



**ST. JOSEPH'S UNIVERSITY**

**BENGALURU**

**MODELLING THE PASSENGER CHOICE**  
**BEHAVIOUR OF PURCHASING RAILWAY**  
**TICKETS IN INDIA**

**Submitted by:**

Thomas Samuel

[thomassamuel702@gmail.com](mailto:thomassamuel702@gmail.com)

222ECST38

**Under the Guidance of:**

Prof. Raisa Elsa Josphe

Assistant Professor,

Department of Economics,

St. Joseph's University, Bengaluru



# **ST. JOSEPH'S UNIVERSITY**

## **BENGALURU**

### **CERTIFICATE**

This is to certify that the term paper “Modelling the passenger choice behaviour of purchasing railway tickets in India” by Thomas Samuel under the guidance of Prof. Raisa Elsa Joseph for the fulfilment of the requirement of bachelor’s degree in Economics and Statistics is a record of the candidate’s personal efforts.

Prof. Raisa Elsa Joseph

Assistant Professor

Department of Economics



# **ST. JOSEPH'S UNIVERSITY**

## **BENGALURU**

### **DECLARATION**

I, Thomas Samuel, declare that this Term Paper entitled “Modelling the passenger choice behaviour of purchasing railway tickets in India” was carried out in this year 2024 in St. Joseph's University under the guidance of Prof Raisa Elsa Joseph.

I declare that the study has not been submitted to any other institution for the award of any degree or diploma.

Thomas Samuel

222ECST38

## **ACKNOWLEDGEMENT**

As part of my three-year Bachelor's degree in Economics, and Statistics, I was required to write a term paper on "Modelling the passenger choice behaviour of purchasing railway tickets in India." I did the best I could to deliver the material.

I would like to express my gratitude to the college for giving me the chance to work on this term paper. I want to express my sincere gratitude to my guide, Prof. Raisa Elsa Joseph, for her advice and unwavering support throughout the drafting of the paper.

I also want to express my gratitude to my family for always providing me with emotional support. Finally, I would want to express my gratitude to my friends for serving as an inspiration and a source of ideas that have been invaluable to me

Yours Sincerely,

Thomas Samuel

222ECST38

## **TABLE OF CONTENTS**

Introduction	
Literature review	
Sp survey and Data	
Discussion	
Conclusion	
References	

## **1. Introduction**

In India the capital expenditure on railways for the year 2023-24 was Rs 2,60,200 crore. In the same year about 94% of the revenue was earned from freight and passenger trains.[1] Passenger trains account for about 28% of the revenue receipts. While freight trains contribute 68%.[1] The passenger traffic is expected to grow with passing time. Some of the reasons for a lower share of passenger train contribution are inadequate speed of trains, unmet demands of passengers, inadequate capacity, high fares, competition from roads which have better connectivity to name a few.[2] The railway system is the backbone of the Indian economy and one of the primary methods of connectivity in our country. Owing to this it is necessary to take a deeper look at how passengers spend their money purchasing tickets. The main reasons for losses in passenger services stem from two factors: (i) when passenger fares are set lower than operational costs, and (ii) when concessions are provided to specific passenger categories such as senior citizens and National award winners. These concessions are categorized by Railways as social service obligations.[3] To improve the revenue of the Indian Railways it is important to understand the consumer choice behaviours. This way we can price tickets to appeal to passengers better and increase revenue. Passengers of trains will face various choices in terms of tickets and the way passengers choose tickets will have a direct impact on the revenue of the Indian Railways. In the context of India, a wide scope of research has been covered on the revenue of the Indian Railways, operating costs, and a cost benefit analysis of the trains presently in use. However, there is not a lot of research about how consumer choice would impact the revenue. The passenger choice is what would contribute to a healthy revenue, so it is important to understand how both the personal attributes and train ticket attributes play an important role in a passenger's choice of buying a particular ticket. The influencing factors can be a person's socioeconomic conditions, the cost of the train, the time of departure, type of train or any other attributes related to the train. Another important aspect is the data on present consumer preferences. This will give an insight on the types of trains that are frequented by passengers and passenger travel experience is a valuable insight into understanding how that passenger will travel in the future.

