land

MDPI

Article

How Rail Transit Makes a Difference in People's Multimodal Travel Behaviours: An Analysis with the XGBoost Method

Lixun Liu [1,2], Yujiang Wang [3,*] and Robin Hickman [4]

[1] School of Architecture and Planning, Hunan University, Changsha 410082, China; liulixun@hnu.edu.cn
[2] Hunan Key Laboratory of Sciences of Urban and Rural Human Settlements in Hilly Areas, Hunan University, Changsha 410082, China
[3] Department of Engineering Science, University of Oxford, Oxford OX1 2JD, UK
[4] Bartlett School of Planning, University College London, London WC1E 6BT, UK; r.hickman@ucl.ac.uk
* Correspondence: yujiang.wang@eng.ox.ac.uk

Abstract: The rail transit system was developed in Chinese large cities to achieve more efficient and sustainable transport development. However, the extent to which the newly built rail transit system can facilitate people's multimodality still lacks evidence, and limited research examines the interrelationship between trip stages within a single trip. This study aims to explore the interrelations between trip stage characteristics, socio-demographic attributes, and the built environment. It examines how rail transit is integrated as part of multimodal trips after it is introduced. The data are extracted from the Chongqing Urban Resident Travel Survey from 2014, three years after the new rail transit network was established. It applies an XGBoost model to examine the non-linear effect. As a result, the separate trip stage characteristics have more of an impact than the general trip characteristics. The non-linear effects revealed by the machine learning model show changing effects and thresholds of impact by trip stage characteristics on people's main mode choice of rail transit. An optimal radius of facility distribution along the transit lines is suggested accordingly. Synergistic effects between variables are identified, including by groups of people and land use characteristics.

Keywords: Urban mobility; multimodality; rail transit; travel behaviour; travel mode choice; machine learning

check for updates

Citation: Liu, L.; Wang, Y.; Hickman, R. How Rail Transit Makes a Difference in People's Multimodal Travel Behaviours: An Analysis with the XGBoost Method. *Land* **2023**, *12*, 675. https://doi.org/10.3390/land12030675

Academic Editor: Thomas W. Sanchez

Received: 12 February 2023
Revised: 7 March 2023
Accepted: 11 March 2023
Published: 13 March 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the aim of achieving sustainable transport, rail transit systems were developed as an important contribution to achieve more effective transport to accommodate mass travel demand and move towards less car-dependent lifestyles. A prevailing part of previous research focused on investigating the potential of rail transit to substitute motorized vehicles [1,2]. However, in order to improve sustainable transport, the integration of existing transport facilities with urban transit systems should play an increasing role in the future, especially for the so-called 'multimodal' trip chain [3]. Multimodality, which is defined as the flexible use of various modes of transport within a certain time period [4], received increasing research attention in recent years [5].

The aim of the paper is to examine the specific contribution of rail transit to the multimodal travel behaviours. It requires an assessment of each stage of the trip to increase the understanding of how travel patterns are composed in diversified ways. Therefore, it is useful to distinguish the main trip from the other trip stages, such as the starting and ending trip, often known as the 'first mile' and 'last mile' (irrespective of actual distance). Little previous research quantitatively examines the interrelations of different trip stages within a single trip. The paper examines the different trip stages in relation to how the main trip mode is impacted by other interrelated trip characteristics, in which the role of rail transit in the multimodal trip chain is identified.

In order to achieve the research aim, the paper addresses the following questions: First, how the different trip stages are interrelated and contribute to the main trip, which is further elaborated and explained in relation to socio-demographics and urban settings. Second, to examine how rail transit combines with other travel modes and forms part of the multimodal patterns of daily transport behaviours. Third, how the impact varies in a non-linear relationship with the variables in order to give evidence on planning and policies, which are aimed at promoting rail transit use and multimodal transport. The analysis utilizes data extracted from the Chongqing Urban Household Official Travel Survey, 2014. The data were collected three years after the formation of the early stage of the rail transit network in the city. XGBoost is utilized to explore the non-linear relationship of the impact variables with the main trip mode choice. In previous studies, logistic regression was widely used in category prediction of travel mode choice [6–9]. In this study, it is used as a baseline to compare the model performance and explanatory power with that of XGBoost and other machine learning models.

This study hence contributes to the literature in three ways: (1) Concerning multimodal trips, it explores how the trip stages are interrelated and the main trip mode choice is impacted by other trip stage characteristics. The separate trip stages contribute more accuracy to the prediction of main trip mode choice than general trip features; (2) the synergistic effects between variables show how rail transit is embedded into people's multimodal trips after a new rail transit network is established; and (3) the non-linear relationships between travel mode choice and its correlates, revealed by the machine learning models, present varying predictions with changes of the variables, which contribute to potential spatial planning requirements and can be reflected in spatial strategies.

The remainder of the paper is organized as follows: The following section discusses the emerging literature on multimodal travels. This is followed by the applied methodology, including data and the XGBoost method. The result section describes the nonlinear and synergistic effects of variables. The conclusions discuss the results and give recommendations for planning practice.

2. Literature Review

Recent research indicated a growing interest in analysing the variability of behaviour of the same people, showing a shift from studying behaviour differences between people [10,11]. Despite being a concept used for a long time among scholars and practitioners, multimodality is currently receiving a renewed interest [12,13]. One part of the research contributes to correlating social or built environment attributes with multimodality, explaining which groups of people are more likely to be multimodal, why people tend to be multimodal, and what people experience when they become multimodal [3,14,15]. Further research provides useful insights into how the multimodal level changes with people's life course, such as through moving house or other key events [12,16,17].

2.1. Identify the Role of Rail Transit in Multimodal Travel

There are generally two ways to measure the level of multimodality. One way is to classify individuals' multimodal behaviour into nominal categories. This is often based on the combination of transport modes they use [3,16] or the perception and attitudes they hold [18]. In this way, the role of a certain kind of transport is exhibited as a single category or a hybrid one. Molin [3] identified five multimodal travel groups based on the frequency of mode use in combination, in which the role of rail transit is integrated into the three groups of car multimodal, bike multimodal, and public transport. The probability of belonging to a specific multimodal travel group is predicted by the attitudes of travel they hold. Olafsson [18] combines travel mode and travel purpose together to form five clusters in order to identify the role of cycling in multimodal transport behaviours.

The second way is measuring multimodality quantitatively by calculating statistical indices. In this way, the choice of a specific transport mode is integrated into the calculation of a certain index indiscriminately with other modes. The number of modes/stage [19], the

difference in percentage of use between primary and secondary modes [20], Herfindahl–Hirschman index (HHI) [20], Shannon's entropy [16,21], etc., are all indices to measure multimodality. The number of modes and frequency, as indicators which describe the extent of variability in mode use, are mostly used to calculate above the indices [16,22]. For example, Heinen [22] finds that only when controlling for trip distances, the indices of multimodality are associated with lower CO2 emissions.

However, previous studies seldom look into a single trip and quantitatively examine the interrelationship among the travel behaviours of different trip stages. It is useful to study the characteristics of separate trip stages by distinguishing the main trip from the other trip stages. In particular, how the characteristics of the starting trip and ending trip influence the interconnected main trip travel behaviours is inadequately studied. Furthermore, as an important public investment towards solving traffic congestion and promoting sustainable transport [23], some research focused on investigating the potential of rail transit to substitute motorized vehicles in some circumstances [24]. However, the extent to which the newly built rail transit system can perform as a part and integrate into people's travel chains needs further research. How it facilitates people's multimodal travel behaviour formation still lacks evidence.

2.2. Predictors of Multimodality

Previous research demonstrated that multimodality is related to several socio-demographic attributes and built environment elements, such as age, gender, car availability, public transport access or density [19,20]. Among socio-demographic attributes, the most varying factor is age. Some studies find that multimodality is disproportionately high among adolescents as well as seniors, especially if retired [4], indicating a U-shaped distribution [3,20]. Meanwhile, the middle-aged segment is strongly dependent on labour and family related constraints, and therefore relies more often solely on the private car [12], depending of course on the context. In a typical continental European urban context, 'older people' are found to comprise a segment characterized by multimodal transport use and a less car-dependent lifestyle [25]. While in the US context, higher age tends to a higher single-mode usage [26].

Another factor, which shows an ambivalent result, is income. Whereas some studies find that higher income leads to more multimodality [26], others report the opposite [27] or no significant relation between income and combined mode choice at all [12]. It is worth noting that, following the literature, the availability and ownership of specific travel modes is critical for the formation of certain multimodality, especially that access to a car leads to monomodal car use in most cases [20]. In terms of the impact from employment, it was shown that fully employed people are less multimodal than part-time workers or unemployed persons [4,16,26]. An increase in the number of children also makes a shift to more multimodality unlikely [3,4,12].

In terms of physical conditions, an improvement to the public transport system in the neighbourhood increases multimodality, and vice versa. Reduced parking space availability also increases multimodality [16]. Meanwhile, extensive research showed that urban land use characteristics are significantly interrelated with travel behaviour [28–30], and urban settings, such as adjacent to the centre, high population density, and land use mix, contribute to a higher level of multimodality [12,31,32]. Zhao [33] employed a deep neural network to identify regional land use characteristics and quantify land use intensity using ridership data of bicycle sharing. A deep neural network is established and trained based on the processed ridership features and land use labels.

2.3. Non-Linear Relationships in Multimodal Travels

In the previous literature, it is assumed that the predictors and travel behaviours follow a pre-defined or parametric relationship. More recent studies began to explore the non-linear relationships between built environment variables and travel behaviours [34–37]. Most statistical models predetermine a model structure that requires the input data to satisfy as-

sumptions, such as random utility maximization theory [38], while many machine learning methods rely on computers to probe the data for its structure. The machine learning models allow for forming more flexible model structures to reduce the model's incompatibility with the empirical data, which can often lead to higher predictive capability [39]. For example, Cheng [40] evaluated the relative importance of explanatory variables of the random forest model to help provide important insights into formulating transport policies. Liu utilized XGBoost to explore the non-linear association between the built environment and active travel for walking and shopping at both origins and destinations [41]. Zhao [39] carried out a comprehensive comparison of the predictive performance and travel behavioural outputs of logit models with those of machine learning models. The best performing machine learning model, random forest, has significantly higher predictive accuracy than multinomial logit and mixed logit models.

