Tsinghua Science and Technology

Volume 23 | Issue 2 Article 5

2018

Comparison of Travel Mode Choice Between Taxi and Subway Regarding Traveling Convenience

Li Li
Department of Automation, BNRist, Tsinghua University, Beijing 100084, China

Shuofeng Wang
Department of Automation, Tsinghua University, Beijing 100084, China.

Meng Li
Department of Civil Engineering, Tsinghua University, Beijing 100084, China.

Jiyuan Tan
Beijing Key Laboratory of Urban Road Intelligent Traffic Control, North China University of Technology, Beijing 100144, China.

Recommended Citation

Li Li, Shuofeng Wang, Meng Li et al. Comparison of Travel Mode Choice Between Taxi and Subway Regarding Traveling Convenience. Tsinghua Science and Technology 2018, 23(2): 135-144.

TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 03/09 pp135–144
DOI: 10.26599 / TST.2018.9010059
Volume 23, Number 2, April 2018

Comparison of Travel Mode Choice Between Taxi and Subway Regarding Traveling Convenience

Li Li, Shuofeng Wang, Meng Li*, and Jiyuan Tan

Abstract: In this study, we investigate travel mode choice behavior between taxi and subway with an emphasis on the influence of traveling convenience. In the first stage, we examine the Origin-Destination (OD) points of Beijing taxi trips and compare these locations with the respective nearest subway station. Statistics reveal several interesting conclusions. First, for approximately 24.89% of all trips, no convenient subway connections exist between the OD pairs. As such, a taxi becomes the only viable choice. Second, for 80.23% of the remaining 75.11% of trips (equivalent to 60.26% of all trips), access distance from either the origin or the destination to the nearest subway station is greater than 500 meters. This phenomenon indicates that walking distance plays an important role in travel mode choice. In the second stage, we examine groups of taxi trips with similar travel distances and travel times to reveal common features. We establish a preference rule in terms of travel distance and travel time. This determines whether an individual driver will take a taxi or the subway, using a pairwise comparison-based preference regression model. Tests indicate that more than 95% of taxi trips can be correctly predicted by this preference rule. This conclusion reveals that traveling convenience dominates the travel model choice between taxi and subway. All these findings shed light on the factors that influence travel mode choice behavior.

Key words: travel mode choice; behaviors; taxi; subway; preference modeling

1 Introduction

Rail travel is extensively viewed to be more sustainable than road-based transport and an appropriate solution to relieve traffic congestion[1−4]. However, many residents still prefer to use cars rather than express rail/subways for traveling[5,6]. Some residents will choose a taxi for traveling when private cars are not available. As a result, cars remain the dominant travel mode in most cities. The continuous increase in car use has placed considerable pressure on many local governments.

To attract more residents to take rail/subways, it is necessary to investigate the factors that influence the choice of travel mode, before we reshape and balance travel demand and supply[7−9]. Different models have been proposed to mine travel mode data from GPS (Global Position System) trajectory data[10−13] or smart card boarding data[14]. To understand why people choose particular travel modes, researchers have studied several factors which could influence the choice of travel mode. These factors include travel distance[15], travel time[16], travel costs[17,18], perceptions[19], habits[20], trip purpose[14], desires for personal space or for safety[15,21], and identities or prestige[22,23]. Studies on the choice of

• Li Li is with Department of Automation, BNRist, Tsinghua University, Beijing 100084, China
• Shuofeng Wang is with Department of Automation, Tsinghua University, Beijing 100084, China.
• Meng Li is with Department of Civil Engineering, Tsinghua University, Beijing 100084, China. E-mail: mengli@tsinghua.edu.cn.
• Jiyuan Tan is with Beijing Key Laboratory of Urban Road Intelligent Traffic Control, North China University of Technology, Beijing 100144, China.
* To whom correspondence should be addressed.
Manuscript received: 2016-11-20; revised: 2017-04-21; accepted: 2017-05-21

travel model often concentrate on comparisons between private cars and public transport, cycling and walking. Recently, some comparisons have been made with taxis and subways. Oil prices, taxes on taxi use[5,17,18], and the price of rail/subway journeys[6,24] were found to alter the use rate of taxis, but not the governing factors. The reasons why people choose taxis and not the subway need to be further explored.

We carried out a questionnaire survey about travel mode choice between subway and taxi in Beijing, China. Almost all the interviewees said that "taking a taxi is more convenient than taking the subway" to explain their choice. According to the survey, convenience may be a mixture of factors, such as less walking, less travel time, fewer transfers, protection of privacy, having a seat, and so on.

In this study, we investigate a special case of the general problem of whether traveling convenience influences travel mode choice between taxi and subway. We found that we can explain the reasons for why most (more than 90%) passengers choose taxis by regarding traveling convenience as a mixture of travel time and travel distance (including walking distance). According to Occam's Razor, we simply regard convenience as a mixture of travel distance (including walking distance) and travel time in this study. Considering that the Beijing subway fare was very low during our investigation, then economic cost was not the reason for preferring taxis for traveling and here, economic cost is not taken as a component of convenience. Our study is based on taxi records collected in Beijing, a large city that has a rich network of subways and low subway fares. The study involves two stages.

