Palacký University Olomouc

Transactions on Transport Sciences
Peer-Reviewed Open Access Journal
Vol. 1/2024 DOI: 10.5507/tots.2023.019
journal homepage: www.tots.upol.cz

Transportation Mode Choice Behavior with Multinomial Logit Model: Work and School Trips

YESHITILA DENEKE[a], ROBEL DESTA[b*], ANTENEH AFEWORK[b], JÁNOS TÓTH[b]

a. Department of Civil Engineering, Institute of Technology, Hawassa University. P.O.Box: 05, Hawassa, Ethiopia.
b. Department of Transport Technology and Economics, Faculty of Transportation Engineering and Vehicle Engineering, Budapest University of Technology and Economics, Muegyetem rkp. 3, H-1111 Budapest, Hungary.

ABSTRACT: This paper focuses on modeling transportation mode choice for commuting to work and school. The study employs a Multinomial Logit (MNL) model to examine how individuals choose their modes of transportation for work and school trips in Hawassa city. Additionally, the study aims to predict both the current and future distribution of transportation modes. To construct the model, surveys were distributed across the city's seven sub-cities, involving inspections of workplace and school travel. Primary data were collected through site visits to key transportation hubs, and travel cost data were obtained from the city's Road and Transportation Bureau. The data used for the MNL model describe the travel behaviors of employees and students, and these behaviors are integrated into the statistical analysis to formulate utility functions. The choice of transportation mode for a trip is treated as the dependent variable, while independent variables include factors like out-of-vehicle travel time, in-vehicle travel time, and travel cost. The model's validity and accuracy were assessed by examining the direction of parameter signs and comparing them to the fundamental properties of the MNL model. The study revealed that factors such as average monthly income, in-vehicle travel time, out-of-vehicle travel time, total travel cost, and comfort during the journey significantly influence individuals' choices of transportation modes. The results from the travel behavior forecast, which examines how employees choose their transportation modes, highlight the pressing need for implementing effective policy measures to incentivize the adoption of more sustainable transportation modes and promote a modal shift. Without such measures, employees are likely to increasingly favor motorcycles as their preferred mode of transportation, potentially exacerbating issues related to fuel consumption and congestion. It's evident that students tend to favor public transportation over motorcycles when selecting their mode of travel. The findings of this study offer valuable insights to decision-makers and transportation planners, shedding light on the critical factors influencing travel patterns, as well as providing estimates of existing and future market shares. These findings can serve as a foundation for crafting targeted policy adjustments to encourage sustainable transportation choices in a comprehensive manner.

KEYWORDS: Mode choice; multinomial logit model; travel behaviour; utility function; work and school trips

1. INTRODUCTION

Urbanization has emerged as a pivotal global phenomenon as an increasing portion of the world's population resides in cities. This phenomenon presents a complex challenge, particularly in metropolitan areas, due to the diverse transportation options available, the variety of starting and ending points, and the substantial volume and diversity of traffic (Rodrigue, 2020). To effectively plan, design, assess, and regulate transportation and supply chain systems, it's imperative to accurately predict overall passenger and freight demand. Furthermore, understanding the competitive and cooperative interactions among various transportation modes is essential (Tskeris & Tsekeris, 2011). The choice of transportation mode is influenced by several crucial factors, including the characteristics of the traveler, the purpose of the trip, and the attributes of the transportation infrastructure (Essam & Sadi, 2013; Tyrinopoulos & Antoniou, 2013; Ortúzar & Willumsen, 2011; Limtanakool, Dijst, & Schwanen, 2006).

The transportation planning process, a critical tool for decision-making, typically involves an iterative modeling approach that considers the interplay between travel demand and the efficiency of multi-modal transport networks. Traditional transportation planning follows a four-step process comprising separate models for estimating trip generation, distribution, modal choice, and traffic assignment (Nair, Dias, Ruiz-Juri, Kuhr, & Bhat, 2018; Ahmed, 2012). Travel demand forecasting models are essential for predicting future traffic patterns and for assessing the need for additional road capacity, alterations in public transit services, and changes in land use policies and trends. These models rely on a series of mathematical models designed to replicate human travel behavior (Ahmed, 2012). However, given the significant uncertainty surrounding the operation of transportation systems in the future, particularly with the advent of autonomous and connected vehicle technologies, there is a growing belief that conventional methods for executing these four planning steps may not be suitable for forecasting traffic conditions beyond a decade from now (Hasnine & Nurul Habib, 2021; Nair et al., 2018).

The choices people make when it comes to their daily commutes are influenced by several factors, with education level, commute distance, and the potential financial gains from their job locations being among the most significant. The emergence of the SARS-CoV-2 pandemic in developing countries has introduced various mobility restrictions between regions that are functionally connected, altering the structure and dynamics of labor markets. Considering this, policymakers should carefully manage the evolving dynamics of the transportation system while also prioritizing commuter safety (Carriel, Lufin, & Pérez-Trujillo, 2022; Mayo, Maglasang, Moridpour, & Taboada, 2022). A plausible approach for addressing issues related to driving-related

* *Corresponding author: e-mail: robeldesta@edu.bme.hu*

fatigue involves a two-pronged strategy. Firstly, interventions should focus on improving the work environment to reduce environmental and behavioral stresses experienced by commuters. Secondly, there should be an emphasis on education and awareness campaigns aimed at combating fatigue and enhancing road safety. For instance, levels of stress associated with commuting tend to decrease when commute distances and trip lengths are shortened. Additionally, implementing environmental improvements, such as promoting the use of non-motorized transportation modes for daily commuting, can have a positive impact (Cendales, Llamazares, & Useche, 2023).

Statistical models serve as invaluable tools for policymakers and decision-makers, providing valuable insights into the mobility behavior of commuters and travelers. They enable the development of adaptable and more efficient strategies for managing mobility (Tyrinopoulos & Antoniou, 2013). Understanding the choice of transportation mode is crucial because it directly affects travel efficiency, the urban area required for transportation services, and the range of mode options available to users. Mode choice, a pivotal step in transportation planning, has a direct impact on policy decisions (Eom, Lee, Ko, & Lee, 2022; Ali et al., 2021; Sekhar, 2014). Discrete choice models are commonly used to analyze and predict a decision maker's selection of one alternative from a finite set of mutually exclusive and collectively exhaustive options (Koppelman & Bhat, 2006). Historically, the Discrete Choice model has been the dominant approach in travel mode choice modeling. However, with advancements in computing techniques, machine learning has also gained prominence in the analysis of travel behavior. Machine learning methods such as Neural Networks, Random Forest, Decision Trees, and Support Vector Machines have become popular tools in choice modeling (Ali et al., 2021; Cheng, Chen, De Vos, Lai, & Witlox, 2019).