Finding the right range or thresholds of variable impact can be cost-effective [42]. Scholars exploited the explanatory power of the non-linear model in finding the changing impact with the change in a certain range of the variables, which can be especially useful in providing advice on planning implications [43–45]. Yue and Ma utilized a transformer prediction model to estimate transfer passenger flow and demonstrate the performance of deep learning models in multimodal transportation forecasting [46]. However, still few studies explored the non-linear effect on specific multimodal travel behaviours in relation to social demographic attributes. The limitation of research in giving a more detailed inspection on the varying and nuanced impact of the interrelated trip stages, especially the starting and ending trip, on people's main mode choice, leads to a lack of clear evidence to direct planning implication.

This study aims to address the following research gaps: First, it aims to examine the interrelations among different trip stages within the multimodal trip itself to provide evidence on how the main mode choice is influenced by the starting and ending trip. Second, it investigates how the rail transit mode is embedded into the multimodal trips as a part of, and whether it facilitates, people's multimodal travel. Thirdly, non-linear relationships of the variables with the main trip mode are explored with machine learning methods, as well as synergic effects between travel characteristics and the socio-demographic and land use attributes. It aims to reveal the varying influence hid behind the averaged coefficients provided by the commonly used statistical methods. It helps to find the targeted demographic groups and urban settings, aiming to provide more concise and practical suggestions to planning implications and policies.

3. Methodology

3.1. Data

3.1.1. Survey Method

Data are extracted from Chongqing Urban Resident Travel Survey, 2014, which took place three years after a 4-line network of rail transit was established. It is officially carried out by the Planning Bureau of Chongqing as part of the nation-wide Resident Travel Survey in 2014. A stratified probability method is used to sample. In the survey, the central urban area of Chongqing was divided into 25 transport zones (Figure 1). The sample size of each zone was based on its population size, representing a 1% sampling of the total population. The whole survey contained a sample of 80,000 persons, in 28,000 households, in the main city region. For this study, data were extracted for analysis from zone 1 and zone 2, and zone 10 and zone 11, to represent the central area of the city. The total sample size is 5110, belonging to 1926 households, and contains information of 11,729 trips in total. Data were collected by staff with an electronic tablet equipped with specially designed GPS software to record people's location. Respondents were asked to report their commuting information on an ordinary weekday. Information collected in the survey contains socio-demographic attributes, travel information, and the original and destination land use characteristics, which is explained in the following section.

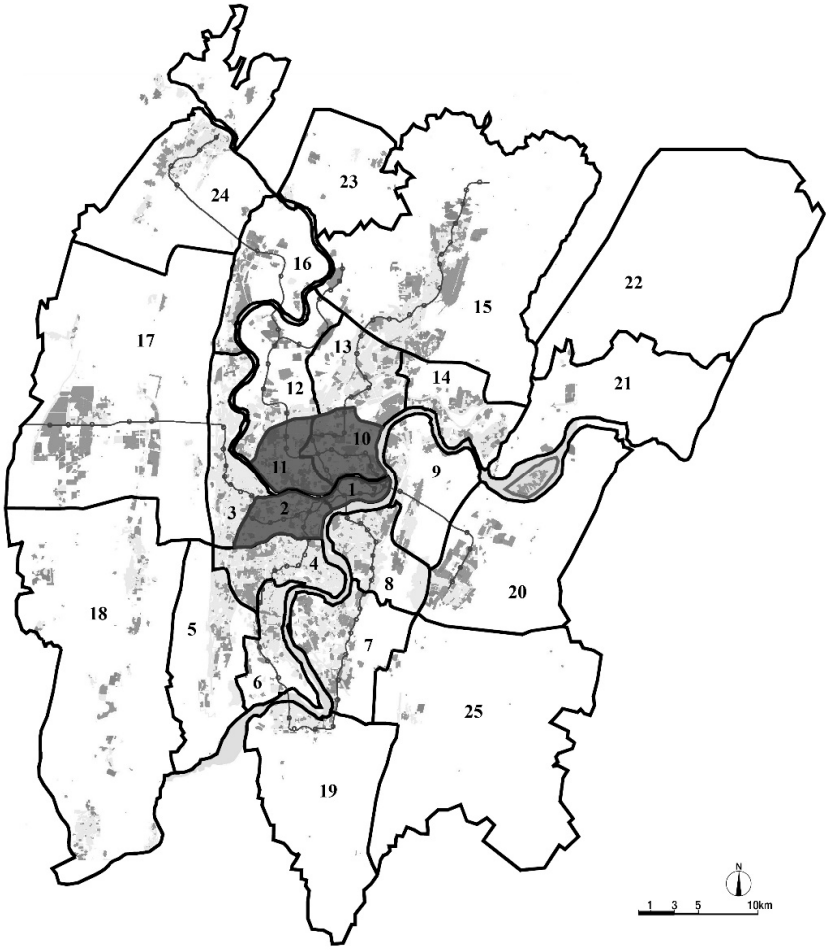


Figure 1. Transport zones for the Urban Household Travel Survey (by authors).

3.1.2. Variables and Predictors

Table 1 presents the variables used in the study. A total of 11,664 trips are valid with complete travel information for different trip stages. For this study, the dependent variable is their main travel mode choice. Independent variables are categorized into trip characteristics, socio-demographic attributes, and built environment elements. In the survey, respondents were asked to record their trip stages within one trip. The travel-related information they recorded includes the starting time, arriving time, total duration, distance, travel purpose, travel mode, and duration of each trip stage. The questionnaire provided five trip stages to fill, while nearly most of the respondents (98%) only have three trip stages. Therefore, in convenience for comparison, the trip stage of the longest duration is identified as the main trip. The trips before and after the main trip are considered as the starting and ending trip. Therefore, we kept the three trips for analyses. If both of the starting and ending trips are vacant, the trip is considered as unimodal. There are 7244 unimodal trips, and the remaining 4400 cases are considered as multimodal trips. We then categorize the exact starting and arriving time to five different time slots. The total travel distance is calculated using Euclidian distance with their XY coordinates. These data are recalled from memory, so the total time duration is slightly different from the added time of each trip stages. As income amount is difficult to collect directly from people, car ownership, car consumption plan (aspiration to buy a car), residential property level, and parking place are collected as substitutes of income. As shown in the descriptive statistical table, the gender percentage of male is slightly lower than the citywide census data (50.55% male, 49.45% female). Built environment elements contain land use characteristics of both a starting and arriving point. It is inferred by relating the geographic coordinates and the land use codes of planning.

Table 1. Definitions and descriptive characteristics of variables. Variable description and descriptive statistics (N = 11,664).

Variable	Description	Category	%	Min	Max	Mean	SD
Dependent Variable							
MA_mode	Travel mode of the main trip	0 = walk	55.2%				
		1 = bus	23.6%				
		2 = rail transit	6.2%				
		3 = private vehicles	15.0%				
Independent Variables							
Trip characteristics							
TT	Travel time of the whole trip (min)			2	630	31.9	22.3
MA_time	Travel time of the main trip (min)			2	240	22.4	15.6
ST_time	Travel time of the starting trip (min)			0	60	3.1	5.0
ED_time	Travel time of the ending trip (min)			0	44	1.7	3.9
TD	Distance between the start point and the ending point of an outbound journey, based on the road network in the city (km)			0	61	4.4	5.9
ST_mode	Travel mode of the starting trip	0 = walk	30.1%				
		1 = bus	5.9%				
		2 = rail transit	1.7%				
		3 = private vehicles	0.3%				
		4 = vacant	62.1%				
ED_mode	Travel mode of the ending trip	0 = walk	19.0%				
		1 = bus	1.8%				
		2 = rail transit	0.8%				
		3 = private vehicles	0.1%				
		4 = vacant	78.4%				
ST_ct	Starting time categories	1 = before 10:00 am,	40.7%				
		2 = 10:00 am-12:00 am	7.9%				
		3 = 12:00 am-15:00 pm	10.7%				
		4 = 15:00 am-18:00 pm	20.1%				
		5 = after 18:00 pm	20.5%				
AR_category	Arriving time categories	1 = before 10:00 am,	38.9%				
		2 = 10:00 am-12:00 am	8.4%				
		3 = 12:00 am-15:00 pm	11.5%				
		4 = 15:00 am-18:00 pm	14.0%				
		5 = after 18:00 pm	27.1%				
TP	Trip purpose	1 = go back home	43.4%				
		2 = commute to work	30.6%				
		3 = go to school	4.7%				
		4 = visiting friends, dining, recreation, shopping etc.	18.4%				
		5 = go to hospital	0.6%				
		6 = others	2.3%				

Table 1. *Cont.*

Variable	Description	Category	%	Min	Max	Mean	SD
Socio-demographic attributes							
Age	Age of the respondent in years			6	94	39.30	15.64
H_size	Number of members in the respondent's household			1	8	3.11	0.95
Child_No	Number of young child (below 18) in the respondent's household			0	2	0.21	0.42
Car_own	Number of household -owned cars			0	2	0.33	0.52
Gender		1 = male	47.6%				
		2 = female	52.4%				
Occupation		1 = student	9.5%				
		2 = company employees	25.3%				
		3 = service staff	15.6%				
		4 = civil servant	7.9%				
		5 = industrial workers	4.9%				
		6 = retired people	16.9%				
		7 = free-lanced workers	11.3%				
		8 = others	8.6%				
Hukou		1 = local hukou holders	73.3%				
		2 = outbound resident who do not have local hukou	26.7%				
Car_plan	Household car consumption plan	1 = plan to buy car before 2015	5.5%				
		2 = plan to buy cars between 2016 and 2020 (five years after the rail transit system established)	25.5%				
		3 = no plan to buy cars before 2020	69.0%				
Parking	Whether or not respondent's household has a parking place of their own	1 = no	69.5%				
		2 = yes	30.5%				
RP_level	Including upscale new residential buildings, town houses, and villas	1 = poor condition/old properties/public rental properties	22.8%				
		2 = ordinary properties/gated communities	74.9%				
		3 = luxury properties	2.3%				

Table 1. *Cont.*

Variable	Description	Category	%	Min	Max	Mean	SD
Build environment elements							
O_land	Origin land use coding	1 = residential land	45.3%				
D_land	Destination land coding	2 = office land	12.7%				
		3 = commercial land	17.8%				
		4 = education land	7.6%				
		5 = hospital land	1.5%				
		6 = industrial and storage land	2.1%				
		7 = transportation land	0.5%				
		8 = others	12.5%				

The thresholds of multimodal travel behaviour are defined as the frequency of using two or more modes, illustrated by Buehler and Hamre [26] by using data from the US National Travel Survey [12]. Walking is excluded in some studies on multimodality [27], but included in other studies, as it plays a major role in the transport system [16]. In some studies, walking is identified together with cycling as active travel [12,20]. However, because of the mountain topology of Chongqing, very few people take cycling (less than 1%), but walk instead. Therefore, we keep walking as one travel mode category.

