Transportation Research Part D 93 (2021) 102776

Contents lists available at ScienceDirect

Transportation Research Part D

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

Exploring nonlinear effects of the built environment on ridesplitting: Evidence from Chengdu

Meiting Tu [a,c,d], Wenxiang Li [b,*], Olivier Orfila [d], Ye Li [a], Dominique Gruyer [d]

[a] The Key Laboratory of Road and Traffic Engineering, Ministry of Education, 4800 Cao'an Road, Shanghai 201804, PR China
[b] Business School, University of Shanghai for Science and Technology, 516 Jungong Road, Shanghai 200093, PR China
[c] College of Automatic, University Paris-Saclay, 91190 Saint Aubin, Paris, France
[d] PICS-L, IFSTTAR, University Gustave Eiffel, 25 allée des Marronniers, 78000 Versailles, France

A R T I C L E I N F O

Keywords:
Ridesplitting
Ridesourcing
GBDT
Machine learning
Built environment
Travel behavior

A B S T R A C T

Ridesplitting, a form of ridesourcing services that matches riders with similar routes to the same driver, is a high occupancy travel mode that can bring considerable benefits. However, the current ratio of ridesplitting in the ridesourcing services is relatively low and its influencing factors remain unrevealed. Therefore, this paper uses a machine learning method, gradient boosting decision tree (GBDT) model, to explore the nonlinear effects of built environment on the ridesplitting ratio of origin–destination pairs (census tract to census tract). The GBDT model also provides the relative importance ranking of all the built environment factors. The results indicate that distance to city center, land use diversity and road density are the key influencing factors of ridesplitting ratio. In addition, the non-linear thresholds of built environment factors are identified based on partial dependence plots, which could provide policy implications for the government and transportation network companies to promote ridesplitting.

1. Introduction

On-demand ridesourcing services, operated by transportation network companies, match passengers and drivers through intelligent mobile phone applications (Rayle et al., 2016). Ridesourcing services have become increasingly popular due to their convenience. As of 2020, Uber has provided services in over 890 cities in 71 countries (Wyatt, 2020). As of 2018, Lyft has provided services in 200 cities in the US with over 315,000 drivers (Lyft, 2020). DiDi Chuxing has provided mobility services to more than 550 million users all over China. On average, there are more than 30 million travel orders on the DiDi Chuxing platform per day, representing more than 350 million vehicle kilometers per day nationwide (DiDi Chuxing, 2018).

Ridesplitting, one form of ridesourcing services, matches riders with similar routes to the same driver, such as "UberPool" and "Lyft Line" (Shaheen et al., 2016a). There are some debates claiming that on-demand ridesourcing services could increase congestion and pollution (Anair et al., 2020). However, ridesplitting, a new shared mobility service, is a more sustainable travel mode for improving traffic efficiency and reducing traffic congestion and air pollution problems. Policy makers have realized the benefits of ridesplitting services and encouraged ridesourcing companies to launch ridesplitting services for passengers. For DiDi Chuxing, the number of passengers who choose ridesplitting services has continued to increase since its launch at the end of 2015. By the end of 2019, the cumulative number of users reached 2.9 billion, with a compound annual growth rate of 143.3%. The passenger volume using

* Corresponding author.
E-mail address: liwx@usst.edu.cn (W. Li).

https://doi.org/10.1016/j.trd.2021.102776

Available online 8 March 2021
1361-9209/

ridesplitting services was equivalent to 1.2 times the civil aviation passenger volume in 2019 (DiDi Chuxing, 2020).

However, the current ridesplitting ratio in key ridesourcing services provided in the city of Chengdu is low, only 7% (Li et al., 2019). The average time usage rate of private cars in China is only 7% (Accenture, 2016), while the rate is only 4% in the United Kingdom (Bates & Leibling, 2012) and 17% in the United States (Santos et al, 2011). This low utilization rate leads to a serious waste of resources. With the development of technology for autonomous driving, idle private cars could be used for autonomous ridesourcing services or ridesplitting services in the future. Tu et al. (2019) showed that such ridesplitting services could potentially have great benefits in the city of Chengdu, specifically by improving the percentage of cost savings and time savings to 17% and 23%, respectively.

The existing literature has proven that the built environment has a strong relationship with travel behaviors (Ding et al., 2018; Ding et al., 2019; Durning and Townsend, 2015; Tao et al., 2020). However, to the best of our knowledge, no existing research has investigated how the built environment impacts ridesplitting services. To fill this gap, the aim of this study is to explore the relationship between the built environment and origin–destination (OD) ridesplitting ratio using observed ridesourcing data of Chengdu, China. Specifically, this study involves three main tasks. First, we use a machine learning method, gradient boosting decision tree (GBDT) model, to identify the important features of the built environment at the origins and destinations of ridesplitting services. Second, we explore the nonlinear relationship between the ridesplitting service and key explanatory variables by creating partial dependence plots. Third, we provide insightful results that can help transportation network companies improve existing ridesplitting services and have policy implications for urban planners seeking to better understand how the built environment, demographic factors and travel time impact ridesplitting services.

The remainder of this paper is organized as follows. Section 2 summarizes the related research. Section 3 presents the data and variables used in this study. The GBDT model is described in Section 4. Section 5 presents the results in detail and discusses the related insights regarding how the built environment features, demographic factors and travel time impact ridesplitting services. The final section summarizes the main findings, presents some policy implications and proposes future research directions.

2. Literature review

2.1. Studies on ridesourcing services

Most studies on ridesourcing focus on three specific aspects: (1) identifying the characteristics of passengers who use ridesourcing services and their preferences (Shaheen et al, 2016b; Nielsen et al., 2015; Wang et al., 2020); (2) exploring the impacts of ridesourcing services on passengers, such as waiting time (Rayle et al., 2014; Hughes & MacKenzie, 2016), drivers, such as income (Angrist et al., 2017; Chen et al., 2019; Hall et al., 2018), the environment (Kent, 2014), other traffic modes (Zhang & Zhang, 2018; Chan & Shaheen, 2012); and (3) optimizing ridesplitting operations, such as matching and dispatching strategies (Tang et al., 2019; TU et al., 2019). The literature suggests that ridesourcing services have been increasingly important in the entire transportation system. However, few studies have explored the relationship between ridesourcing and the built environment.

2.2. Studies on the relationship between the built environment and travel behaviors

Many studies that identify the relationship between the built environment and travel behaviors attract people's attention because the findings can be used to design our built environment better and thus make the whole transportation system more efficient. In these studies that are published in mainstream academic journals, the terms used to reflect built environment features often start with a "D". Cervero and Kockelman (1997) proposed the initial "three Ds": density (e.g., population), diversity (e.g., land use mixture), and design (e.g., road density or percentage of 4-way intersections). These features were extended to the "five Ds" by adding destination accessibility and distance to transit (Ewing & Cervero, 2001). Demand management (e.g., dynamic parking charges) is the sixth D, and demographics is the seventh D (Ewing & Cervero, 2010).

We summarized the literature to show the different effects of built environment features on various travel behaviors, which can be classified by different traffic modes, as shown in Table 1.

For rails and metros, Durning and Townsend (2015) find that population density, intersection density, and three main land use ratios, industrial/commercial/residential ratios, have significant positive associations with the ridership of rapid rail transit. Ding et al.

Table 1
Literature review exploring the relationship between built environment and travel behaviors.

Traffic modes	Travel behaviors	Methodologies
Rail/Metro	Ridership of rail rapid transit systems (Durning and Townsend, 2015)	Bootstrapped ordinary least squares regression model
	Metro ridership at station level (Ding et al., 2019)	GBDT
Bus	Bus ridership in Montreal (Chakour & Eluru, 2016)	Ordered regression model
Bike	Bike sharing demand in Toronto (El-Assi et al., 2017)	Ordinary Least Square (OLS) regression model
Walking	Walking distance to transit (Tao et al., 2020)	GBDT
Private vehicles	Driving distance of private vehicles in Oslo (Ding et al, 2018)	GBDT
Ridesourcing	The number of OD ridesourcing ridership in Chicago (Yan et al., 2020)	Random forest model
	Uber demand (Sabouri et al., 2020)	Mixed linear regression model
	Spatial variation of ridesourcing demand in Austin, Texas (Yu & Peng, 2019)	Geographically Weighted Poisson regression model

(2019) apply a GBDT model and show that the number of bus stops, compact and mixed land use development, and car ownership play a key role in determining station-level Metrorail ridership. With respect to buses, Chakour & Eluru (2016) find that transit facilities, the presence of parks and public transport services have a positive effect on ridership, while the presence of highways has a negative effect. For bikes, El-Assi et al. (2017) identify that network intersection density and bicycle infrastructure, such as bike lanes and path situations, have significant effects on bike-sharing demands. For walking, Tao et al. (2020) confirm that job density, population density, land use diversity, and pedestrian road density have strong nonlinear relationships with walking distance to transit. For private vehicles, Ding et al. (2018) identify the strong nonlinear relationship between driving distance and built environment factors, as well as demographics in Oslo.

For ridesourcing, Yan et al. (2020) find that socioeconomic and demographic variables, travel impedance, built environment characteristics, and transit-related factors have strong correlations with the ridesourcing demand in Chicago. Sabouri et al. (2020) reveal that Uber demand is positively correlated with the land use mix or entropy, total population and employment, activity density, and transit stop density of a census block group. In addition, intersection density and destination accessibility have negative impacts on Uber demand. Yu and Peng (2019) identify that there are strong relationships between ridesourcing demand and built environment factors such as density, land use, infrastructure, and transit accessibility in Austin, Texas, using a geographically weighted Poisson regression model.

In summary, the above studies have confirmed the strong relationship between the built environment and travel behaviors for different traffic modes, namely rail and metro, bus, bike, walking, private vehicles and ridesourcing services. However, to the best of our knowledge, there is no existing research exploring the relationship between ridesplitting and the built environment, which should be investigated. By exploring the relationship between the built environment and ridesplitting, urban planners can better understand how the features of the built environment impact ridesplitting and improve traffic efficiency by way of a more reasonable design and planning of built environment.

2.3. Models used in related studies

Most studies use statistical modeling to explore the relationship between the built environment and travel behaviors, whereas it should be noted that modeling is just one way to explore the relationship. The widely used models can be summarized as follows: ordinary least squares (OLS) models; discrete choice models, such as logit models and probit models; and geographically weighted regression (GWR) models. For example, Durning and Townsend (2015) applied a bootstrapped OLS regression model to identify the association between station boardings of the Canadian rapid rail transit and 44 socioeconomic, built environment, and system attributes. Alemi et al. (2018) used a binomial logit model to analyze the key factors, such as the built environment and travel distance, affecting the travel demand of ridesourcing services using comprehensive survey data on 2400 residents in California. Choudhury et al. (2018) applied a nested logit model using data from the Stated Preference (SP) survey in Lisbon, Portugal to analyze the influencing factors of travelers' departure times and ridesharing mode choices. However, OLS and discrete choice models can be used for only global estimations that do not consider spatial heterogeneity (Brunsdon et al., 1996).

The GWR is a regression model that uses local estimation rather than global estimation to address spatially nonstationary processes. Specifically, GWR involves the generation of a regression for each location rather than using a global regression for the entire research area. GWR models have been used to analyze the relationship between the built environment and travel behaviors. For example, Torun et al. (2020) applied a logistic GWR model to explore how urban form factors and sociodemographic factors impact transportation walking. However, this kind of model assumes a prior hypothesis of a particular function, such as a linear or log linear relationship, which could lead to modeling biases.

Due to the development of new technology for big data analyses, increasingly more intelligent algorithms are used to explore the characterization of travel behaviors. In recent years, machine learning models such as GBDT models have performed well when applied to explore the characterization of different transport modes. Friedman developed the first GBDT model, which could provide highly robust, outstanding, and interpretable regression and classification problems (Friedman, 2001). Gan et al. (2020) used a GBDT model to examine the nonlinear effects of the built environment on OD metro ridership in Nanjing, China; this model performed better than the multiplicative model. Ding et al. (2018) employed a GBDT model to investigate the relationship between the built environment and driving distance in Oslo. GBDT is a machine learning method without a prior hypothesis of a particular function and can predict the target variable more accurately, even the nonlinear relationships with explanatory variables.

3. Data and variables

3.1. Study area

Chengdu is the capital of Sichuan Province and is one of the three most populous cities in Western China. As of 2020, the total area in Chengdu was 14,335 square kilometers, the local population was 16.58 million, the urban population was 12.34 million, and the urbanization rate was 74.42% (Chengdu Bureau of Statistics, 2020).

Since the launch of ridesplitting services on the DiDi Chuxing platform in 2015, the number of passengers who chose ridesplitting services has continued to increase. Chengdu is one of the leading cities that actively advocate for and promote ridesplitting services. Chengdu government keeps taking active part in developing the ridesourcing and ridesplitting services with Didi Chuxing to mitigate the congestion level and air pollution problems. In 2018, the users of ridesplitting services provided by the DiDi platform shared more than 800 million kilometers with other riders across China, which is equivalent to reducing fuel consumption by 43 million liters and reducing CO_2 emissions by 97,000 tons (DiDi Chuxing, 2018). For 2018 and 2019, the DiDi Chuxing platform reduced carbon dioxide emissions across China by a total of 1.303 million tons, and Chengdu city ranked first. The ridesharing and ridesplitting services provided by this platform reduced carbon dioxide emissions by 913 thousand tons and 158 thousand tons, respectively (DiDi Chuxing, 2020). Thus, we choose the municipal districts of Chengdu as the study area, as shown in Fig. 1.

3.2. Data sources

Ridesourcing order and GPS data. These data were obtained from the DiDi platform through the DiDi GAIA Initiative, which compiles an open dataset (DiDi Chuxing, 2017). The dataset includes the complete trajectory and order data of DiDi Express and DiDi Premier Services in Chengdu, China, from November 1st to November 30th, 2016. The fields of the trajectory dataset are anonymous driver ID, order ID, longitude, latitude, and timestamp, with an average interval time of 3 s. The fields of the order dataset are order ID, pick-up and drop-off locations, start timestamp and end timestamp. We cleaned the data and removed outliers such as order data with travel times less than 120 s or more than 7200 s. Furthermore, we divide the ride orders into single rides and shared rides based on the ridesplitting trip identification algorithm proposed by Li et al. (2019). Then, the ridesplitting ratio of each OD pair can be calculated.

GIS layers. We divide the municipal districts of Chengdu into 162 census tracts based on the administrative boundaries and obtain the road networks data from OpenStreetMap. The city center is Tianfu Square (Yu & Gao, 2011), which is not only Chengdu's economic, cultural and commercial center but also a comprehensive transportation hub.

Built environment data. The built environment is measured by the four Ds, namely, density, design, land use diversity and distance to the city center. Density measures include population density, local population density, park density, public transportation density, public services density, point of interest (POI) density, and the ratios of industrial, residential and commercial land use (Li et al., 2017). Population density and local population density are from the sixth nationwide population census survey in China. Other density factors are calculated by dividing the number of corresponding facilities by the area. The number of corresponding facilities is obtained by using the POI search API of the Gaode Map Web Services (Gaode Map, 2021). Public service density is measured by the number of health care, government, and education institutions per unit area. Public transportation density is measured by the number of airports, train stations, metro stations and bus stops per unit area. The commercial ratio is measured by the number of financial institutions, shopping places, restaurants, entertainment facilities, and hotels per unit area. Design is measured by the road density, which is calculated by dividing the road length by the area. Land use diversity is measured by the entropy index, which ranges between 0 and 1, where 0 indicates that the land use is homogenous, and 1 indicates that all land use types are equally distributed (Cervero and Kockelman, 1997).

Demographic data. Demographic characteristics are employed to allow for the exploration of the effects of the demographic variables on the ridesplitting ratio. These variables include the population density who are younger than 14 years old and older than 64 years old, per capita gross domestic product (PGDP), and average housing price. They are measured at the scale of census tract. The population density of individuals younger than 14 years old and older than 64 years old is calculated by dividing the population of

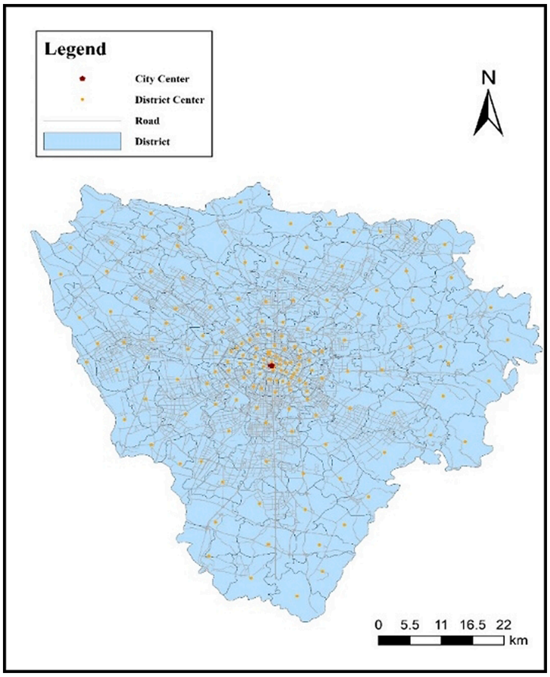

Fig. 1. Study area.

individuals younger than 14 years old and older than 64 years old by the area. PGDP data and average housing prices are obtained from the sixth nationwide population census survey in China and the Lianjia website in Chengdu (http://cd.lianjia.com/).

Travel time is used to explore the effects of travel impedance factors on the ridesplitting ratio and is an important factor for passengers deciding whether to choose ridesplitting services. It is the average travel time of all OD pairs from the DiDi order dataset.

3.3. Variables description

Based on the data available, we calculate the dependent variable and several explanatory variables at a census tract level. The ridesplitting ratio of each OD pair (census tract to census tract) is chosen as the dependent variable. Then, we divide the explanatory variables into 4 categories, namely, the built environment at the origin locations, the built environment at the destination locations, demographic factors and travel time. Thus, a descriptive analysis of all the variables was carried out, as shown in Table 2. The spatial distributions of some explanatory variables are shown in Fig. 2 .

4. Methodology

We choose the built environment features that are measured by the 'four Ds', namely, density, land use diversity, design, and distance to city center. Then, we use the GBDT model to predict the OD ridesplitting ratio and investigate the relative importance ranking of all the explanatory variables. Furthermore, we explore the nonlinear relationship between the ridesplitting ratio and key explanatory variables by generating partial dependence plots. We find some useful thresholds from the partial dependence plots, which could provide a reference for policy makers.

4.1. GBDT model

A GBDT model is an additive model that uses decision trees *f(x)* to estimate the function *F(x)*. The objective of the algorithm is to minimize the loss function $L(\mathrm{y}, F(\mathrm{x})) = (y - F(x))^2$.

$$F(x) = \sum_{m=1}^{M} f_m(x) = \sum_{m=1}^{M} \beta_m h(x; a_m) \tag{1}$$

where × is a set of explanatory variables (i.e., built environment features, demographic factors and travel time) and F(x) is an approximation function of the dependent variable (i.e., OD ridesplitting ratio). a_m is the mean of split locations and the terminal node for each splitting variable for each tree. M represents the m$^{\mathrm{th}}$ iteration.

Table 2
Variable definitions and statistics.

Variables	Variable Description	Mean	S.D.	Min	Max
Orders					
Ridesplitting orders	Number of ridesplitting orders from November 1st to 30th in 2016	26.58	63.81	0	1597
Ridesourcing orders	Number of ridesourcing orders from November 1st to 30th in 2016	431.71	984.20	1	33,477
Ridesplitting ratio	Number of ridesplitting orders/number of ridesourcing orders	21.85%	33.32%	0	99.94%
Built environment					
Population density	Population/area size (person per km^2)	11142.95	14545.89	155.07	67557.29
Local population density	Local population/area size (person per km^2)	6132.88	8789.07	151.72	41861.45
Park density	Number of parks and square/area size (park per km^2)	0.60	0.85	0	5.11
Industrial ratio	Number of industrial locations/number of POI	3.22%	5.03%	0	39.74%
Transportation density	Number of transportation facilities/area size (facility per km^2)	5.14	4.41	0.33	17.75
Residential ratio	Number of residential locations/number of POI	8.78%	5.56%	0	22.78%
Commercial ratio	Number of commercial locations/number of POI	67.66%	13.52%	28.57%	91.76%
Public density	Number of public service facilities/area size (facility per km^2)	29.43	41.48	0.30	174.02
POI density	Number of POI/area size (facility per km^2)	221.96	328.50	0.46	1699.44
Land use diversity	Entropy index of the land use mix	0.80	0.10	0.44	0.99
Area	Area of every tract (km^2)	22.70	20.08	0.64	82.31
Road density	Length of the road/area size (km/km^2)	4.28	3.01	0.22	12.28
Distance to city center	Straight line distance from city center (km)	20.92	14.50	1.07	54.80
Demographic characteristics					
Population density (less than 14 years old)	Population less than 14 years old/area size (person per km^2)	1000.03	1244.20	25.20	6588.39
Population density (more than 64 years old)	Population more than 64 years old/area size (person per km^2)	1131.36	1697.28	20.64	9473.61
PGDP	Gross domestic product/population (10^4 RMB per capita)	10.79	2.87	7.71	15.75
Housing price	Average price per square meter (RMB/m^2)	13656.02	4951.05	4661	31,245
Travel Impedance variables					
Travel time	Average travel time for OD ridesourcing orders (s)	2178.00	1339.14	120	7200

Note: Samples: 14,121 OD pairs.

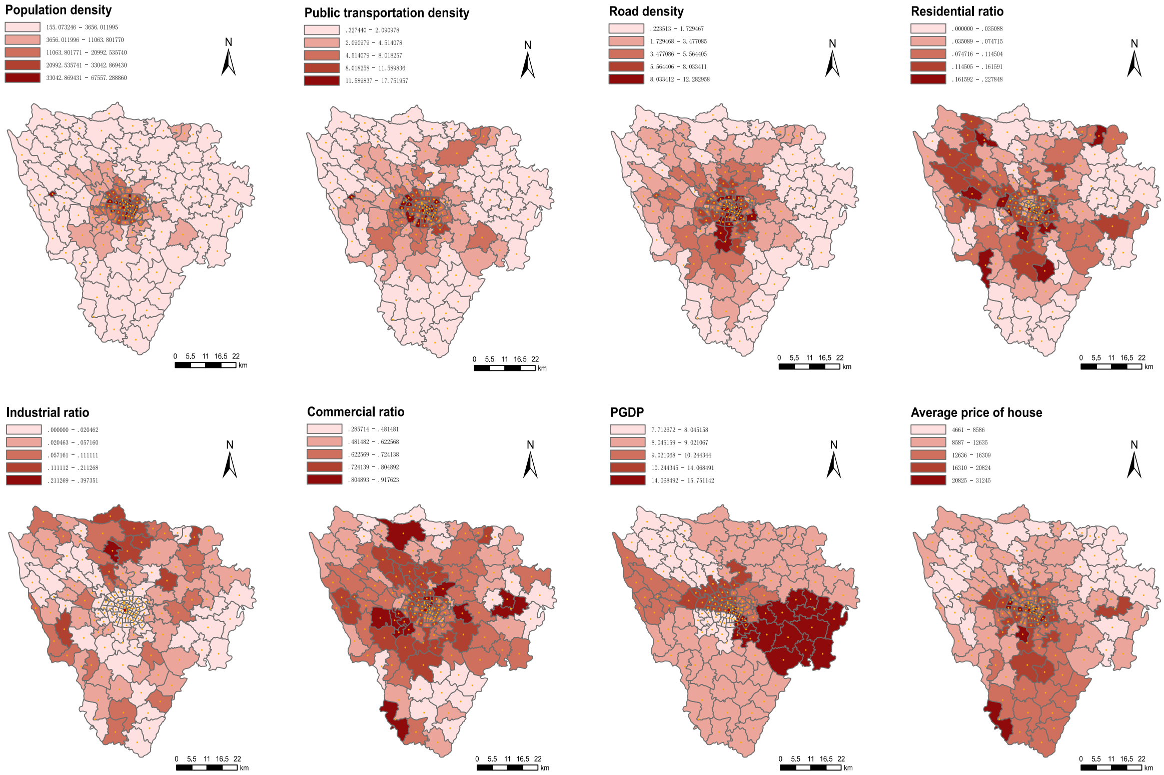

Fig. 2. Distributions of some explanatory variables.

The function F(x) is updated according to the gradient descent direction as follows:

$$F_m(x) = F_{m-1}(x) + \beta_m h(x; a_m)F_m(x) = F_{m-1}(x) + \beta_m h(x; a_m) \tag{2}$$

$$\beta_m = \underset{\beta}{argmin} \sum_{i=1}^{n} L(\mathrm{y}_i, F_{m-1}(\mathrm{x}_i) + \beta h(\mathrm{x}_i; a_m)) \tag{3}$$

where β_m is expected to minimize the value of the loss function.

A shrinkage parameter ξ ($0<\xi<1$), also called the learning rate, can be used to avoid the overfitting problem (Friedman, 2001) by means of scaling the contributions of each tree.

$$F_m(x) = F_{m-1}(x) + \xi^* \beta_m h(x; a_m) \tag{4}$$

4.2. Relative importance and partial dependence of the explanatory variables

Compared to most machine learning algorithms regarded as 'black-box' algorithms, the GBDT model can be used to determine the relative importance ranking of all explanatory variables. The relative importance of variable $x_i$can be obtained as follows:

$$I_{X_i}^2 = \frac{1}{M} \sum_{m=1}^{M} I_{X_i}^2 (T_m) \tag{5}$$

$$I_{X_i}^2 (T_m) = \sum_{j=1}^{J-1} d_j \tag{6}$$

where J is the number of leaves on each tree and $d_j$represents the improvement in the squared error term by making the jth split based on the variable x_i.

The GBDT model can also generate a partial dependence plot to illustrate the relationships between the explanatory variables and dependent variables. The partial dependence plot shows the marginal effect of one or two features on the predicted response variable (Friedman, 2001). The partial dependence of F(x) on x_S can be defined as follows (Hastie et al., 2009):

$$F_{x_S}(x_S) = E_{x_C}[F(x_S, x_C)] = \int F(x_S, x_C)dP(x_C) \tag{7}$$

$$F_{x_S}(x_S) = \frac{1}{n}\sum_{i=1}^{n} F(x_S, x_C^i) \tag{8}$$

wherex_S are the features that we want to know the specific effect on the OD ridesplitting ratio and x_C are the other features used in the study. n represents the number of samples.

5. Results and discussion

5.1. Performance of the GBDT model

We use a five-fold cross-validation procedure to obtain the optimal parameter settings and a robust result. We fit the models with different numbers of trees (5000, 10000, 15000), shrinkage (0.005, 0.05, 0.01, 0.1) and tree complexity (1, 3, 5, 7) based on the experimental findings of previous studies. We fit the model 240 times and find that the best performance is obtained when the number of trees, shrinkage and tree complexity are set as 5000, 0.005 and 7, respectively.

As shown in Table 3, the values of testing pseudo-R^2, mean absolute error (MAE) and mean squared error (MSE) for the GBDT model are 0.7415, 0.1070 and 0.0293, respectively. The corresponding values for the traditional linear regression model are 0.3449, 0.2104 and 0.0727. The pseudo-R^2 is improved by 114.99%, and the MAE and MSE are decreased by 49.14% and 59.70%, respectively.

5.2. Relative importance of the explanatory variables

Fig. 3 shows the relative importance of the explanatory variables and their rankings in predicting the OD ridesplitting ratio. The total importance of all explanatory variables equals 100% because it is measured in a relative way.

As shown in Table 4, the built environment features at the origin locations contribute 29.97%, whereas the contribution of the features at the destination locations is 23.98%. The built environment features at the origin locations have larger effects on the OD ridesplitting ratio than the destination locations. In addition, the collective contribution of the demographic factors is 35.31%. Moreover, travel time contributes 10.74%, which is also an important feature.

For the built environment features at the origin locations, distance to city center, road density, commercial ratio, land use diversity, and public transportation density are the top five important explanatory variables, which contribute to 10.35%, 2.75%, 2.56%, 2.20% and 1.92%, respectively. The three least important features at the origin locations are local population density, population density and public service density, which contribute to 0.59%, 0.75% and 0.90%, respectively.

For the built environment features at the destination locations, distance to city center, public service density, road density, public transportation density, and land use diversity are the five variables with the highest relative importance, which contribute to 9.26%, 3.92%, 1.97%, 1.83% and 1.24%, respectively. The three least important features at the destination locations are local population density, park density and population density, which contribute to 0.41%, 0.55% and 0.55%, respectively.

For the demographic factors, PGDP at the origin locations is the most important factor in all the explanatory variables, accounting for 22.56%. The average price of the house at the origin locations contributes to 6.07%, which is the second most important demographic feature to predict the ridesplitting ratio. The least important demographic factor is the population density of people who are more than 64 years old, which contributes to 0.42%.

5.3. Nonlinear effects of the key explanatory variables

5.3.1. Features of the built environment

Fig. 4 demonstrates the effects of four key built environment variables of the origin locations on the ridesplitting ratio. Fig. 4(a) shows that the distance to the city center at the origin locations has a positive association with the OD ridesplitting ratio. People are more likely to use ridesplitting services when they depart from areas far away from the city center than those near the city center. People are reluctant to use ridesplitting services when the distance to the city center at the origin locations is less than 5 km. However, the OD ridesplitting ratio increases substantially when the distance to the city center is more than 32 km.

Fig. 4(b) displays the relationship between public transportation density at the origin locations and the OD ridesplitting ratio, which shows a significant nonlinear pattern: first decreasing then steady and then increasing slightly. The OD ridesplitting ratio decreases with increasing public transportation density at the destination locations from 0 to 6 per km^2, and the ridesplitting ratio differs

Table 3
Comparison of the GBDT model and basic linear regression model.

Metrics	GBDT	Linear regression	Percentage of improvement
R^2	0.7415	0.3449	+114.99%
MAE	0.1070	0.2104	−49.14%
MSE	0.0293	0.0727	−59.70%

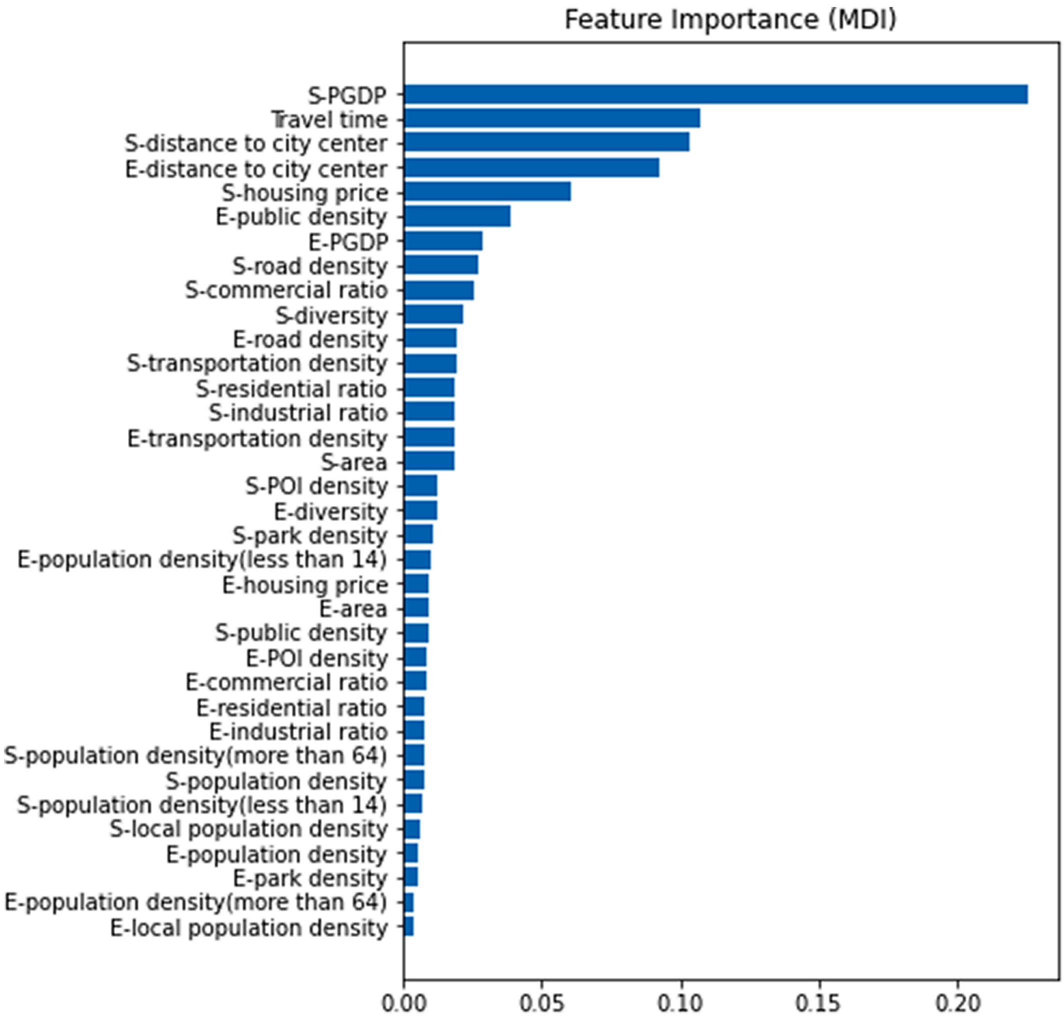

Fig. 3. The importance ranking of all the explanatory variables, Note: 'S-' represents features at the origin locations and 'E-' represents features at the destination locations.

by approximately 4%. The ridesplitting ratio remains steady when the public transportation density is between 6 and 13 per km^2. The ridesplitting ratio increases slightly when the public transportation density at the origin locations is greater than 13 per km^2. The above results show that the ridesplitting ratio is relatively high in the districts where public transportation density is less than 6 per km^2. Therefore, transportation network companies should provide enough ridesplitting services to satisfy ridesplitting demand and guarantee the service quality in areas with low accessibility to public transportation. This threshold could also provide a reference for the government to achieve a good balance of ridesplitting services and public transportation services.

Fig. 4(c) demonstrates that the road density at the origin locations is negatively correlated with the OD ridesplitting ratio. First, the ridesplitting ratio remains high and steady when the road density is less than 2.5 km/km^2. Then, the ratio decreases with an increase in road density between 2.5 km/km^2 and 6 km/km^2. The OD ridesplitting ratio remains almost stable when road density exceeds 6 km/km^2.

Fig. 4(d) generally shows a positive association between land use diversity at the origin locations and the ridesplitting ratio. Land use diversity at the origin locations has almost no effect when it is less than 0.8. However, the OD ridesplitting ratio increases substantially when it reaches 0.85. Then, the ratio remains constant when the land use diversity exceeds 0.85.

To summarize the above results in the origin locations, we can say that the ridesplitting ratio is higher when:

(a) The origin area is close to the city center (>32 km)
(b) The public transportation services at the origin locations have a low density (<6.0 number/km^2)
(c) The road density at the origin locations is low (<4.0 km/km^2)
(d) The land use diversity at the origin locations is high (>0.85)

Table 4
The relative importance ranking of all the explanatory variables.

Built environment features at origin locations	Relative importance	Ranking
S-distance to city center	10.35%	3
S-road density	2.75%	8
S-commercial ratio	2.56%	9
S-diversity	2.20%	10
S-transportation density	1.92%	12
S-residential ratio	1.89%	13
S-industrial ratio	1.87%	14
S-area	1.82%	16
S-POI density	1.26%	17
S-park density	1.10%	19
S-public density	0.90%	23
S-population density	0.75%	29
S-local population density	0.59%	31
Sum	29.97%	–
Built environment factors at destination locations	Relative importance	Ranking
E-distance to city center	9.26%	4
E-public density	3.92%	6
E-road density	1.97%	11
E-transportation density	1.83%	15
E-diversity	1.24%	18
E-area	0.95%	22
E-POI density	0.88%	24
E-commercial ratio	0.83%	25
E-residential ratio	0.79%	26
E-industrial ratio	0.78%	27
E-population density	0.55%	32
E-park density	0.55%	33
E-local population density	0.41%	35
Sum	23.98%	–
Demographic factors	Relative importance	Ranking
S-PGDP	22.56%	1
S-housing price	6.07%	5
E-PGDP	2.84%	7
E-population density (less than 14)	1.04%	20
E-housing price	0.96%	21
S-population density (more than 64)	0.77%	28
S-population density (less than 14)	0.66%	30
E-population density (more than 64)	0.42%	34
Sum	35.31%	–
Travel impedance feature	Relative importance	Ranking
Travel time	10.74%	2

Fig. 5 demonstrates the effects of four key built environment variables of the destination locations on the ridesplitting ratio. The partial dependence plot for the distance to the city center at the destination locations is presented in Fig. 5(a). The plot shows that the distance to the city center at the destination locations is positively associated with the OD ridesplitting ratio. First, the difference in the ridesplitting ratio could reach 20% between the destinations near the city center and those 33 km away from the city center. Distance to the city center at the destination locations has almost no effect when it is less than 20 km. Within this range, the ridesplitting ratio only differs by approximately 5%. However, the OD ridesplitting ratio increases substantially when it exceeds 20 km.

Fig. 5(b) indicates that public transportation density at the destination locations is negatively associated with the ridesplitting ratio, which is similar to the relationship between the public transportation density at the origin locations and the ridesplitting ratio. The OD ridesplitting ratio decreases with increasing public transportation density at the destination locations ranging from 0 to 2 per km^2, and the ridesplitting ratio differs by approximately 5%. The ridesplitting ratio remains steady when the public transportation density is between 2 and 4 per km^2. The ridesplitting ratio decreases slightly when the public transportation density at the destination locations is 4 per km^2 and remains steady beyond this range. The above results show that the ridesplitting ratio is higher in districts where the public transportation density is lower than 2 per km^2.

Fig. 5(c) demonstrates that the road density at the destination locations is negatively correlated with the OD ridesplitting ratio, which is similar to the relationship between the road density at the origin locations and the ridesplitting ratio. First, the ridesplitting ratio remains high and steady when the road density is less than 2.5 km/km^2. Then, the ratio decreases with an increase in the road density between 2.5 km/km^2 and 4 km/km^2. The OD ridesplitting ratio remains almost stable when the road density exceeds 4 km/km^2. Fig. 4(c) and Fig. 5(c) demonstrate that road density is negatively correlated with the ridesplitting ratio at both the origin and destination locations. This result can be explained by other studies finding that road density has a positive correlation with public

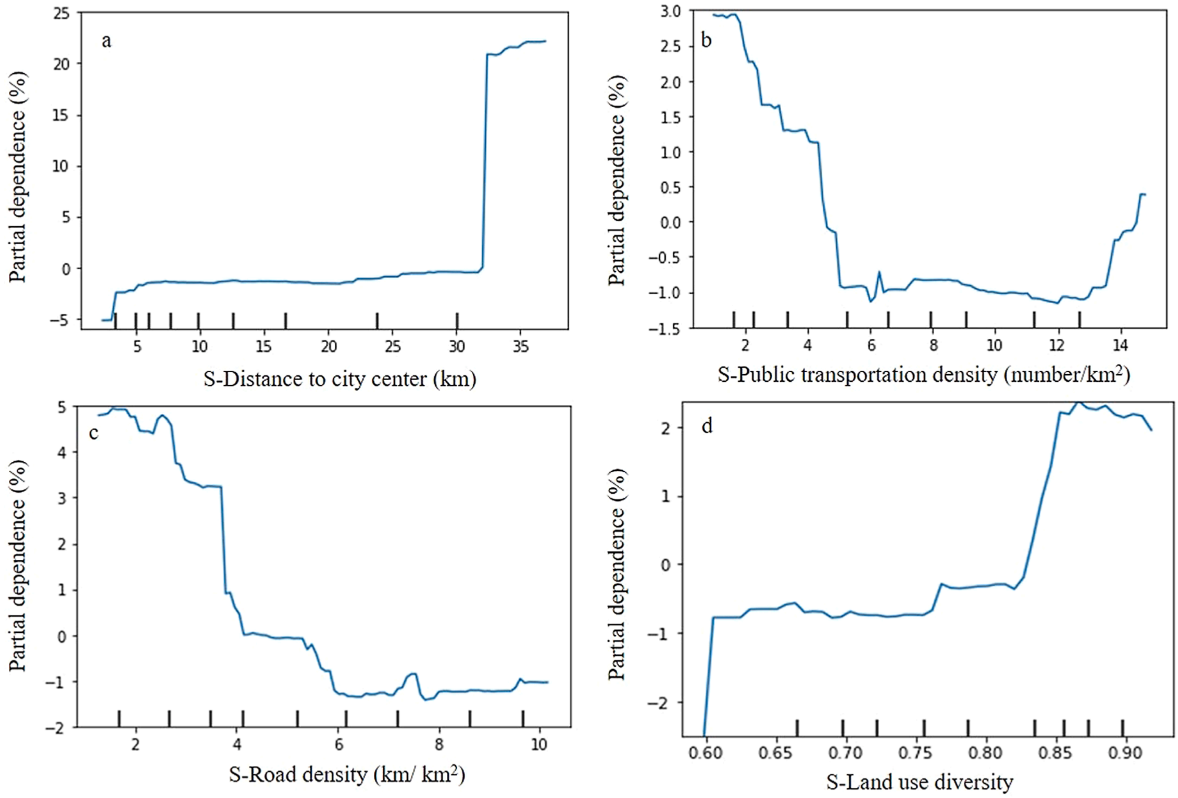

Fig. 4. The effects of the key built environment variables of the origin locations on the ridesplitting ratio.

transportation services such as buses and metros (Tu et al., 2018; Zhao et al., 2014; Gan et al., 2020). Therefore, the share of ridesplitting services can be replaced by public transportation services in areas with high road density.

Fig. 5(d) shows a positive association between the land use diversity at the destination locations and the ridesplitting ratio in general. The land use diversity at the origin locations has almost no effect when it is less than 0.65. However, the OD ridesplitting ratio keeps increasing when the land use diversity increases from 0.65 to 0.9.

To summarize the above results in the destination locations, we can say that the ridesplitting ratio is higher when:

(a) The destination area is far from the city center (>20 km)
(b) The public transportation services at the destination locations have a low density (<4 number/km^2)
(c) The road density at the destination locations is low (<4.0 km/km^2)
(d) The land use diversity at the destination locations is high (>0.85)

To be specific, we choose two typical census tracts for further analysis. In census tract No. 77, the distance to the city center is 44.39 km, the density of public transportation is 0.8 number/km^2, the road density is 1.70 km/km^2 and the land use diversity is 0.93. For all ridesourcing trips departing from the No. 77 census tract, the average real ridesplitting ratio is 74.66%. For all ridesourcing trips arriving at No. 77 census tract, the average real ridesplitting ratio is 53.00%. In census tract No. 137, the distance to the city center is 7.41 km, the density of public transportation is 14.93 number/km^2, the road density is 8.03 km/km^2 and the land use diversity is 0.72. The number of ridesourcing trips departing from No.137 census tract is 81054, and the average real ridesplitting ratio is zero. For all ridesourcing trips arriving at No.137 census tract, the average real ridesplitting ratio is 19.61%. The real ridesplitting situation related to these two census tracts is consistent with the result provided by the proposed model.

5.3.2. Other explanatory variables

Fig. 6 demonstrates the effects of key demographic variables on the OD ridesplitting ratio. Fig. 6(a) displays the relationship between PGDP at the origin locations and OD ridesplitting ratio. The partial dependence plot indicates that the ridesplitting ratio is relatively low when the PGDP is less than 100,000 RMB/year. However, the OD ridesplitting ratio increases substantially when the PGDP increases to 100,000 RMB/year. The ridesplitting ratio remains steady from 100,000 to 125,000 RMB/year and then decreases slightly between 125,000 and 140,000 RMB/year. Then, the ridesplitting ratio increases from 140,000 to 155,000 RMB/year. The difference in the ridesplitting ratio can reach 30% between areas where the PGDP is 155,000 RMB/year and those where the PGDP is approximately 90,000 RMB/year.

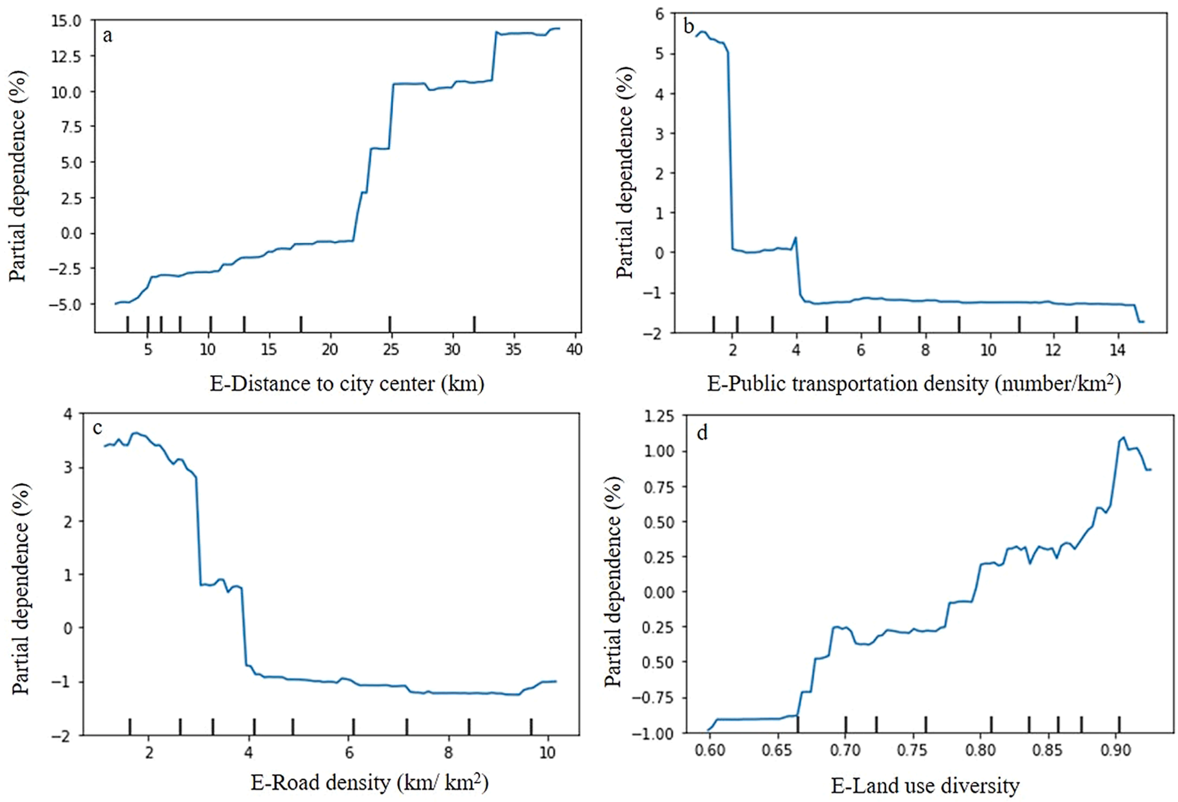

Fig. 5. The effects of the key built environment variables of the destination locations on the ridesplitting ratio.

Fig. 6(b) demonstrates the relationship between the PGDP at the destination locations and the OD ridesplitting ratio. The per capita GDP at the destination locations has little effect on the ridesplitting ratio.

The partial dependence plot for the average price of housing at the origin locations is presented in Fig. 6(c). In general, this plot indicates that the average price of housing is negatively correlated with the OD ridesplitting ratio. The OD ridesplitting ratio decreases when the average price of housing at the origin locations increases from 6000 to 10,000 RMB per m^2. This result indicates that people who live in houses with relatively low prices prefer to choose ridesplitting services. This is consistent with reality because these homeowners want to use ridesplitting services due to the low price. Fig. 6(d) demonstrates that the average price of housing at the destination locations has little effect on the ridesplitting ratio.

The partial dependence plot demonstrates the relationship between travel time and OD ridesplitting ratio in Fig. 7. The results show that travel time has a significant and positive effect on the OD ridesplitting ratio in general. A subpeak exists when the travel time is approximately 20 min. Travel time has almost no effect when it is less than approximately 35 min, while OD ridesplitting ratio increases gradually when travel time exceeds 35 min. Beyond this range, people are more likely to choose the ridesplitting service with the increase of the travel time. Furthermore, the ridesplitting ratio remains nearly the same when the travel time increases to 70 min.

5.4. Multi-predictor partial dependence plots

Fig. 8 demonstrates the combined impacts of the residential ratio and commercial ratio of the origin locations on the OD ridesplitting ratio. The result indicates that when the residential ratio ranges from 0.16 to 0.18 and the commercial ratio ranges from 0.55 to 0.65, the ridesplitting ratio reaches its highest value. This result could provide a significant reference for both policy makers such as urban planners and commercial managers such as transportation network companies.

Fig. 9 depicts the impacts of POI density and land use diversity of the origin locations on the OD ridesplitting ratio. The OD ridesplitting ratio increases with increasing land use diversity at the origin locations. When the land use diversity increases to 0.85, the ridesplitting ratio increases by 5%. The result indicates that when land use diversity ranges from 0.85 to 1 and the POI density is less than 600 per km^2, the ridesplitting ratio remains relatively steady and high.

6. Conclusions and implications

To explore the effects of built environment features, demographic factors and travel time on the OD ridesplitting ratio, this paper employs a GBDT model using observed ridesplitting data of Chengdu obtained from DiDi's open data project. GBDT is a machine

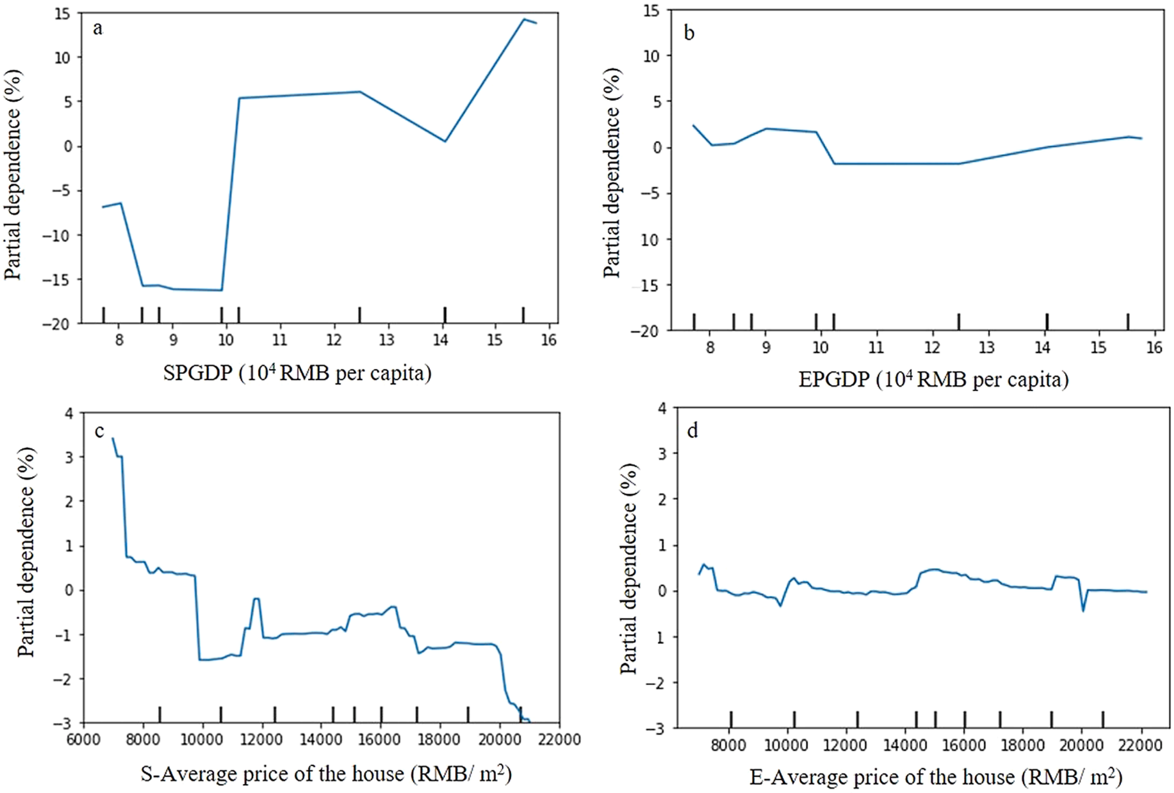

Fig. 6. The effects of the key demographic variables on the OD ridesplitting ratio.

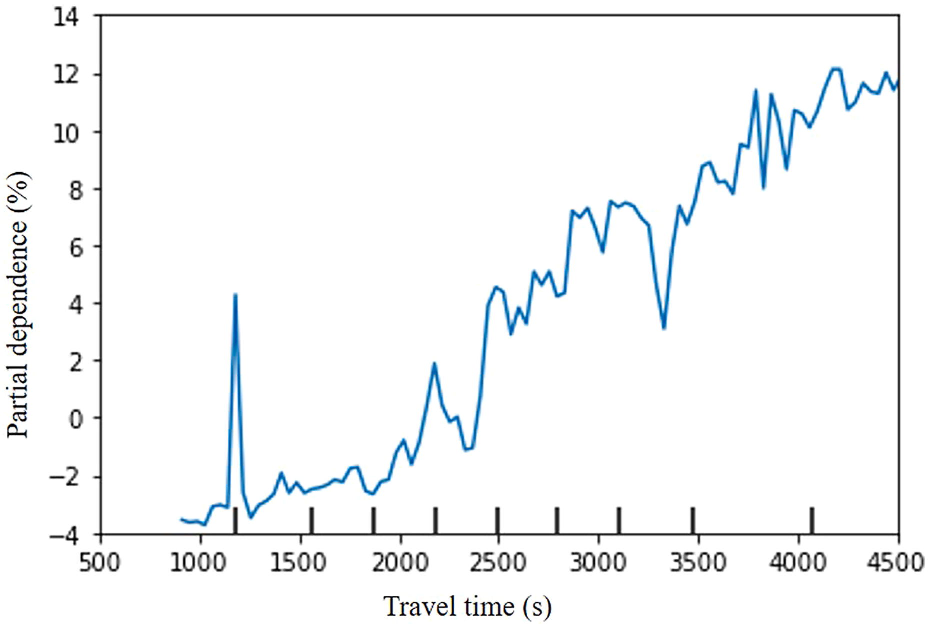

Fig. 7. The effects of the travel time on the OD ridesplitting ratio.

learning method without a prior hypothesis of a particular function such as a linear or log linear relationship. The GBDT model can be used to predict the OD ridesplitting ratio and even the nonlinear relationships with the explanatory variables more accurately. The results of this study verify that the performance of the GBDT model is much better than that of the traditional linear regression model.

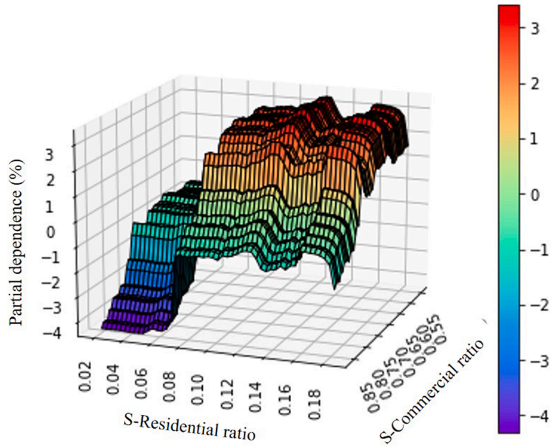

Fig. 8. The combined effects of S-residential ratio and S-commercial ratio on the OD ridesplitting ratio.

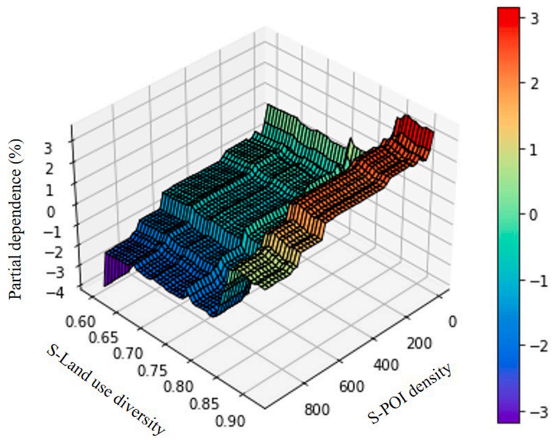

Fig. 9. The combined effects of S-POI density and S-land use diversity on the OD ridesplitting ratio.

Four categories of explanatory variables are used in this study: the built environment at the origin locations and destination locations, demographic factors and travel time. The GBDT model provides the relative importance ranking of all the explanatory variables. The results indicate that the built environment at the origin locations has a larger effect on the OD ridesplitting ratio than that of the destination locations. The collective influence of the built environment is larger than that of demographics and travel time. The built environment at the origin locations contributes 29.97%, whereas the contributions of the features at the destination locations are 23.98%. In addition, the collective contribution of the demographic factors is 35.31%. Travel time contributes to 10.74%.

For the built environment features at the origin locations, distance to city center, road density, commercial ratio, land use diversity, and public transportation density are the five most important explanatory variables. For the built environment features at the destination locations, distance to city center, public service density, road density, public transportation density, and land use diversity are the five most important variables. It can be found that distance to city center, land use diversity, road density and public transportation density rank highest in terms of importance for both the origin and destination locations, indicating that they play key roles in impacting the ridesplitting ratio.

Distance to the city center has a positive association with the ridesplitting ratio at both the origin locations and destination locations. People are more likely to use ridesplitting services when they depart from areas far away from the city center than those near the city center. The difference in the ridesplitting ratio could reach 20% between the destinations near the city and those 33 km away from the city center. In addition, road density is identified to be negatively correlated with the ridesplitting ratio at both the origin and destination locations. In general, land use diversity is positively associated with the ridesplitting ratio both at the origin locations and at the destination locations. Therefore, urban planners should pay more attention to the planning of land use and residential/commercial/industrial ratios, and develop reasonable designs based on the specific thresholds identified in the partial dependence plots. The insightful findings could enhance their understanding of ridesplitting services and be used for land use planning to design a ridesplitting-friendly city.

Public transportation density also plays an essentially important role in the impact on the ridesplitting ratio at both the origin and destination locations. A negative association between public transportation and the ridesplitting ratio can be seen both at the origin locations and at the destination locations. The results show that the ridesplitting ratio is relatively high in the districts where public transportation density is less than 4 per km^2. The government could set a threshold of the available ridesplitting services nearby for transportation network companies to guarantee the service quality, especially in areas with low accessibility to public transportation. This threshold could also provide a reference for the government to achieve a good balance of ridesplitting services and public transportation services.

With respect to demographics, PGDP and the average price of housing at the origin locations are the two most important factors impacting the ridesplitting ratio. The results demonstrate that a positive association exists between the PGDP at the origin locations and the ridesplitting ratio, whereas a negative association exists between the average price of housing at the origin locations and the ridesplitting ratio.

The findings of this study could help to better understand how the built environment, demographic factors and travel time impact ridesplitting services and provide policy implications for decision makers. For urban planners, the insightful results could enhance their understanding of ridesplitting services and be used for land use planning, design, and related features. For the transportation network companies, the results could help them optimize their long-term ridesplitting capacity to satisfy ridesplitting demand and improve the transportation efficiency. These companies should provide enough ridesplitting services in areas with high ridesplitting demand to guarantee service quality and reduce ridesplitting capacity in areas with low ridesplitting demand to reduce empty travel by vehicles.

The GBDT model provides a more reasonable way to explore the nonlinear relationship between the built environment features and ridesplitting ratio without hypothesizing a prior function. However, this study does have some limitations, which could be improved by further research in the future. In this study, we use ridesourcing data for only one city and only one month. The proposed method can be easily applied to other cities. We need more data from other cities and other time durations to compare and validate the relationship between the features of the built environment and the ridesplitting ratio. Furthermore, we consider only the effects of the built environment features, demographic factors and travel time on ridesplitting, while more travel impedance variables such as delays and detours, travel behaviors over a week (weekdays and weekends), and travel behaviors during one day (peak and nonpeak hours), may also have impacts on the ridesplitting ratio; therefore, these variables need to be explored. In addition, we divide Chengdu into 162 census tracts based on the administrative boundaries. We can investigate the results of the different scaling cells to explore the impacts of the scaling method. This paper uses the modeling method to explore the relationship between ridesplitting and built environment. It should be noted that there are some qualitative approaches, such as deductive and inductive, which could also be used for research on the built environment.

CRediT authorship contribution statement

Meiting Tu: Conceptualization, Data curation, Methodology, Writing - original draft. **Wenxiang Li:** Conceptualization, Data curation, Writing - review & editing, Funding acquisition. **Olivier Orfila:** Funding acquisition, Supervision, Writing - review & editing. **Ye Li:** Funding acquisition, Supervision, Writing - review & editing. **Dominique Gruyer:** Project administration, Funding acquisition, Supervision, Writing - review & editing.

Declaration of Competing Interest

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

Acknowledgments

This study was supported by the National Natural Science Foundation of China [Grant numbers: 52002244, 71774118]; Shanghai Pujiang Program [Grant number: 2020PJC083]; Shanghai Planning Office of Philosophy and Social Sciences [Grant number: 2020EGL019]; and the Science and Technology Commission of Shanghai Municipality [Grant number: 20692192200].

References

Accenture, 2016, Car Sharing Service Outlook in China. https://www.accenture.com/t20160701T062629_w_/cn-en/_acnmedia/PDF-24/Accenture-Insight-Car-Sharing-Service-China-v2.pdf.

Alemi, F., Circella, G., Handy, S., Mokhtarian, P., 2018. What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. Travel Behav. Soc. 13, 88–104.

Anair, D., Martin, J., Mourade, M.C.P., Goldman, J., 2020. Ride-Hailing's Climate Risks: Steering a Growing Industry Toward a Clean Transportation Future. Union of Concerned Scientists, Cambridge, MA.

Angrist, J.D., Caldwell, S., Hall, J.V., 2017. Uber vs. taxi: A driver's eye view (No. w23891). National Bureau of Economic Research.

Bates, J., Leibling, D., 2012. Spaced out. Perspectives on parking policy, 9.

Brunsdon, C., Fotheringham, A.S., Charlton, M.E., 1996. Geographically weighted regression: a method for exploring spatial nonstationarity. Geograph. Anal. 28, 281–298.

Cervero, R., Kockelman, K., 1997. Travel demand and the 3Ds: density, diversity, and design. Transport. Res. Part D Transp. Environ. 2, 199–219.

Chakour, V., Eluru, N., 2016. Examining the influence of stop level infrastructure and built environment on bus ridership in Montreal. J. Transp. Geogr. 51, 205–217.

Chan, N.D., Shaheen, S.A., 2012. Ridesharing in North America: past, present, and future. Transp. Rev. 93–112.

Chen, M.K., Rossi, P.E., Chevalier, J.A., Oehlsen, E., 2019. The value of flexible work: evidence from Uber drivers. J. Political Econ. 127 (6), 2735–2794.

Chengdu Bureau of Statistics, 2020. Statistical report on the national economic and social evelopment of Chengdu (accessed August 18th, 2020).
Choudhury, C.F., Yang, L., E Silva, J.D.A., Ben-Akiva, M., 2018. Modelling preferences for smart modes and services: a case study in Lisbon. Transport. Res. Part A: Policy Practice 115, 15–31.
DiDi Chuxing, 2017. What does the GAIA Initiative provide in the first phase? https://outreach.didichuxing.com/research/opendata/en/ (accessed August 18th, 2020).
DiDi Chuxing, 2018. Corporate Citizenship Report (in Chinese). https://www.didiglobal.com/about-didi/responsibility (accessed August 18th, 2020).
DiDi Chuxing, 2020. Green Travel White Paper on Didi Platform (in Chinese).
Ding, C., Cao, X., Liu, C., 2019. How does the station-area built environment influence Metrorail ridership? Using gradient boosting decision trees to identify non-linear thresholds. J. Transp. Geogr. 77, 70–78.
Ding, C., Cao, X.J., Næss, P., 2018. Applying gradient boosting decision trees to examine non-linear effects of the built environment on driving distance in Oslo. Transport. Res. Part A: Policy Practice 110, 107–117.
Durning, M., Townsend, C., 2015. Direct ridership model of rail rapid transit systems in Canada. Transp. Res. Rec. 2537, 96–102.
El-Assi, W., Mahmoud, M.S., Habib, K.N., 2017. Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. Transportation 44 (3), 589–613.
Ewing, R., Cervero, R., 2001. Travel and the built environment: a synthesis. Transp. Res. Rec. 1780, 87–114.
Ewing, R., Cervero, R., 2010. Travel and the built environment: a meta-analysis. J. Am. Plan. Assoc. 76, 265–294.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232.
Gan, Z., Yang, M., Feng, T., Timmermans, H.J., 2020. Examining the relationship between built environment and metro ridership at station-to-station level. Transport. Res. Part D: Transport Environ. 82, 102332.
Gaode map, 2021. Gaode API. https://lbs.amap.com/api/webservice/guide/api/search (accessed January 13th, 2021).
Hall, J.D., Palsson, C., Price, J., 2018. Is Uber a substitute or complement for public transit? J. Urban Econ. 108, 36–50.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Science & Business Media.
Hughes, R., MacKenzie, D., 2016. Transportation network company wait times in Greater Seattle, and relationship to socioeconomic indicators. J. Transp. Geogr. 56, 36–44.
Kent, J.L., 2014. Carsharing as active transport: what are the potential health benefits? J. Transport Health 1 (1), 54–62.
Li, W., Li, Y., Fan, J., Deng, H., 2017. Siting of carsharing stations based on spatial multi-criteria evaluation: a case study of Shanghai EVCARD. Sustainability 9, 152. https://doi.org/10.3390/su9010152.
Li, W., Pu, Z., Li, Y., Ban, X.J., 2019. Characterization of ridesplitting based on observed data: a case study of Chengdu, China. Transport. Res. Part C: Emerg. Technol. 100, 330–353.
Lyft, Economic Impact Report, 2020, Lyft. https://www.lyftimpact.com (accessed August 18th, 2020).
Nielsen, J.R., Hovmøller, H., Blyth, P.-L., Sovacool, B.K., 2015. Of "white crows" and "cash savers:" A qualitative study of travel behavior and perceptions of ridesharing in Denmark. Transport. Res. Part A: Policy Practice 78, 113–123.
Rayle, L., Dai, D., Chan, N., Cervero, R., Shaheen, S., 2016. Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. Transp. Policy 45, 168–178.
Rayle, L., Shaheen, S.A., Chan, N., Dai, D., Cervero, R., 2014. App-based, On-demand Ride Services: Comparing Taxi and Ridesourcing Trips and User Characteristics in San Francisco (No. UCTC-FR-2014-08). Berkeley: University of California Transportation Center.
Sabouri, S., Park, K., Smith, A., Tian, G., Ewing, R., 2020. Exploring the influence of built environment on Uber demand. Transport. Res. Part D: Transp. Environ. 81, 102296.
Santos, A., Mcguckin, N., Nakamoto, H.Y., Gray, D., Liss, S., 2011. Summary of Travel Trends: 2009 National Household Travel Survey. Federal Highway Administration, United States.
Shaheen, S., Cohen, A., Zohdy, I., 2016a. Shared Mobility: Current Practices and Guiding Principles. Federal Highway Administration, United States.
Shaheen, S.A., Chan, N.D., Gaynor, T., 2016b. Casual carpooling in the San Francisco Bay Area: understanding user characteristics, behaviors, and motivations. Transp. Policy 51, 165–173.
Tang, X., Qin, Z., Zhang, F., et al., 2019. A deep value-network based approach for multi-driver order dispatching. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 1780–1790.
Tao, T., Wang, J., Cao, X., 2020. Exploring the non-linear associations between spatial attributes and walking distance to transit. J. Transp. Geogr. 82, 102560.
Torun, A.Ö., Göçer, K., Yeşiltepe, D., Argın, G., 2020. Understanding the role of urban form in explaining transportation and recreational walking among children in a logistic GWR model: a spatial analysis in Istanbul, Turkey. J. Transp. Geography 82, 102617.
Tu, M., Li, Y., Li, W., Tu, M., Orfila, O., Gruyer, D., 2019. Improving ridesplitting services using optimization procedures on a shareability network: a case study of Chengdu. Technol. Forecast. Soc. Chang. 149, 119733.
Tu, W., Cao, R., Yue, Y., Zhou, B., Li, Q., Li, Q., 2018. Spatial variations in urban public ridership derived from GPS trajectories and smart card data. J. Transp. Geogr. 69, 45–57.
Wang, Y., Wang, S., Wang, J., Wei, J., Wang, C., 2020. An empirical study of consumers' intention to use ride- sharing services: using an extended technology acceptance model. Transportation 47, 397–415.
Wyatt, H., 2020. Uber Estimator. <https://uberestimator.com/cities> (accessed August 18th, 2020).
Yan, X., Liu, X., Zhao, X., 2020. Using machine learning for direct demand modeling of ridesourcing services in Chicago. J. Transp. Geogr. 83, 102661.
Yu, B., Gao, W., 2011. Comprehensive development of Tianfu square in Chengdu City. Urban Rapid Rail Transit 1, 8 (in Chinese).
Yu, H., Peng, Z.-R., 2019. Exploring the spatial variation of ridesourcing demand and its relationship to built environment and socioeconomic factors with the geographically weighted Poisson regression. J. Transp. Geogr. 75, 147–163.
Zhang, Y., Zhang, Y., 2018. Exploring the relationship between ridesharing and public transit use in the United States. Int. J. Environ. Res. Publ. Health 15 (8), 1763.
Zhao, J., Deng, W., Song, Y., Zhu, Y., 2014. Analysis of Metro ridership at station level and station-to-station level in Nanjing: an approach based on direct demand models. Transportation 41, 133–155.