Zhenying YAN, Ph.D.[1]
E-mail: yanzhenying@imu.edu.cn
Meiying JIAN, Ph.D.[1]
(Corresponding author)
E-mail: Jianmy321@163.com
Xiaojuan LI, Ph.D.[1]
E-mail: 311987431@imu.edu.cn
Jinxin CAO, Ph.D.[2]
E-mail: imucjx@163.com
[1] Transportation Institute, Inner Mongolia Engineering Research Center for Urban Transportation Data Science and Applications, Inner Mongolia University
No.49, Xilingol South Road, Yuquan District, Hohhot, 010020, China
[2] Inner Mongolia Academy of Science and Technology
No. 70, Zhaowuda Road, Saihan District, Hohhot, 010010, China

Transport Economics
Original Scientific Paper
Submitted: 1 July 2021
Accepted: 16 Nov. 2021

MODELLING THE PASSENGER CHOICE BEHAVIOUR OF BUYING HIGH-SPEED RAILWAY TICKETS

ABSTRACT

Passenger choice behaviour of buying tickets has a great impact on the high-speed rail (HSR) revenue management. It is very critical to find out the sensitive factors that prevent passengers with high willingness to pay for a ticket from buying low-price tickets. The literature on passenger choice behaviour mainly focuses on travel mode choice, choice between a conventional train and a high-speed train and choice among high-speed trains. To extend the literature and serve revenue management, this paper investigates passenger choice behaviour of buying high-speed railway tickets. The data were collected by the stated preference (SP) survey based on Beijing-Hohhot high-speed railway. The conditional logit model was established to analyse influencing factors for business travel and non-business travel. The results show that: business passengers have the higher inherent preference for full-price tickets, while non-business passengers have the higher inherent preference for discount tickets; the number of days booked in advance and frequent passenger points have a significant impact on the ticket choice of business travellers, but not on non-business travellers; passengers are unwilling to buy tickets that depart after 16:00 for non-business travel; factors have different effects on the passengers' choice in business travel and non-business travel. The results can provide parameters for revenue management models and references for the ticket-product design.

KEYWORDS

railway transportation; passenger choice behaviour; conditional logit model; SP survey; revenue management.

1. INTRODUCTION

By the end of 2020, there was 38,000 kilometres of high-speed railway (HSR) in China. According to China Statistical Yearbook 2020, the HSR passenger volume was 2,358.3 million in 2019, accounting for 64.4% of the railway passenger volume; and HSR passenger-kilometres reached 774.7 billion, accounting for 52.7% of the railway passenger-kilometres and 21.9% of the national passenger-kilometres. The HSR has become an important travel mode for passengers in China. However, there are many lines in China with insufficient passenger flow intensity, whose revenue from tickets can hardly cover the cost of train operation and maintenance.

Sufficient revenue is the key to ensuring the sustainable development of HSR. To increase the utilization rate of HSR lines and the revenue, the China State Railway Group Co., Ltd (CR) has introduced revenue management ideas and methods. The flexible fares were introduced to adjust and stimulate passenger demand. HSR passengers would face diversified ticket products, and how passengers choose their tickets will directly affect the effectiveness of revenue management [1]. Therefore, there is an urgent need to study passenger choice behaviour and its influencing factors at the level of ticket products.

Existing studies use ticket data and survey data to study the choice behaviour of HSR passengers at the following three levels. The first level is the choice behaviour of travel mode. In view of the competition among HSR, civil aviation and highways in the short, medium, and long distances, the data from the questionnaire survey are used to study the travel mode choice behaviour of passengers [2, 3]. The influencing factors include individual socio-economic characteristics, travel mode characteristics and psychological variables. The second level is the choice behaviour between a high-speed train and a conventional train. In the transportation corridor, there are both high-speed trains and conventional trains. The binomial logit model is used to study the passenger choice behaviour between a high-speed train and a conventional train [4]. The results show that demographic characteristics, ticket purchase channels, social class and status, departure frequency, departure date and time, travel purpose, distance and other variables have a significant impact on the choice of the train type. The third level is choice behaviour among high-speed trains. Many trains run on a HSR line, and different trains have different departure time. SP survey data and ticket-sale data were used to study how do passengers choose trains [5, 6]. The significant influencing factors include ticket prices, train running time, departure time, on-board service level and income. In addition, the passengers choice behaviour of booking time in different distance markets was explored and compared to the ticket price [7].

To the best of our knowledge, the passenger choice behaviour of buying tickets has not yet been elaborated carefully in the existing literature. When revenue management is applied in the field of HSR, the choices faced by passengers are not travel mode choice and train choice. For example, the second-class seats in a train can be sold as different ticket products with different prices and attributes. The passenger choice behaviour of buying tickets should be modelled and analysed to guide the ticket product design and to provide initial parameters for revenue management models [8].

The research methods of choice behaviour mainly include discrete choice models and machine learning algorithms. Discrete choice models often use the maximum likelihood method for parameter estimation. For situations where certain options are not observed in actual sales data, expectation maximization algorithm, likelihood function improvement algorithm, etc. can be used to estimate parameters [9]. To avoid the truncation of the data and study more attributes, this paper uses the SP survey method to collect data. Then the conditional logit model is selected to investigate the choice behaviour among tickets for HSR passengers and the maximum likelihood method is used to estimate parameters.

In summary, this paper studies the choice behaviour of HSR ticket products under multi-level ticket prices from the aspects of personal attributes, personal travel experience, and product attributes, which analyses the characteristics of passenger ticket demand in a more detailed way. The findings could guide the design of ticket products, provide initial parameters for the application of choice-based revenue management models and provide reference factors for the data collection of the revenue management system.

2. LITERATURE REVIEW

This research is primarily in the field of travel choice behaviour. The choice of the travel mode is an important part of the traffic mode split, route choice and traffic demand forecast [10–12]. Freitas et al. [13] used Swiss household travel survey data to establish a large-scale multi-modal recursive logit path choice model. Luan et al. [14] applied the mixed-logit model to study the choice behaviour of urban transportation modes including walking, cycling, bus and metro, and compare the characteristics and advantages of travel modes in different regions from the perspective of the catchment areas of urban railway transit stations. Weis et al. [15] analysed data from the Mobility and Transport Micro Census to provide parameters for the transportation demand model. Jian et al. [16] used SP survey data to study the travel behaviour of the elderly in Beijing in the future. Great progress has also been made in research methods. Cheng et al. [17] applied robust random forest to study transportation mode choice behaviour and analysed the relative importance of explanatory variables and their influence on mode choice. Cheng et al. [18] used the ensemble-based model composed of a series of multinomial logit models to reduce the uncertainty. Li and Hensher [19] reviewed studies about risky travel choice behaviour and put forward the measures to avoid common errors. Zhao et al. [20] comparatively studied the difference between the logit model and machine learning model in mode selection. The study found that the prediction accuracy of the random forest model is higher than that of the multinomial logit and mixed logit models, but the behavioural robustness is slightly worse. In the mode choice model, when choosing between machine learning and logit models, there seems to be a trade-off between prediction accuracy and behavioural robustness.

Research on the choice behaviour of intercity travel and regional travel has also received a lot of attention. Zhou et al. [21] used the latent category model to segment the Western Australia travellers, and then studied the travel mode choice behaviour. The travel modes include car, bus and two flights. The studied attributes include travel cost, access time, journey time, frequency and seat comfort level. Hess et al. [22] used the choice data on the corridors between two major cities in the United States (Northeast Corridor and Cascade Corridor) to develop a hybrid choice model that takes into account certainty and confidence in travellers' preferences. Random changes explore the relationship between attitude and actual choice. Alternative transportation modes include car, air and train. Li et al. [23] investigated the inter-city choice behaviour of air, HSR, traditional passenger train and express bus for tourists in Xi'an. The relationship between selection and attributes is analysed, and 12 variables including travel distance, ticket price, 100 km intercity travel time, service quality, accessibility of transportation hubs and ticket-sale methods are analysed. Intercity transfers also affect travel mode choice behaviour. Allard and Moura [24] used a combination of SP and RP surveys to collect European inter-city travel data, highlighting the importance of transfer attributes.

Research on travel choice behaviour related to HSR has also emerged. Li et al. [25] used the nested logit model to describe the travel mode choice behaviour between the high-speed train and the conventional train. The upper-level options are aviation and railway, and railways are divided into the categories of conventional train, bullet train and HSR according to different speeds. Ren et al. [26] used survey data to establish binomial logit models to analyse the factors and changes related to passengers choosing conventional trains before and after the emergence of the high-speed trains. The study found that many people still choose conventional trains after the arrival of the high-speed trains. The key factors excluding passengers from using the HSR system include the fare, travel habits and amenities for long-distance trips. Losada-Rojas et al. [27] introduced attitude assessment into the choice of intercity travel mode, used the bivariate ordered probit model to describe the willingness to use trains, and studied passenger attitudes and loyalty to railway transportation. The mode of transportation includes train, drive alone, carpool, airplan and bus. Daly et al. [28] combined attitude and choice models to study transportation choice behaviours, revealing the role of potential attitudes when it comes to the railway travel environment. Yang et al. [29] used the SP survey to obtain data and refine the choice behaviours research to the level of compartment category. They investigated the western corridors of Taiwan, and the mode choice level was car, bus, train, HSR and HSR nested the choice between the reserved seat cabin and non reserved seat cabin.

This article studies the more detailed choice behaviour among HSR ticket products. According to different ticket products formed by different trains and different price levels, the influencing factors of passenger choice behaviour are studied. It is different from the choice of transportation modes, choice of train types and choice of HSR trains. It is intended to provide a reference for the ticket product design in revenue management applications and provide parameters for revenue management models.

The organization of this paper is as follows; Section 3 presents the SP survey and data, Section 4 builds the ticket choice behaviour model for HSR passengers and estimate the parameters; and Section 5 reports discussions on personal socioeconomic attributes, personal travel experience attributes and ticket attributes. Finally, conclusions and possibilities for future works are presented in Section 6.

3. SP SURVEY AND DATA

3.1 Questionnaire design and survey

The sale data accurately record the choice results of buying tickets, but the lack of records of unpurchased tickets leads to data truncation, and the recorded attribute data are limited to the system settings. Questionnaire survey data have the advantages of being complete and the scope of attribute research is flexible, but the reliability of the data is poor. However, the stated preference investigation and analysis of what has not happened is of reference value for the implementation of new transportation services and transportation policies. This article investigates the passenger's choice preference before the implementation of the multi-level fare revenue management strategy and studies the multi-directional attribute data to guide the design of passenger ticket products. The SP survey can provide a certain reference for the implementation of the strategy.