Urban transportation planning and policy analysis primarily focus on work and school trips because these journeys occur regularly, unlike social outings, shopping trips, or long-distance travel. When implementing measures to reduce congestion, it's important to consider that some travelers might switch transportation modes. This can lead to changes in the total distance traveled by each mode. For instance, if less polluting modes of transportation see an increase in travel distance, the overall emissions may decrease, and vice versa (Bull & CEPAL, 2003). Mode choice is a critical factor in estimating the number of trips for each available transportation mode, and understanding the preferences of travelers in selecting modes is essential for accurate quantification. Work and school trips are influenced by a variety of personal characteristics and urban transportation system attributes, including age, gender, income, comfort, travel time, and travel cost. This study specifically targets employees and students in different sub-cities within Hawassa city. Consequently, the survey was designed to capture the factors influencing the mode choice behavior of these two groups. Additionally, using collected data on travel time and waiting time at major road segments and stations, a multinomial logit (MNL) model was developed in SPSS Software to forecast the potential future market share of various transportation modes.

2. LITERATURE REVIEW

Previous research endeavours have broadly explored mode choice behaviors and the factors influencing them. These studies aim to provide valuable insights for decision-makers to formulate effective transportation policies that can mitigate challenges arising from urbanization and rapid population growth. For example, specific investigations have concentrated on developing mode choice models tailored to work trips, analyzing the factors that impact the mode preferences of employed individuals (Essam & Sadi, 2013; Tushara, Rajalakshmi, & Bino, 2013). Other studies have examined travel-to-school mode choice patterns (Müller, Tscharaktschiew, & Haase, 2008), the relationship between land use, spatial configuration of transportation systems, and mode choice for medium and long-distance travel (Šinko, Rupnik, Prah, & Kramberger, 2021; Takahashi, 2019; Limtanakool et al., 2006). There have also been explorations into travel and mode preferences in the context of smart cities, with a specific focus on external commuting trips (Eom et al., 2022). Researchers have also investigated the influence of transportation supply, spatial characteristics, and socio-economic factors on individual mode choice within metropolitan areas (Al-Salih & Esztergár-Kiss, 2021; Harbering & Schlüter, 2020; Khan, Kruger, & Trivedi, 2007). Transit users' preferences for emerging competitive transit options, including fixed and flexible routes, have also been analyzed (Suaa et al., 2022; Saxena, Rashidi, & Auld, 2019; Chavis & Gayah, 2017; Tuan, 2015; Tyrinopoulos & Antoniou, 2013). Furthermore, research has delved into the impact of travel attitudes on mode choice (De Vos et al., 2022) and explored a wide range of related topics, contributing to a comprehensive understanding of the intricate dynamics of transportation mode choice.

The MNL model is a widely used tool in research papers investigating mode choice behaviors, especially for work trips, school trips, and individual choices within households (Harbering & Schlüter, 2020; Tuan, 2015; Essam & Sadi, 2013; Tushara et al., 2013; Müller et al., 2008). For studies focusing on medium- and longer-distance travel, researchers often employ the broader family of logit models, including Nested Logit, Multinomial Logit, and Binary Logit models, or combinations thereof (Al-Salih & Esztergár-Kiss, 2021; Chavis & Gayah, 2017; Shang & Zhang, 2013; Khan et al., 2007; Limtanakool et al., 2006). To balance subjective and objective factors in mode choice, some studies utilize a combination of the Analytic Hierarchy Process (AHP) method and Geographic Information Systems (GIS) tools (Šinko et al., 2021). In emerging smart cities, mixed logit mode choice models have been used to examine mode choice (Eom et al., 2022). For investigations into transit users' choices, researchers have employed Probit and Structural Equation Models (Tyrinopoulos & Antoniou, 2013). A Latent Class Choice Model (LCCM) has been used to study the factors influencing Mode Choice Behavior of Travelers under Transit Service Disruptions (Saxena et al., 2019). Tour-based mode choice modeling serves as a core component of activity-based travel demand modeling frameworks, especially in the study of features related to automated vehicles and Mobility-as-a-Service (Hasnine & Nurul Habib, 2021). Additionally, machine learning methods are gaining traction in predicting users' travel model preferences, often in comparison to traditional discrete choice models (Ali et al., 2021; Richards & Zill, 2019; Sekhar, 2014). These diverse modeling approaches provide a comprehensive toolkit for understanding and predicting transportation mode choices in various contexts.

The possession of a motor vehicle plays a pivotal role in individuals' transportation mode choices (Harbering & Schlüter, 2020). Availability of parking space is another crucial factor (Tyrinopoulos & Antoniou, 2013; Olsson, 2003). However, due to factors like familiarity or personal preference, some individuals consistently choose the same mode regardless of other considerations (Chavis & Gayah, 2017). A mode choice study focused on work trips identified several significant factors affecting mode choice, including total travel time, total cost relative to personal income, ownership of transportation means, distance, age, and average family monthly income (Essam & Sadi, 2013). Findings from a logit model developed for mode choice on railway route corridors for work and school trips revealed that work type (government, private, self-employed), out-of-vehicle time, vehicle travel time, and income are critical factors influencing employees' mode choices. Regarding the Light Rail Transit (LRT) system, users of the Higer bus and Anbesa bus showed the most interest in switching to the LRT mode, followed by taxi users among both employees and students. However, there was minimal interest from private automobile users (Mengistu & Teklu, 2015). In an Asian city, motorcycle mode dominated, suggesting a need for investment in mass rapid transit (MRT) systems (Tuan, 2015). Another study conducted in Calicut city found that 62.8% of respondents preferred two-wheelers for their work trips, with age, gender, income, time, and cost being the most influential factors (Tushara et al., 2013). These findings highlight the multifaceted nature of mode choice decisions and the importance of various factors in shaping transportation preferences.

There remains a notable gap in the existing literature regarding comprehensive mode choice studies that encompass both essential travel types, namely work and school trips, in the context of urban transportation planning and policy analysis. Consequently, this study aims to address this gap by focusing on understanding the factors that influence the mode choice behavior of employed individuals and students. Additionally, it seeks to predict the potential future market share of various transportation modes using the MNL model. By addressing this gap, the present study endeavors to provide valuable insights for more effective urban transportation planning and policymaking.

3. MATERIALS AND METHODS

3.1 Sampling and data collection

In the study area, a multistage sampling method is employed. Initially, the research region is divided into distinct, non-overlapping sections using cluster sampling. Subsequently, governmental employee institutions and university colleges are deliberately selected from these sub-cities due to the importance of minimizing travel time and ensuring punctuality. Lastly, questionnaires are randomly distributed to the chosen government institutions and colleges using a simple random sampling technique. Individuals employed by various organizations, including the Hawassa City Road and Transportation Bureau, SNNPR Transportation Bureau, Hawassa Municipality, Hawassa City Administration, Hawassa Industrial Park, and Hawassa Referral Hospital, were among the primary participants randomly chosen for this research. As for college students in Hawassa City, the study included random selections from Rift Valley College, Info Link College, Zion College, and Pharma Health Science College.