The preferred research methodologies for analysing choice behaviour primarily consist of discrete choice models and machine learning algorithms. Discrete choice models commonly employ the maximum likelihood method for parameter estimation. In cases where specific

options are absent from observed sales data, methods like the expectation maximization algorithm and likelihood function improvement algorithm are utilized to estimate parameters.[4] The paper will use a survey method to collect data from potential clients. Using this data, a conditional logit model will be built. We will try to answer the question: Will an average Indian choose a faster, costlier train or not. This is an important question to answer to better understand the kind of trains that passengers would prefer.

In summary, this paper will study the passenger behaviour choice of passengers in India in to answer the question “Will an average Indian choose a faster, costlier train or not.” The paper will consider various variables that are the passengers personal attributes and attributes of the train that the passenger would choose. The findings of this paper could guide in making trains that would suit the preferences of the common public and thereby increase the revenue earned by passenger trains.

## **2. Literature Review**

This research done in this paper is predominantly in the field of passenger choice behaviour. This paper will examine how passengers choose between trains based on personal attributes and attributes of the train. Utilizing data from the Swiss household travel survey, a comprehensive multi-modal recursive logit path choice model was developed [5] by Freitas et al. This paper gave insights about logit models. The mixed-logit model was employed to examine urban transportation mode choice behaviour encompassing walking, cycling, bus, and metro [6], by Luan et al. This model compared the different modes of transport and drew some very important insights into using the model. Hess et al. [8] utilized choice data from corridors connecting two major cities in the United States, namely the Northeast Corridor and Cascade Corridor. They constructed a hybrid choice model that considers travellers' certainty and confidence in their preferences. Random variations were investigated to analyse the connection between attitude and actual choice. Yan Z et al [9] used a conditional logit model to model the behaviour of passengers buying high speed rail tickets. They used personal attributes, ticket attributes and train attributes to model the data. The research was instrumental in revealing the factors behind low usage of the high-speed railways in China. Several reports about the railways in India helped give insight into the revenue and budget of trains in India. Saket Surya [1] laid out the revenue of the Indian railways and brought to light some important factors that are causing passenger trains to run at a loss.

This paper looks at the Indian context in terms of passenger behaviour choice and aims to find some contributing attributes that makes a passenger choose one train over another. In India while a lot of research has been done about the budget and revenue, there are almost no articles about passenger behaviour choice. This paper will look at some personal attributes and some train attributes that cause a passenger to make a specific choice. This will be done using an SP survey and a conditional logit model to model the behaviour. The intention of this research is to understand passenger choices so that trains and tickets can be designed to better suit the needs of passengers and thereby increase revenue. The paper is organized in the following way: section 3 will talk about the survey and the data. This section will also cover the modelling of the data. Section 4 will have the discussion about the paper and Section 6 will have the conclusion of the paper.



### **3. SP Survey and Data**

#### **3.1. Design of Survey**

Due to a lack of direct consumer choice preference data from the Ministry of Railways, a survey was conducted to understand behaviours better. Data obtained from a questionnaire survey have the advantage of being complete and added to this, the scope of the research is very flexible. However, the reliability of the data acquired is poor. In this situation, the data obtained from a stated preference survey have an advantage over a regular survey, since it poses a question with fixed responses to the participant. This gives more rigor to the research and data obtained from the survey. To understand the average Indian train passenger's choice preference, a Stated preference survey was conducted. The survey posed hypothetical situations, to which the participants would have to either agree or disagree.

The responses were binary for the questions posed. The questions took into consideration various factors and questions were posed to understand the consumer's preference of choosing passenger trains. The factors considered were, the speed of the train, type of seating arrangement in the train, departure time and the cost of the train given an early booking.

Along with this some socio-economic variables were also collected. Data about the passenger's gender, age and income were collected to understand what an average Indian would value most while booking a train. The survey was circulated as a google form through various WhatsApp groups. It was circulated among people of various age groups and various incomes, to get an overall idea of public spending choices and purchasing behaviour.