Before proceeding with the statistical analysis, we first use Pearson's correlation coefficient metric to detect multicollinearity among the all dependent and independent variables. Multicollinearity describes the state where the independent variables exhibit a strong relationship with each other. It will negatively impact the interpretation of the predictors and lead to a large change in feature importance scores. We assume that any features with a correlation coefficient exceeding 0.80 are suspected of causing multicollinearity following the previous research [47,48]. By computing correlation coefficient, we identify that parking availability and car ownership, starting time category, and arriving time category have strong relationships. By comparing the correlation coefficient with other variables, we delete parking availability and arriving time category. We used the Grubbs test [49] to detect outliers of response variables in the dataset and winsorized the outliers by replacing them with the maximum non-outliers. These measures are consistent with the literature [50].

3.2. Modelling Approach

Random forest (RF) was widely used as a machine learning method in previous research [40,51]. Gradient boosting decision trees (GBDT) use ensemble decision trees as RF [43]. However, the main difference between random forest and gradient boosting trees is how the models are trained and how they output decisions. RF is built in the way that each decision tree is used as a parallel estimator. The trees are trained independently. As output, the individual predictions are aggregated into a collective one, recognizing a majority vote of all decision trees for a classification task or the mean value for a regression one. In contrast to RF, GBDT uses a boosting technique to create an ensemble learner, while decision trees are connected sequentially with one tree built at a time. Each tree fits to the residuals from the previous one. In this way, it gradually increases the overall accuracy and robustness of the mode. However, the focus of new trees becomes the detail after some point and the cause of overfitting. Unlike RF, the number of trees in Gradient tree boosting is of crucial importance in terms of overfitting. Therefore, it is critical to find the key point after which each addition covers a detail or noise in the training data.

XGBoost is an improved method of gradient tree boosting proposed by Chen and Guestrin [52] and was utilized in transport behaviour analysis [41]. It has the advantage especially in dealing with sparsity in the dataset and quicker model calculation. An approximate tree-boosting algorithm is used to efficiently find split points on weighted

data. Parallel and distributed computing enables quicker model exploration. A sparsity-aware algorithm for parallel tree learning is introduced to deal with the sparsity in the dataset. It is pretty much useful for the dataset of this study, because for most part of the data, the starting and ending trips are vacant as unimodal trips.

In order to explain the mathematics in XGBoost, it is useful to begin with introducing the regular functions of gradient boosting [53]. Let D represent a dataset with a total of n samples, and each sample has a feature dimension of m, $O = \{(x_i, y_i)\}(|O| = n, x_i \in \mathbb{R}^m, y_i \in \mathbb{R})$. Representing the number of trees, K additive functions are used to predict the output in a tree ensemble model:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in \mathcal{F} \tag{1}$$

where $\mathcal{F} = \left\{ f(x) = \omega_{q(x)} \right\} (q : \mathbb{R}^m \to T, \omega \in \mathbb{R}^T)$, f_k is a function in the functional space of regression trees $\mathcal{F}$. T is the number of leaves in the tree. Each f_k corresponds to an independent tree structure q and leaf weights ω, q gives the decision rules in the trees and classify it into the leaves, and ω_i represents the score on the i-th leaf. The continuous scores in the corresponding leaves are summed up to calculate the final prediction.

To measure how well the model fits the training data, the objective function to be optimized is given by Equation (2):

$$obj = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t)}\right) + \sum_{i=1}^{t} \omega(f_i). \tag{2}$$

Formally, let $\hat{y}_i^{(t)}$ be the prediction of the i-th instance at the t-th iteration, the function is composed of two terms. The first term l, a differentiable convex loss function, measures the difference between the prediction $\hat{y}_i$ and the target y_i. The second term $\omega(f_i)$, as an additional regularization term, helps to smooth the final learnt weights to avoid over-fitting. It is what is improved by XGBoost [52]. The complexity penalized by the regularization term is defined as:

$$\omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \tag{3}$$

f_t is added in Equation (3) to minimize the following objective:

$$obj = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \omega(f_t) + constant. \tag{4}$$

We take the Taylor expansion of the loss function up to the second order, and the objective value with the t-th tree is rewritten as:

$$\begin{aligned} obj^{(t)} &\approx \sum_{i=1}^{n} \left[g_i \omega_{q(x_i)} + \tfrac{1}{2} h_i \omega_{q(x_i)}^2 \right] + \gamma T + \tfrac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \\ &\approx \sum_{j=1}^{T} \left[\left(\sum_{i \in I_j} g_i \right) w_i + \tfrac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) \omega_j^2 \right] + \gamma T \end{aligned} \tag{5}$$

where $I_j = \{i | q(x_i) = j\}$ is the set of indices of data points assigned to the j-th leaf. For a fixed structure $q(x)$, the optimal weight ω_j^* of leaf j can be computed as

$$\omega_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \tag{6}$$

and the corresponding optimal value is:

$$obj^* = -\frac{1}{2} \sum_{j=1}^{T} \frac{\left(\sum_{i \in I_j} g_i \right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T. \tag{7}$$

The quality of a tree structure q can therefore be measured by Equation (7). order to decide when to stop splitting a leaf into two leaves, a loss reduction function is given. Assume that I_L and I_R are the instance sets of left and right nodes after the split, $I = I_L \cup I_R$,

$$Gain = \frac{1}{2}\left[\frac{\left(\sum_{i\in I_L} g_i\right)^2}{\sum_{i\in I_L} h_i + \lambda} + \frac{\left(\sum_{i\in I_R} g_i\right)^2}{\sum_{i\in I_R} h_i + \lambda} - \frac{\left(\sum_{i\in I_j} g_i\right)^2}{\sum_{i\in I_j} h_i + \lambda}\right] - \gamma. \tag{8}$$

If the gain is smaller than zero, we would do better not to add that branch. This formula can be used in practice for evaluating the split candidates to search for an optimal split. An approximate algorithm is introduced in XGBoost in aiding to do so efficiently.

To overcome the interpretability weakness of machine learning models, a variety of machine learning interpretation tools were utilized, including variable importance and partial dependence plots [39,53,54]. The importance of a feature for an entire dataset can be measured as the standard deviation of the partial dependence plot. This study utilizes the partial dependence-based variable importance measure proposed by Greenwell [55]. These measures are consistent with the methods used in the literature [56,57], conductive to comparing the results across different models. The partial dependence plot shows the marginal effect that one or two features have on the predicted outcome of a machine learning model [53]. It can show whether the relationship between the target and a feature is linear, monotonic, or more complex. The partial dependence function for regression is defined as:

$$\hat{f}_s(x_s) = E_{X_c}\left[\hat{f}(x_s, X_c)\right] = \int \hat{f}(x_s, X_c) dP(X_c). \tag{9}$$

The x_s are the features for which the partial dependence function should be plotted and X_c are the other features used in the machine learning model $\hat{f}$, which are here treated as random variables. Partial dependence works by marginalizing the machine learning model output over the distribution of the features in set C, so that the function shows the relationship between the features in set S we are interested in and the predicted outcome. For classification, partial dependence plots measure the influence of a variable x_s on the log odds or probability of choosing a specific travel mode after accounting for the average effects of all other variables, which is the task of this study [53].

4. Results

4.1. Baseline Model

We carried out analysis on the whole dataset, in order to have a comprehensive picture on people's mode choice. A logistic regression test is carried out first in order to set a baseline of the model prediction (Table 2). The chi-square value for the whole model is highly significant, with a Cox and Snell R square value of 0.785. Looking at the likelihood test for each variable, they are all past the 0.05 significant test except for the two variables of residential property level and car consumption plan. However, though the X-standardized B value or Exp (B) of logistic regression test can present the direction of impact of each parameter on the estimated variable, it can only give an average parameter prediction. Therefore, we resort to partial dependence plots to exhibit the influence varying with the change in the input variables.

Table 2. Logistic regression test.

Effect	Likelihood Ratio		
	Chi-Square	**df**	**Sig.**
Constant	0.00	0	
TD	985.73	3	0.000
ST_mode	4417.19	12	0.000
ST_time	121.09	3	0.000
MA_time	223.97	3	0.000
TT	12.57	3	0.006
ED_time	18.65	3	0.000
ED_mode	545.85	12	0.000
ST_ct	47.10	12	0.000
TP	29.43	15	0.014
Car_own	1212.12	3	0.000
Age	378,193.07	3	0.000
Gender	151.24	3	0.000
H_size	41.17	3	0.000
Occupation	148.50	21	0.000
Child_No	23.21	3	0.000
Hukou	19.83	3	0.000
RP_level	11.30	6	0.079
Car_plan	11.19	6	0.083
D_land	63.14	21	0.000
O_land	61.69	21	0.000

4.2. Model Comparison

The machine learning model is interpreted with variable importance measures and partial dependence plots. Five different machine learning models are carried out on the dataset. They are AdaBoost, decision trees, random forest, and XGBoost. Table 3 shows the relative importance (RI) of variables for different models. RI for logistic regression is also calculated for comparison. Table 3 exhibits that generally, ML models have higher model performance than logistic regression, with a higher F1 score, recall, and precision value. In particular, XGBoost has the highest F1 score, of 0.848, among all the models.