Sample size determination: Calculating the required sample size for multinomial logistic regression can be intricate. However, drawing from the insights of Peduzzi et al.'s research, it's possible to propose a guideline for the minimum number of instances to incorporate into the study (Peduzzi et al., 1996). In this context, if p represents the lesser proportion of negative or positive cases within the population—typically 0.5 or 50%—and k stands for the count of covariates (independent variables), the corresponding minimum number of cases can be determined (as indicated in Equation 1):

(1) $N = 10\frac{k}{p}$

For students = $10\frac{k}{p}$ = $10*\frac{8}{0.5}$ = 160 whereas, for employee = $10\frac{k}{p}$ = $10*\frac{9}{0.5}$ = 180. Hence, a minimum sample size of 340 is required for both students and employees. The allocation of sampling percentages for each sub-city within Hawassa city is subsequently determined according to their demographic characteristics, as detailed in Table 1.

S/N	Name of the sub city	Male	Female	Total	Calculated Sample Percentage
1	Misrak subcity	20,694	18,737	39,431	14.87%
2	Menaheria sub city	21,307	20,338	41,645	15.71%
3	Tabor subcity	38,993	35,064	74,057	27.94%
4	Mehal ketema subcity	12,640	12,245	24,885	9.39%
5	Addis ketema subcity	15,150	15,146	30,296	11.43%
6	Haik dar sub city	15,218	14,321	29,539	11.14%
7	Bahel adarash sub city	12,669	12,568	25,237	9.52%
	Total	**136,671**	**128,419**	**265,090**	**100%**

Table 1. Number of the people and sample percentage in each sub-city of the Hawassa city (Gebre & Quezon, 2021).

Study variables: The independent variables considered for employees encompass age, gender, average monthly income, family size, car ownership, travel cost (TC), out-of-vehicle travel time (OVTT), in-vehicle-travel time (INVTT), and comfort. On the other hand, for students, the independent variables consist of age, gender, average monthly income, travel cost (TC), out-of-vehicle travel time (OVTT), in-vehicle-travel time (INVTT), waiting time, and comfort. The dependent variables under examination encompass various transportation service modes, both motorized and non-motorized, which include City bus, Minibus taxi, Damas taxi (a minivan 4-wheeler taxi for 6-8 passengers), Bajaj taxi (a 3-wheeler for 3 passengers), Motorcycle (2-wheeler), Bicycle, and Walking.

Data Collection: This study employs a combination of primary and secondary data sources. Primary data are collected through the distribution of questionnaires to college students and city government employees. Secondary data, on the other hand, are provided by the Hawassa city road and transportation bureau. Additionally, travel surveys are conducted using the stopwatch method to measure the time passengers spend waiting for a vehicle to arrive as well as the time it takes for the vehicle to depart from the station.

The collected data is subjected to analysis using SPSS, with the aim of identifying the most significant factors influencing passengers' choices of transportation modes. This determination is accomplished through the utilization of the relative importance index method, as outlined in Equation 2. When constructing a model for passenger mode choice, it's imperative to give particular attention to factors with higher importance indexes. These indexes are computed on a four-point scale, ranging from 0 (indicating a very low degree of influence) to 4 (indicating a very high degree of influence), based on their effects (Aibinu & Jagboro, 2002; Lim & Alum, 1995).

(2) $\mathrm{RII} = \frac{4n_1+3n_2+2n_3+1n_4+0n_5}{4N}$

Where:
N = the total number of the respondents,
n_1 = the number of frequencies for a very high (extremely high) significant response,
n_2 = the number of frequencies for highly significant response,
n_3 = the number of frequencies for moderately significant response,
n_4 = the number of frequencies for low (slightly) significant response,
n_5 = the number of frequencies for a very low (extremely low) significant response, i.e., not significant

The analysis of data, including the testing and estimation of the utility function, is facilitated through multinomial logistic regression for commuting to work and school within the study area. The mathematical foundation of the logit models is rooted in the theory of utility maximization (Ben-Akiva, Lerman, & Lerman, 1985), and this technique is widely recognized and established for modeling travel mode choices (Hensher & Johnson, 2018). In accordance with the work of Koppelman and Bhat (2006), the utility experienced by an individual (denoted as n) for a specific mode (represented as j), denoted as V_{jn}, is formulated as a linear function of the explanatory variables (as illustrated in Equation 3).

$$(3) \quad V_{jn} = \beta_{oj} + \beta_{1j}X_{1n} + \beta_{2j}X_{2n} + \ldots\ldots\ldots\ldots\ldots + \beta_{nj}X_{qn}$$

Where:
β_{oj} = the alternative specific constant for mode j,
β_{1j}, β_{2j} ... β_{nj} = the coefficients associated with explanatory variables,
X_{1n}, X_{2n} ...., X_{qn} = the explanatory variables for individual n, and
q = the number of the explanatory variables included in model

The utility functions are thoroughly examined and assessed to determine the most appropriate model specification for the study area. A range of methods is employed to evaluate the coefficient of determination, helping to pinpoint the most suitable model specification for the studied region.

3.2 Model development

The SPSS program is utilized for the analysis of discrete choice data. The results obtained from the MNL model help ascertain whether a significant relationship exists between the mode choice, which serves as the dependent variable, and the set of independent variables.

The likelihood ratio test is employed to determine whether the model performs better than a null model in terms of fitting the data. It is typically preferable to hypothesize that there is no significant difference between the null and final models. The chi-square statistic represents the difference between the -2 log-likelihoods of the Null and Final models. The null hypothesis is rejected when the significance value is less than 0.05, indicating that the final model fits the data appropriately. In the Multinomial Logistic Regression procedure, both the Pearson and Deviance goodness-of-fit indices are calculated. When assessing these indices, the null hypothesis is rejected if the significance value is less than 0.05. However, if the significance value is greater than 0.05, the null hypothesis is accepted, indicating that the model is adequately fitting the data (Lewis, Butler, & Gilbert, 2011).

The Pseudo R-square, calculated with two independent variables, measures how effectively it explains the variance in the dependent variables. It falls within a range of zero to one, where zero signifies no variation explained between the dependent and independent variables, while one indicates a perfect fit, with all variation accounted for. In the context of the best models, higher Pseudo R-square values denote stronger explanatory power and associations within the model. Cox and Snell R-square, similar to R-square in ordinary linear regression, assesses model fit, with higher values suggesting better fit. However, Nagelkerke introduced a modification to this metric, expanding its range from 0 to 1. According to the literature, Nagelkerke R-square values of 0.2 to 0.4 are considered indicative of an excellent model fit (Domencich & McFadden, 1975).