#### **3.2. Data**

The survey was sent out to approximately 300 people and 167 responses were obtained. To assure quality of data the responders that have not used trains in the near past were removed. The responses were screened and filtered. After removing the invalid responses, 94 positive responses were attained. Out of which the responses by male and female participants were approximately even. Coming to the age groups of the participants, 10.6% of the participants were <20, 24.46% were of the age 20–35, 30.85% of responses were in the age group 35-50, 31.9% of the responses were from the age group 50-65 and finally from the age group <65 approximately 2% of the responses were collected. A probable reason for extremely low responses from the last age group could be due to unfamiliarity with online surveys. The distribution of income of the responders were reasonable. The income distribution was: <1

Lakh INR accounted for about 28.7% of the responses, 1-5 Lakhs INR accounted for 25.53% of responses, 5-10 Lakhs INR accounted for 18.08% of the responses, 10-15 Lakh INR accounted for 10.63% of the responses and finally <15 Lakhs INR accounted for 17.02% of the responses. Now the income distribution was a little higher than the average income of residents in Bangalore, but this difference could be since the survey was answered mostly by respondents with college graduation.

*Table 1: Distribution of data collected*

Attribute	Category	Proportion
Gender	Male	0.41
	Female	0.59
Age	<20	0.11
	20-35	0.24
	35-50	0.31
	50-65	0.32
	>65	0.02
Income	<1 Lakh INR	0.29
	1-5 Lakhs INR	0.25
	5-10 Lakhs INR	0.18
	10-15 Lakhs INR	0.11
	>15 Lakhs INR	0.17
Choice of train Speed and Cost	Faster, costlier train	0.80
	Slower, cheaper train	0.20
Choice of seating arrangement	Sleeper	0.69
	Chair car	0.31
Choice of departure time for journeys >6 hours	Night	0.88
	Day	0.12
Choice of train if ticket is booked 1 month prior	Cheaper train	0.61
	Quicker train	0.39

### 3.3. Passenger Behaviour Choice Model

This study considers various variables from personal attributes to attributes of trains to understand consumer choice behaviour. To model this, considering the mixed situation of the two variables, we will use a conditional logit model. The model is based on the theory of random utility, as a result it assumes that consumers will make choices to maximize their utility. The passenger ticket choice is given by the following utility function given in equation 1.

$$u_{hi} = v_{ni} + \varepsilon_{ni} \quad (1)$$

Here,  $u_{ni}$  is the utility function for passengers to choose ticket category  $i$ ,  $v_{ni}$  is the system utility of the choice of passenger  $n$  and  $\varepsilon_{ni}$  is the random part of the utility function. In this the system utility given by  $v_{ni}$  is assumed to be in linear combination with the attribute variables,  $x_{nk}$  and  $z_{nj}$  given by the equation 2.

$$v_{ni} = \sum_{k=1}^k \beta_k x_{ni_k} + \sum_{j=1}^j \alpha_j z_{n_j}$$

Using this equation, the parameters to estimate are  $\beta_k$  and  $\alpha_j$ . Here,  $x_{nik}$  is the ticket choice attribute of the passenger and  $z_{nj}$  is the personal attribute of the passenger. We will assume that the error part is white noise and hence follows standard normal i.i.d. To understand choice behaviours, we considered the attributes that would cause a passenger to purchase a ticket in a faster, costlier train.

The model output is displayed below:

Concordance= 0.724 (se = 0.057 )

Likelihood ratio test= 12.96 on 6 df, p=0.04

Wald test = 11.2 on 6 df, p=0.08

Score (logrank) test = 12.75 on 6 df, p=0.05

When the significance level is set to 0.1, we see that there is a significant impact of the attributes on the passenger choice behaviour since the Wald test returned a score of 0.08 which is less than 0.1. Since the size of the dataset is small, the rigor, accuracy and efficiency of the model is weak. However, it gives us a basis to better understand passenger choice behaviour.

#### **4. Discussion**

The conditional logit model shows that income had a significant impact on whether the passenger would choose a faster, costlier train. The coefficient was found to have significant impact at 0.58 with a significance of 0.0253. This would mean that the trains that are premium quality would be out of reach for lower income groups. While people from higher income groups would prefer to spend more to save time. This is a natural understanding as not everyone will be will to spend more to reach their destination quicker. Given their personal income backgrounds people would choose to not spend more than necessary when travelling. Taking this into consideration, the need for affordable quick transport is one of the biggest takeaways. The conditional probability of a passenger choosing a faster, costlier ticket given income is found to 0.37 for low-income groups while it is found to be 0.89 for passengers with a higher disposable income.