In the first stage, we examine the Origin-Destination (OD) locations of 2 753 418 valid taxi trips in Beijing and compare them with the respective nearest subway stations. Statistics revealed several interesting conclusions. First, for approximately 24.89% of all trips, no convenient subway connection existed between the ODs. As such, the taxi was the only choice. Second, for 80.23% of the remaining 75.11% trips (equivalent to 60.26% of all trips), either the boarding or destination location was more than 500 m from the nearest subway station. This interesting phenomenon indicates that walking distance plays an important role in travel mode choice. Third, for 30.89% of the remaining 14.85% trips (equivalent to 4.59% of all trips), either the boarding or destination locations were at airports or railway stations. Surveys reveal that these long-journey passengers prefer taxis to subways mainly because they either carried heavy baggage or were unfamiliar with the subways in the city. Fourth, for 26.48% of the remaining 10.26% trips (equivalent to 2.72% of all trips), either the boarding or destination location was around famous tourist areas such as the Imperial Palace, the Summer Palace, and so on. These passengers may be tourists who took taxis because they were unfamiliar with the subways in the city.

In summary, for more than 85% of taxi trips, travel by taxi is more convenient than travel by subway. This finding indicates that travel convenience (particularly, a mixture of travel time and walking distance) is the dominant factor that drives people to select taxis rather than subways.

In the second stage, we examine groups of taxi trips with similar travel distances and travel times to reveal common features. In particular, we use a pairwise comparison-based regression model to establish a preference rule in terms of travel distance and time to determine whether a driver will take a taxi or the subway. Tests indicate that more than 95% of taxi trips can be correctly predicted by this preference rule. This conclusion reveals that traveling convenience dominates the travel mode choice between taxi and subway.

The rest of this paper is organized as follows. In Section 2, we explain the data used in this study and use it to provide a detailed presentation of our findings. In Section 3, we explain our definition of travel convenience and present the statistical results of the ODs of taxi trips. In Section 4, we explain how pairwise comparison-based preference regression model is used to analyze the taxi trips. Finally, we conclude the entire study in Section 5.

2 Data

2.1 Beijing city and its subway systems

The Beijing metropolitan area covers 16 410 km^2 and, since 2010 has more than 20 million residents, making it one of the most populous cities in the world.

The Beijing government began constructing the subway system in the early 1960s to simplify the daily lives of residents. By June 2014, the entire Beijing subway system comprised 17 subway lines and 233 unique stations (see Fig. 1).

The distance between two consecutive subway stations on the same line is usually less than 2 km. As such, the distance between any OD location and the nearest subway station is usually less than 1 km.

In the time period studied, the Beijing subway adopted a fixed charge policy for each subway trip. Every trip costs 2 Chinese Yuan (approximately 0.32 US dollars),

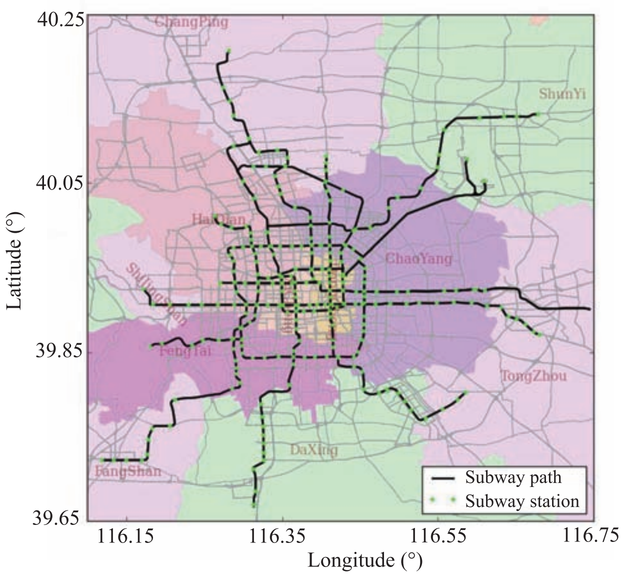

Fig. 1 Map of the Beijing metropolitan area and the layout of subways.

regardless of the ODs. Considering that this fare is very low, economic cost is not the reason that residents take taxis for traveling.

2.2 Taxis in Beijing and the collected data

At present, about 66 000 taxis are running in Beijing. The charge for a taxi trip is mainly calculated based on the distance of the trip. In particular, the base fare is 13 Chinese Yuan (approximately 2.09 US dollars) for the first 3 km, and the passenger needs to pay 2.3 Chinese Yuan (approximately 0.37 US dollars) for each kilometer that he/she travels beyond that. Moreover, the low-speed ($\leqslant$ 12 km/h) surcharge policy requires a passenger to pay 4.6 Chinese Yuan (approximately 0.74 US dollars) per 5 min during rush hours or 2.3 Chinese Yuan (approximately 0.37 US dollars) per 5 min during other periods of their travel.

As such, we determined that taking a taxi is usually more expensive than taking the subway if a resident wants to travel in Beijing. Thus, the price is not a major factor affecting the choice of taxi passengers. Therefore we do not take economic costs as a component of convenience.

In this study, we analyzed 3 681 718 taxi trip records collected from approximately 32 000 taxis that served in Beijing from June 1, 2014 to June 26, 2014. The weather in Beijing in June is good, and passengers do not take taxis to protect themselves from rain or cold weather. Each record in this taxi dataset consists of the following information: taxi ID, timestamp, current GPS position (longitude and latitude), and operation state of the investigated taxi. The taxi state contains the basic information useful for taxi management. In particular, the state of load/unload (whether this taxi is serving passengers or not) is labeled in each record. The sampling time interval for the same taxi is about 60 s.