Parameter estimations play a crucial role in simplifying the analytical process and in selecting the most appropriate model. By designating one mode as a reference category, the results of the parameter estimations become more interpretable. Parameters with significant negative coefficients are indicative of a reduced likelihood of the corresponding response category compared to the reference category. Conversely, parameters with positive coefficients are indicative of an increased likelihood of the corresponding response category. This approach simplifies the interpretation of the model's findings and its practical implications.

3.3 Verification and validation of the multinomial logit model

In line with Koppelman and Bhat's work (2006), the developed MNL model possesses three crucial properties. The model's validity and reliability are confirmed by assessing whether it adheres to these three properties characteristic of the MNL model.

The sigmoid or S shape: The MNL model's S-shaped probabilities represent the probability as a function of its own utility while keeping the utilities of the other alternatives constant. This sigmoid or S-shaped function illustrates the likelihood of selecting various alternatives within the model. To convert any expected probability into a value between 0 and 1, the sigmoid function can be effectively utilized.

Equivalent differences property: The choice probability equations remain unaltered when the same incremental value, denoted as α, is added to the utility of each alternative in the existing probability function (as shown in Equation 4).

$$(4) \quad \Pr(i) = \frac{\exp(Vi)}{\exp(Vi)+\exp(Vj)+\exp(Vk)}$$

Where, i is the alternative for which the probabilities are being computed.

When α is added to the systematic components of modes i, j, and k, the resulting probability remains identical to the probability obtained when α is not added to any of the utilities. This holds true for any value of α, as illustrated in Equation 5.

$$(5) \quad \Pr(i) = \frac{\exp(Vi+\alpha)}{\exp(Vi+\alpha)+\exp(Vj+\alpha)+\exp(Vk+\alpha)}$$

Independence of Irrelevant Alternatives Property: The ratio of the probabilities of choosing two alternatives remains unaffected by the presence or characteristics of any other alternative when an individual makes a choice. This assumption is based on the idea that the other alternatives have no bearing on the decision between the two alternatives in the pair being considered.

$$(6) \quad \boxed{\begin{array}{c} \Pr(i) = \frac{\exp(Vi)}{\exp(Vi)+\exp(Vj)+\exp(Vk)} \qquad \Pr(j) = \frac{\exp(Vj)}{\exp(Vi)+\exp(Vj)+\exp(Vk)} \\ \Pr(k) = \frac{\exp(Vk)}{\exp(Vi)+\exp(Vj)+\exp(Vk)} \end{array}}$$

The ratios of probabilities for each pair of alternatives described in Equation 6 are equivalent to the exponential of the utility function, as indicated in Equation 7.

$$(7) \quad \frac{p(i)}{p(j)} = \frac{exp\ (Vi)}{exp\ (Vj)}, \ \frac{p(i)}{p(k)} = \frac{exp\ (Vi)}{exp\ (Vk)} \text{ and } \frac{p(j)}{p(k)} = \frac{exp\ (Vj)}{exp\ (Vk)}$$

4. RESULT AND DISCUSSION

4.1 Mode choice pattern of Employees

Importance of factors that affect mode choice of employees: The factors influencing employees' mode choice have been organized in a ranking order using relative importance indexes, as indicated in Table 2.

Parameter estimates for the model: Parameter estimates were computed to facilitate a comparative analysis, with the minibus category serving as the reference point, and the minimum acceptable model was subsequently established (refer to Table 3). When all other variables in the model are held constant, a positive parameter B signifies a preference for the specified mode of transportation over the minibus. However, it's important to note that the Wald test statistic for the predictor and its associated p-value (Sig.) exceeding 0.05 indicate that the regression coefficient is statistically insignificant and should not be relied upon. For example, in Table 3, the positive parameter sign for government employees suggests that they have a 1.754 times higher preference for private cars over minibuses when all other variables are held constant. Nevertheless, the Wald test statistic for the predictor "government employee" is recorded as 0.000, with a p-value significantly greater than 0.05. This outcome implies that the regression coefficient for government employees is statistically insignificant and should not be considered valid.

When all other variables in the model are held constant, a negative parameter B signifies a preference for the specified mode of transportation less than that of the minibus. In this case, the regression coefficient is statistically significant and reliable, as evidenced by a Wald test statistic and an associated p-value that are both less than 0.05. For instance, when considering all other variables constant in the model, the negative parameter sign for government employees in relation to City bus indicates that employees are 0.140 times less likely to prefer the city bus over the minibus. The Wald test statistic for the predictor "out of vehicle time" is 5.431, with a p-value of 0.02, which is less than 0.05. This outcome suggests that "out of vehicle time" is statistically significant and valid in this context.

Factor	RII	Rank
Travel cost	0.74	1
Out of the vehicle travel time	0.73	2
Average monthly income	0.69	3
Comfort	0.67	4
In vehicle travel time	0.66	5
Weather conditions	0.54	6
ITS	0.52	7
ATM	0.51	8
Age	0.48	9
Gender	0.38	10

Table 2. Relative importance index and rank of the factors that affect mode choice.

Given that the Wald test for the predictor and its associated p-value are both below 0.05, it's appropriate to incorporate the positive sign of parameter B as an addition and the negative sign of parameter B as a subtraction in the empirical utility functions. In these utility functions, the reference category is taken as the minibus. The empirical utility functions for various modes can be formulated as follows:

City bus = − 4.993 + 0.097 ***OVTT*** + 0.125 ***INVTT*** − 0.14 ***TC***
Minivan = − 3.268 + 0.133 ***OVTT*** − 0.086 ***INVTT*** + 0.012 ***TC***
3-Wheeler = −6.89 + 0.413 ***OVTT*** − 0.26 ***INVTT*** + 0.028 ***TC***
Motorcycle = − 8.983 + 0.31 ***INVTT*** + 0.142 ***TC***

Transport Mode	Attributes	B	Std. Error	Wald	Df	Sig.	Exp (B)
Private car	Intercept	-97.928	30973.063	.000	1	.997	
	OVTT	1.754	1490.632	.000	1	.999	5.777
	INVTT	1.685	2742.755	.000	1	1.000	5.393
	TC	.160	78.049	.000	1	.998	1.173
City Bus	Intercept	-4.993	3.618	1.904	1	.168	
	OVTT	.097	.142	.459	1	.050	1.101
	INVTT	.125	.206	9.181	1	.002	1.868
	TC	-.140	.017	5.431	1	.020	.961
Minivan	Intercept	-3.268	3.737	4.896	1	.027	
	OVTT	.133	.146	.829	1	.042	1.142
	INVTT	-.086	.186	2.376	1	.023	1.331
	TC	.012	.012	1.048	1	.031	1.012
3-Wheeler	Intercept	-6.890	2.950	5.455	1	.020	
	OVTT	.413	.114	13.038	1	.000	1.511
	INVTT	-.260	.154	2.872	1	.040	.771
	TC	.028	.009	9.760	1	.002	1.028
Motorcycle	Intercept	-8.983	3.629	6.126	1	.013	
	OVTT	-.009	.133	.004	1	.948	.991
	INVTT	.310	.202	2.503	1	.014	1.377
	TC	.142	.010	5.699	1	.017	1.024
Bicycle	Intercept	10.743	23731.809	.000	1	1.000	
	OVTT	-1.547	464.560	.000	1	.997	.213
	INVTT	1.376	1603.519	.000	1	.999	3.958
	TC	-.275	75.339	.000	1	.997	.759
Walking	Intercept	25.241	.000	.	1	.	
	OVTT	-.196	1659.201	.000	1	1.000	.822
	INVTT	-3.306	5931.786	.000	1	1.000	.037
	TC	-.088	428.541	.000	1	1.000	.916

Table 3. Parameter estimates and their statistical significance.