This would mean that though faster, costlier trains are available to everyone, only a select few have the luxury of choosing this option. From the model we can also understand that sleeper trains are preferred more by female passengers. The conditional probability of choosing a sleeper train given it is female is 0.78 while the conditional probability of a male passenger choosing a sleeper train is an approximately even 0.44. The model also brought forward that age played a significant impact of a passenger's choice on the speed of the train when the ticket is booked 1 month prior. From the conditional probability of age given the choice of a quicker or cheaper train, the participants of the survey chose the quicker train as age increased. The most significant of this was seen in the age group of 50-65 where the proportion who would choose a quicker train over a cheaper one was 0.86.

The data also showed that the preferred time of departure for most passengers was night when the travel was 6 hours or longer. This means that even for journeys between 2 close cities like Bangalore and Hyderabad or Chennai to Mysore, passengers would prefer to use a train that leaves at night instead of a train that leaves in the day. The responses showed that conditional probability of a passenger buying a costlier ticket would also prefer a train that is a sleeper train which leaves at night.

Some important things to consider when understanding the data better is there is no classification between business and non-business travel. The difference between business travel and non-business travel is that the cost of the journey is taken up by the company in the case of business travel. The purpose of the travel would affect the purchasing behaviour in

terms of the cost of the ticket, the speed of the train and the departure time of train. All of these will play an important role in the decision process of buying a train ticket. Added to this, the lack of frequent high-speed trains in India between cities make trains an unpopular choice for business travel. Trains in India as of right now are slow and cater mostly to passengers travelling for leisure or passengers who cannot afford flights. The sample of data collected may not be representative of the population at the whole since the sample size is very small. This may also cause the model to be slightly inefficient and imperfect.

## **5. Conclusion**

This paper uses a conditional logit model to model the purchasing behaviour of passengers buying train tickets in India. From the research we can see that personal attributes, and train attributes have different effects on the purchasing behavior of consumers. While the study was conducted on a very small data set, the inferences are still of importance. The research shows that income, departure time and type of transport are all important when buying a train ticket for the average Indian. While affluent passengers will prefer to buy tickets for a faster, costlier train, the average Indian cannot afford to do so. Income is one of the most significant factors considered when purchasing a ticket. Added to this, the research showed that choices change if the ticket is booked in advance. Older passengers prefer booking tickets for a quicker train than a cheaper train.

The study shows that income being a major factor for passengers purchasing behaviour, most passengers choose to travel in cheaper trains, which often tend to be slower than their costlier counterparts. The result of the research is to better understand what consumers deem important when purchasing tickets to travel. Consumer choice behaviour gives key insights to the factors that cause passengers to choose various trains and understanding consumer behaviour can play an important role in increasing revenue earned by the Ministry of Railways of India. While implementing new trains and deciding pricing, it is important to also consider the personal and train attributes to maximize revenue and increase profits. The research is incomplete since it did not examine the data on a large scale and this leaves potential for further research.

## 6. References

1. “Demand for Grants 2023-24 Analysis: Railways” *PRS Legislative Research, Institute for Policy Research Studies, 2023*
2. “Reform, Perform, and Transform,” *Indian Railways, July 2017*
3. “Reviewing the Impact of “Social Service Obligations” by Indian Railways”, *NITI Aayog*
4. “Estimation of choice-based models using sales data from a single firm. Manufacturing & Service Operations Management.” *Newman JP, et al.2014*
5. “Modelling intermodal travel in Switzerland: A recursive logit approach. Transportation Research Part A: Policy & Practice.” *Freitas LM, Becker H, Zimmermann M, Axhausen KW, 2019*
6. “Better understanding the choice of travel mode by urban residents: New insights from the catchment areas of rail transit stations.” *Sustainable Cities and Society. Luan X, Cheng L, Song Y, Zhao J, 2020*
7. “Hettrakul P, Cirillo C. Accommodating taste heterogeneity in railway passenger choice models based on internet booking data.” *Journal of Choice Modelling. Hettrakul P, Cirillo C. 2013*
8. “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. Hess S, Spitz G, Bradley M, Coogan M. 2018*
9. “Modelling the Passenger Choice Behaviour of Buying High-Speed Railway Tickets” *Yan Z, Jian M, Li X, Cao J. 2022*
10. “India Transport Report: Moving India to 2032” *Volume II, National Transport Development Policy Committee, June 2014*