Based on the literature and survey with the railway companies, the factors influencing HSR passenger choice behaviour of buying tickets are determined. They are divided into three types: personal socioeconomic attributes, personal travel experience attributes and ticket attributes. Personal socioeconomic attributes include gender, age, education and income. Personal travel experience attributes refer to indicators related to the personal travel experience. We introduce 4 indicators that may influence ticket purchase behaviour, including times of taking (Beijing to Hohhot) trains in the past year, number of flights taken (Beijing-Hohhot) in the past year, the experience of taking HSR and favourite departure time. The first two indicators can reflect how familiar passengers are with the corridor. Ticket attributes include fare, booking days in advance, railway frequent passenger points, refund fee, operation hours, and departure time.

The options of passenger ticket products are set as full-price tickets, discount tickets and no purchase. The value of each ticket product attribute is set according to the Beijing-Hohhot HSR line, as shown in *Table 1*, where the non-purchase option is the zero option. There are 4,374 scenarios for the combination of various options and different attribute levels. Considering the reasonable limit on the number of questions answered by the respondents, ensuring the accuracy of the survey data and the rationality of the modelling, the 27 scenarios obtained through orthogonal design are divided into 3 questionnaires, and the surveys were conducted separately for business travel and non-business travel. The main difference between business travel and non-business travel is that the former is paid by the employer and the latter is paid by the traveller himself.

3.2 Data

Taking the HSR line from Beijing to Hohhot as the research background, the survey was carried out by selecting passengers with travel experience among cities in this corridor as effective samples, including Beijing, Hohhot, Baotou and Ordos. A total of 729 questionnaires were collected by a professional online survey company in August 2019. The company conducts surveys based on an online platform wjx.cn which is used by many researchers in China [16, 30]. After invalid questionnaires were screened out, 630 valid questionnaires were obtained. The subjects of the investigation are lim-

Table 1 – Options and properties

Option	Attribute of ticket	Attribute value
Full-price ticket	Fare	1. Full fare (RMB 220)
	Booking days in advance	1. Unlimited (the value is 0)
	Railway frequent passenger points	1. 5 times of price (1110 points)
	Refund fee	1. 5%; 2. 10%; 3. 20%
	Operation hours	1. 2.5h; 2. 3h; 3. 3.5h
	Departure time	0. 7:00~10:00 (base); 1. 10:00~16:00 (variable name: departure time-1); 2. 16:00~22:00 (variable name: departure time-2)
Discount ticket	Fare	1. 10% off (RMB 198); 2. 20% off (RMB 176); 3. 40% off (RMB 132)
	Booking days in advance	1. 7 days; 2. 14 days; 3. 21 days
	Railway frequent passenger points	1. 5 times of price; 2. equal to price; 3. No points
	Refund fee	1. 50%; 2. 100%
	Operation hours	1. 2.5h; 2. 3h; 3. 3.5h
	Departure time	0. 7:00~10:00 (base); 1. 10:00~16:00 (variable name: departure time-1); 2. 16:00~22:00 (variable name: departure time-2)
No purchase	-	-

Note: RMB is the Chinese currency "Renminbi". US$ 1 approximates RMB 6.462 as of September 29, 2021.

ited to people living in Beijing, Hohhot, Baotou and Ordos, to ensure that they are within the scope of potential passengers.

The basic data description of the valid questionnaire is shown in *Table 2*. The number of males and females in the survey data is relatively evenly distributed; the proportion of people under 18 in the age distribution is relatively low; the proportions of all income levels covered by the sample is equal and the favourite departure time is evenly distributed.

In the interview conducted by the railway company CR, we learned that in 2017, the railway company conducted a passenger survey on 11 railway bureaus. A total of 30 trains were investigated and 33,271 valid questionnaires were obtained on the spot, and 175,909 valid questionnaires on the website of 12,306 were obtained. The age distribution in the field survey is: 1.27% under 18 years of age, 7.98% over 50 years of age; age distribution in the online survey is 6.09% under 18 years of age and 7.08% over 50 years of age. In comparison, our survey data are similar in age distribution to their survey. The income distribution in the field survey is: 27.09% under RMB 3,000, 39.23% between RMB 3,000 and 5,000, 18.99% between RMB 5,001 and 10,000, and 14.68 over RMB 10,000; the income distribution in the online survey is: 30.08% under RMB 3,000, 32.21% between RMB 3,000 and 5,000, 26.93% between RMB 5,001 and 10,000, and 10.06 over RMB 10,000. Chinese average disposable income is RMB 2,561 per month in 2019, which is RMB 397 higher than 2017. The disposable income of residents in the four cities

Table 2 – Data description

Type	Attribute	Category	Variable Name	Proportion (%)
Personal socioeconomic attributes	Gender	Male	Gender	46.2
		Female	(Base)	53.8
	Age	<18	Age-1	2.1
		18~44	Age-2	84.7
		>44	(Base)	13.2
	Education	Senior high school and below	Education-1	11.4
		Junior college	Education-2	16.5
		Bachelor or above	(Base)	72.1
	Income (RMB)	<3,000	(Base)	21.8
		3,000~5,000	Income-1	25.8
		5,001~8,000	Income-2	25.2
		>8,000	Income-3	27.3
Personal travel experience attributes	Times of taking (Beijing to Hohhot) trains in the past year	0	(Base)	35.9
		1~2 times	Times of train-1	41.5
		3~6 times	Times of train-2	17.9
		>6 times	Times of train-3	4.7
	Number of flights taken (Beijing-Hohhot) in the past year	0	(base)	61.5
		1~2 times	Times of air-1	28.5
		3~6 times	Times of air-2	8.3
		>6 times	Times of air-3	1.7
	Have you ever taken a high-speed train?	Yes	(base)	82.4
		No	Never took HSR	17.6
	Favourite departure time	7:00-10:00	(base)	38.2
		10:00-16:00	Favourite departure time-1	37.4
		16:00~22:00	Favourite departure time-2	34.4

we surveyed (Beijing, Hohhot, Baotou and Ordos) is RMB 5,646 per month, RMB 3,192 per month, RMB 3,729 per month and RMB 3,447 per month in 2019. They are all higher than national average value. It can be seen that the income distribution of the sample is reasonable. More than 70% of the interviewees had a university education. The cause of this may be that the highly educated groups have responded more positively to surveys based on online platforms.

4. PASSENGER CHOICE BEHAVIOUR MODEL

4.1 Conditional logit model

The attribute variables studied in this paper include not only the personal attributes that do not change with the program but also the attributes of ticket products that change with the program. Considering the mixed situation of two attribute variables, a conditional logit model is established for the analysis. The model is based on the theory of random utility and assumes that passengers make choices based on the principle of maximizing utility. The utility function expression of passenger ticket choice is shown in *Equation 1*

$$U_{ni} = V_{ni} + \varepsilon_{ni} \tag{1}$$

where U_{ni} is the utility function for passengers to choose ticket category $i(i \in I)$; V_{ni} is the system utility of the choice for the passenger and n; ε_{ni} is the random part of the utility of the option for passenger n. The system utility V_{ni} of the passenger's choice is linear in relation to the attribute variables x_{nik} and z_{nj}, as shown in *Equation 2*

$$V_{ni} = \sum_{k=1}^{K} \beta_k x_{nik} + \sum_{j=1}^{J} \gamma_j z_{nj} \tag{2}$$

where β_k and γ_j are the parameters to be estimated, x_{nik} is the attribute variable related to the ticket and z_{nj} represents personal socioeconomic attribute variables and travel experience attribute variables that do not change with the ticket. V_{n0} indicates the utility of not buying any product, as a base option. Assuming that the utility random item ε follows the independent and identical Gumbel distribution, the probability that the passenger n chooses the ticket i is shown in *Equation 3*.

$$P_{ni} = \frac{e^{V_{ni}}}{e^{V_{n0}} + \sum_{j \in I} e^{V_{nj}}} \tag{3}$$

4.2 Parameter calibration

The variables in *Tables 1 and 2* are introduced into the model, and the maximum likelihood method is used to estimate the parameters. The calibration results of the conditional logit model of business travel and non-business travel are shown in *Tables 3 and 4*.

Table 3 – Model tests

Item	Business Travel	Non-business Travel
Number of sample observations	5,940	5,940
Number of individuals in the sample	660	660
Wald chi2(40)	406.28	432.26
Prob > chi2	0.00	0.00
Log likelihood	-3,420.36	-4,516.67

According to *Table 3*, the chi-square test statistic Wald chi2 and the corresponding P-value show that the joint significance of all the coefficients of the entire model is very high and indicate that both models are valid. When the significance level is set to 0.05, the italics of P-values in *Table 4* indicate significant impact. The results in *Table 4* show that there are differences in the significance of the attribute variables between the business travel model and the non-business travel model.

5. DISCUSSION

5.1 Individual socioeconomic attributes

The conditional logit model of business travel shows that gender and education level have a significant influence on the choice behaviour of buying tickets. Compared with female passengers, the probability of buying full-price tickets for male passengers is 31.47% lower while other factors remain unchanged and there is no significant difference when buying discounted tickets. Compared with people with bachelor's degree or above, the probability of buying full price tickets and discount tickets for passengers with a high school degree or below decreased by 77.58% and 71.57%, respectively. Passengers with junior college education are more interested in discount tickets. Income level and age did not have a significant impact on ticket purchase choices. Since the ticket costs for business travel were generally not paid by individuals themselves, the income level did not have a significant impact

Table 4 – Parameter estimation for passenger choice model of HSR

Option	Variable Name	Business Travel			Non-business Travel		
		ODS	Coefficient	P-Value	ODS	Coefficient	P-Value
Full-price ticket	Gender	0.6853	-0.3780*	0.0240	0.8200	-0.1985	0.0900
	Education-1	0.2242	-1.4952*	0.0000	0.7824	-0.2453	0.1670
	Education-2	1.3124	0.2719	0.3210	1.2637	0.2340	0.1570
	Income-1	1.0363	0.0357	0.8760	0.6619	-0.4126*	0.0120
	Income-2	0.7824	-0.2454	0.2930	0.6817	-0.3831*	0.0320
	Income-3	1.1553	0.1443	0.5790	0.8989	-0.1066	0.5750
	Age-1	1.8455	0.6127	0.2010	4.5902	1.5239*	0.0000
	Age-2	0.9560	-0.0450	0.8310	0.9450	-0.0566	0.7170
	Never took HSR	0.5205	-0.6530*	0.0000	0.5187	-0.6565*	0.0000
	Favorite departure time-1	1.8102	0.5934*	0.0030	0.7173	-0.3322*	0.0120
	Favorite departure time-2	0.8935	-0.1127	0.5640	0.7056	-0.3487*	0.0170
	Times of train-1	1.0285	0.0281	0.8800	1.0615	0.0597	0.6450
	Times of train-2	2.6269	0.9658*	0.0080	1.8392	0.6093*	0.0030
	Times of train-3	0.4859	-0.7218*	0.0310	2.7571	1.0142*	0.0020
	Times of air-1	1.1599	0.1483	0.5130	1.3828	0.3241*	0.0350
	Times of air-2	1.0908	0.0869	0.7940	0.8409	-0.1732	0.4520
	Times of air-3	2.4551	0.8981	0.3860	0.6473	-0.4349	0.3180
	Constant	428.9039	6.0612*	0.0000	32.8218	3.4911*	0.0000
Discount ticket	Gender	0.7241	-0.3229	0.0670	0.6593	-0.4165*	0.0000
	Education-1	0.2843	-1.2576*	0.0000	0.5217	-0.6506*	0.0000
	Education-2	1.8104	0.5936*	0.0350	1.2033	0.1851	0.2350
	Income-1	1.3903	0.3295	0.1700	0.4625	-0.7711*	0.0000
	Income-2	0.9829	-0.0173	0.9440	0.6354	-0.4535*	0.0060
	Income-3	1.2378	0.2133	0.4350	0.8787	-0.1293	0.4670
	Age-1	2.4431	0.8933	0.0760	1.3223	0.2794	0.5290
	Age-2	0.9095	-0.0949	0.6720	1.1500	0.1398	0.3410
	Never took HSR	0.8738	-0.1349	0.4900	0.4658	-0.7641*	0.0000
	Favorite departure time-1	2.2518	0.8117*	0.0000	1.0734	0.0708	0.5650
	Favorite departure time-2	1.0411	0.0402	0.8460	0.9722	-0.0282	0.8370
	Times of train-1	0.9959	-0.0041	0.9830	1.1875	0.1719	0.1520
	Times of train-2	1.7556	0.5628	0.1340	1.7581	0.5643*	0.0040
	Times of train-3	0.2215	-1.5073*	0.0000	1.3535	0.3027	0.3300
	Times of air-1	0.8871	-0.1198	0.6140	1.4516	0.3727*	0.0100
	Times of air-2	0.8427	-0.1712	0.6270	1.2188	0.1979	0.3470
	Times of air-3	7.8886	2.0654*	0.0490	0.7448	-0.2947	0.4650
	Constant	70.9936	4.2626*	0.0000	63.3962	4.1494*	0.0000
	Fare	0.9902	-0.0098*	0.0000	0.9946	-0.0054*	0.0000
	Advance booking	0.9814	-0.0188*	0.0020	0.9936	-0.0064	0.2010
	Score	1.0002	0.0002*	0.0050	0.9999	-0.0001	0.3350
	Operation time	0.8475	-0.1655*	0.0030	0.8725	-0.1364*	0.0090
	Departure time-1	1.1003	0.0956	0.0980	1.0258	0.0255	0.6200
	Departure time-2	1.0234	0.0231	0.6830	0.8909	-0.1156*	0.0230

** Significant at the 0.05 level.*

on ticket purchases. Women prefer full price tickets more than men when they travel because of work. When designing passenger ticket products, value-added services can be added to the results for these passengers, such as setting up female-only carriages, placing luggage and providing blankets.

The conditional logit model of non-business travel shows that gender, education level, income and age all have a certain influence on ticket purchase behaviour. When purchasing full-price tickets, there is no significant difference between male and female travellers. When travelling for non-business purposes, female passengers prefer discount tickets more than male passengers. The passengers with high school degrees or below and those with junior college education are not significantly different from those with bachelor's degree or above when purchasing full-price tickets, while the passengers with high school degree or below are more likely to purchase discount tickets. When the monthly income exceeds RMB 8,000, the effect on ticket choice behaviour is not significant. Passengers with a monthly income of more than RMB 3,000 and less than RMB 8,000 are less inclined to choose HSR than those with an income of less than RMB 3,000. This may be because they have other travel modes to choose from. For example, if they have a car, it is also a good choice to drive with their family when travelling privately. However, the coefficients of Income-1 and Income-2 for discount tickets are smaller than those for full-price tickets. It can be concluded that compared to passengers earning less than RMB 3,000, passengers with an income of RMB 3,000-8,000 dislike discount tickets more than full-price tickets. This may be due to too many restrictions on discount tickets. Improving the attractiveness of HSR to these groups is worthy of attention. Compared with passengers over 44 years of age, there is no significant difference in the purchase of discount tickets for passengers of other age groups. Passengers younger than 18 years of age are more willing to buy full-price tickets than passengers older than 44 years of age when traveling on non-business trips. Those under the age of 18 were born after 2000 and they have a more positive attitude for consumption than those over the age of 44. After all, the disposable income of urban residents in 2000 exceeded 18 times of that in 1978. In fact, such groups do not have the opportunity to travel on business and the coverage of their expenses comes from their families.

5.2 Personal travel experience attributes

Whether passengers have the experience of taking HSR has a significant impact on the purchase of full-price tickets, and the corresponding coefficient is negative, indicating that passengers with HSR experience are more willing to buy full-price tickets. The travel experience of HSR has no significant impact on the purchase of discount tickets. The survey finds that business travellers who have taken the HSR will have a good impression of the HSR because of the speed, punctuality and comfort of the HSR, and they tend to continue to travel by HSR. Compared with passengers who like to depart at 7:00 to 10:00 in the morning, those who prefer to travel at 10:00 to 16:00 are more inclined to choose HSR for business travel. The probability of buying full-price tickets for the latter is 1.81 times that of the former, and the probability of buying discounted tickets for the latter is 2.25 times that of the former. The number of trains and airplanes taken in the survey interval in the past year can reflect the familiarity of passengers with the two travel modes between Beijing and Huhhot. The model calibration results show that compared with passengers who did not travel by train between Beijing and Hohhot in the past year, the passengers who did travel from Beijing to Hohhot for 3-6 times in the past year have a significant impact on the purchase of full-price tickets for HSR. They are more willing to purchase full-price tickets than passengers who did not travel by train in the past year. Passengers who travelled with Beijing to Hohhot trains more than 6 times in the past year have a significant impact on the purchase of HSR discount tickets, and compared with passengers who did not travel by train in the past year, they are less willing to buy discounted HSR tickets. Air travel frequency has no significant influence on purchase of full-price HSR tickets for business travel. Passengers who travelled more than 6 times by air between Beijing and Hohhot in the past year are more willing to purchase discounted HSR tickets for business travel than those who did not travel. Passengers who have taken the HSR, prefer to travel from 10:00 to 16:00, or travel with Beijing-Huhhot conventional trains with a frequency of 3-6 times, prefer full fare tickets. This enlightens railway operating companies to give extra care to frequent travellers to increase their loyalty.

Passengers who have taken HSR are also more inclined to buy HSR tickets when traveling on non-business trips. Compared with passengers who like to travel between 7:00 and 10:00 in the morning, those who like to travel in other times are more reluctant to buy full-priced HSR tickets. When buying discount tickets, the effect of travel time preference is not significant. Compared with passengers who did not take this type of train in the past year, passengers who did take 3 or more trains are more willing to purchase full-price tickets for HSR trips, and the higher the frequency, the greater the probability of choice. Passengers who travelled more than 6 times by air between Beijing and Hohhot in the past year are also more willing to buy discounted HSR tickets. Compared with passengers who did not take this type of flights in the past year, passengers who did take them 1-2 times are more inclined to switch to HSR. The probability of purchasing full-price tickets increases to 1.38 times and the probability of purchasing discount tickets increases to 1.45 times. Passengers who have taken more than 3 flights of this type have no significant influence on the HSR ticket choice behaviour. Passengers who have taken the HSR or prefer to travel from 7:00 to 10:00, whose travel frequency between Beijing-Hohhot by conventional trains is more than 3 times, or whose air travel frequency is 1-2 times, prefer full fare tickets when traveling on non-business trips.

5.3 Ticket attributes

Passengers travelling on business trips have a higher inherent preference for full-price tickets. The full-price tickets designed in the questionnaire allow passengers to change their tickets once free of charge. They are more flexible and business travellers prefer full-price tickets. The significant factors among ticket attributes include fare, booking in advance, railway frequent passenger points and operating hours. Among them, the fare, booking in advance and operating time have a negative impact on the ticket purchase choice. It means that the longer the discount tickets need to be booked in advance, the more business travellers will lose interest in discount tickets. Railway frequent passenger points have a positive impact on ticket choice, so increasing the points of full-price tickets will attract more business travellers. The departure time has no significant influence on ticket choices for passengers travelling from Beijing to Hohhot.

Passengers travelling on non-business have a higher inherent preference for discount tickets. The factors that have a significant impact on passenger ticket choice include fare, operating time and departure time, while the effects of booking in advance and railway frequent passenger points are not significant. Compared to tickets for trips departing from 7:00 to 10:00 in the morning, the probability of non-business travellers choosing a ticket after 16:00 will be reduced by 10.91%.

6. CONCLUSIONS

This paper establishes conditional logit models of passenger ticket purchase choice at the level of HSR ticket products. The research shows that personal socioeconomic attributes, travel experience attributes and ticket attributes have different effects on passenger ticket purchase behaviour for business and non-business travel. It is recommended to add statistical information of passengers' socio-economic attributes and travel experience attributes in the ticket system to improve the accuracy of using ticket data to model the passengers' ticket purchase behaviour.

Booking days in advance and railway frequent passenger points have a significant impact on business travellers' choice. This has implications for the design of passenger ticket products. For the Beijing-Hohhot HSR line, it is recommended to increase the booking days in advance for discount tickets, reduce railway frequent passenger points for discount tickets and increase railway frequent passenger points for full-price tickets during the design of HSR ticket products, to limit the demand for full-price tickets, to transfer purchase discount tickets and ensure the successful implementation of revenue management measures.

For business travel and non-business travel, the train operating time has a significant impact on ticket choice behaviour. For the HSR line from Beijing to Hohhot, higher discounts can be set for long-running trains when designing ticket products. Non-business passengers are unwilling to choose tickets that depart after 16:00. It is suggested for the discount to be adopted in order to improve the attraction of this kind of ticket products.

Passengers who have taken HSR trains are more willing to choose full-price tickets when traveling on business trips and are more willing to buy discount tickets when traveling on non-business trips. Passengers who have taken HSR prefer to choose

HSR. For newly opened operating lines, publicity and discount measures can be adopted to attract more passengers to take high-speed trains to cultivate passenger flow.

With questionnaire surveys to study passenger ticket choice behaviour, the research conclusions can promote a more reasonable design of passenger ticket products and provide initial parameters for the revenue management model. The results in this paper can also provide references for other HSR lines. With the implementation of flexible fares, in the future, more accurate characterization of ticket choice behaviour should be explored with sale data.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China [72061028 and 71961024] and the Key Technology Research Plan of Inner Mongolia Autonomous Region [2019GG287]. The authors are very grateful for the support.

闫振英，博士[1]
邮箱：yanzhenying@imu.edu.cn
（通讯作者）
訾美英，博士[1]
邮箱：Jianmy321@163.com
李晓娟，博士[1]
邮箱：311987431@imu.edu.cn
曹瑾鑫，博士[2]
邮箱：imucjx@163.com
[1] 内蒙古大学交通学院，内蒙古自治区城市交通数据科学与应用工程技术研究中心，
中国呼和浩特市玉泉区锡林郭勒南路49号，
邮编：010020
[2] 内蒙古科学技术研究院，中国呼和浩特市赛罕区昭乌达路70号，邮编：010010

高铁旅客购票行为建模

摘 要

旅客购票选择行为直接影响着高铁收益管理的收益提升效果。找出阻碍支付意愿高的旅客转移购买低价客票产品的敏感因素至关重要。目前关于旅客选择行为的研究主要集中在出行方式选择、普速列车与高速列车间的选择行为和高速列车之间的选择行为。本文采用*SP*调查研究客票产品层面的高铁旅客购票选择行为，既拓展了现有旅客选择行为研究的深度，也能够为收益管理模型提供参数。本文以北京与呼和浩特之间的高铁线路为背景，利用*SP*调查问卷搜集数据。针对公务出行和非公务出行分别建立条件*Logit*模型，采用极大似然法估计参数，并进行影响因素分析。研究发现：公务旅客对高铁全价票具有较高的固有偏好，而非公务旅客对高铁折扣票具有较高的固有偏好；提前预定天数和常旅客积分对公务旅客的购票选择影响显著，而对非公务旅客的购票选择影响不显著；旅客在非公务出行时不愿意购买*16:00*之后的车票；在公务出行和非公务出行中，各种因素对旅客的选择行为具有不同的影响。研究结果可为收益管理模型提供参数，也可为客票产品设计提供参考。

关键词

铁路运输；旅客选择行为；条件*Logit*模型；*SP*调查；收益管理.

REFERENCES

[1] Strauss AK, Klein R, Steinhardt C. A review of choice-based revenue management: Theory and methods. *European Journal of Operational Research.* 2018;271(2): 375-387. doi: 10.1016/j.ejor.2018.01.011.

[2] Rui H, Wu Q. Medium and long-distance travel mode decision between high-speed rail and civil aviation. *China Journal of Highway and Transport*. 2016;29(3): 134-141. doi: 10.19721/j.cnki.1001-7372.2016.03.017.

[3] Ye Y, Han M, Chen J. Intercity passenger travel mode choice behavior based on trip chain. *Journal of Tongji University (Nature science)*. 2018;46(09): 77-83. doi: 10.11908/j.issn.0253-374x.2018.09.011.

[4] Cao W, et al. Investigating passengers choice behavior of intercity rails with large-scale ticketing data. *System Engineering Theory and Practice*. 2020;40(4): 989-1000. doi: 10.12011/1000-6788-2018-2198-12.

[5] Zhao P, Zhai R, Song W. Passengers choice behavior of high-speed railway considering individual heterogeneity. *Journal of Beijing Jiaotong University*. 2019;43(02): 121-127. doi: 10.11860/j.issn.1673-0291.2017.06.008.

[6] Sun Y, et al. Analyzing high speed rail passengers' train choices based on new online booking data in China. *Transportation Research Part C: Emerging Technologies*. 2018;97: 96-113. doi: 10.1016/j.trc.2018.10.015.

[7] Hetrakul P, Cirillo C. Accommodating taste heterogeneity in railway passenger choice models based on internet booking data. *Journal of Choice Modelling*. 2013;6: 1-16. doi: 10.1016/j.jocm.2013.04.003.

[8] Yan Z, Li X, Zhang Q, Han B. Seat allocation model for high-speed railway passenger transportation based on flexible train composition. *Computers and Industrial Engineering*. 2020;142. doi: 10.1016/j.cie.2020.106383.

[9] Newman JP, et al. Estimation of choice-based models using sales data from a single firm. *Manufacturing & Service Operations Management.* 2014;16(2): 184-197. doi: 10.1287/msom.2014.0475.

[10] Wu J, Yang M, Sun S, Zhao J. Modeling travel mode choices in connection to metro stations by mixed logit models: A case study in Nanjing, China. *Promet – Traffic&Transportation.* 2018;30(5): 549-561. doi: 10.7307/ptt.v30i5.2623.

[11] Lai, X, Li J, Li Z. A subpath-based logit model to capture the correlation of routes. *Promet – Traffic&Transportation.* 2016;28(3): 225-234. doi: 10.7307/ptt.v28i3.1808.

[12] Ramírez HG, et al. Travel time and bounded rationality in travellers' route choice behaviour: A computer route choice experiment. *Travel Behaviour and Society*. 2021;22: 59-83. doi: 10.1016/j.tbs.2020.06.011.

[13] Freitas LM, Becker H, Zimmermann M, Axhausen KW.

Modelling intermodal travel in Switzerland: A recursive logit approach. *Transportation Research Part A: Policy & Practice*. 2019;119: 200-213. doi: 10.1016/j.tra.2018.11.009.

[14] Luan X, Cheng L, Song Y, Zhao J. Better understanding the choice of travel mode by urban residents: New insights from the catchment areas of rail transit stations. *Sustainable Cities and Society.* 2020;53: 10196. doi: 10.1016/j.scs.2019.101968.

[15] Weis C, et al. Surveying and analysing mode and route choices in Switzerland 2010–2015. *Travel Behaviour and Society*. 2021;22: 10-21. doi: 10.1016/j.tbs.2020.08.001.

[16] Jian M, Shi J, Liu Y. Dependence of the future elderly on private cars: A case study in Beijing. *Promet – Traffic&Transportation*. 2018;30(1): 45-55. doi: 10.7307/ptt.v30i1.2364.

[17] Cheng L, et al. Applying a random forest method approach to model travel mode choice behaviour. *Travel Behaviour and Society*. 2019;14: 1-10. doi: 10.1016/j.tbs.2018.09.002.

[18] Cheng L, et al. Applying an ensemble-based model to travel choice behavior in travel demand forecasting under uncertainties. *Transportation Letters*. 2020;12(6): 375-385. doi: 10.1080/19427867.2019.1603188.

[19] Li Z, Hensher D. Understanding risky choice behaviour with travel time variability: A review of recent empirical contributions of alternative behavioural theories. *Transportation Letters.* 2020;12(8): 580-590. doi: 10.1080/19427867.2019.1662562.

[20] Zhao AX, Yan B, Yu A, van Hentenryck P. Prediction and behavioral analysis of travel mode choice: A comparison of machine learning and logit models. *Travel Behaviour and Society.* 2020;20: 22-35. doi: 10.1016/j.tbs.2020.02.003.

[21] Zhou H, et al. Analysing travel mode and airline choice using latent class modelling: A case study in Western Australia. *Transportation Research Part A: Policy and Practice*. 2020;137: 187-205. doi: 10.1016/j.tra.2020.04.020.

[22] Hess S, Spitz G, Bradley M, Coogan M. Analysis of mode choice for intercity travel: Application of a hybrid choice model to two distinct US corridors. *Transportation Research Part A: Policy and Practice.* 2018;116: 547-567. doi: 10.1016/j.tra.2018.05.019.

[23] Li X, Tang J, Hu X, Wang W. Assessing intercity multimodal choice behavior in a touristy city: A factor analysis. *Journal of Transport Geography*. 2020;86: 1-13. doi: 10.1016/j.jtrangeo.2020.102776.

[24] Allard RF, Moura F. Effect of transport transfer quality on intercity passenger mode choice. *Transportation Research Part A: Policy and Practice*. 2018;109: 89-107. doi: 10.1016/j.tra.2018.01.018.

[25] Li H, Wang K, Yu K, Zhang A. Are conventional train passengers underserved after entry of high-speed rail? -Evidence from Chinese intercity markets. *Transport Policy*. 2020;95: 1-9. doi: 10.1016/j.tranpol.2020.05.017.

[26] Ren X, et al. Impact of high-speed rail on social equity in China: Evidence from a mode choice survey. *Transportation Research Part A: Policy and Practice*. 2020;138: 422-441. doi: 10.1016/j.tra.2020.05.018.

[27] Losada-Rojas LL, Gkartzonikas C, Pyrialakou VD, Gkritza K. Exploring intercity passengers' attitudes and loyalty to intercity passenger rail: Evidence from an on-board survey. Transport Policy. 2019;73: 71-83. doi: 10.1016/j.tranpol.2018.10.011.

[28] Daly A, et al. Using ordered attitudinal indicators in a latent variable choice model: A study of the impact of security on rail travel behaviour. *Transportation*. 2012;39(2): 267-297. doi: 10.1007/s11116-011-9351-z.

[29] Yang CW, Tsai MC, Chang CC. Investigating the joint choice behavior of intercity transport mode and high-speed rail cabin with a strategy map. *Journal of Advanced Transportation.* 2014;49: 297-308. doi:10.1002/atr.1264.

[30] Yang X, et al. Car ownership policies in China: Preferences of residents and influence on the choice of electric cars. *Transport Policy.* 2017;58(8), 62-71. doi: 10.1016/j.tranpol.2017.04.010.