Mode alternatives	Utility		Exponent	Probability
	Expression	Value		
City bus	– 4.993+0.097*15+0.125*20 – 0.140*7.3	– 2.071	0.126	28.83%
Minivan	– 3.268+0.133*15– 0.086*12+0.012*9.1	– 2.196	0.111	25.4%
3-Wheeler	– 6.89+0.413*15 – 0.26*12+0.028*10	– 3.615	0.027	6.2%
Motorcycle	– 8.983+0.31*15 + 0.142*18.18	– 1.753	0.173	39.58%
			0.437	

Table 4. Yielded survey data from the travelers.

Prediction of the probability for the modal split alternatives: An MNL model is a suitable approach for characterizing how people in the study area select between transportation modes such as City bus, Minivan, 3-Wheeler, and Motorcycle. The available dataset originates from a travel survey conducted in the study area, and the details are documented in Table 4. Within this dataset, the daily travel cost (TC) is computed by dividing the average monthly travel expenses of employees by 22, which presumably represents the typical number of working days in a month.

Beyond assessing the current distribution of transportation modes in the city, the developed model also has the capability to forecast the likelihood of the city's future modal distribution. To gain insight into the model's behavior in the future, using data gathered by the relevant organization, let's consider a hypothetical scenario for the upcoming year. In this hypothetical scenario, it's assumed that due to increased levels of fuel consumption and traffic congestion, the variables INVTT, OVTT, and TC will each experience an increase of 5 minutes, 3 minutes, and 1 birr, respectively (as outlined in Table 6).

Mode	INVTT	OVTT	Average TC (daily amount in ETB)
City Bus	20-25	13-18	8.3
Minivan	13-17	13-18	10
3-Wheeler	11-17	13-18	11
Motorcycle	13-20	0	19.18

Table 6. The yielded data determined for the hypothetical future modal split.

The projected probability for the future distribution of modal choices (as presented in Table 7) offers insights into the anticipated developments based on the current pattern of mode shares. Consequently, unless a greater emphasis is placed to prioritize more sustainable transit modes, the likelihood of travelers choosing the Motorcycle as their preferred mode will be significantly higher than that for those considering the City bus, Minivan, and 3-Wheeler options, respectively.

Verification and validation of the model with the Sigmoid or S-shape Probabilities of MNL model: The S-shaped curve illustrates how the probability of the current modal distribution varies. It spans from a low of 0.062 when the utility of the 3-Wheeler is notably lower than other options to a high of 0.39 when the utility of the motorcycle is considerably superior to other alternatives. As utility increases, so does the probability, reaching a maximum point where the curve's slope is steepest (refer to Figure 1). The gradual slope of the motorcycle's utility curve at extreme values becomes much steeper as its utility approaches a point where the choice probability is nearly zero. This means that as shown in Figure 1, the likelihood of future mode share for motorcycles increases as utility approaches zero, but only up to a certain point in the future until it reaches its peak. After reaching this maximum capacity, employees may transition to other private modes of transportation, such as private cars, as people's utility and satisfaction are limitless and not constrained.

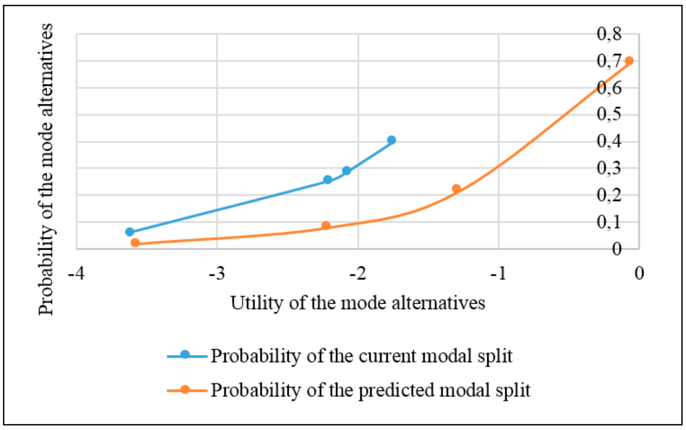

Figure 1: The Sigmoid or S-shape Probabilities of MNL model for employee's modal split.

Verification and validation of the model with the Equivalent Differences Property: The equivalent differences property, as described by Koppelman and Bhat (2006), states that the choice probability equations remain unaffected when a uniform incremental value, such as 1, is added to the utility score of each alternative within the equation. This property highlights the notion that relative differences in utility values among alternatives are more critical for choice probabilities than the absolute utility values themselves.

Mode alternative	Utility		Exponent	Probability
	Expression	Value		
City bus	– 4.993+0.097*18+0.125*25– 0.140*8.3	– 1.284	0.277	21.5%
Minivan	– 3.268+0.133*18–0.086*17+0.012*10	– 2.215	0.109	8.04%
3-Wheeler	– 6.89+0.413*18– 0.26*17+0.028*11	– 3.568	0.028	2.06%
Motorcycle	– 8.983+0.31*20 +0.142*19.18	– 0.059	0.942	69.46%
			1.356	

Table 7. Prediction the probabilities of the hypothetical future modal split with MLM.

The computations supporting this assertion (refer to Table 8) suggest that the probability of each alternative, based on the initial set of modal utilities (as shown in Table 5) before adding one, is identical to the probability of each alternative after adding one unit to the modal utility values. This property demonstrates that such uniform increments do not alter the relative probabilities among the alternatives.

Mode alternative	Utility		Exponent	Probability
	Expression	Value		
City bus	– 2.071	– 1.071	0.343	28.84%
Minivan	– 2.196	– 1.196	0.302	25.4%
3-Wheeler	– 3.615	– 2.615	0.073	6.2%
Motorcycle	– 1.753	– 0.753	0.471	39.59%
			1.189	

Table 8: Probability of each alternative on the initial set of modal utilities after adding one.

Verification and validation of the model with the Independence of Irrelevant Alternatives Property: This property, stating that the probability ratio of choosing two alternatives by an individual remains consistent regardless of the presence or attributes of any other alternatives, is as expected. The probability ratios for each pair of alternatives and the ratio of their corresponding exponent utility values are constant, and they are solely determined by the attributes of those specific alternatives. These ratios are not influenced by the attributes of a third alternative, which remains consistent whether third alternative is part of the choice set (as demonstrated in Table 9).