The ranking of the most significant variables among all the models are generally similar, with the first seven trip characteristic variables (except for starting time category and travel purpose) filled within the top 10. The greatest contribution is consistent with previous research of Liu [41]. For XGBoost, travel characteristic variables collectively contributed to approximately 74% of the predictive power for the main mode choice. Trip distance, starting trip mode, starting trip time, and main trip time have much greater predictive power than other variables, with 66% loading in total. Socio-demographic attributes only contribute to 25%, among which, car ownership is the most important predictor of the main mode choice. However, because of the inaccessibility of most of the built environment data, the built environment elements of origin and destination land use only contribute to less than 1% in total.

4.3. Nonlinear Associations between Predictors and Travel Mode Choice

The one variable PDP plots show the nonlinear effect of trip stage characteristics on the probability of choosing each category of the main mode.

Table 3. Comparison of relative importance (RI) of predictors and predictive accuracy of models.

Predictor Variables	XGBT		Random Forest		Decision Tree		AdaBoost		Logistic Regression	
	RI	**Rank**	**RI**	**Rank**	**RI**	**Rank**	**RI**	**Rank**	**RI**	**Rank**
Trip characteristics										
TD	26.56%	1	20.17%	2	20.80%	1	5.23%	6	16.39%	2
ST_mode	21.18%	2	20.81%	1	19.92%	2	12.99%	4	15.22%	3
ST_time	10.05%	4	8.06%	5	9.61%	5	0.79%	13	6.67%	5
MA_time	8.59%	5	8.58%	4	8.00%	6	17.87%	2	11.38%	4
TT	2.61%	7	4.76%	6	14.03%	4	2.01%	9	2.39%	11
ED_time	2.60%	8	4.09%	7	6.32%	7	2.73%	8	4.08%	7
ED_mode	1.95%	9	3.29%	8	0.76%	10	12.00%	5	4.12%	6
ST_ct	0.56%	13	1.00%	16	0.25%	12	0.49%	15	1.22%	16
TT	0.13%	19	0.99%	17	0.00%	18	0.26%	18	1.48%	15
Socio-demographic attributes										
Car_own	17.35%	3	15.57%	3	14.21%	3	4.07%	7	19.53%	1
Age	4.58%	6	2.95%	9	4.40%	8	18.35%	1	2.82%	8
Gender	0.90%	10	1.34%	12	0.00%	20	0.27%	17	1.09%	17
H_size	0.71%	11	0.84%	18	0.00%	16	17.23%	3	2.49%	10
Occupation	0.60%	12	1.62%	11	1.01%	9	1.60%	11	2.36%	12
Child_No	0.55%	14	1.64%	10	0.00%	17	0.00%	19	2.78%	9
Hukou	0.23%	17	0.50%	19	0.02%	15	0.27%	16	0.46%	20
RP_level	0.16%	18	1.03%	15	0.53%	11	0.00%	20	0.55%	18
Car_plan	0.11%	20	0.36%	20	0.07%	14	0.63%	14	0.47%	19
Build environment elements										
D_land	0.32%	15	1.32%	13	0.00%	19	1.44%	12	2.18%	14
O_land	0.24%	16	1.09%	14	0.08%	13	1.78%	10	2.31%	13
F1 score	0.848		0.791		0.753		0.726		0.696	
Recall	0.829		0.762		0.744		0.721		0.684	
Precision	0.874		0.858		0.772		0.733		0.757	

4.3.1. Trip Distance

Trip distance takes the most importance in all the variables. The impact from trip distance is disparate across different modes in Figure 2. When trip distance increases from 0 to 10 km, the probability of taking transit (Figure 2c) increases about 20 percent and then remains stable. However, that of taking private vehicles (Figure 2d) increases about 30 percent when trip distance experiences the same changes, while that of walking (Figure 2a) decreases 40 percent. However, the probability of taking the bus (Figure 2b) only shows an indiscernible change when trip distance changes. It suggests that 10 km is the threshold that people's interest in choosing transit increases if they travel longer, beyond which the incentive to take transit for a longer trip is not so effective.

4.3.2. Travel Modes of the Starting Trip Stages

The travel mode of the starting trip takes about 21% of the relative importance and ranks the second in all the variables. However, looking at the PDP plots in Figure 3, the variance in main mode choice mostly results from the same variable of walking. If people take walking as the starting trip mode, their probability of taking bus, transit, and private vehicles as their main trip mode is much higher than other modes as the starting trip mode. In contrast, the ending trip mode does not exhibit such a variance in influencing people's main mode choice (for saving place it is not presented here). It means there is not much evidence of the interrelations between mode choice of different trip stages, except when walking is the starting trip mode.

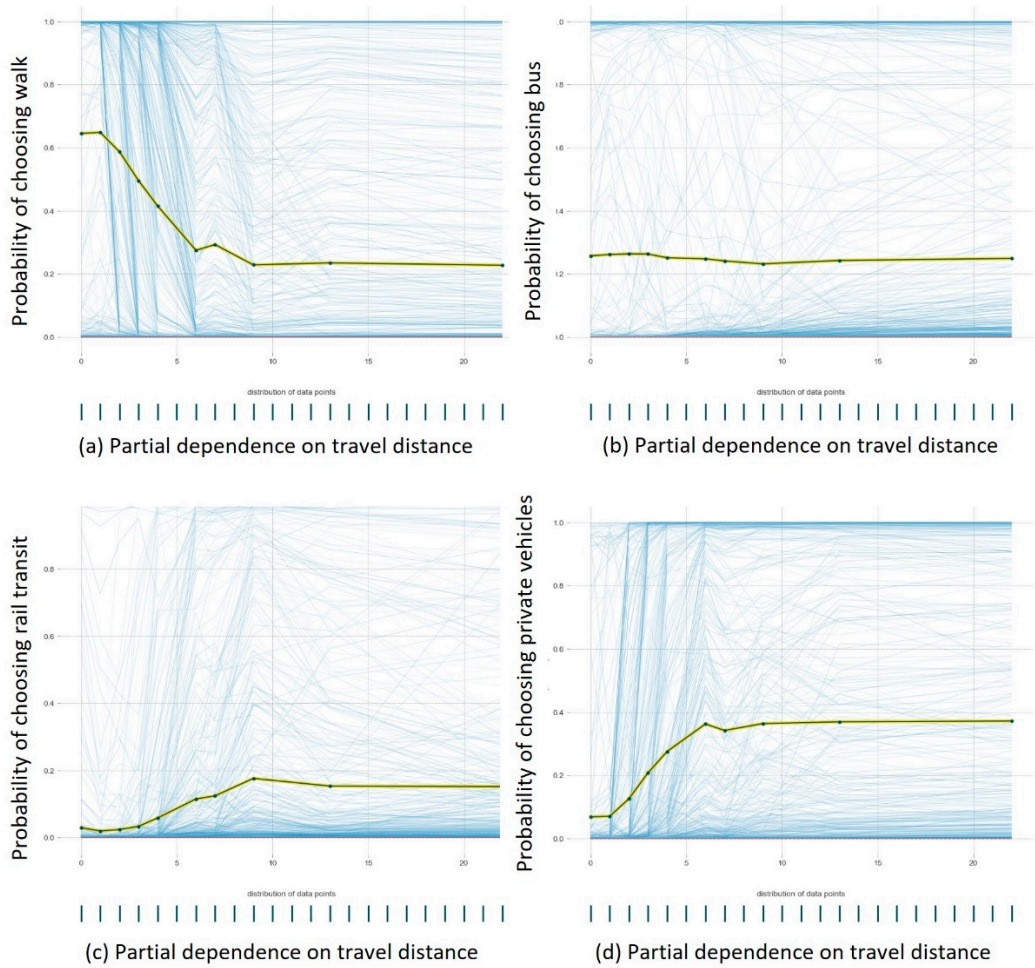
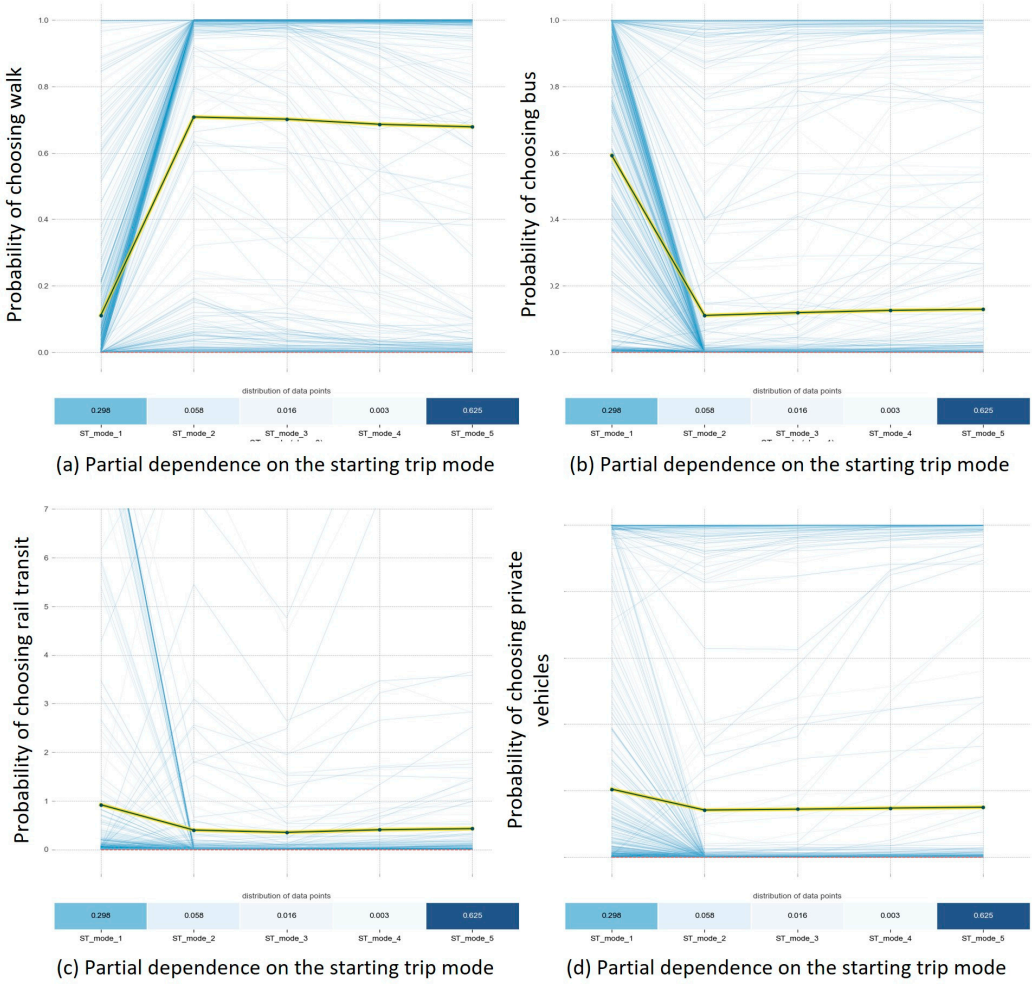


Figure 2. Partial dependence plots of trip distance for the main trip mode.