By detecting the change in the load/unload state of each taxi, we can easily extract the ODs of every taxi trip. We filtered out trip data collected between 00:00−05:59 and 22:01−23:59 as the subway lines are closed at those times. We also filtered out any abnormal raw trip data with travel time less than 1 min or distance less than 1 km. Finally, 2 753 418 trips were extracted and investigated in this study.

Figure 2 shows the density map of all sampled taxi trip records in red. If an area contains more training samples, then its color is darker. We observe that only 1.5% taxi trips have a travel distance above 40.9 km and only 1.5% taxi trips have a travel time above 61.3 min. This indicates that a passenger may prefer other travel modes if either his/her travel distance is more than 40.9 km or his/her travel time will consume more than 61.3 min, under a confidence limit of 97%.

2.3 Questionnaire survey

We carried out a questionnaire survey on travel mode choice between subway and taxi in Beijing, China. We received 460 validated questionnaires. 60% of interviewees were women and 40% were men. 60% were company employees, and 7% were under graduates, and 5.7% were civil servants, and 5.4% were teachers.

According to our survey, the top 3 factors that influence interviewees travel mode choice are convenience, travel cost, and travel speed as shown in Table 1. We investigated the reasons to take a taxi and the reason to take the subway.

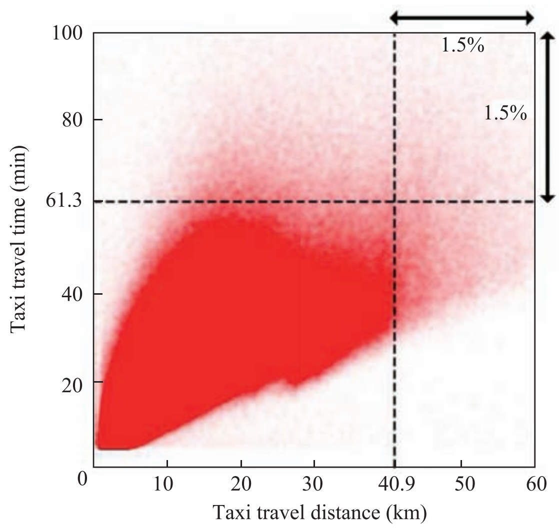

Fig. 2 Density map of taxi trip records.

From Table 2 we can see that the top 3 influencing factors to take a taxi and not the subway are walking distance, transfer number, and travel time. From Table 3 we can see that the top 3 influencing factors to take the subway and not a taxi are travel cost, walking distance, and travel time. Because the corresponding questions in Tables 1−3 are multiple choices questions, the sum of sample percentages is not equal to 1.

According to our survey, 71.2% of interviewees would only accept walking distance below 500 m when taking the subway. 93.3% interviewees would only accept a maximum of three transfers when taking subway. 96.5% of interviewees would only accept waiting times below 10 minutes when taking the subway. 94.9% of interviewees would only accept travel times less than three times that of the alternative taxi travel time.

3 Statistics on the Possible Influence of Travel Convenience

3.1 Disciplines for traveling convenience comparison

Notably, so-called traveling convenience refers to a certain measure of service quality over several sets of travel between the investigated OD. According to the survey, convenience might be a mixture of factors such as less walking, less travel time, less transfer, protection of privacy, having a seat, and so on. In this study, we particularly take traveling convenience as a mixture of travel distance (including walking distance) and travel time. As mentioned before, we do not take economic cost as a component of convenience.

Table 1 Major factors concerned in travel mode choice.

Factor	Sample percent (%)
Convenience	72.0
Travel cost	57.5
Travel speed	47.3
Comfort	30.8
Traffic condition	25.4
Departure time	17.14

Table 2 Reasons to take a taxi, not the subway.

Reason	Sample percent (%)
Walking distance of subway is too long	53.0
Transfer number of subway is too large	48.9
Travel time of taxi is less	32.4
Subway is close	29.5
Subway is crowded	27.6
No seat in subway	7.3
No privation in subway	1.9

Table 3 Reasons to take the subway, not a taxi.

Reason	Sample percent (%)
Subway is cheaper	50.8
Origins/destinations are around subway stations	50.5
Travel time of subway is less	53.3
There is congestion in road	42.3
There is a through subway	27.6
No taxi is available	7.3

Naturally, if more than one travel mode is available, then travelers will take the travel mode that is most convenient. Although we cannot determine the exact form of traveling convenience for each traveler, we can derive several important disciplines for traveling convenience comparison.

The first discipline can be presented as follows: If taking Travel Mode A consumes significantly less travel time than taking Travel Mode B, then Travel Mode A is more convenient than Travel Mode B.

The second discipline can be presented as follows: If taking Travel Mode A requires a significantly shorter walking transfer than taking Travel Mode B, then Travel Mode A is more convenient than Travel Mode B.

If the travel convenience (particularly the mixture of travel time and walking distance due to Occam's Razor) is the dominant factor that drives several people to select taxis rather than subways, the statistical results of the collected taxi trips would show that the travel time or walking distance to the corresponding possible subway alternatives for most OD pairs is unacceptable. According to the survey, we take 500 m as the threshold to determine whether the walking distance is acceptable. If the corresponding subway travel time is more than three times the taxi travel time, we say the corresponding subway travel time is unacceptable.