For City Bus and Minivan	For City Bus and 3-Wheeler
$\frac{p(bus)}{p(damas)} = \frac{\exp(Vbus)}{\exp(Vdamas)}$	$\frac{p(bus)}{p(bajaj)} = \frac{exp\ (Vbus)}{exp\ (Vbajaj)}$
$\frac{0.2883}{0.2540} = \frac{0.126}{0.111}$	$\frac{0.2883}{0.0620} = \frac{0.126}{0.027}$
1.13 = 1.13	4.65 = 4.66
For City Bus and Motorcycle	**For Minivan and 3-Wheeler**
$\frac{p(bus)}{p(motorcycle)} = \frac{exp\ (Vbus)}{exp\ (Vm/cycle)}$	$\frac{p(damas)}{p(bajaj)} = \frac{exp\ (damas)}{exp\ (Vbajaj)}$
$\frac{0.2883}{0.3959} = \frac{0.126}{0.173}$	$\frac{0.254}{0.062} = \frac{0.111}{0.027}$
0.73 = 0.73	4.11 = 4.11
For Minivan and Motorcycle	**for 3-Wheeler and Motorcycle**
$\frac{p(damas)}{p(m/cycle)} = \frac{exp\ (Vbdamas)}{exp\ (Vm/cycle)}$	$\frac{p(bajaj)}{p(motorcycle)} = \frac{exp\ (Vbajaj)}{exp\ (Vm/cycle)}$
$\frac{0.2540}{0.3959} = \frac{0.111}{0.173}$	$\frac{0.0620}{0.3959} = \frac{0.0126}{0.173}$
0.64= 0.64	0.157=0.157

Table 9. The ratios of each pair of probabilities.

4.2 Mode choice pattern of students

Importance of factors that affect mode choice: The RII serves as a tool for evaluating the relative significance of different factors affecting mode choice behaviors. This index is calculated by examining the frequency distribution of each factor that has an impact on mode choice behavior, as presented in Table 10.

Parameter estimates: When all other variables in the model are kept constant, a negative parameter B (as seen in Table 11) signifies that the mode of transportation is favored less than the minibus. Notably, the regression coefficient for this parameter has been determined to be statistically significant and reliable, as indicated by a Wald test statistic and its associated p-value (Sig.) being below the 0.05 threshold.

Identified Factors	RII	Rank
Average monthly income	0.66	1
Out of the vehicle travel time	0.64	2
Travel cost	0.64	3
In vehicle travel time	0.63	4
Comfort	0.61	5
Weather conditions	0.49	6
ITS	0.45	7
ATM	0.40	8
Age	0.38	9
Gender	0.36	10

Table 10: Relative importance index and ranking of the factors that affect mode choice.

Mode of transport	Attributes	B	Std. Error	Wald	df	Sig.	Exp(B)
City Bus	Intercept	-5.503	7.023	.005	1	.943	
	OVTT	.135	.284	3.029	1	.042	1.640
	INVTT	.162	.255	4.979	1	.026	1.767
	TC	-.224	.056	5.066	1	.024	.881
Minivan	Intercept	-7.725	6.159	1.573	1	.210	
	OVTT	.575	.204	.005	1	.041	.985
	INVTT	-.258	.363	1.592	1	.027	.633
	TC	.079	.030	7.174	1	.007	1.082
3-Wheeler	Intercept	-2.490	3.800	.429	1	.512	
	OVTT	.233	.127	3.375	1	.036	.792
	INVTT	-.314	.220	7.818	1	.005	.541
	TC	.085	.022	14.493	1	.000	1.089
Motor cycle	Intercept	2.078	6.557	.220	1	.639	
	OVTT	-.435	.213	6.297	1	.112	.586
	INVTT	-.515	.589	5.777	1	.016	.243
	TC	.095	.023	16.453	1	.000	1.100
Bicycle	Intercept	12.403	.000	.	1	.	
	OVTT	-4.426	3806.078	.000	1	.999	.012
	INVTT	1.498	3047.056	.000	1	1.000	4.471
	TC	-.150	388.585	.000	1	1.000	.861
Walking	Intercept	23.733	30716.378	.000	1	.999	
	OVTT	.336	2360.155	.000	1	1.000	1.399
	INVTT	.340	1320.445	.000	1	1.000	1.405
	TC	-.579	255.507	.000	1	.998	.561

Table 11. Parameter estimates for school trips.

The empirical utility functions for various modes should incorporate the positive parameter B as an addition, while for modes with a negative parameter B, it should be subtracted, with the minibus serving as the reference category. Consequently, the empirical utility functions for different modes are as follows:

City bus = – 5.503 + 0.135 OVTT + 0.162 INVTT – 0.224 TC
Minivan = – 7.725 + 0.575 OVTT – 0.258 INVTT + 0.079 TC
3-Wheeler = – 2.49 + 0.233 OVTT – 0.314 INVTT + 0.085 TC
Motorcycle = 2.078 – 0.515 INVTT + 0.095 TC

Prediction of the probability for the modal split mode choice alternatives: A travel survey conducted among trip makers in the study area to gather information about their travel experiences resulted in the following dataset (refer to Table 12).

Mode	INVTT	OVTT	Monthly TC	Average TC (daily amount in ETB)
City Bus	15-20	10-15	80-140	5.5
Minivan	8-12	10-15	120-150	6.8
3-Wheeler	6-12	10-15	140-180	8.5
Motorcycle	8-15	0	250-400	18.18

Table 12. Yielded survey data from the traveler.

In the present distribution of transportation modes for school trips within Hawassa city, travelers show a higher likelihood of selecting the City bus as their preferred choice. Following the City bus, the preferences are in the order of Minivan, Motorcycle, and 3-Wheeler, as indicated by their respective probabilistic percentages (refer to Table 13).

Hypothesizing that in the upcoming year, there will be an increase in fuel consumption and congestion levels, leading to a rise in INVTT, OVTT, and TC by 5 minutes, 3 minutes, and 1 birr, respectively (as specified in Table 14), the model predicts changes in future modal splits. This suggests that the expected alterations in these factors will likely impact the choices individuals make in terms of transportation modes.

According to the prediction probabilities for the future modal split (as outlined in Table 15), it is anticipated that travelers will have a higher likelihood of choosing the City bus as their preferred mode of transportation, compared to the alternatives of Minivan, 3-Wheeler, and Motorcycle.

Verification and validation of the model with the Sigmoid or S shape of MNL Probabilities: The validation of mode choice patterns is conducted by assessing the reasonableness of parameter signs and the quality of predictions. This verification process involves comparing the model's results to the fundamental properties associated with the MNL model. This ensures that the model aligns with the expected behavior and statistical properties inherent to MNL models.