4.3.3. Duration of Different Trip Stages

The relative importance of the starting trip (10.05%) and main trip (8.59%) duration is higher than the whole trip duration (2.61%) (Table 3). Furthermore, the probability of main mode choice is impacted much more by the starting trip than the ending trip (2.6%). For the mode of bus (Figure 4b), the probability increases about 20% when the starting trip duration increases from 0 to 12 min and then decreases. In contrast, the probability of taking private vehicles (Figure 4d) decreases about 15 percent when the starting trip duration experiences the same change and remains stable after that. However, for the mode of transit (Figure 4c), the impact is not so obvious, which slightly decreases when the starting trip duration increases from 0 to 12 min, and then steadily increases. It means that when the starting trip duration increases within a certain range, 12 min in this case, it increases the probability of taking bus. However, it shows the increased probability is compensated by the decrease in probability of taking transit or private vehicles. While the time exceeds this threshold range, people are more likely to resort to transit. It suggests that for those who choose transit as the main travel mode, they are more willing to spend longer time for the starting trip. As to the interrelation with main trip time (Figure 5), there exhibits a threshold of 15 min for the main trip time, within which probability of travel by transit (Figure 5c) increases with time duration increases, and a threshold of 20 min above which the probability does not change accordingly. The probability of taking a bus (Figure 5b) experiences similar but inverse changes while it decreases first and then increases when exceeding the same threshold of 15 min. However, the probability of taking private vehicles (Figure 5d) does not change much with the main time duration. This trend also suggests a network effect of people's

main trip mode choice. It indicates that the increased transit usage might be attracted from bus and walking rather than from the private vehicles.



(a) Partial dependence on the starting trip mode (b) Partial dependence on the starting trip mode

(c) Partial dependence on the starting trip mode (d) Partial dependence on the starting trip mode

Figure 3. Partial dependence plots of the starting trip mode for the main trip mode.

4.4. Synergistic Effects between Variables

4.4.1. Synergistic Effects between Age and Trip Stage Characteristics

The two variable PDP plots (Figure 6) reveal the synergistic effect of age when interacts with the changes of separate trip stage characteristics. From the 2 variable PDP plot we can see that the highest probability of taking transit (Figure 6a) accumulates at the twenties when travelling for about 10 kms. With travel distance increases from 5 km to 10 km, which is the radius of the central urban area, the probability steeply increases. It means the adolescences more likely utilize rail transit as their main travel mode, as the probability increases much faster than older people. Similar pattern is exhibited in the mode of private vehicles (Figure 6b), while the highest probability of age is around thirties, which indicates that the middle-aged people more likely choose private vehicles when travel distance increases.

The synergistic effect between age and the starting trip mode varies for the mode choice of rail transit (Figure 6c) and cars. However, the synergistic effect is not distinct for walking and taking bus (for saving space not exhibited here). Especially for people who take a walk for the starting trip, their probability of taking transit as main mode decreases almost 10% (from 0.16 to 0.06) when age increases from early twenties to fifties. For other modes to and from the transit station, the probability decreases only about 4% as age has the same change. It suggests the walking environment of the last mile between origination/destination and transit station is very important especially for those aged to

take transit. The synergistic effect with the starting trip time (Figure 6d) shows people between 25 and 45 have the highest probability of taking private vehicles. The probability decreases more slowly than elder people even when the starting trip gets longer. It means the habit cultivated by the middle aged is not so easy to change.

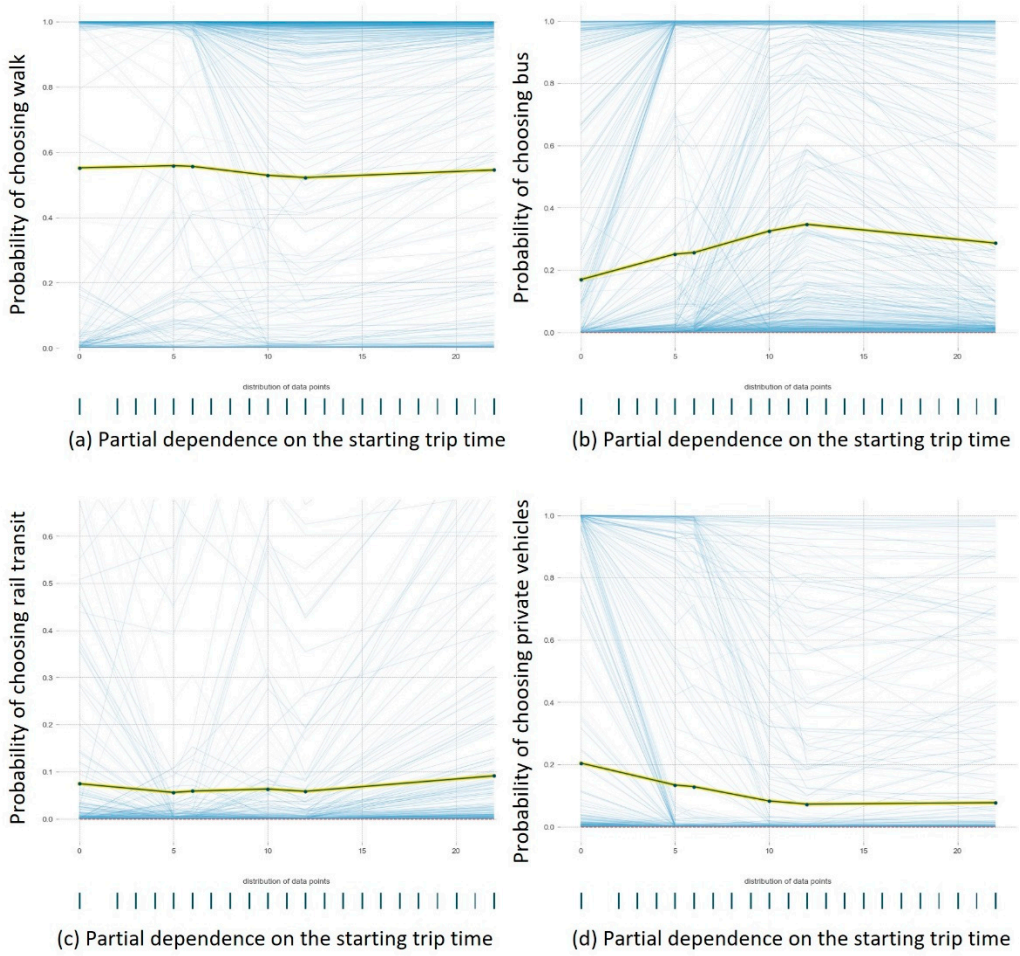


Figure 4. Partial dependence plots of the starting trip time for the main trip mode.

The probability of taking transit decreases when main trip time increases above 15 min (Figure 6e). However, the speed of decrease is distinct among different ages. The probability of the elderly to taking transit is decreasing from a relatively lower probability to a much lower level than the younger people, especially the early twenties, with a faster speed. The probability of fifties decreases from 0.125 to almost 0 (12.5 percent) when the main trip time increases from 15 min to 30 min, while that of the early twenties only decreases from 0.15 to 0.075 (7.5 percent) when the main trip time experiences the same change. It means when the main trip duration surpasses a certain threshold and gets longer, the older people are more likely give up transit. Meanwhile, when it is compared with the pattern of probability of taking bus (Figure 6f), the probability of people above sixties to take bus increases faster than the younger people. It means the older people are more inclined to give up transit and shift to take bus when the main trip time increases. It may because that the bus service provides more comfort riding experience for longer trip, a cheaper ticket price, and is more accessible from the origin point.