3.2 Travel time and travel distance estimation for subway travel

We need to determine the travel time when taking subways and travel distance (especially the associated walking transfers) to validate the aforementioned disciplines based on taxi data. In this study, the travel time of taking the subway for a certain OD is calculated as follows:

$$t_{\text{subway}}^{\text{OD}} = t_{\text{walking}}^{O,O_1} + t_{\text{subway}} + t_{\text{walking}}^{D_1,D} \tag{1}$$

where $t_{\text{walking}}^{O,O_1}$ and $t_{\text{walking}}^{D_1,D}$ indicate the time consumed walking from the original location to the corresponding nearest subway station (station O_1) and time consumed walking from the appropriate subway station (station D_1) to the destination of a certain trip, respectively (see Fig.

3). t_{subway} denotes the time needed to transfer from station O_1 to station D_1 by taking the subway. By checking the subway timetables, we can estimate the value of t_{subway} for each OD.

The estimation of walking transfer time cannot be accurate because of the heterogeneous features of people and streets. In this study, we calculate the approximate walking time from the original location to station O_1 as follows:

$$t_{\text{walking}}^{O,O_1} = D_{\text{walking}}^{O,O_1} / v_{\text{walking}} \tag{2}$$

where $D_{\text{walking}}^{O,O_1}$ denotes the walking distance between the two places, and $v_{\text{walking}} = 6\ \text{km/h}$ is the velocity of walking.

We assume that the positions of the original location and station O_1 are $[x_O, y_O]$ and $[x_{O_1}, y_{O_1}]$, where x and y denote the horizontal and vertical coordinates, respectively, and subscripts O and O_1 indicate the original location and station O_1, respectively. Then, the walking distance can be approximately estimated as follows:

$$D_{\text{walking}}^{O,O_1} = \sqrt{(x_O - x_{O_1})^2 + (y_O - y_{O_1})^2} \tag{3}$$

The travel time t_{taxi}^{OD} of each taxi trip and the travel distance of each taxi trip D_{taxi}^{OD} can be directly measured from the dataset.

3.3 Statistical analysis results

We investigate the distribution of $t_{\text{subway}}^{\text{OD}}/t_{\text{taxi}}^{\text{OD}}$ and $\max(D_{\text{walking}}^{O,O_1}, D_{\text{walking}}^{D_1,D})$ to verify whether the travel time or walking distance of the corresponding possible subway alternatives for most OD pairs is unacceptable. The distribution of $t_{\text{subway}}^{\text{OD}}/t_{\text{taxi}}^{\text{OD}}$ and $\max(D_{\text{walking}}^{O,O_1}, D_{\text{walking}}^{D_1,D})$ are shown in Fig. 4. From Fig. 4a, we can see that for 24.89% OD pairs the corresponding subway travel time $t_{\text{subway}}^{\text{OD}}$ is estimated to be three times higher than the corresponding taxi travel time $t_{\text{taxi}}^{\text{OD}}$; hence, the travel time of the corresponding possible subway alternatives for those OD pairs is unacceptable. In Fig. 4b, we can see that for 83.73% of OD pairs the walking distance between OD locations and corresponding nearest subway stations is more than 500 m and then the walking distance of the corresponding possible subway alternatives for those OD pairs is unacceptable. It should be noted that for 60.26% of OD pairs the travel time of the corresponding possible subway alternative is acceptable but the walking distance is unacceptable.

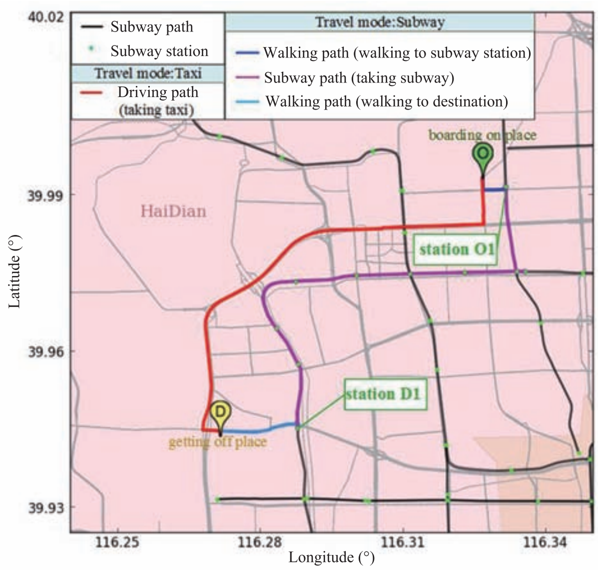

Fig. 3 Illustration of the comparison between taxi route and subway route.

In summary, for 85.15% of the OD pairs in the investigated taxi trips, the travel time or walking distance taking the subway as the travel mode goes beyond the acceptable threshold. This finding proved that travel convenience is a dominant factor that affects the selection of travel mode.

Furthermore, among the undiscussed 14.85% of all trips (408 770 trips), either the original location or the destination of 126 259 trips (approximately 30.89% of those 408 770 trips, equivalent to 4.59% of all the trips) was located in an airport or railway station. The reasons for these passengers taking a taxi may be that needed to take many bags with them or were unfamiliar with the subways in the city.

For the remaining 10.26% of all trips (282 511 trips), either the original location or the destination of 74 800 trips (approximately 26.48% of those 282 511 trips, equivalently 2.72% of all the trips) were in famous tourist areas such as the Imperial Palace and the Summer Palace. These passengers may be tourists who took taxis because they were unfamiliar with the subways in the city.

For the remaining 7.54% of all trips (207 711 trips), the reason for taking a taxi remains uncertain and is partly related to special personal preference (see Fig. 5).