The S-shaped curve in the MNL model illustrates the probabilities of choosing different alternatives (as shown in Figure 2). This graph is constructed by plotting the probability of a mode alternative on the y-axis and the utility function associated with that alternative on the x-axis. The S-shape characterizes how changes in utility values affect the likelihood of selecting each alternative, showcasing the typical behavior observed in MNL models.

The Equivalent Differences Property: The estimates that facilitate this comparison (as presented in Tables 16 and 17) suggest that the probability of each alternative, based on the initial set of modal utilities before adding one unit, remains identical to the probability of each alternative after adding one unit. This finding demonstrates that the addition of a consistent increment to utility values does not alter the relative probabilities among the alternatives, which aligns with the principles of the MNL model.

Independence of Irrelevant Alternatives Property: The probability ratios for each pair of alternatives and the ratio of their respective exponent utility values are consistent (as indicated in Table 18). Importantly, these ratios are solely dependent on the attributes of the two alternatives being compared and are unaffected by the attributes of a third alternative, regardless of whether that third alternative is part of the choice set or not. This underscores the fundamental property of the MNL model, where the relative probabilities among alternatives are determined by the attributes of the alternatives themselves and not influenced by the presence or attributes of other alternatives in the choice set.

Mode alternative	Utility		Exponent	Probability
	Expression	Value		
City bus	– 5.503+0.135*15+0.162*20 – 0.224*5.5	– 1.471	0.2299	35.92%
Minivan	– 7.725+0.575*15– 0.258*12+0.079*6.8	– 1.658	0.1903	29.74%
3-Wheeler	– 2.49+0.233*15 – 0.314*12+0.085*8.2	– 2.066	0.1267	19.79%
Motorcycle	3.078 – 0.515*15 + 0.095*18.18	– 2.374	0.0930	14.53%
			0.6399	

Table 13. Prediction the probabilities of the current modal split with MLM.

Mode	INVTT	OVTT	Average TC (daily amount in ETB)
City Bus	20-25	13-18	6.8
Minivan	13-17	13-18	7.8
3-Wheeler	11-17	13-18	9.2
Motorcycle	13-20	0	19.18

Table 14. The yielded data as determined for hypothetical future modal split.

Mode alternative	Utility		Exponent	Probability
	Expression	Value		
City bus	– 5.503+0.135*18+0.162*25 – 0.224*6.5	– 0.479	0.6194	61.94%
Minivan	– 7.725+0.575*18– 0.258*17+0.079*7.8	– 1.145	0.3183	31.83%
3-Wheeler	– 2.49+0.233*18 – 0.314*17+0.085*9.2	– 2.852	0.0577	5.77%
Motorcycle	3.078 – 0.515*20 + 0.095*19.18	– 5.399	0.0045	0.45%
			0.9999	

Table 15. Prediction the probabilities of the hypothetical future modal split with MLM.

Mode alternative	Utility		Exponent	Probability
	Expression	Value		
City bus	– 1.471	– 1.471	0.6244	35.90%
Minivan	– 1.658	– 1.658	0.5178	29.76%
3-Wheeler	– 2.066	– 2.066	0.3443	19.79%
Motorcycle	– 2.374	– 2.374	0.2531	14.54%
			1.7396	

Table 16: Probability of each alternative at initial set of modal utilities.

Mode alternative	Utility		Exponent	Probability
	Expression	Value		
City bus	– 1.471	– 0.471	0.2299	35.91%
Minivan	– 1.658	– 0.658	0.1903	29.75%
3-Wheeler	– 2.066	– 1.066	0.1267	19.79%
Motorcycle	– 2.374	– 1.374	0.0930	14.53%
			0.6399	

Table 17: Computation after each of the modal utilities are increased by one.

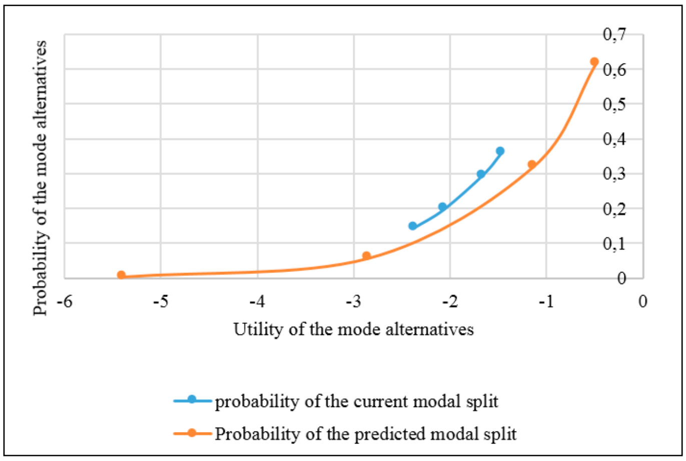

Figure 2: The Sigmoid shape of MNL Probabilities of the current student modal split.

For City bus and Minivan	**For City bus and 3-Wheeler**
$\frac{p(bus)}{p(damas)} = \frac{exp\ (Vbus)}{exp\ (Vdamas)}$	$\frac{p(bus)}{p(bajaj)} = \frac{exp\ (Vbus)}{exp\ (Vbajaj)}$
$\frac{0.3590}{0.2976} = \frac{0.6244}{0.5178}$	$\frac{0.3590}{0.1979} = \frac{0.6244}{0.3443}$
1.21 = 1.21	1.81 = 1.81
For City bus and Motorcycle	**For Minivan and 3-Wheeler**
$\frac{p(bus)}{p(motorcycle)} = \frac{exp\ (Vbus)}{exp\ (Vm/cycle)}$	$\frac{p(damas)}{p(bajaj)} = \frac{exp\ (damas)}{exp\ (Vbajaj)}$
$\frac{0.3590}{0.1454} = \frac{0.6244}{0.2531}$	$\frac{0.2976}{0.1979} = \frac{0.5178}{0.3443}$
2.46 = 2.46	1.50 = 1.50
For Minivan and m/cycle	**for 3-Wheeler and Motorcycle**
$\frac{p(damas)}{p(m/cycle)} = \frac{exp\ (Vbdamas)}{exp\ (Vm/cycle)}$	$\frac{p(bajaj)}{p(motorcycle)} = \frac{exp\ (Vbajaj)}{xp\ (Vm/cycle)}$
$\frac{0.2976}{0.1454} = \frac{0.5178}{0.2531}$	$\frac{0.1974}{0.1454} = \frac{0.3443}{0.2531}$
2.05= 2.05	1.36=1.36

Table 18: Ratios of each pair of probabilities.

4.3 Comparison on the mode choice patterns of employees and students within the city

Utilizing the MNL model, the collective prediction accuracy for employees stands at 81.5%. Surprisingly, despite the increased fuel consumption, motorcycle usage remains high among employees, accounting for approximately 39.58% of the mode share in Hawassa city, according to the probability-based assessment of the current mode distribution. This suggests that many employees opt for personalized modes of transportation to enhance their utility by minimizing travel and waiting times associated with public transit.