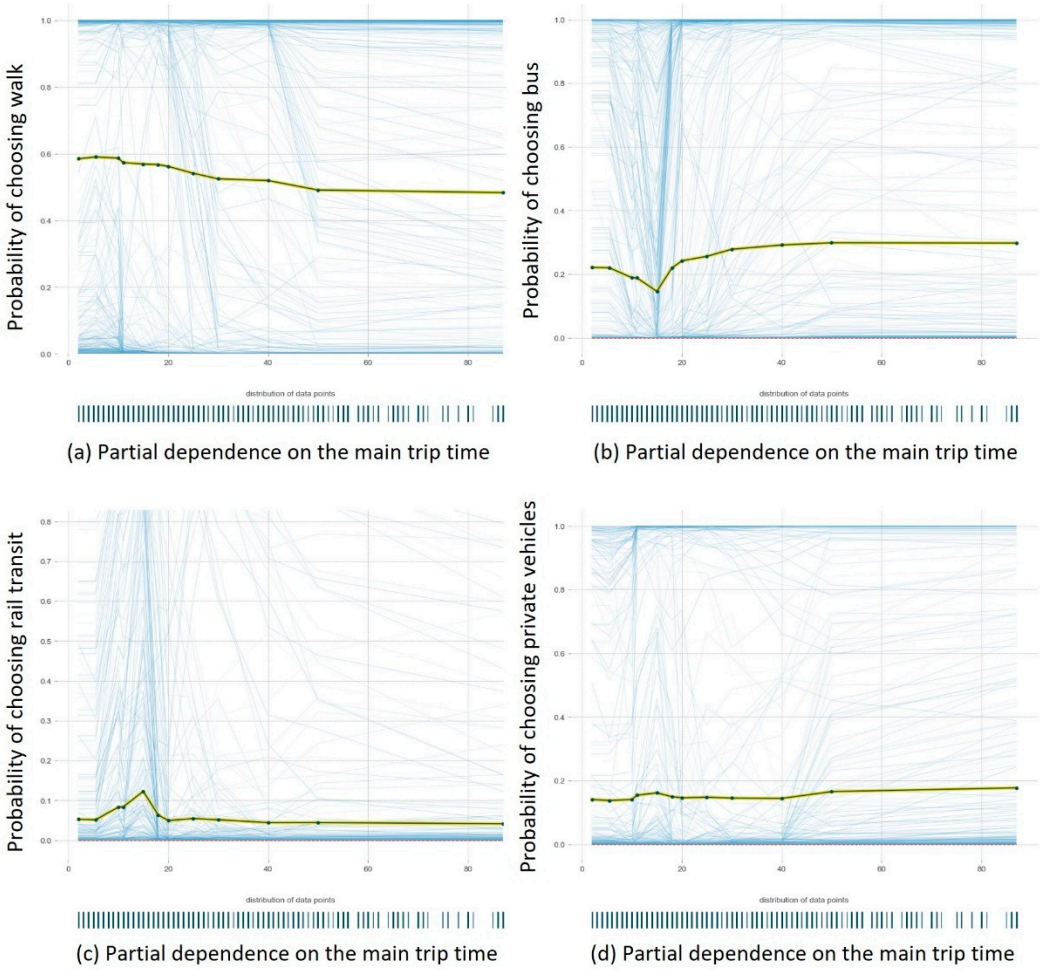


Figure 5. Partial dependence plots of the main trip time for the main trip mode.

4.4.2. Synergistic Effects between Other Socio-Demographic Attributes and Trip Stage Characteristics

To be concise, we mainly focus on an explanation of the synergistic effects of main trip time with other socio-demographic and attributes on the main mode choice, while there are synergistic effects existing between other travel characteristics and socio-demographic variables.

The synergistic effect of main trip time and car ownership categories shows the variance in transit use probability mainly comes from those who don't have cars (Figure 7c). The probability of the non-car owners to choose transit increases faster than those who have cars when the main trip time changes. However, the probability of choosing private vehicles (Figure 7d) for different groups is comparatively stable when the main time changes, that people who owns cars have higher probability choosing private vehicles and those who owns more than one car especially higher. Meanwhile, the probability of walking decreases faster and that of taking bus increases more slowly for those car owners compared to non-car owners when the main trip time experience the same change. It indicates that when main trip time changes, those people who don't have cars are more likely to switch to the use of transit, while those car owners keep inertia in their established driving habit or more likely to give up public transport and resort private vehicles.

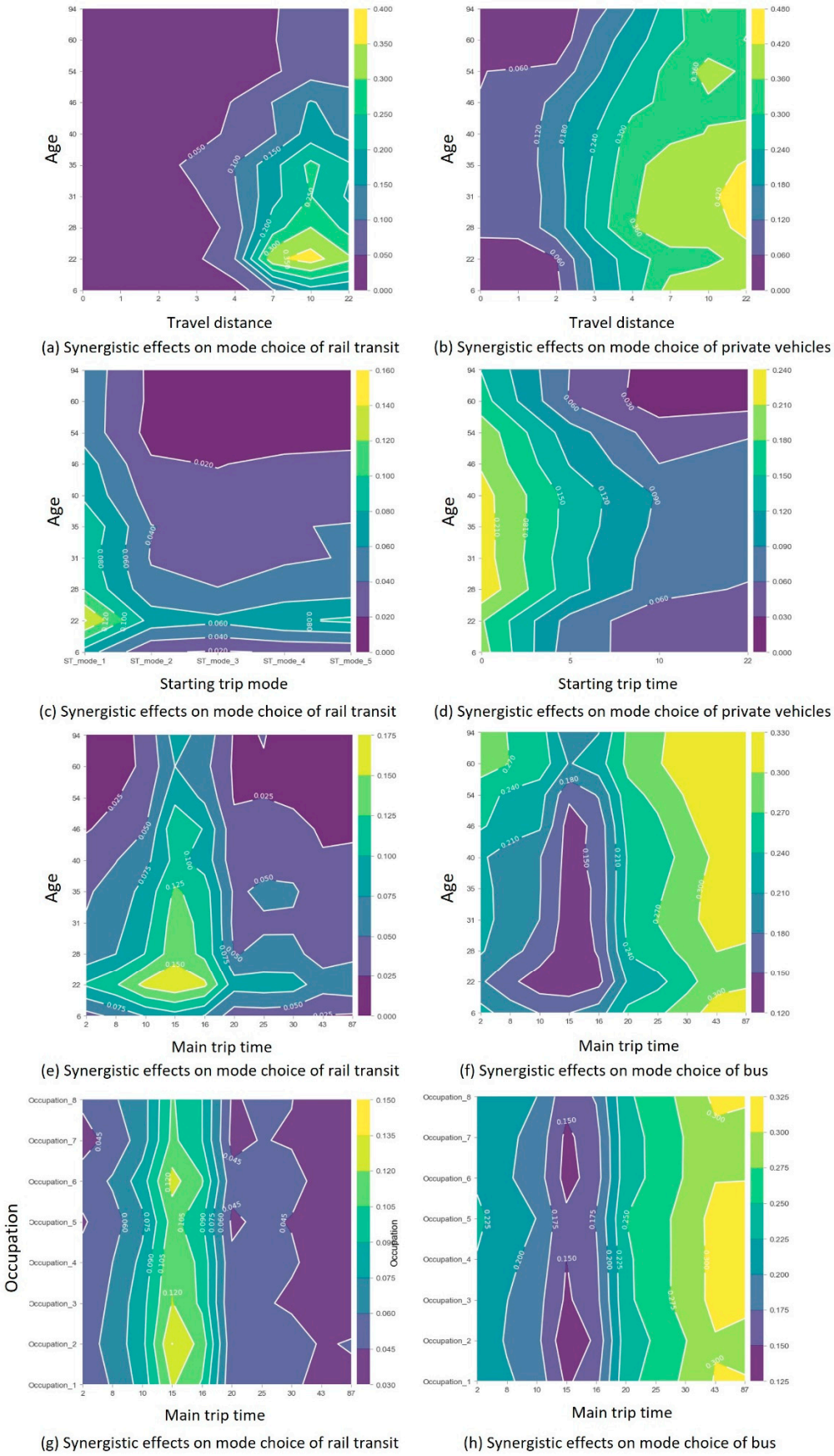


Figure 6. Synergistic effects of trip characteristics with age and occupation on main trip choice.

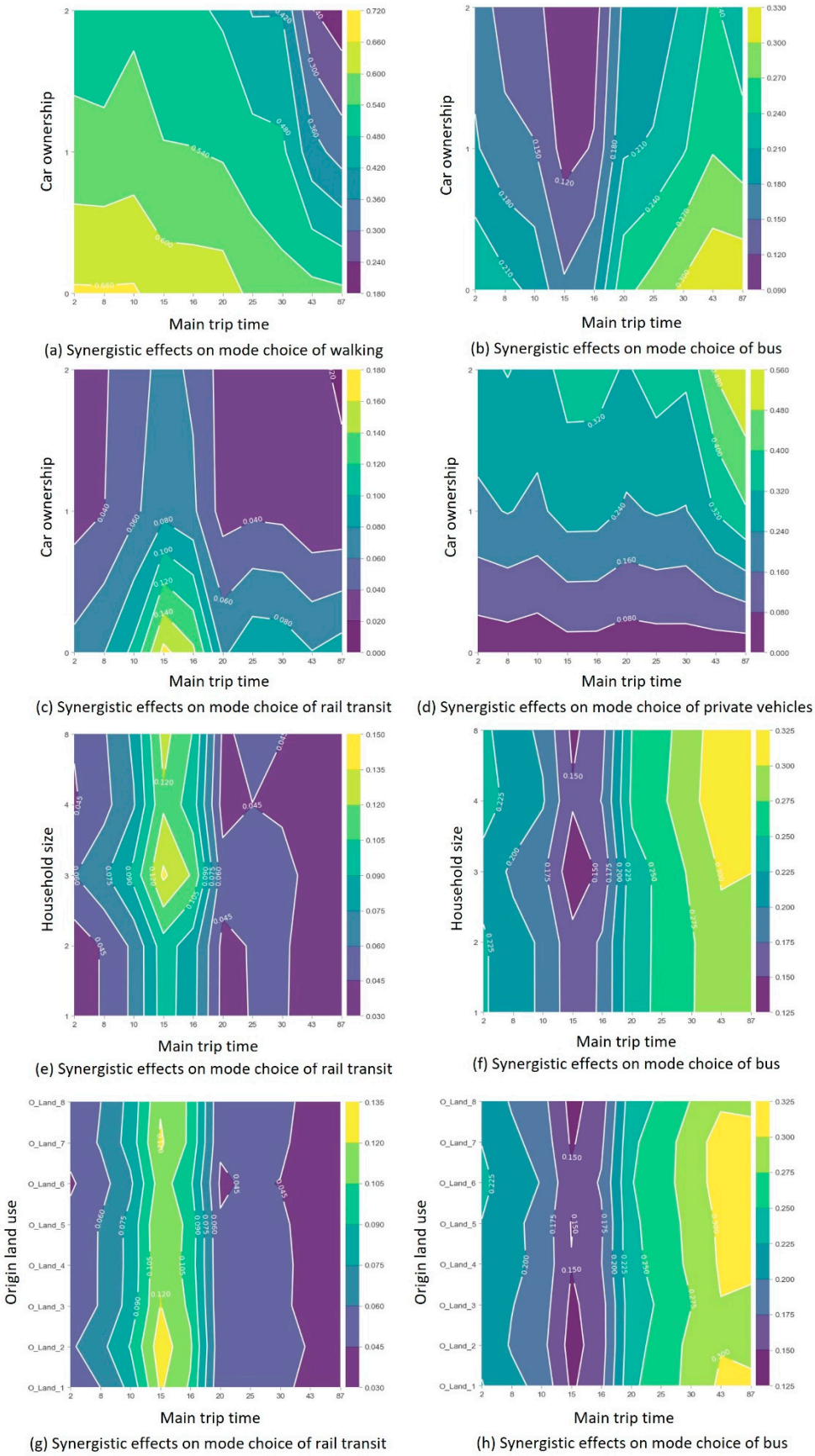


Figure 7. Synergistic effects of trip characteristics with other socio-demographic and built environment attributes.