We also examine the possible influence of sampling time. Figure 6 shows the statistical analysis result for 345 656 taxi trips sampled during rush hours (7:30−8:30 and 17:30−18:30). During rush hours, the subway is crowded and passengers may choose taxis for comfort rather than travel time or walking distance.

From Fig. 6 we can see that the travel time or walking distance of taking the subway as the travel mode goes beyond the acceptable threshold for 86.29% taxi trips during rush hour. This indicates that the aforementioned

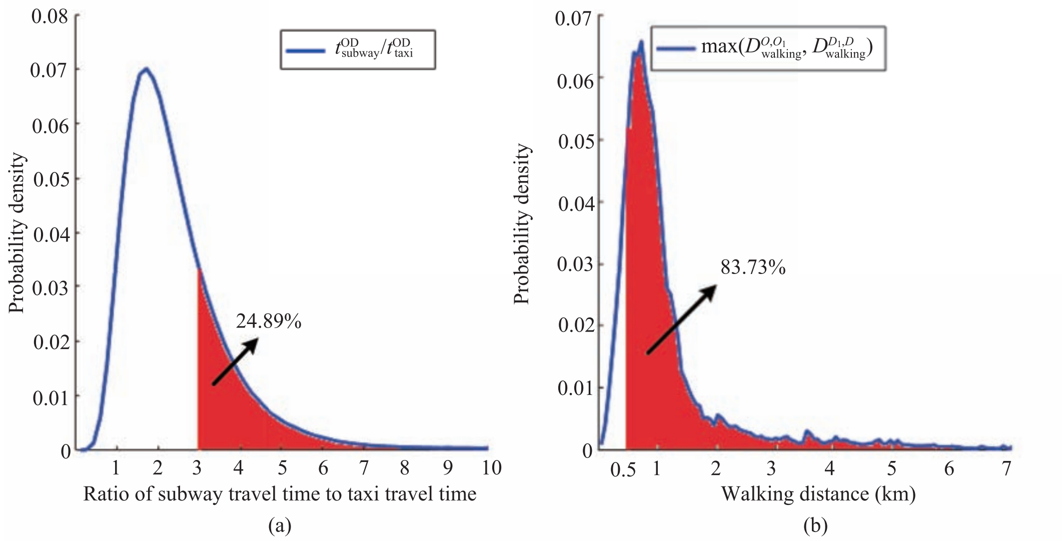

Fig. 4 Cumulative distribution: (a) the ratio of subway travel time to taxi travel time; (b) walking distance.

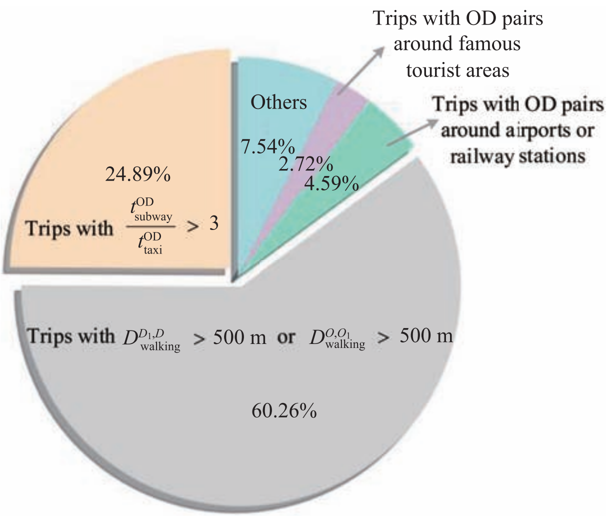

Fig. 5 Proportion of different types of trips.

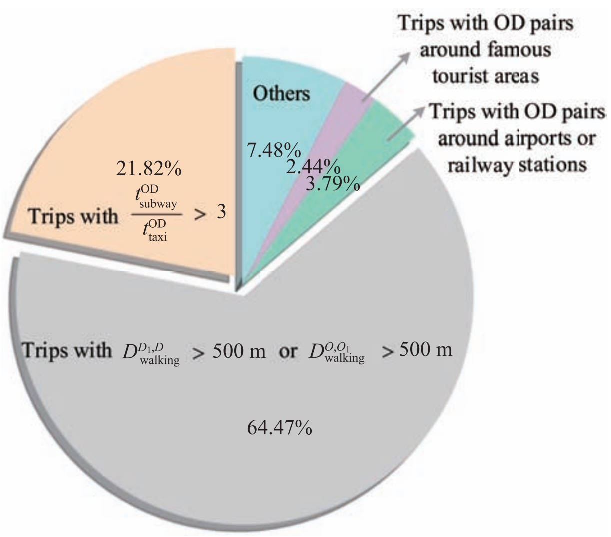

Fig. 6 Proportion of different types of trips during rush hours.

conclusions also hold for trips during rush hour. Travel convenience is also a dominant factor in the selection of travel mode during rush hour. Tests for trips sampled solely on weekdays, weekends, daytime, or nighttime also support the aforementioned conclusions.

4 Travel Mode Preference Modeling

4.1 Preference decision rule model

To further characterize the relationship between travel mode choice and travel convenience (particularly travel distance and travel time), we applied a pairwise comparison-based logistic regression model[25] to establish a preference decision rule whether passengers prefer taxi or subway travel when facing an OD pair. We define (A_i, B_i) as the corresponding option pair for each OD pair OD_i, where $A_i = (A_{i1}, A_{i2})$ and $B_i = (B_{i1}, B_{i2})$. A_{i1} and A_{i2} are the corresponding taxi travel distance and taxi travel time, respectively. B_{i1} and B_{i2} are the corresponding subway travel distance and subway travel time, respectively. In this study, the unit of travel distance is 10 km, and the unit of time is one hour. $P(A_i > B_i)$ denotes the probability of taking a taxi, not the subway, for OD pair OD_i. Given the option pair (A_i, B_i), if we calculate $P(A_i > B_i) > 0.5$, then the passenger would prefer to take a taxi, not the subway.