In the city, the current mode share for the City bus is around 28.83%, possibly indicating the need for transit priority solutions. This information holds significance for policymakers and transportation planners who face ongoing challenges with the task of offering viable public transportation alternatives. The Minivan has a current mode share of approximately 25.4%, while the 3-Wheeler's share is below 6.2%. This implies that the 3-Wheeler mode is less favored among employees due to its higher travel costs, which aligns with expectations based on actual travel behaviors.

The MNL model has been effective in predicting the modal choices of students, achieving an overall prediction level of 86.3%. Notably, the City bus emerges as the preferred mode for students in Hawassa city, accounting for approximately 35.92% of the current mode share. This suggests that a significant majority of students opt for the City bus to maximize their utility while keeping their travel costs in check.

Within the city, the Minivan captures a mode share of around 29.74%, while the 3-Wheeler mode accounts for approximately 19.79% of the current modal split. These statistics provide valuable guidance for decision-makers and transportation planners, indicating the potential for promoting more sustainable mass transit options to a greater extent. Conversely, the motorcycle mode exhibits a lower current mode share probability of 14.53%. This indicates that students are less inclined to choose motorcycles as a mode of transportation, likely due to the higher travel costs associated with motorcycles, primarily stemming from their elevated fuel consumption.

4.4 Limitations of the study and further research

This study focuses on modeling the transportation choices of employers commuting from home to work and college students traveling from home to school within Hawassa city. It's important to note that this research specifically addresses work and school trips and does not encompass other types of trips, such as social or recreational journeys. Similar modeling approaches could be applied to those trip types in future research.

Furthermore, future studies could enhance the reported findings by adopting more sophisticated methodological designs, including the utilization of machine learning techniques to analyze mode choice patterns. Additionally, given the evolving landscape of transportation, researchers may want to explore the incorporation of emerging modes such as Autonomous Vehicles (AVs) into their analyses to gain insights into their potential impact on travel behavior.

5. CONCLUSIONS

This study undertook an estimation and comparison of mode choice patterns for work and school trips, aiming to understand passenger travel behavior using the Multinomial Logit (MNL) model. Significant factors influencing mode choice behavior were identified through the analysis of primary and secondary data. The developed model underwent validation and verification, aligning with fundamental properties of the MNL model and undergoing parametric assessments.

Based on their relative importance index, the most influential factors affecting employees' mode choice were revealed to be travel cost, out-of-vehicle travel time, average monthly income, comfort, and in-vehicle travel time. Presently, the prediction pattern for employee mode share suggests a distribution of 39.58% for motorcycles, 28.83% for City bus, 25.4% for Minivan, and 6.2% for 3-Wheeler. In a hypothetical future scenario, this mode share is projected to shift to 69.46% for motorcycles, 21.5% for City bus, 25.4% for Minivan, and 6.2% for 3-Wheeler based on the developed model. This underscores the urgency of implementing effective policy measures to incentivize the adoption of more sustainable transportation modes and promote a modal shift. Without such measures, employees are likely to increasingly favor motorcycles as their preferred mode of transportation, potentially exacerbating issues related to fuel consumption and congestion.

As indicated by their relative relevance index, the most crucial factors influencing students' mode choices are average monthly income, out-of-vehicle travel time, travel cost, in-vehicle travel time, and comfort. According to the developed model's probability predictions, the current mode share for students is distributed as 35.92% for City bus, 29.74% for Minivan, 19.79% for 3-Wheeler, and 14.53% for motorcycles. However, in a hypothetical future scenario with assumed parameter changes, the estimated mode share shifts to 61.94% for City bus, 31.83% for Minivan, 5.77% for 3-Wheeler, and 0.45% for motorcycles. Evidently, students' mode choice behavior appears to be closely aligned with public transportation options rather than a preference for motorcycles.

The findings of this study offer valuable insights to decision-makers and transportation planners, shedding light on the critical factors influencing travel patterns, as well as providing estimates of existing and future market shares. These findings can serve as a foundation for crafting targeted policy adjustments to encourage sustainable transportation choices in a comprehensive manner.

ACKNOWLEDGEMENTS

This research was supported by the Hungarian Government and cofinanced by the European Social Fund through the project "Talent management in autonomous vehicle control technologies" (EFOP-3.6.3-VEKOP-16-2017-00001).

DATA AVAILABILITY

All data and models generated or used during the study are available from the corresponding author by request.

CONFLICT OF INTEREST

The authors declare that they have no competing interest.

REFERENCES

Ahmed, B. (2012). The traditional four steps transportation modeling using a simplified transport network: A case study of Dhaka City, Bangladesh. *International Journal of Advanced Scientific Engineering and Technological Research, 1*(1), 19-40.

Aibinu, A. A., & Jagboro, G. O. (2002). The effects of construction delays on project delivery in Nigerian construction industry. *International journal of project management, 20*(8), 593-599.

Ali, N. F. M., Sadullah, A. F. M., Majeed, A. P., Razman, M. A. M., Zakaria, M. A., & Nasir, A. F. A. (2021). Travel Mode Choice Modeling: Predictive Efficacy between Machine Learning Models and Discrete Choice Model. *The Open Transportation Journal, 15*(1).

Al-Salih, W. Q., & Esztergár-Kiss, D. (2021). Linking mode choice with travel behavior by using logit model based on utility function. *Sustainability, 13*(8), 4332.

Ben-Akiva, M. E., & Lerman, S. R. (1985). Discrete choice analysis: theory and application to travel demand, 9.

Bull, A., & CEPAL, N. (2003). Traffic Congestion: The Problem and how to Deal with it. *ECLAC.*

Carriel, V., Lufin, M., & Pérez-Trujillo, M. (2022). Do workers negative self-select when they commute? Evidence for the Chilean case of long-distance commuting. *The Annals of regional science, 69*(1), 255-279.

Cendales, B., Llamazares, F. J., & Useche, S. A. (2023). Are subjective outcomes a "missing link" between driving stress and risky driving behaviors of commuters? Assessing the case of a LMIC. *Safety Science*, 158, 105996.

Chavis, C., & Gayah, V. V. (2017). Development of a mode choice model for general purpose flexible-route transit systems. *Transportation Research Record, 2650*(1), 133-141.

Cheng, L., Chen, X., De Vos, J., Lai, X., & Witlox, F. (2019). Applying a random forest method approach to model travel mode choice behavior. *Travel behaviour and society, 14*, 1-10.

De Vos, J., Singleton, P. A., & Gärling, T. (2022). From attitude to satisfaction: introducing the travel mode choice cycle. *Transport Reviews, 42*(2), 204-221.

Domencich, T. A., & McFadden, D. (1975). Urban travel demand-a behavioral analysis (No. Monograph).

Eom, J. K., Lee, K. S., Ko, S., & Lee, J. (2022). