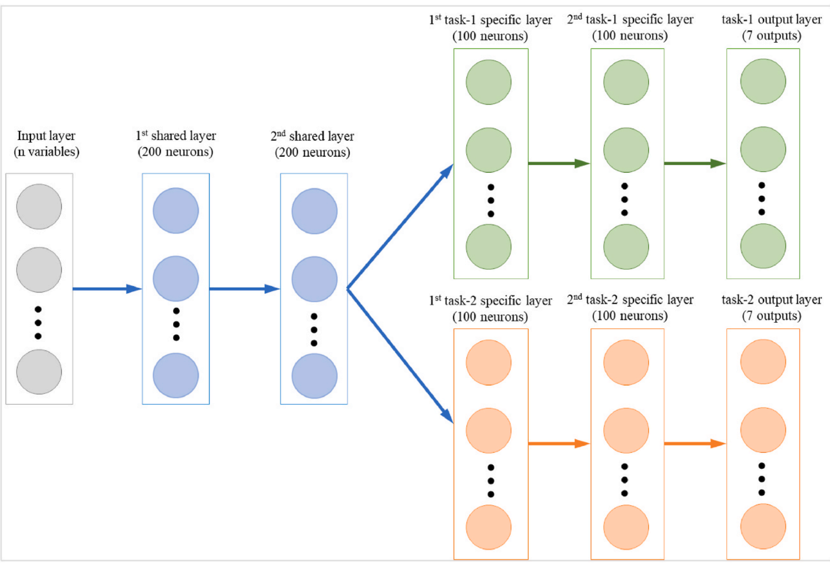

Fig. 11. Structure of MTL model.

best model performance. The models are trained over 2000 iterations. The model parameters are updated in each iteration through a random subset of the training dataset, also known as a mini-batch. The search space for the batch size was between 10 and 100, and a batch size of 100 was selected for all models after hyperparameter tuning. Although a larger batch size can improve computational efficiency (Masters and Luschi, 2018), this batch size is chosen because it achieved better performance during hyperparameter tuning.

The dropout regularization method is employed in both MLP and MTL models to reduce overfitting. Dropping out means temporarily removing a neuron and all its connections from the network (Srivastava et al., 2014). During training, a neuron can be randomly removed with a probability of p. This is to introduce noise and break up the potential co-adaptations between layers, so that the model predictions do not rely too much on a few particular neurons. The dropout regularization is only applied during training. When testing the model, the weights of the retained neurons are rescaled by multiplying a probability of $1-p$ to offset the effect of the dropout (Ma and Zhang, 2020; Srivastava et al., 2014). The search space for the dropout probability p was between 0.1 and 0.5, and a value of 0.25 was ultimately applied to all models after hyperparameter tuning.

An extensive number of variables can be extracted from the collected data. Because of the automatic feature learning ability of the DNN, the variable selection process was not conducted through a specific method before developing the models. All variables believed to be relevant to mode choice decisions are used in the models, including variables related to socio-economic attributes, trip characteristics of each mode, and attitudes towards the pandemic. The attitudinal variables are treated as continuous variables ranging from 1 to 5, with 1 being strongly disagree with the attitudinal statement and 5 being strongly agree with the attitudinal statement.

Each of the SiSTM-1 and SiSTM-2 datasets is randomly split into a training set and a testing set. The ratio between the training and testing set is 5:1. A 5-fold stratified cross-validation is applied to the training set for hyperparameter tuning and out-of-sample validation. Eventually, the performance of the tuned model is evaluated using the testing set. When developing the models, it was discovered that the unbalanced datasets could negatively impact the performances of the models, as the model tends to favour the dominant mode to minimize loss. To mitigate this issue, the random over-sampling technique is applied to the training datasets, so that all modes have an equal number of observations. Since resampling techniques can only be conducted based on one dependent variable, the under-represented alternatives are randomly over-sampled based on the during-pandemic mode choices.

4.1.4. Availability of mode choice alternatives in DNN models

The set of available modal alternatives (i.e., choice set) is often different across individuals. For example, the auto-drive mode is not expected to be available for an individual who does not have a driver's license or access to a private vehicle. Therefore, identifying the choice set is an important step in mode choice modelling. In discrete choice models, the availability of a modal alternative is usually addressed through the use of a binary indicator when calculating the probability of choosing each mode. However, the identification of choice sets is less investigated in studies using ML methods to model mode choice. Most ML studies assume that all modal alternatives are available to all trip-makers. To the best of our knowledge, only two studies attempted to incorporate choice set identification in their DNN models by modifying the softmax function shown in Equation (5) and excluding the unavailable modes (Nam et al., 2017; Wang et al., 2020a).

This study proposes an alternative approach by exploiting the learning ability of the DNN model. The SP design used in this study assumes that all modes are available for all survey respondents, except for auto-drive and auto-passenger. For respondents who do not have a driver's license or a private vehicle in the household, the SP design automatically excludes auto-drive and auto-passenger in their choice sets. When developing the MLP and MTL models, a binary variable is added as a feature to reflect the availability of these modes. The models are expected to learn the relationships between these variables and the choices and discover that auto-drive or auto-passenger cannot be chosen when this variable is zero. A preliminary comparison was conducted between this method and the approach of modifying the softmax function. The results show that there was no obvious difference in terms of prediction accuracies, and the former took less time to train the models. Nevertheless, to fully investigate the effectiveness of this approach, future studies should conduct more in-depth analysis using different datasets with various choice-set availability conditions.

4.2. Discrete choice models

To improve the robustness of this study, discrete choice models are also estimated to compare with the MTL models. The discrete choice models are to provide a benchmark reference to the interpretation of the MTL models. In particular, integrated choice and latent variable (ICLV) models are developed to account for the effects of attitudes on mode choice decisions. There are two components in an ICLV model: a discrete choice model and a latent variable model (Ben-Akiva et al., 2002). The MNL model is utilized as the discrete choice component of the ICLV models. The systematic utility function of a choice alternative k for individual a is presented in Equation (5).

$$V_k = \beta_k x_a + \mu_k LV_a + \varepsilon_k \tag{5}$$

where x_a is a set of variables, including socio-economic and level-of-service variables, and LV_a is a set of latent variables to account for attitudinal influence. The β_k and μ_k are the estimated coefficients, and ε_k is a random error term with Type I Extreme Value distribution.

The multiple indicators multiple causes (MIMIC) model is applied to establish the relationship between the latent variables, socio-economic variables observed values of the attitudinal variables. The formulation is shown in Equations (6) and (7) (McFadden, 1986; Train et al., 1987).

$$LV_a = \lambda S_a + \eta_a \tag{6}$$

where LV_a is the latent variable, S_a is a set of socio-economic variables with estimated coefficients λ, and η_a is a random error term that has a standard normal distribution.

The attitudinal statements are the indicators of their corresponding latent variable and are expressed through a measurement equation, as shown in Equation (7) (McFadden, 1986; Train et al., 1987).

$$I_a = \gamma LV_a + v_a \tag{7}$$

where I_a is the observed value of the indicator (i.e., attitudinal statement), LV_a is the latent variable with an estimated coefficient γ, and v_a is the random error with a standard normal distribution.

5. Modelling results

5.1. Model performance

This section presents the results of ICLV, MTL and MLP models estimated for each survey cycle. Fig. 12 shows the workflow of each model for each survey cycle. In total, five models are estimated for each survey cycle. ICLV and MLP models are estimated and trained for each DP and PP scenario, separately. MTL model jointly takes both DP and PP scenario in the shared layer but make prediction. Both models take socio-economic variables as input. The ICLV uses latent variables calculated using Equation (6). Conversely, MTL and MLP directly use attitudinal variables as input variables.

Table 1 summarizes the overall prediction accuracies of the best-performed models. The prediction accuracy of each model is calculated on the testing datasets. Overall, the MTL models outperform the ICLV and MLP models by significant margins in both datasets. Similar observations were also reported by Wang et al. (2020a), who found that MTL models can outperform nested logit models by around 5% when predicting mode choices using RP-SP data. Notably, the MTL models consistently demonstrate higher accuracy compared to the MLP models. One of the significant factors contributing to the performance improvement of the MTL model is its ability to utilize correlated choices. The comparison in this study is made in a carefully designed controlled setting. Both MTL and MLP models are very alike. The models were trained on the same dataset with turned hyperparameters to achieve the highest possible accuracy. The only difference is that MTL is equipped with shared layers aiming to capture any shared information between DP and PP scenarios. In this study, induced correlation from the same respondents is the major known source of the shared information that can be utilized by the MTL. This correlation is induced purposely by the experimental design. Indeed, in the data-driven method, it is challenging to conclude with certainty the improved efficiency is entirely brought by utilizing shared correlated choices. This proposition is derived from the model performance and rationing model structure difference between MTL and MLP. Future studies could perform rigorous investigations such as activation mapping to verify the proposition.

Given the higher prediction accuracies of the MTL models, their model performances are discussed in detail in the following sections. The confusion matrices of the MTL model results are presented in Table 2. A confusion matrix includes the observed and predicted numbers of observations of each modal alternative, along with their recall and precision. Equations (8) and (9) show the computation for recall and precision, respectively (Liu et al., 2021). The recall value of an alternative can be interpreted as its prediction accuracy.

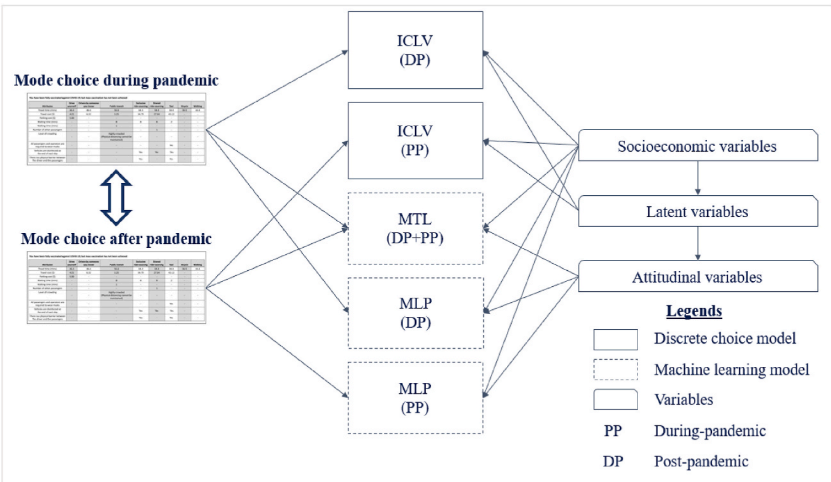

Fig. 12. Graphic illustrations of models for each survey cycle.

Table 1
Summary of model performance.

Model	SiSTM-1 (DP)	SiSTM-1 (PP)	SiSTM-2 (DP)	SiSTM-2 (PP)
ICLV	50.2%	46.4%	48.5%	45.6%
MLP	51.5%	45.7%	56.2%	51.5%
MTL	65.0%	61.2%	69.1%	70.6%

*DP: during-pandemic, PP: post-pandemic.

Table 2
Confusion matrices of MTL models.

SiSTM-1 MTL	Predicted Mode (During-Pandemic)										Overall Accuracy
Observed Mode	Walk	AD	AP	PT	ERS	SRS	Taxi	Cycling	Total	Recall	
Walk	19	11	3	1	1	0	0	0	35	54.3%	65.0%
AD	8	163	15	5	2	1	1	1	196	83.2%	
AP	2	16	36	4	1	0	0	1	60	60.0%	
PT	5	14	1	19	0	2	0	1	42	45.2%	
ERS	0	7	2	4	10	1	1	1	26	38.5%	
SRS	0	2	3	0	3	2	0	0	10	20.0%	
Taxi	0	1	2	1	1	2	1	0	8	12.5%	
Bicycle	1	3	4	0	0	0	1	3	12	25.0%	
Total	35	217	66	34	18	8	4	7	389	–	
Precision	54.3%	75.1%	54.5%	55.9%	55.6%	25.0%	25.0%	42.9%	–	–	
SiSTM-1 MTL	Predicted Mode (Post-Pandemic)										Overall Accuracy
Observed Mode	Walk	AD	AP	PT	ERS	SRS	Taxi	Cycling	Total	Recall	
Walk	11	7	4	3	1	0	0	0	26	42.3%	61.2%
AD	2	150	9	6	1	1	1	2	172	87.2%	
AP	1	25	32	10	0	1	0	1	70	45.7%	
PT	6	15	6	32	3	0	0	1	63	50.8%	
ERS	2	11	0	6	6	2	1	0	28	21.4%	
SRS	1	4	2	1	4	5	1	0	18	27.8%	
Taxi	0	0	0	0	2	0	0	0	2	0.0%	
Bicycle	3	2	0	1	1	1	0	2	10	20.0%	
Total	26	214	53	59	18	10	3	6	389	–	
Precision	42.3%	70.1%	60.4%	54.2%	33.3%	50.0%	0.0%	33.3%	–	–	
SiSTM-2 MTL	Predicted Mode (During-Pandemic)										Overall Accuracy
Observed Mode	Walk	AD	AP	PT	ERS	SRS	Taxi	Cycling	Total	Recall	
Walk	23	3	2	7	0	0	0	2	37	62.2%	69.1%
AD	2	126	6	9	1	1	0	0	145	86.9%	
AP	2	11	24	3	2	0	1	0	43	55.8%	
PT	9	9	4	54	1	0	0	0	77	70.1%	
ERS	0	4	1	7	6	2	0	0	20	30.0%	
SRS	0	2	2	4	3	3	1	0	15	20.0%	
Taxi	1	1	2	0	0	1	4	0	9	44.4%	
Bicycle	2	1	0	1	2	0	0	11	17	64.7%	
Total	39	157	41	85	15	7	6	13	363	–	
Precision	59.0%	80.3%	58.5%	63.5%	40.0%	42.9%	66.7%	84.6%	–	–	
SiSTM-2 MTL	Predicted Mode (Post-Pandemic)										Overall Accuracy
Observed Mode	Walk	AD	AP	PT	ERS	SRS	Taxi	Cycling	Total	Recall	
Walk	18	2	2	6	0	1	0	1	30	60.0%	70.6%
AD	1	119	5	6	2	0	2	0	135	88.1%	
AP	3	9	33	2	3	0	0	0	50	66.0%	
PT	2	12	6	58	4	4	0	1	87	66.7%	
ERS	2	4	0	6	5	0	0	0	17	29.4%	
SRS	0	1	1	1	3	3	0	0	9	33.3%	
Taxi	0	2	1	3	0	0	4	0	10	40.0%	
Bicycle	2	0	0	3	1	3	0	17	26	65.4%	
Total	28	149	48	85	18	11	6	19	364	–	
Precision	64.3%	79.9%	68.8%	68.2%	27.8%	27.3%	66.7%	89.5%	–	–	

*AD: auto-drive, AP: auto-passenger, PT: public transit, ERS: exclusive ride-sourcing, SRS: shared ride-sourcing.

$$Recall = \frac{samples\ that\ are\ accurately\ predicted\ as\ alternative\ N}{all\ samples\ of\ alternative\ N} \tag{8}$$

$$Precision = \frac{samples\ that\ are\ accurately\ prediected\ as\ alternative\ N}{all\ samples\ that\ are\ predicted\ as\ alternatvie\ N} \tag{9}$$

The confusion matrices from the two datasets reveals the same information. The prediction accuracies of the auto drive, auto passenger, and public transit are relatively higher than the other modes. This is caused by the significant imbalance between the sample sizes of the choice alternatives. None of the other modes has a market share higher than 10%. Even though the resampling technique was applied to balance the training data, there is much less information that the models can learn for the underrepresented modes. The modal alternatives that generally have the lowest prediction accuracies are exclusive and shared ride-sourcing and taxi. Several of these mode choices are incorrectly predicted as auto drive or public transit. The reason may be that they share similar travel time as auto trips or similar behavioural patterns as transit trips. Nevertheless, compared with several other studies that included fewer modes (Golshani et al., 2018; Lee et al., 2018), the prediction accuracies of the MTL models are still relatively promising.

5.2. Estimated parameters of ICLV models

The final specifications of the ICLV models for the SiSTM-1 and SiSTM-2 datasets are summarized in Table 3 and Table 4, respectively. For the ICLV models, factor analysis was utilized to identify the latent variables based on the attitudinal statements presented in the surveys. Two latent attitudinal variables are identified and included in the ICLV models: 1) perception of increased risk associated with using shared mobility during the pandemic, and 2) feeling safe using ride-sourcing services during the pandemic. The indicator variables corresponding to each latent variable are summarized in Table A1 in Appendix A.

In Tables 3 and 4, variables that are statistically significant at the 95% confidence level and have the expected sign are retained in the final model. Although the critical value of t-stat is 1.96, some parameters with lower t-stat values are also kept in the final model for comparison purposes. Given the length of the tables, we only show the estimated parameters for the choice model component in

Table 3
ICLV model specification for SiSTM-1.

		SiSTM-1 (during-pandemic)		SiSTM-1 (post-pandemic)	
Variable	Mode	Coefficient	t-stat	Coefficient	t-stat
Choice model component					
Alternative specific constant	Auto-drive (AD)	0	NA	0	NA
	Auto-passenger (AP)	−1.4575	−8.4811	−1.1046	−5.8464
	Public transit (PT)	−2.2741	−8.7120	−1.1395	−5.3481
	Exclusive ride-sourcing (ERS)	−2.0296	−5.3855	−2.2518	−9.1727
	Shared ride-sourcing (SRS)	−3.1743	−7.5306	−1.9271	−3.3496
	Taxi	−3.0239	−7.4987	−3.1709	−8.6621
	Cycling	−3.3390	−4.3020	−2.6379	−3.3619
	Walk	−2.3340	−2.9859	−1.8590	−2.5200
In-vehicle travel time (min)	All motorized modes	−0.0004	−1.1627	−0.0114	−2.4376
Travel cost ($)	All motorized modes	−0.0273	−2.9099	−0.0103	−1.0265
Travel time (min)	Cycling	−0.0187	−1.2090	−0.0381	−2.1103
Travel time (min)	Walk	−0.0198	−1.3650	−0.0201	−1.6192
Use as a commute mode pre-pandemic and it is a commute trip	Auto-drive (AD)	1.4436	7.3717	1.2827	6.9528
	Auto-passenger (AP)	1.0810	3.7583	1.1158	3.8449
	Public transit (PT)	0.7832	3.2391	1.0624	4.8267
	Exclusive ride-sourcing (ERS)	0.4224	0.9521	1.0215	2.4758
	Shared ride-sourcing (SRS)	1.0089	1.9203	0.8337	1.1588
	Cycling	1.5021	2.3794	2.5510	4.5955
	Walk	1.3128	3.9166	1.3501	4.1641
Number of bikes in household	Cycling	0.6420	3.9117	0.5777	4.3683
Has a transit pass	Public transit (PT)	0.3005	1.3444	–	–
Age	Exclusive ride-sourcing (ERS)	−0.0106	−1.4282	−0.0213	−1.8068
	Walk	0.0169	2.1171	0.0085	1.0284
Perception of increased risk associated with using shared mobility during the pandemic	Walk	0.4275	2.7928	–	–
Feeling safe using ride-sourcing services during the pandemic	Auto-passenger (AP)	−0.3206	−2.4372	−0.4456	−2.7102
	Public transit (PT)	0.4567	3.2419	0.2104	1.2405
	Exclusive ride-sourcing (ERS)	0.5133	3.1593	0.4687	2.5580
	Shared ride-sourcing (SRS)	0.7596	4.0349	0.7008	2.4721
	Taxi	0.7478	2.7625	0.6856	2.4128

Table 4
ICLV model specification for SiSTM-2.

		SiSTM-2 (during-pandemic)		SiSTM-2 (post-pandemic)	
Variable	Mode	Coefficient	t-stat	Coefficient	t-stat
Choice model component					
Alternative specific constant	Auto-drive (AD)	0	NA	0	NA
	Auto-passenger (AP)	−1.2939	−5.8256	−1.5093	−6.5763
	Public transit (PT)	−1.4093	−6.5810	−1.0267	−5.5931
	Exclusive ride-sourcing (ERS)	−2.0360	−2.8540	−2.4822	−8.5657
	Shared ride-sourcing (SRS)	−1.0142	−1.7420	−1.2161	−1.6178
	Taxi	−1.8704	−2.6936	−0.6839	−0.9698
	Cycling	−2.4728	−3.7886	−2.3682	−4.3833
	Walk	−1.2658	−1.9211	−0.3104	−0.4875
In-vehicle travel time (min)	All motorized modes	−0.0034	−0.6770	−0.0139	−2.9792
Parking cost ($)	Auto-drive (AD)	−0.0188	−1.2253	−0.0012	−2.1109
Travel time (min)	Cycling	−0.0186	−0.9876	−0.0194	−1.1954
Travel time (min)	Walk	−0.0301	−2.1999	−0.0495	−3.3265
Use as a commute mode pre-pandemic and it is a commute trip	Auto-drive (AD)	1.6596	9.7029	1.3348	7.3205
	Auto-passenger (AP)	1.7950	6.4580	1.6798	6.1313
	Public transit (PT)	1.4683	7.6666	1.7202	9.0290
	Exclusive ride-sourcing (ERS)	2.0237	6.5662	1.6134	4.4285
	Shared ride-sourcing (SRS)	1.0221	2.6294	1.5148	3.8171
	Taxi	1.2751	2.6614	1.2480	2.2905
	Cycling	2.3755	4.2451	2.1541	4.4539
	Walk	1.7638	5.3892	1.3267	3.3116
Number of bikes in household	Cycling	0.3949	3.0123	0.4866	3.2936
Age	Exclusive ride-sourcing (ERS)	−0.0199	−1.4146	–	–
	Shared ride-sourcing (SRS)	−0.0563	−4.3659	−0.0514	−3.0544
	Taxi	−0.0278	−1.5953	−0.0595	−3.2079
	Walk	0.0133	1.5052	–	–
Household annual income > $100,000	Cycling	−1.5234	−2.2363	−1.2859	−2.2984
Perception of increased risk associated with using shared mobility during the pandemic	Public transit (PT)	−0.6328	−3.6921	−0.6140	−3.2000
	Exclusive ride-sourcing (ERS)	−0.4280	−2.0138	−0.3903	−1.6320
Feeling safe using ride-sourcing services during the pandemic	Public transit (PT)	0.7214	3.4587	0.5956	2.9309
	Exclusive ride-sourcing (ERS)	1.3857	3.8306	1.1495	3.4655
	Shared ride-sourcing (SRS)	1.5742	3.8604	1.2532	3.3577
	Taxi	0.2632	0.9076	0.5449	1.9394

Table 2 and Table 3. The estimated results of the structural model and measurement model components of the ICLV models for SiSTM-1 and SiSTM-2 are included in Table A2 and Table A3 in Appendix A.

The estimated coefficients of the SiSTM-1 ICLV models show that the LOS variables (i.e., travel time and cost) have negative signs, as expected. One latent attitudinal variable is found to be statistically significant for both the exclusive and shared ride-sourcing modes, which is "feeling safe using ride-sourcing services during the pandemic". This variable has a strong and positive effect on both ride-sourcing modes, which is reasonable because people who feel safe using ride-sourcing services are more likely to choose such modes. The significance of this latent attitudinal variable demonstrates the impact of attitudes and perceptions of risk on the decision to use ride-sourcing.

By comparing the same parameters between the during- and post-pandemic models, it is observed that the statistical significance of in-vehicle travel time is higher in the post-pandemic model than the during-pandemic model, whereas the opposite pattern is noted for the latent attitudinal variable. Such an observation suggests that the trip-makers may be more concerned about the perceived safety of ride-sourcing modes during the pandemic, whereas travel time becomes more relevant for post-pandemic mode choices. Nevertheless, the latent attitudinal variable is statistically significant at a 95% confidence level for both the during- and post-pandemic models, but in-vehicle travel time is only statistically significant at around the 70% confidence level in the during-pandemic model.

The SiSTM-2 ICLV models reveal the same behavioural patterns as the SiSTM-1 models. In addition, another latent attitudinal variable – "perception of increased risk associated with using shared mobility during the pandemic" is shown to be statistically significant for the exclusive ride-sourcing mode, especially in the during-pandemic model. This variable has a strong and negative effect on using exclusive ride-sourcing services, which is also intuitive. Exclusive ride-sourcing services can be considered as a shared mobility service, and trip-makers who believe there is more risk associated with using shared mobility are less likely to use them.

However, this latent attitudinal variable is found to be insignificant for the shared ride-sourcing mode.

6. Model interpretation and implications

Model interpretation derives the effects of factors influencing mode choice probabilities. In this section, two economic interpretation methods are presented: marginal effects of the variables and marginal rate of substitution between variables. This section highlights the similarities and differences between the interpretation of results from MTL models and ICLV models.

6.1. Marginal effects of level-of-service variables

The marginal effects are computed for two sets of determinants identified for the ride-sourcing modes: the LOS variables and the attitudinal variables. The marginal effect of a variable reflects the change in the probability of choosing an alternative given a unit change in the variable (Zhao et al., 2020). For the MTL models, the marginal effects are estimated as the partial derivative of the choice probabilities with respect to the variables, as proposed by Wang et al. (2020b). Table 5 presents the average marginal effects of the LOS variables with respect to exclusive (ERS) and shared ride-sourcing (SRS) modes across all samples in the testing set.

The directionality of effects between ICLV and MTL agreed. All the marginal effect values from the ICLV models are negative, as expected. Most of the marginal effect values from the MTL models are negative as well, albeit very few of them have positive effects. In particular, the magnitude of the effect of travel cost in SiSTM-1 is consistent between both methods. Albeit the overall trend aligns, the exact values of the marginal effects of in-vehicle travel time differ between MTL and ICLV models. This may be a result of the local irregularity issue that was discussed by Wang et al. (2020b). When the sample size is not large enough for a complex model, it is possible to observe counter-intuitive behavioural patterns. The negative signs are reasonable because people are less likely to choose a mode when its travel time and cost increase. In general, the interpretations of LOS variables from the MTL models agree with those from the ICLV models in terms of directional effect. The same conclusion was drawn by other studies as well, such as Wang et al. (2020b) and Zhao et al. (2020).

6.2. Marginal effects of attitudinal variables

Table 6 presents the average marginal effects of the attitudinal variables with respect to exclusive and shared ride-sourcing modes across all samples in the testing set. The interpretation of latent variables in the ICLV model is straightforward. The marginal effects of the latent attitudinal variables in the ICLV models have expected signs. The perception of increased risk of using shared mobility during the pandemic has negative effects on the probabilities of choosing exclusive ride-sourcing services, whereas the variable of feeling safe using ride-sourcing services during the pandemic has positive effects on the probabilities of choosing exclusive and shared ride-sourcing services. In terms of magnitude, it is observed that the marginal effects of the latent attitudinal variables are often much larger than those of the LOS variables. It appears that trip-makers may be more sensitive to a unit change in the perceived risk and safety of ride-sourcing services than a unit change in travel time or cost. This is a reasonable observation given the pandemic context. The perceived risk and safety associated with using ride-sourcing services during the pandemic are most likely derived from the fear of COVID-19 infection when being in contact with others. Given that such pandemic-related attitudes are relevant to the health of the trip-makers, it is not surprising that the latent attitudinal variables are more impactful on the mode choice probabilities than travel time and cost. However, it is worth noting that, since the perceived risk and safety of using ride-sourcing services are latent factors that cannot be directly observed, it is nontrivial to comprehend a unit change in a latent attitude, and it may be subjective to context.

Unlike LOS variables, it is difficult to draw the same conclusion that the effects of attitudes agreed between ICLV and MTL. In fact, it is difficult to draw meaningful interpretations from the MTL model. In general, the marginal effects of the same attitudinal variable in the MTL models vary across the two datasets and pandemic periods, making it difficult to conclude an overall behavioural pattern. The MTL (or data-driven approach) is trained to identify (hidden) patterns in the data. These models do not incorporate theoretical constructs, which are necessary for the interpretation of latent variables such as attitudes or perceptions. Latent variables are not directly observable, making them more abstract and complex than observable variables such as travel time and cost. The relationship between latent variables and the observable variables used to measure them can be complex and multidimensional, which makes interpretation more challenging. As such, their interpretation requires a theoretical understanding of the construct they represent and

Table 5
Marginal effects of LOS variables.

Variable	SiSTM-1				SiSTM-2			
	ERS (DP)	ERS (PP)	SRS (DP)	SRS (PP)	ERS (DP)	ERS (PP)	SRS (DP)	SRS (PP)
ICLV models								
In-vehicle travel time (mins)	−2.52E-05	−6.00E-04	−1.14E-05	−4.11E-04	−1.73E-04	−8.28E-04	−1.35E-04	−5.14E-04
Travel cost ($)	−1.72E-03	−5.39E-04	−7.81E-04	−3.69E-04	–	–	–	–
MTL models								
In-vehicle travel time (mins)	−2.46E-04	1.90E-04	−3.12E-04	−8.28E-05	−8.21E-04	−4.33E-04	1.19E-04	1.48E-04
Travel cost ($)	−1.12E-03	−4.44E-04	−5.51E-04	−1.83E-04	−2.75E-04	4.71E-05	−3.04E-04	−1.02E-04

*DP: during-pandemic, PP: post-pandemic, ERS: exclusive ride-sourcing, SRS: shared ride-sourcing.

Table 6
Marginal effects of attitudinal variables.

Variable	SiSTM-1				SiSTM-2			
	ERS (DP)	ERS (PP)	SRS (DP)	SRS (PP)	ERS (DP)	ERS (PP)	SRS (DP)	SRS (PP)
ICLV models								
Perception of increased risk of using shared mobility during-pandemic (LV1)	–	–	–	–	−2.17E-02	−2.32E-02	–	–
Feel safe using ride-sourcing services during-pandemic (LV2)	3.23E-02	2.46E-02	2.17E-02	2.52E-02	6.98E-02	6.84E-02	6.22E-02	4.63E-02
MTL models								
I believe there are more risks associated with leaving my home than before the pandemic (LV1*)	1.93E-03	1.22E-04	−1.51E-03	−3.55E-04	−2.50E-03	−1.87E-03	1.53E-04	7.79E-04
I believe there is more risk associated with using ride-sourcing services than before the pandemic (LV1*)	1.30E-03	4.68E-04	−1.13E-03	−1.29E-03	7.72E-04	1.32E-04	8.72E-04	1.30E-04
I believe there is more risk associated with using taxi services than before the pandemic (LV1*)	−1.20E-03	−6.33E-04	−1.29E-05	−8.05E-04	−2.32E-03	−1.48E-03	−6.05E-04	−5.85E-05
I believe there is more risk associated with carpooling than before the pandemic (LV1*)	1.31E-03	8.90E-04	−1.01E-04	−2.15E-05	−1.15E-03	−1.14E-03	−1.93E-04	8.12E-04
I believe there is more risk associated with using car-sharing services than before the pandemic (LV1*)	−1.95E-06	1.47E-03	−3.36E-04	−1.44E-03	3.45E-04	−1.89E-04	1.48E-04	5.62E-04
I believe there is more risk associated with using bicycle sharing services than before the pandemic (LV1*)	2.39E-03	−8.96E-04	−1.92E-03	6.19E-04	1.53E-03	−7.60E-04	−3.68E-04	1.42E-03
I am less willing to spend time outside of my home than I was before the pandemic (LV1*)	−3.07E-03	−4.13E-04	1.26E-03	3.34E-04	2.22E-03	4.29E-04	5.85E-04	9.68E-04
I prefer to stay away from others when I am travelling (LV1*)	2.13E-03	6.54E-04	−1.21E-03	−5.50E-05	5.85E-04	6.82E-04	−1.01E-04	5.43E-06
I would like to return to my daily routine once the current restrictions are lifted, even while COVID-19 is still considered a public health threat (LV2*)	−2.45E-03	−4.18E-06	1.26E-03	−5.72E-05	−1.79E-04	3.61E-05	3.63E-04	−9.14E-05
I will go back to my daily routine once the current restrictions are lifted, even while COVID-19 is still considered a public health threat (LV2*)	−1.97E-03	−8.73E-04	5.18E-04	1.96E-04	8.38E-04	1.22E-03	1.12E-03	−1.91E-04
I would feel safe using a ride-sourcing service while COVID-19 is still considered a public health threat (LV2*)	5.60E-04	−4.85E-04	1.18E-03	1.99E-03	−5.01E-04	−2.42E-05	1.62E-05	−2.98E-05
I would feel safe riding in a taxi while COVID-19 is still considered a public health threat (LV2*)	2.76E-03	−5.30E-04	1.02E-03	1.99E-03	−2.85E-04	−1.22E-03	5.68E-05	3.42E-04

*DP: during-pandemic, PP: post-pandemic, ERS: exclusive ride-sourcing, SRS: shared ride-sourcing.
The variables noted as LV1 in the MTL models are the same variables used in the measurement model of latent variable LV1 in the ICLV models.
The variables noted as LV2 in the MTL models are the same variables used in the measurement model of latent variable LV2 in the ICLV models.

the factors that influence it, namely, the casual effect. Therefore, interpreting latent variables can be more difficult than interpreting socioeconomic and cost variables. Modellers can exploit the power of MTL models allowing them to self-identify the relationship between attitudinal variables and achieve the highest prediction accuracy. However, meaningful variable interpretation on attitudinal variables is proved to be challenging through this study. There is a paradox between adding proper representation of latent variables in ML models and the underlying concept of the data-driven approach. However, this remains a challenge that must be addressed in order to successfully integrate data-driven approaches into the field of choice modelling.

6.3. Marginal rate of substitution between variables from ICLV models

The marginal rate of substitution is also important for quantifying variable effects and policy evaluation. They are computed as the ratio between the parameters of two variables of interest (Wang et al., 2020b). It reflects the value of time savings or willingness-to-pay for an improvement in choice marking context. In this study, the marginal rates of substitution between LOS variables and the latent attitudinal variables in the ICLV models are generated to analyze their trade-offs.

The marginal rate of substitution can be understood as the equivalent value of travel time or cost for the perceived risk or safety associated with using ride-sourcing services given the influence of the pandemic. For example, the marginal rate of substitution between "perception of increased risk associated with using shared mobility during the pandemic" and in-vehicle travel time can reflect the additional units of travel time needed to decrease one unit of the perceived risk. The marginal rate of substitution between "feeling safe using ride-sourcing services during the pandemic" and travel cost can reflect the equivalent decrease in travel cost given one unit increase in the perceived safety of ride-sourcing services is equivalent. The absolute number of the ratio is not meaningful because one unit of the latent variables does not have any intrinsic meaning. Instead, a comparison between the ratios from during-pandemic and post-pandemic scenarios could reflect the impact of the pandemic-provoked risk on travel behaviours.

Table 7 presents the results. Overall, the equivalent travel time of the latent attitudinal variables during the pandemic is much larger than in the post-pandemic period. The decrease in marginal rates of substitution between travel time and latent attitudinal variables in the post-pandemic model shows that the perceived risk and safety of using ride-sourcing services are less influential in ride-

Table 7
Marginal rate of substitution between variables in ICLV models.

		SiSTM-1				SiSTM-2			
Latent attitudinal variable	LOS variable	ERS (DP)	ERS (PP)	SRS (DP)	SRS (PP)	ERS (DP)	ERS (PP)	SRS (DP)	SRS (PP)
Perception of increased risk of using shared mobility during-pandemic	In-vehicle travel time	–	–	–	–	126	28	–	–
Feel safe using ride-sourcing services during-pandemic	In-vehicle travel time	−1 283	−41	−1899	−61	−408	−83	−463	−90
Feel safe using ride-sourcing services during-pandemic	Travel cost	−19	−46	−28	−68	–	–	–	–

*DP: during-pandemic, PP: post-pandemic.

sourcing mode choices after the pandemic.

For travel cost, the absolute value for the ratio is larger in the post-pandemic scenario compared to the during-pandemic. This means that post the pandemic, travellers demand a larger decrease in their travel costs in return for the same unit of increase in their perceived safety. In other words, they are willing to trade off less money for their perceived safety in the post-pandemic scenario compared to the during-pandemic scenario.

The above observation shows that pandemic-provoked attitudes were influential factors in the early stages of the COVID-19 pandemic. More importantly, their effects diminished as individuals believed they entered the post-pandemic era. The marginal rate of substitution in the SiSTM-2 during-pandemic model is much smaller than in the SiSTM-1 during-pandemic model. The SiSTM-2 survey was conducted in July 2021, one year later than the SiSTM-1. The modelling results indicate that trip-makers have started to pick up their trust in ride-sourcing services as the pandemic develops.

7. Discussion & conclusion

This study uses both data-driven (machine learning) and theory-driven (discrete choice) methods to examine the factors influencing the decision to use ride-sourcing services during and after the pandemic. Three models are developed in this study: the novel multi-task learning (MTL) deep neural network framework, the classical multi-layer perceptron (MLP) deep neural network framework and the integrated choice and latent variable (ICLV) model. The MTL models show the highest model performance in terms of prediction accuracies than MLP and ICLV. This demonstrates that the sophisticated architecture of multi-task learning deep neural networks supports and is suitable for leveraging shared information across multiple mode choices made by the same individual in various choice contexts. In this study, the shared information between travel mode choices is carefully induced through SP experimental design. Individuals were given the same hypothetical choice scenario asking them to make decisions considering the during-pandemic and post-pandemic context. The application of a similar SP design and the utilization of the MTL structure can be extended to other choice modelling studies aiming to examine variations in choice decisions for specific scenarios with a single controlling variable, such as investigating the impact of raising violence on the decision to use public transit.

In addition, this study evaluates the conformity and complementary of the modelling results between data-driven and theory-driven methods. The marginal effects of the LOS and attitudinal variables are computed for MTL and ICLV models. The results show that the directional effects of LOS variables are largely consistent between the two modelling methods. This showcases the consistency of economic information for observable variables between data-driven and theory-driven methods. It aligns with previous work from Wang et al. (2020b) and Zhao et al. (2020). However, the marginal effects of the latent attitudinal variables in the ICLV models exhibit more explainable behavioural patterns compared to those in the MTL models. The ICLV model demonstrates a significant decrease in individuals' perceived pandemic-provoked risk associated with using shared mobility in the post-pandemic context. On the other hand, meaningful interpretation of attitudinal variables cannot be extracted using the MTL model.

The difficulty of interpreting attitudinal variables from the MTL model emphasizes the challenge that needs to be addressed in order to achieve successful integration of data-driven methods in choice modelling. Evaluation of modelling results from this study raised the question of interpreting attitudinal variables in data-driven models. It has been overlooked by literature that compares two modelling paradigms (Wang et al., 2020b; Zhao et al., 2020). From this perspective, the ICLV approach serves as an invaluable and complementary counterpart to any existing data-driven approach. This study, together with previous literature such as the work of Vij and Walker (2016), highlights the usefulness of latent variables in choice modelling. ICLV models incorporate theoretical-backed latent constructs. This allows for the interpretation of the impact of latent variables on decision-making behaviour, providing a deeper understanding of the underlying mechanisms that drive choice-marking behaviour. It is particularly useful and essential in the pandemic context when unobservable latent variables dominate the choice process. Successfully integrating data-driven approaches in choice modelling should resolve the interpretability of attitudes. There exists a paradox in incorporating the representation of latent variables within machine learning models while adhering to the fundamental concept of data-driven approaches. This is a valid open research question for future efforts that integrate these two modelling paradigms.

Like any research, this study has the following limitations. The MTL model structure presented in this study was only applied to data collected from a specific SP design. The generalization of the MTL model structure to incorporate a broader range of SP designs would require further testing and investigation. Future studies are also recommended to fully investigate the effectiveness of including binary choice-set variables in DNN models to represent choice-set availability. Different datasets with various choice-set availability conditions can be utilized to test the performance and generalization of this method, in comparison with modifying the softmax function to restrict choice-set availability. Moreover, due to data limitations, this study only analyzes the mode choice decisions during- and post-pandemic periods. Future studies can also compare pre-pandemic mode choice decisions with the during- and post-pandemic results and explore the possibility of post-pandemic modal preference returning to the pre-pandemic situation.

Author statement

Yicong Liu: Conceptualization, Methodology, Formal analysis, Investigation, Software, Roles/Writing – original draft. **Patrick Loa:** Conceptualization, Data curation, Visualization, Roles/Writing – reviewing & editing. **Kaili Wang:** Conceptualization, Methodology, Investigation, Roles/Writing – reviewing & editing. **Khandker Nurul Habib:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors do not have any conflicts of interest to disclose.

Data availability

The data that has been used is confidential.

Acknowledgment

The study was funded by an NSERC Discovery Grant and Percy Edward Hart Professorship Grant. The authors bear sole responsibility for all results, interpretations, and comments made in the paper.

Appendix A. Structural Model and Measurement Model Components of the ICLV Models

The indicator variables used to identify the latent variables are summarized in Table A1 with their factor loadings. The estimated results of the structural model and measure model components of the ICLV models for SiSTM-1 and SiSTM-2 are included in Table A2 and Table A3, respectively.

Table A1
Definition of latent variables in the ICLV models

Latent Variable	Indicator Variables	Factor Loading
Perception of increased risk of using shared mobility during-pandemic	I believe there are more risks associated with leaving my home than before the pandemic	0.784
	I believe there is more risk associated with using ridesourcing services than before the pandemic	0.791
	I believe there is more risk associated with using taxi services than before the pandemic	0.795
	I believe there is more risk associated with carpooling than before the pandemic	0.844
	I believe there is more risk associated with using car-sharing services than before the pandemic	0.793
	I believe there is more risk associated with using bicycle sharing services than before the pandemic	0.542
	I am less willing to spend time outside of my home than I was before the pandemic	0.569
	I prefer to stay away from others when I am travelling	0.658
wish to maintain daily routine and feel safe using ride-sourcing during-pandemic	I would like to return to my daily routine once the current restrictions are lifted, even while COVID-19 is still considered a public health threat	0.637
	I will go back to my daily routine once the current restrictions are lifted, even while COVID-19 is still considered a public health threat	0.716
	I would feel safe using a ridesourcing service while COVID-19 is still considered a public health threat	0.876
	I would feel safe riding in a taxi while COVID-19 is still considered a public health threat	0.832

Table A2
Structural and measurement model components of SiSTM-1 ICLV models

	SiSTM-1 (during-pandemic)		SiSTM-1 (post-pandemic)	
Variable	Coefficient	t-stat	Coefficient	t-stat
Structural model for latent variable "Perception of increased risk and concern of using ride-sourcing services post-pandemic"				
Household size	–	–	−0.1449	−4.2517

(continued on next page)

Table A2 (*continued*)

Variable	SiSTM-1 (during-pandemic) Coefficient	SiSTM-1 (during-pandemic) t-stat	SiSTM-1 (post-pandemic) Coefficient	SiSTM-1 (post-pandemic) t-stat
Have a driver's licence	–	–	0.2969	1.8977
Structural model for latent variable "Feel safe using ridesourcing during pandemic"				
Household annual income >80,000	−0.1049	−1.1179	–	–
Age (log)	−0.1072	−4.4737	−0.0552	−1.0396
Gender: Male	0.2364	2.4909	–	
Part-time employed	0.5854	3.6731	–	–
Full-time employed	0.4687	4.8643	0.3671	–
Never used ride-sourcing services pre-pandemic	−0.1342	−1.2031	−0.3495	−1.5733
Used ride-sourcing services more than once a week pre-pandemic	0.5719	4.1932	0.4597	1.5267
Being a student	0.6743	5.4294	0.5333	1.8503
Measurement model for latent variable "Perception of increased risk of using shared mobility during-pandemic"				
I believe there are more risks associated with leaving my home than before the pandemic	0.8069	14.0336	0.9360	10.7974
I believe there is more risk associated with using ridesourcing services than before the pandemic	0.9383	20.1558	1.0956	13.7123
I believe there is more risk associated with using taxi services than before the pandemic	1.0023	22.9698	1.1661	14.3490
I believe there is more risk associated with carpooling than before the pandemic	0.9668	23.0612	1.1112	16.0868
I believe there is more risk associated with using car-sharing services than before the pandemic	0.8882	17.5263	1.0406	13.7008
I believe there is more risk associated with using bicycle sharing services than before the pandemic	0.7691	13.7134	0.9068	11.8588
I am less willing to spend time outside of my home than I was before the pandemic	0.6486	9.9780	0.7533	8.8815
I prefer to stay away from others when I am travelling	0.5523	7.7618	0.6488	7.7544
Measurement model for latent variable "Wish to maintain daily routine and feel safe using ridesourcing services during-pandemic"				
I would like to return to my daily routine once the current restrictions are lifted, even while COVID-19 is still considered a public health threat	0.8151	15.5396	0.8368	8.2788
I will go back to my daily routine once the current restrictions are lifted, even while COVID-19 is still considered a public health threat	0.8193	15.6372	0.8288	7.6968
I would feel safe using a ridesourcing service while COVID-19 is still considered a public health threat	0.9878	27.8102	0.9269	13.5257
I would feel safe riding in a taxi while COVID-19 is still considered a public health threat	0.9973	25.4734	0.9421	11.2885

Table A3
Structural and measurement model components of SiSTM-2 ICLV models

Variable	SiSTM-2 (during-pandemic) Coefficient	SiSTM-2 (during-pandemic) t-stat	SiSTM-2 (post-pandemic) Coefficient	SiSTM-2 (post-pandemic) t-stat
Structural model for latent variable "Perception of increased risk of using shared mobility during-pandemic"				
Household annual income >80,000	0.2210	1.8926	0.1316	0.9141
Age (log)	−0.0377	−2.0493	−0.2062	−2.7050
Structural model for latent variable "Feel safe using ridesourcing during pandemic"				
Full-time employed	0.5476	4.5044	0.4202	4.0404
Never used ride-sourcing services pre-pandemic	−0.5726	−5.7732	−0.4861	−5.1641
Used ride-sourcing services more than once a week pre-pandemic	–	–	0.6041	6.0339
Being a student	0.1032	1.0365	–	–
Measurement model for latent variable "Perception of increased risk of using shared mobility during-pandemic"				
I believe there are more risks associated with leaving my home than before the pandemic	0.8801	16.2422	0.8870	15.0788
I believe there is more risk associated with using ridesourcing services than before the pandemic	0.9683	20.5434	0.9791	18.7862
I believe there is more risk associated with using taxi services than before the pandemic	0.9819	19.3278	0.9918	17.7333
I believe there is more risk associated with carpooling than before the pandemic	0.9517	17.4572	0.9635	16.6462
I believe there is more risk associated with using car-sharing services than before the pandemic	0.9796	18.0999	0.9844	16.7425
I believe there is more risk associated with using bicycle sharing services than before the pandemic	0.7185	10.2469	0.7217	10.0238
I am less willing to spend time outside of my home than I was before the pandemic	0.7493	11.7645	0.7577	11.1917
I prefer to stay away from others when I am travelling	0.7614	11.7056	0.7674	11.0500
Measurement model for latent variable "Wish to maintain daily routine and feel safe using ridesourcing services during-pandemic"				
I would like to return to my daily routine once the current restrictions are lifted, even while COVID-19 is still considered a public health threat	0.6576	7.6192	0.6596	8.8671
I will go back to my daily routine once the current restrictions are lifted, even while COVID-19 is still considered a public health threat	0.7796	9.1189	0.7832	10.8101
I would feel safe using a ridesourcing service while COVID-19 is still considered a public health threat	1.0755	9.9291	1.1043	14.8550
I would feel safe riding in a taxi while COVID-19 is still considered a public health threat	1.0293	9.4879	1.0422	13.4367

References

Ben-Akiva, M., Walker, J., Bernardino, A.T., Gopinath, D.A., Morikawa, T., Polydoropoulou, A., 2002. Integration of choice and latent variable models. In: Perpetual Motion. Elsevier, pp. 431–470. https://doi.org/10.1016/B978-008044044-6/50022-X.

Bensaoud, A., Kalita, J., 2022. Deep multi-task learning for malware image classification. J. Inf. Secur. Appl. 64, 103057 https://doi.org/10.1016/j.jisa.2021.103057.

Bergstra, J., Bengio, Y., 2012. Random search for hyper-parameter optimization. J. Mach. Learn. Res. 13, 281–305.

Bogage, J., 2020. Uber Rides Are Down 70 Percent in Some Markets amid Coronavirus Pandemic [WWW Document]. Wash. Post. URL. https://www.washingtonpost.com/technology/2020/03/19/uber-coronavirus-analyst-call/ (accessed 7.September.22).

Brown, A.E., Williams, R., 2021. Equity implications of ride-hail travel during COVID-19 in California. Transp. Res. Rec. J. Transp. Res. Board, 036119812110372. https://doi.org/10.1177/03611981211037246.

Budd, L., Ison, S., 2020. Responsible Transport: a post-COVID agenda for transport policy and practice. Transp. Res. Interdiscip. Perspect. 6, 100151 https://doi.org/10.1016/j.trip.2020.100151.

Caruana, R., 2004. Multitask learning. Mach. Learn. https://doi.org/10.1023/A:1007379606734.

Chen, S., Zhang, Y., Yang, Q., 2021. Multi-task learning in natural language processing: an overview. https://doi.org/10.48550/arXiv.2109.09138.

City of Toronto, 2021. The Transportation Impacts of Vehicle-For-Hire in the City of Toronto. October 2018 to July 2021. Toronto.

Clewlow, R.R., Mishra, G.S., 2017. Disruptive transportation: the adoption, utilization, and impacts of ride-hailing in the United States. UC Davis Inst. Transp. Stud. 44, 1307–1323. https://doi.org/10.1139/gen-44-3-401.

Costa, C.S., Pitombo, C.S., Souza, F.L.U. de, 2022. Travel behavior before and during the COVID-19 pandemic in Brazil: mobility changes and transport policies for a sustainable transportation system in the post-pandemic period. Sustainability 14, 4573. https://doi.org/10.3390/su14084573.

de Haas, M., Faber, R., Hamersma, M., 2020. How COVID-19 and the Dutch 'intelligent lockdown' change activities, work and travel behaviour: evidence from longitudinal data in The Netherlands. Transp. Res. Interdiscip. Perspect. 6, 100150 https://doi.org/10.1016/j.trip.2020.100150.

De Vos, J., 2020. The effect of COVID-19 and subsequent social distancing on travel behavior. Transp. Res. Interdiscip. Perspect. 5, 100121 https://doi.org/10.1016/j.trip.2020.100121.

Dzisi, E.K.J., Obeng-Atuah, D., Ackaah, W., Tuffour, A.Y., Aidoo, N.E., 2021. Uptake in on-demand ride-hailing for intracity long distance trip making during COVID-19. Urban Plan. Transp. Res. 1–12 https://doi.org/10.1080/21650020.2021.1872415.

Erhardt, G.D., Roy, S., Cooper, D., Sana, B., Chen, M., Castiglione, J., 2019. Do transportation network companies decrease or increase congestion? Sci. Adv. 5 https://doi.org/10.1126/sciadv.aau2670.

Golshani, N., Shabanpour, R., Mahmoudifard, S.M., Derrible, S., Mohammadian, A., 2018. Modeling travel mode and timing decisions: comparison of artificial neural networks and copula-based joint model. Travel Behav. Soc. 10, 21–32. https://doi.org/10.1016/j.tbs.2017.09.003.

Hensher, D.A., Ton, T.T., 2000. A comparison of the predictive potential of artificial neural networks and nested logit models for commuter mode choice. Transplant. Res. Part E Logist. Transp. Rev. 36, 155–172. https://doi.org/10.1016/S1366-5545(99)00030-7.

Kamargianni, M., Dubey, S., Polydoropoulou, A., Bhat, C., 2015. Investigating the subjective and objective factors influencing teenagers' school travel mode choice – an integrated choice and latent variable model. Transplant. Res. Part Policy Pract. 78, 473–488. https://doi.org/10.1016/j.tra.2015.06.011.

Khamis, A., 2021. Smart mobility during the COVID-19 pandemic and in the post-pandemic world. In: Smart Mobility, pp. 131–146. https://doi.org/10.1007/978-1-4842-7101-8_6.

Kingma, D.P., Ba, J., 2015. Adam: A Method for Stochastic Optimization. ICLR.

Kong, H., Zhang, X., Zhao, J., 2020. How does ridesourcing substitute for public transit? A geospatial perspective in Chengdu, China. J. Transport Geogr. 86, 102769 https://doi.org/10.1016/j.jtrangeo.2020.102769.

Kuang, Z., Li, Z., Zhao, T., Fan, J., 2017. Deep multi-task learning for large-scale image classification. In: 2017 IEEE Third International Conference on Multimedia Big Data (BigMM). Presented at the 2017 IEEE Third International Conference on Multimedia Big Data (BigMM), pp. 310–317. https://doi.org/10.1109/BigMM.2017.72.

Lee, A., 2020. Uber and Lyft Suspend Pool Rides in U.S.. Canada to reduce spread of coronavirus [WWW Document] CTV News. https://www.ctvnews.ca/health/coronavirus/uber-and-lyft-suspend-pool-rides-in-u-s-canada-to-reduce-spread-of-coronavirus-1.4856507 (accessed 7.September.22).

Lee, D., Derrible, S., Pereira, F.C., 2018. Comparison of four types of artificial neural network and a multinomial logit model for travel mode choice modeling. Transp. Res. Rec. J. Transp. Res. Board 2672, 101–112. https://doi.org/10.1177/0361198118796971.

Liu, Y., Miller, E., Habib, K.N., 2021. Detecting transportation modes using smartphone data and GIS information: evaluating alternative algorithms for an integrated smartphone-based travel diary imputation. Transp. Lett. 0, 1–11. https://doi.org/10.1080/19427867.2021.1958591.

Loa, P., Habib, K.N., 2021. Examining the influence of attitudinal factors on the use of ride-hailing services in Toronto. Transplant. Res. Part A 146, 13–28. https://doi.org/10.1016/j.tra.2021.02.002.

Loa, P., Hossain, S., Liu, Y., Habib, K.N., 2021a. How has the COVID-19 pandemic affected the use of ride-sourcing services? An empirical evidence-based investigation for the greater Toronto area. Transplant. Res. Part Policy Pract. 155, 46–62. https://doi.org/10.1016/j.tra.2021.11.013.

Loa, P., Hossain, S., Liu, Y., Habib, K.N., Nurul Habib, K., 2021b. How have ride-sourcing users adapted to the first wave of the COVID-19 pandemic? Evidence from a survey-based study of the Greater Toronto Area. Transp. Lett. 13, 404–413. https://doi.org/10.1080/19427867.2021.1892938.

Loa, P., Hossain, S., Liu, Y., Mashrur, SkMd, Ong, F., Habib, K.N., 2021c. Exploring the Impacts of COVID-19 on Ride-Sourcing Use in the Greater Toronto Area: Results from the Second Cycle of the SiSTM Satellite Survey (Toronto).

Loa, P., Mashrur, SkMd, Nurul Habib, K., 2021d. What influences the substitution of ride-sourcing for public transit and taxi services in Toronto? An exploratory structural equation model-based study. Int. J. Sustain. Transp. 0, 1–14. https://doi.org/10.1080/15568318.2021.1978018.

Ma, Y., Zhang, Z., 2020. Travel mode choice prediction using deep neural networks with entity embeddings. IEEE Access 8, 64959–64970. https://doi.org/10.1109/ACCESS.2020.2985542.

Marsden, G., Anable, J., Docherty, I., Brown, L., 2021. At a Crossroads: Travel Adaptations during Covid-19 Restrictions and where Next (Oxford).

Mashrur, SkMd, Wang, K., Loa, P., Hossain, S., Nurul Habib, K., 2022. Application of protection motivation theory to quantify the impact of pandemic fear on anticipated postpandemic transit usage, 036119812110654 Transp. Res. Rec. J. Transp. Res. Board. https://doi.org/10.1177/03611981211065439.

Masters, D., Luschi, C., 2018. Revisiting small batch training for deep neural networks. https://doi.org/10.48550/arXiv.1804.07612.

Matson, G., McElroy, S., Lee, Y., Circella, G., 2021. Longitudinal Analysis of COVID-19 Impacts on Mobility: an Early Snapshot of the Emerging Changes in Travel Behavior (Davis).

McFadden, D., 1986. The choice theory approach to market research. Market. Sci. 5, 275–297. https://doi.org/10.1287/mksc.5.4.275.

Moons, E., Wets, G., Aerts, M., 2007. Nonlinear models for determining mode choice. In: Neves, J., Santos, M.F., Machado, J.M. (Eds.), Progress in Artificial Intelligence, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 183–194. https://doi.org/10.1007/978-3-540-77002-2_16.

Nair, V., Hinton, G.E., 2019. Rectified Linear Units Improve Restricted Boltzmann Machines. Presented at the ICML.

Nam, D., Kim, H., Cho, J., Jayakrishnan, R., 2017. A Model Based on Deep Learning for Predicting Travel Mode Choice. In: Proceedings of the transportation research board 96th annual meeting transportation research board, Washington, DC, USA, pp. 8–12.

Ozbilen, B., Slagle, K.M., Akar, G., 2021. Perceived risk of infection while traveling during the COVID-19 pandemic: insights from Columbus, OH. Transp. Res. Interdiscip. Perspect. 10, 100326 https://doi.org/10.1016/j.trip.2021.100326.

Rahimi, E., Shabanpour, R., Shamshiripour, A., Kouros, Mohammadian, A., 2021. Perceived risk of using shared mobility services during the COVID-19 pandemic. Transport. Res. F Traffic Psychol. Behav. 81, 271–281. https://doi.org/10.1016/j.trf.2021.06.012.

Shamshiripour, A., Rahimi, E., Shabanpour, R., Mohammadian, A., Kouros), 2020. How is COVID-19 reshaping activity-travel behavior? Evidence from a comprehensive survey in Chicago. Transp. Res. Interdiscip. Perspect. 7, 100216 https://doi.org/10.1016/j.trip.2020.100216.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15, 1929–1958.

Tirachini, A., 2020. Ride-hailing, travel behaviour and sustainable mobility: an international review. Transportation 47, 2011–2047. https://doi.org/10.1007/s11116-019-10070-2.

Train, K.E., McFadden, D.L., Goett, A.A., 1987. Consumer attitudes and voluntary rate schedules for public utilities. Rev. Econ. Stat. 69, 383–391.

Van Cranenburgh, S., Wang, S., Vij, A., Pereira, F., Walker, J., 2022. Choice modelling in the age of machine learning - discussion paper. J. Choice Model. 42, 100340 https://doi.org/10.1016/j.jocm.2021.100340.

Vij, A., Walker, J.L., 2016. How, when and why integrated choice and latent variable models are latently useful. Transp. Res. Part B Methodol. 90, 192–217. https://doi.org/10.1016/j.trb.2016.04.021.
Vuchic, V.R., 2005. Urban Transit: Operations, Planning, and Economics. John Wiley & Sons, Inc., Hoboken.
Wang, S., Wang, Q., Zhao, J., 2020a. Multitask learning deep neural networks to combine revealed and stated preference data. J. Choice Model. 37, 100236 https://doi.org/10.1016/j.jocm.2020.100236.
Wang, S., Wang, Q., Zhao, J., 2020b. Deep neural networks for choice analysis: extracting complete economic information for interpretation. Transport. Res. C Emerg. Technol. 118, 102701 https://doi.org/10.1016/j.trc.2020.102701.
Weiss, A., Salehin, M.F., Habib, K.N., 2019. A Joint RP-Off_SP Survey to Understand the Impacts of Autonomous Vehicle on Travel Mode Choices in the Greater Toronto Area. Transp. Res. Board 98th Annu. Meet. Wash. DC.
Yang, H., Zhai, G., Yang, L., Xie, K., 2022. How does the suspension of ride-sourcing affect the transportation system and environment? Transplant. Res. Part Transp. Environ. 102, 103131 https://doi.org/10.1016/j.trd.2021.103131.
Zhang, Y., Yang, Q., 2018. An overview of multi-task learning. Natl. Sci. Rev. 5, 30–43. https://doi.org/10.1093/nsr/nwx105.
Zhang, Z., Ji, C., Wang, Y., Yang, Y., 2020. A customized deep neural network approach to investigate travel mode choice with interpretable utility information. J. Adv. Transport. 2020, e5364252 https://doi.org/10.1155/2020/5364252.
Zhao, X., Yan, X., Yu, A., Van Hentenryck, P., 2020. Prediction and behavioral analysis of travel mode choice: a comparison of machine learning and logit models. Travel Behav. Soc. 20, 22–35. https://doi.org/10.1016/j.tbs.2020.02.003.