We further assume that the probability of taking a taxi rather than the subway is logistically distributed and can be expressed as follows:

$$P(A_i > B_i) = \frac{1}{1+\exp(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\phi}(B_i) - \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\phi}(A_i))} \tag{4}$$

where $\boldsymbol{\phi}$ is the mapping function, and $\boldsymbol{\beta}$ is the regression parameter that needs to be determined. In this study, we selected the nonlinear mapping function ϕ to determine the distribution features of taxi trips as discussed in Section 2, which can be expressed as follows:

$$\boldsymbol{\phi}(A_i) = (A_{i1}^2, \sqrt{2}A_{i1}A_{i2}, A_{i2}^2)^{\mathrm{T}},$$
$$\boldsymbol{\phi}(B_i) = (B_{i1}^2, \sqrt{2}B_{i1}B_{i2}, B_{i2}^2)^{\mathrm{T}} \tag{5}$$

We calculated the likelihood function $L(\boldsymbol{\beta})$ for each option pair (A_i, B_i) to determine the regression parameter $\boldsymbol{\beta}$ as follows:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{i=N} P(A_i, B_i) \tag{6}$$

By applying the maximum likelihood estimator[26,27], we can derive the best estimation of $\boldsymbol{\beta}$ as $\hat{\boldsymbol{\beta}}$.

After several mathematical derivations, the decision surface can be written as the following equation:

$$\hat{\boldsymbol{\beta}}^{\mathrm{T}}(\boldsymbol{\phi}(B_i) - \boldsymbol{\phi}(A_i)) = 0 \tag{7}$$

As such, we can simplify the preference decision rule $P(A_i > B_i) > 0.5$ as follows: If $\hat{\boldsymbol{\beta}}^{\mathrm{T}}(\boldsymbol{\phi}(B_i) - \boldsymbol{\phi}(A_i)) > 0$, then the passenger prefers to take the subway. Otherwise, the passenger prefers to take a taxi rather than the subway.

The assumed decision surface $\hat{\boldsymbol{\beta}}^{\mathrm{T}}(\boldsymbol{\phi}(B_i) - \boldsymbol{\phi}(A_i)) = 0$ consists of a pair of hyperbolic curves. However, only a part of the upper curve is valid because the travel time and distance are positive. The decision rule indicates that any corresponding subway option located above this curve will not be preferred.

4.2 Results

We randomly divided the total taxi trip records and their corresponding possible subway alternatives into 3 parts. We derive 3 estimations of β as $\hat{\boldsymbol{\beta}}_1 = [2.24, -5.97, -12.03]$, $\hat{\boldsymbol{\beta}}_2 = [1.91, -3.77, -5.03]$, $\hat{\boldsymbol{\beta}}_3 = [2.77, -5.85, -12.87]$ based on the 3 parts of the data.

We designed a type of validation test to prove the effectiveness of the proposed preference rule. For example, we assume taxi option A_i with travel distance 14 km and travel time 24 min. We observed few taxi trips with these exact parameters. As such, we relaxed our sampling requirements and selected all the observed taxi trips $\bar{A}_i$ whose taxi travel distance falls within $A_{i1} \in [13\ \mathrm{km}, 15\ \mathrm{km}]$ and taxi travel time falls within $A_{i2} \in [21\ \mathrm{min}, 27\ \mathrm{min}]$. Then, we examined the corresponding subway options $\bar{B}_i$ and checked whether all these $\bar{B}_i$ located above the decision surface.

Figure 7a plots three different decision surfaces $\hat{\boldsymbol{\beta}}^{\mathrm{T}}(\boldsymbol{\phi}(B_i) - \boldsymbol{\phi}(A_i)) = 0$ with different values of $\hat{\boldsymbol{\beta}}$. We can see that these decision surfaces are very similar in the first quadrant.

Figure 7b plots all 45 934 corresponding subway options $\bar{B}_i$, where the 45 838 subway options above the decision surfaces are denoted by blue stars and the 96 subway options below the decision surfaces are denoted

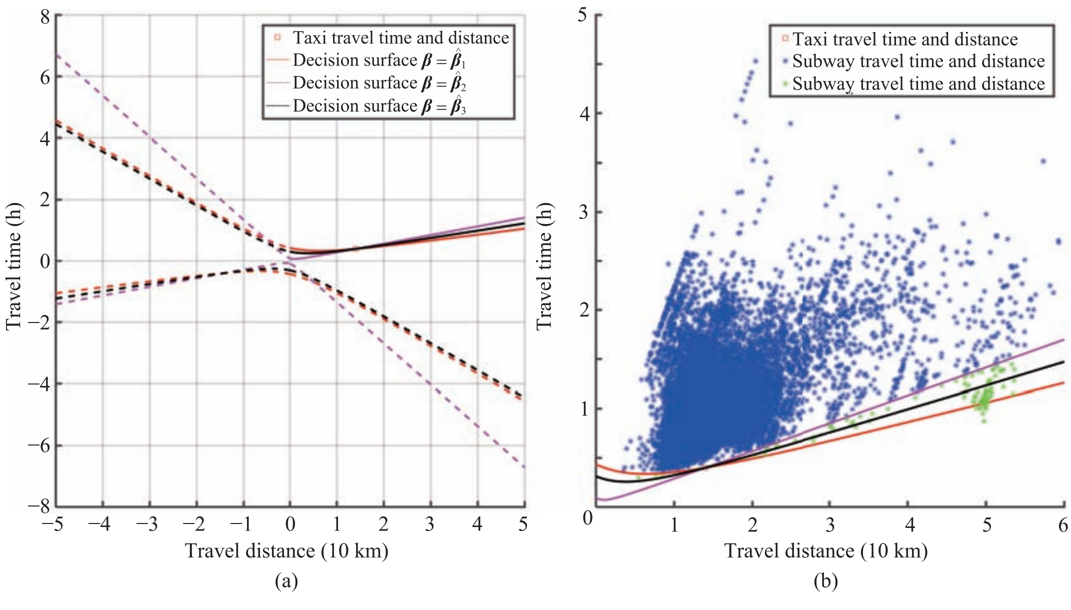

Fig. 7 Illustration of the decision surface given a taxi option $A_{i1} \in [13\ \mathrm{km}, 15\ \mathrm{km}]$ and $A_{i2} \in [21\ \mathrm{min}, 27\ \mathrm{min}]$: (a) the decision surface contains a pair of hyperbolic curves, but only the solid curve part is valid because the travel time and distance are positive and (b) the preference prediction results.

by green stars. For the decision surface with $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_1$, there are 13 subway options above it. For the decision surface with $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_2$, there are 92 subway options above it. For the decision surface with $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_3$, there are 44 subway options above it. In other words, at least 99.79% of subway options $\bar{B}_i$ are located above the decision surface. This high accuracy indicates that the proposed preference decision rule works well.

We choose $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}_1$ as the best estimation of $\boldsymbol{\beta}$. Similarly, we randomly selected 100 groups of data to validate the decision correctness of the obtained preference rule. The results indicate that more than 95% of the observed taxi trip records can be successfully predicted for each group of data. This finding proves that the proposed preference rule can determine when a passenger takes a taxi rather than the subway, if the travel distances and times are given. This conclusion also reveals that convenience (particularly travel distance and travel time) dominate the travel model choice between taxi and subway.

5 Conclusion

The subway is regarded as an economical and appropriate solution to relieve traffic congestion. Government also encourages passengers to take subways. However, in several cities, including Beijing, many people still travel by taxi even when subway travel is cheaper. Understanding the factors that influence travel mode choice is critical to attract more passengers to rail/subways in the future.

In this paper, we examine taxi trip records collected in Beijing to reveal the dominant factors influencing passengers travel mode choice between taxi and subway. Tests on taxi travel records show that traveling convenience (particularly, a mixture of travel time and walking distance) is the dominant factor. For more than 85% of investigated taxi trips, the travel time or walking distance involved in taking the subway goes beyond an acceptable threshold.

Further, we established a preference decision rule in terms of traveling convenience (a mixture of travel time and distance) to determine whether a passenger will take a taxi or the subway, based on a preference regression model. Tests indicate that more than 95% of taxi trips can be correctly predicted by this preference decision rule.

These findings shed light on ways to attract more passengers to rail/subways. Governments should try to reduce the walking distance to subway stations. Building more subway stations is an apparent solution. In addition, the rise in bike sharing, such as the Ofo bike and the Mobike, may also help reduce the walking distance. Governments should also try to reduce the travel time in the subway itself by attempting to increase through subway lines and decrease the number of transfers.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (Nos. 61603005 and 61503007) and the Beijing Municipal Science and Technology Project (No. D171100000317002).

References

[1] P. Knoppers and T. Muller, Optimized transfer opportunities in public transport, *Transportation Science*, vol. 29, no. 1, pp. 101–105, 1995.

[2] A. Spencer and A. Wang, Light rail or busway? A comparative evaluation for a corridor in Beijing, *Journal of Transport Geography*, vol. 4, no. 4, pp. 239–251, 1996.

[3] C. Nie, H. Z. Wen, and X. F. Fan, The spatial and temporal effect on property value increment with the development of urban rapid rail transit: An empirical research, *Geographical Research*, vol. 29, no. 5, pp. 801–810, 2010.

[4] D. Salon, J. Wu, and S. Shewmake, Impact of bus rapid transit and metro rail on property values in Guangzhou, China, *Transportation Research Record*, no. 2452, pp. 36–45, 2014.

[5] Y. C. Chiou, R. C. Jou, and C. H. Yang, Factors affecting public transportation usage rate: Geographically weighted regression, *Transportation Research Part A: Policy and Practice*, no. 78, pp. 161–177, 2015.

[6] J. Eagling and T. Ryley, An investigation into the feasibility of increasing rail use as an alternative to the car, *Transportation Planning and Technology*, vol. 38, no. 5, pp. 552–568, 2015.

[7] W. Bohte and K. Maat, Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands, *Transportation Research Part C: Emerging Technologies*, vol. 17, no. 3, pp. 285–297, 2009.

[8] D. N. Tang, M. Yang, and M. H. Zhang, Travel mode choice modeling: A comparison of Bayesian networks and neural networks, *Applied Mechanics and Materials*, vols. 209–211, pp. 717–723, 2012.

[9] Z. Li and D. Hensher, Prospect theoretic contributions in understanding traveller behaviour: A review and some comments, *Transport Reviews*, vol. 31, no. 1, pp. 97–115, 2011.

[10] A. Bolbol, T. Cheng, I. Tsapakis, and J. Haworth, Inferring hybrid transportation modes from sparse GPS data using a moving window SVM classification, *Computers, Environment and Urban Systems*, vol. 36, no. 6, pp. 526–

537, 2012.

[11] H. Gong, C. Chen, E. Bialostozky, and C. T. Lawson, A GPS/GIS method for travel mode detection in New York City, *Computers, Environment and Urban Systems*, vol. 36, no. 2, pp. 131–139, 2012.

[12] Y. Liu, C. Kang, S. Gao, Y. Xiao, and Y. Tian, Understanding intra-urban trip patterns from taxi trajectory data, *Journal of Geographical Systems*, vol. 14, no. 4, pp. 463–483, 2012.

[13] G. Xiao, Z. Juan, and C. Zhang, Travel mode detection based on GPS track data and Bayesian networks, *Computers, Environment and Urban Systems*, vol. 54, pp. 14–22, 2015.

[14] T. Kusakabe and Y. Asakura, Behavioural data mining of transit smart card data: A data fusion approach, *Transportation Research Part C: Emerging Technologies*, vol. 46, pp. 179–191, 2014.

[15] N. McDonald, Active transportation to school: Trends among US schoolchildren, 1969-2001, *American Journal of Preventive Medicine*, vol. 32, no. 6, pp. 509–516, 2007.

[16] J. Anable and B. Gatersleben, All work and no play? The role of instrumental and affective factors in work and leisure journeys by different travel modes, *Transportation Research Part A Policy and Practice*, vol. 39, nos. 2&3, pp. 163–181, 2005.

[17] Z. J. Wang, X. H. Li, and H. Chen, Impact evaluation of a mass transit fare change on demand and revenue utilizing smart card data, *Transportation Research Part A: Policy and Practice*, vol. 77, pp. 213–224, 2015.

[18] B. Leng, H. Du, J. Wang, L. Li, and Z. Xiong, Analysis of taxi drivers' behaviors within a battle between two taxi apps, *IEEE Transaction on Intelligent Transportation Systems*, vol. 17, no. 1, pp. 296–300, 2016.

[19] C. Lee, X. Zhu, J. Yoon, and J. Varni, Beyond distance: Children's school travel mode choice, *Ann. Behav. Med.*, vol. 45, no. S1, pp. 55–67, 2013.

[20] B. Lanken, H. Aarts, A. V. Knippenberg, and C. V. Knippenberg, Attitude versus general habit: Antecedents of travel mode choice, *Journal of Applied Social Psychology*, vol. 24, no. 4, pp. 285–300, 1994.

[21] E. Mann and C. Abraham, The role of affect in UK commuters' travel mode choices: An interpretative phenomenological analysis, *British Journal of Psychology*, vol. 97, no. 2, pp. 155–176, 2006.

[22] N. Murtagh, B. Gatersleben, and D. Uzzell, Multiple identities and travel mode choice for regular journeys, *Transportation Research Part F Traffic Psychology and Behaviour*, vol. 15, no. 5, pp. 514–524, 2012.

[23] D. Lois, J. A. Moriano, and G. Rondinella, Cycle commuting intention: A model based on theory of planned behaviour and social identity, *Transportation Research Part F Traffic Psychology and Behaviour*, vol. 32, pp. 101–113, 2015.

[24] S. Flügel, A. H. Halse, J. D. D. Ortúzar, and L. I. Rizzi, Methodological challenges in modelling the choice of mode for a new travel alternative using binary stated choice data - The case of high speed rail in Norway, *Transportation Research Part A: Policy and Practice*, vol. 78, pp. 438–451, 2015.

[25] J. P. Arias-Nicolás, C. J. Pérez, and J. Martín, A logistic regression-based pairwise comparison method to aggregate preferences, *Group Decision and Negotiation*, vol. 17, no. 3, pp. 237–247, 2008.

[26] E. L. Lehmann and G. Casella, *Theory of Point Estimation.* Springer, 1999.

[27] G. Casella and R. L. Berger, *Statistical Inference, 2nd edition.* Pacific Grove, CA, USA: Duxbury and Thomson Learning, 2002.

Li Li is currently an associate professor with Department of Automation, Tsinghua University, Beijing, China, working in the fields of complex and networked systems, intelligent control and sensing, intelligent transportation systems, and intelligent vehicles. He received the PhD degree in systems and industrial engineering from University of Arizona, USA, in 2005. Dr. Li had published over 70 SCI indexed international journal papers and over 70 international conference papers as a first/corresponding author.

Shuofeng Wang is currently a PhD student studying at Department of Automation, Tsinghua University, China. He mainly studies travel behavior analysis. He received the MS degree in control science and engineering from Tsinghua University, China, in 2017.

Meng Li received the PhD degree in transportation engineering from Nagoya University, Japan, in 2010. He is currently an associate professor with Department of Civil Engineering, Tsinghua University, Beijing, China, working in the fields of intelligent transportation systems and advanced traffic operation and management.

Jiyuan Tan received the PhD degree in control science and engineering from Tsinghua University, Beijing, China, in 2014. He is currently an associate professor with Department of Transportation, North China University of Technology, Beijing, China, working in the fields of intelligent transportation systems, advanced traffic management, and traffic safety.