As main trip time increases from 0 to 15 min, for the groups of company employees and those who are retired, their probability of taking transit (Figure 6g) is higher than other groups. Correspondingly, the probability of taking bus for these two groups of people is also the lowest when the main trip time experienced the same change. Time is the main concern of these two groups and they are likely to be attracted from bus use (Figure 6h). However, their choice of transit is still subject to a duration threshold of 15 min to 20 min. The company employees are more time strict, while the retired may balance the cost and comfort the system provides.

For synergistic effect with household size, the most distinct disparity exists in the 3-person family (family with a child), which takes about 50% percent in the dataset. When the main trip time increases from 0 to 15 mins, the probability of taking transit (Figure 7e) increases from 0.06 to 0.15. However, for single or 2-person household, the probability only increases from 0.03 to 0.09 when the main trip time experience the same change. It maybe because of the comfort to travel with child, the reasonable cost, and the ride experience the system provides for a typical household with a child. A correspondingly similar change is shown for bus trips (Figure 7f), where the probability of taking bus of the 3-person family experience the similar decrease in probability when the main trip time experience the same change. It means when the trip time increases, the probability of the three-person family than other households are more likely to resort to transit use and abandon the bus.

4.4.3. Synergistic Effects between Land Use and Trip Stage Characteristics

The synergic effects between land use and main trip time are not so distinguished. Trips from business/office land, and transport land exhibit a higher probability of choosing transit (Figure 7g) when the main trip time increases to 15 mins. Correspondingly, the probabilities to take bus (Figure 7h) from these two kinds of land use reach the lowest level when the main trip time experience the same change. However, the probability of walk and private vehicles remains stable (for saving space not exhibited here). It also indicates there is a reciprocal effect between rail transit and bus. The similar reciprocal effect is also found with destination land use, though not so obvious as original land use, trips to schools and commercial centres have higher probability of choosing transit (not exhibited here for space reasons).

5. Conclusions

As with newly introduced public transport in some urban areas, the rail transit development aims to broaden the mobility options and support sustainable transport. Investment promoting rail transit use needs to be based on a better understanding of the complex interrelation between social demographic attributes and travel characteristics, in order to promote multimodal travel behaviours for more efficient use of urban transport. The novelty of this paper lies in identifying the role of the rail transit system in multimodality by correlating the main mode choice with characteristics of separate trip stages. The varying non-linear and synergistic effects revealed by the machine learning models provides detailed interpretations of impacts with the change in the variables, which contributes to more adaptive planning strategies. This paper contributes to the research field in the following aspects.

Firstly, with the insight into the multimodal travel behaviours, we explored how the main trip mode is impacted by other interrelated trip stages. As a result, the separate trip stage characteristics have more of an impact than the general trip characteristics, in that the starting/ending and main trip time has a higher relative importance in predicting the main mode choice. It demonstrates that research in differentiating different trip stages reveals more intrinsic interrelation than treating the trip as a whole in travel behaviour analysis.

Secondly, the non-linear effects revealed by the machine learning models show varying impact on people's main mode choice, in which the role of rail transit is identified in multimodal travels. The ML model provides more accurate estimates than traditional models. The impact of variables on travel mode choice is more effective at a certain range

of these variables than other ranges of these variables. There are thresholds of variable impact by trip stage duration and travel distance on main mode choice, within which the probability of choosing certain travel modes is increased, while beyond which the probability is stable or decreased. For instance, the threshold of main trip duration impact suggests an optimal 15–20 min radius of allocation of functional utilities accessible by transit. The results have implications for spatial planning. It emphasizes the importance of accessibility to utilities within a certain time duration from the residential areas along the transit lines, which is about 7–10 stations.

Thirdly, the synergistic effects between variables revealed by ML models provide more effective suggestions on the targeted groups of people and land use characteristics, which should be concerned by the planning strategies [58]. For example, the synergistic effect between the starting trip mode, the main trip time, and age suggests the importance of creating walkable environments and an increasing accessibility level of the first/last mile connection to the transit stations, particularly for the transit use of the elderly. It adds more detailed information on the travel behaviour of the aged to the literature, that their multimodal choice varies with the change in trip stage characteristics [25,59]. In general, company employees and the retired, non-car owners, and three-person families are more likely to switch to the use of transit when the main trip time increases and the travel distance expands within a certain threshold. They are the groups of people who are willing to integrate the new rail transit system into their trips to expand their activities. The varying relationship with occupation and car ownership adds more detailed evidence to the literature [4,16,20,26]. The evidence of the three-person family indicates that the birth of a child means the role of transit is embedded into people's multimodal travels in multiple ways, which provides contradicting evidence to the previous literature [3,4,12]. Therefore, planning practice and policies may take into consideration of the needs of these groups.

Meanwhile, there is a reciprocal effect between public transport modes of rail transit and bus mode choice. The land use characteristics exhibit a marginal, but still apparent, impact. Trips from business/office land, transport land, and trips to schools and business/office land exhibit a higher probability of choosing transit when main trip time and travel distance increases within the thresholds. This suggests that if transit stations are planned on these sites, people are more likely to be attracted to use transit. However, the increased transit ridership is largely compensated by the decreased use of bus. A similar reciprocal effect is also found in the synergistic effect of travel characteristics with occupation and household size. This indicates that the increased transit usage might be attracted from bus and walking rather than from private vehicles. There is a network effect between public transport modes of people's main trip.

This paper has some limitations. A citywide travel survey is only carried out periodically in Chinese cities. At present, the only accessible dataset for research is from 2014. Nevertheless, the results can be considered as an assessment of the effect of a newly established transit system. Further research will seek to access more recent data available, and potentially carry out analysis in comparison. Furthermore, because of data accessibility, data of more built environment elements are not available in this study, such as density, diversity, etc., which makes a low relative importance for the model prediction. Further research could include more detailed built environment data for a more complete picture of the impact on multimodality.

Rail transit systems are developed with significant investment in aiming to accommodate mass travel demands and sustainable transportation. However, the goal can be achieved only when integrated planning is implemented, including transportation planning. This requires synergistic land use and urban design strategies, such as facility allocation along the lines, walkable environments, and convenient connections to the stations, as well as specific policies for the targeted population groups. Otherwise, the mode choice might be attracted from other public transport modes rather than the private vehicles.

Author Contributions: Conceptualization, L.L. and R.H.; Methodology, L.L.; Formal analysis, Y.W.; Writing—original draft, L.L.; Writing—review & editing, R.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The Fundamental Research Funds for the Central Universities (531118010792).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dai, F.; Diao, M.; Sing, T.F. Effects of rail transit on individual travel mode shares: A two-dimensional propensity score matching approach. *Transp. Res. Part D Transp. Environ.* **2020**, *89*, 102601. [CrossRef]
2. Luan, X.; Cheng, L.; Song, Y.; Zhao, J. Better understanding the choice of travel mode by urban residents: New insights from the catchment areas of rail transit stations. *Sustain. Cities Soc.* **2020**, *53*, 101968. [CrossRef]
3. Molin, E.; Mokhtarian, P.; Kroesen, M. Multimodal travel groups and attitudes: A latent class cluster analysis of Dutch travelers. *Transp. Res. Part A: Policy Pract.* **2016**, *83*, 14–29. [CrossRef]
4. Nobis, C. Multimodality: Facets and causes of sustainable mobility behavior. *Transp. Res. Rec.* **2007**, *2010*, 35–44. [CrossRef]
5. Kuhnimhof, T.; Buehler, R.; Wirtz, M.; Kalinowska, D. Travel trends among young adults in Germany: Increasing multimodality and declining car use for men. *J. Transp. Geogr.* **2012**, *24*, 443–450. [CrossRef]
6. Gadepalli, R.; Tiwari, G.; Bolia, N. Role of user's socio-economic and travel characteristics in mode choice between city bus and informal transit services: Lessons from household surveys in Visakhapatnam, India. *J. Transp. Geogr.* **2020**, *88*, 102307. [CrossRef]
7. Du, M.; Cheng, L.; Li, X.; Yang, J. Factors affecting the travel mode choice of the urban elderly in healthcare activity: Comparison between core area and suburban area. *Sustain. Cities Soc.* **2020**, *52*, 101868. [CrossRef]
8. Dėdelė, A.; Miškinytė, A.; Andrušaitytė, S.; Nemaniūtė-Gužienė, J. Dependence between travel distance, individual socioeconomic and health-related characteristics, and the choice of the travel mode: A cross-sectional study for Kaunas, Lithuania. *J. Transp. Geogr.* **2020**, *86*, 102762. [CrossRef]
9. Chen, X.; Guo, Y.; Yang, C.; Ding, F.; Yuan, Q. Exploring essential travel during COVID-19 quarantine: Evidence from China. *Transp. Policy* **2021**, *111*, 90–97. [CrossRef]
10. Faber, R.M.; Jonkeren, O.; de Haas, M.C.; Molin, E.J.E.; Kroesen, M. Inferring modality styles by revealing mode choice heterogeneity in response to weather conditions. *Transp. Res. Part A Policy Pract.* **2022**, *162*, 282–295. [CrossRef]
11. Deschaintres, E.; Morency, C.; Trépanier, M. Measuring Changes in Multimodal Travel Behavior Resulting from Transport Supply Improvement. *Transp. Res. Rec.* **2021**, *2675*, 533–546. [CrossRef]
12. Klinger, T. Moving from monomodality to multimodality? Changes in mode choice of new residents. *Transp. Res. Part A: Policy Pract.* **2017**, *104*, 221–237. [CrossRef]
13. Duca, G.; Trincone, B.; Bagamanova, M.; Meincke, P.; Russo, R.; Sangermano, V. Passenger Dimensions in Sustainable Multimodal Mobility Services. *Sustainability* **2022**, *14*, 12254. [CrossRef]
14. Oostendorp, R.; Gebhardt, L. Combining means of transport as a users' strategy to optimize traveling in an urban context: Empirical results on intermodal travel behavior from a survey in Berlin. *J. Transp. Geogr.* **2018**, *71*, 72–83. [CrossRef]
15. An, Z.; Heinen, E.; Watling, D. The level and determinants of multimodal travel behavior: Does trip purpose make a difference? *Int. J. Sustain. Transp.* **2023**, *17*, 103–117. [CrossRef]
16. Scheiner, J.; Chatterjee, K.; Heinen, E. Key events and multimodality: A life course approach. *Transp. Res. Part A Policy Pract.* **2016**, *91*, 148–165. [CrossRef]
17. Heinen, E. Are multimodals more likely to change their travel behaviour? A cross-sectional analysis to explore the theoretical link between multimodality and the intention to change mode choice. *Transp. Res. Part F Traffic Psychol. Behav.* **2018**, *56*, 200–214. [CrossRef]
18. Olafsson, A.S.; Nielsen, T.S.; Carstensen, T.A. Cycling in multimodal transport behaviours: Exploring modality styles in the Danish population. *J. Transp. Geogr.* **2016**, *52*, 123–130. [CrossRef]
19. Heinen, E.; Mattioli, G. Does a high level of multimodality mean less car use? An exploration of multimodality trends in England. *Transportation* **2019**, *46*, 1093–1126. [CrossRef]
20. Heinen, E.; Chatterjee, K. The same mode again? An exploration of mode choice variability in Great Britain using the National Travel Survey. *Transp. Res. Part A Policy Pract.* **2015**, *78*, 266–282. [CrossRef]
21. Diana, M.; Pirra, M. A comparative assessment of synthetic indices to measure multimodality behaviours. *Transp. A Transp. Sci.* **2016**, *12*, 771–793. [CrossRef]
22. Heinen, E.; Mattioli, G. Multimodality and CO_2 emissions: A relationship moderated by distance. *Transp. Res. Part D Transp. Environ.* **2019**, *75*, 179–196. [CrossRef]
23. MoT. *12th Five-Year Plan for Transportation*; Ministry of Transport of the People's Republic of China: Bejing, 2011.
24. Kirtonia, S.; Sun, Y. Evaluating rail transit's comparative advantages in travel cost and time over taxi with open data in two U.S. cities. *Transp. Policy* **2022**, *115*, 75–87. [CrossRef]
25. Haustein, S.; Siren, A. Older people's mobility: Segments, factors, trends. *Transp. Rev.* **2015**, *35*, 466–487. [CrossRef]

26. Buehler, R.; Hamre, A. The multimodal majority? Driving, walking, cycling, and public transportation use among American adults. *Transportation* **2015**, *42*, 1081–1101. [CrossRef]
27. Kroesen, M. Modeling the behavioral determinants of travel behavior: An application of latent transition analysis. *Transp. Res. Part A Policy Pract.* **2014**, *65*, 56–67. [CrossRef]
28. Li, S.; Lyu, D.; Huang, G.; Zhang, X.; Gao, F.; Chen, Y.; Liu, X. Spatially varying impacts of built environment factors on rail transit ridership at station level: A case study in Guangzhou, China. *J. Transp. Geogr.* **2020**, *82*, 102631. [CrossRef]
29. Zhou, M.; Wang, D.; Guan, X. Co-evolution of the built environment and travel behaviour in Shenzhen, China. *Transp. Res. Part D Transp. Environ.* **2022**, *107*, 103291. [CrossRef]
30. Bautista-Hernández, D.A. Mode choice in commuting and the built environment in México City. Is there a chance for non-motorized travel? *J. Transp. Geogr.* **2021**, *92*, 103024. [CrossRef]
31. Blumenberg, E.; Pierce, G. Multimodal travel and the poor: Evidence from the 2009 National Household Travel Survey. *Transp. Lett.* **2014**, *6*, 36–45. [CrossRef]
32. McLaren, A.T. Families and transportation: Moving towards multimodality and altermobility? *J. Transp. Geogr.* **2016**, *51*, 218–225. [CrossRef]
33. Zhao, J.; Fan, W.; Zhai, X. Identification of land-use characteristics using bicycle sharing data: A deep learning approach. *J. Transp. Geogr.* **2020**, *82*, 102562. [CrossRef]
34. Tao, T.; Cao, J. Examining motivations for owning autonomous vehicles: Implications for land use and transportation. *J. Transp. Geogr.* **2022**, *102*, 103361. [CrossRef]
35. Cheng, L.; Yang, J.; Chen, X.; Cao, M.; Zhou, H.; Sun, Y. How could the station-based bike sharing system and the free-floating bike sharing system be coordinated? *J. Transp. Geogr.* **2020**, *89*, 102896. [CrossRef]
36. Bai, S.; Jiao, J. Using Deep Learning to Understand Travel Demands in Different Urban Districts. *Transp. Res. Rec.* **2021**, *2675*, 117–128. [CrossRef]
37. Cheng, L.; Wang, K.; De Vos, J.; Huang, J.; Witlox, F. Exploring non-linear built environment effects on the integration of free-floating bike-share and urban rail transport: A quantile regression approach. *Transp. Res. Part A Policy Pract.* **2022**, *162*, 175–187. [CrossRef]
38. Xie, C.; Lu, J.; Parkany, E. Work travel mode choice modeling with data mining: Decision trees and neural networks. *Transp. Res. Rec.* **2003**, *1854*, 50–61. [CrossRef]
39. Zhao, X.; Yan, X.; Yu, A.; Van Hentenryck, P. Prediction and behavioral analysis of travel mode choice: A comparison of machine learning and logit models. *Travel Behav. Soc.* **2020**, *20*, 22–35. [CrossRef]
40. Cheng, L.; De Vos, J.; Zhao, P.; Yang, M.; Witlox, F. Examining non-linear built environment effects on elderly's walking: A random forest approach. *Transp. Res. Part D Transp. Environ.* **2020**, *88*, 102552. [CrossRef]
41. Liu, J.; Wang, B.; Xiao, L. Non-linear associations between built environment and active travel for working and shopping: An extreme gradient boosting approach. *J. Transp. Geogr.* **2021**, *92*, 103034. [CrossRef]
42. Frank, L.D.; Pivo, G. Impacts of mixed use and density on utilization of three modes of travel: Single-occupant vehicle, transit, and walking. *Transp. Res. Rec.* **1994**, *1466*, 44–52.
43. Ding, C.; Cao, X.; Liu, C. How does the station-area built environment influence Metrorail ridership? Using gradient boosting decision trees to identify non-linear thresholds. *J. Transp. Geogr.* **2019**, *77*, 70–78. [CrossRef]
44. Ding, C.; Cao, X.; Næss, P. Applying gradient boosting decision trees to examine non-linear effects of the built environment on driving distance in Oslo. *Transp. Res. Part A Policy Pract.* **2018**, *110*, 107–117. [CrossRef]
45. Szaruga, E.; Załoga, E. Sustainable Development Programming of Airports by Identification of Non-Efficient Units. *Energies* **2022**, *15*, 932. [CrossRef]
46. Yue, M.; Ma, S. LSTM-Based Transformer for Transfer Passenger Flow Forecasting between Transportation Integrated Hubs in Urban Agglomeration. *Appl. Sci.* **2023**, *13*, 637. [CrossRef]
47. Slinker, B.; Glantz, S. Multiple regression for physiological data analysis: The problem of multicollinearity. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **1985**, *249*, R1–R12. [CrossRef]
48. Abdelsalam, O.; El-Masry, A.; Elsegini, S. Board composition, ownership structure and dividend policies in an emerging market: Further evidence from CASE 50. *Manag. Financ.* **2008**, *34*, 953–964.
49. Grubbs, F.E. Procedures for detecting outlying observations in samples. *Technometrics* **1969**, *11*, 1–21. [CrossRef]
50. Yin, C.; Cao, J.; Sun, B.; Liu, J. Exploring built environment correlates of walking for different purposes: Evidence for substitution. *J. Transp. Geogr.* **2023**, *106*, 103505. [CrossRef]
51. Lhéritier, A.; Bocamazo, M.; Delahaye, T.; Acuna-Agost, R. Airline itinerary choice modeling using machine learning. *J. Choice Model.* **2019**, *31*, 198–209. [CrossRef]
52. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
53. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
54. Goldstein, A.; Kapelner, A.; Bleich, J.; Pitkin, E. Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *J. Comput. Graph. Stat.* **2015**, *24*, 44–65. [CrossRef]
55. Greenwell, B.M.; Boehmke, B.C.; McCarthy, A.J. A Simple and Effective Model-Based Variable Importance Measure. *arXiv* **2018**, arXiv:1805.04755. [CrossRef]

56. Karhade, A.V.; Schwab, J.H.; Bedair, H.S. Development of machine learning algorithms for prediction of sustained postoperative opioid prescriptions after total hip arthroplasty. *J. Arthroplast.* **2019**, *34*, 2272–2277. [CrossRef]
57. Ghosh, S.M.; Behera, M.D. Aboveground biomass estimates of tropical mangrove forest using Sentinel-1 SAR coherence data-The superiority of deep learning over a semi-empirical model. *Comput. Geosci.* **2021**, *150*, 104737. [CrossRef]
58. Liu, L.; Dennett, A.; Hickman, R. Social equity impacts of rail transit investments using community perceptions: Evidence from Chongqing. *Transp. Plan. Technol.* **2023**, *46*, 1–21. [CrossRef]
59. Becker, H.; Ciari, F.; Axhausen, K.W. Measuring the car ownership impact of free-floating car-sharing–A case study in Basel, Switzerland. *Transp. Res. Part D: Transp. Environ.* **2018**, *65*, 51–62. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.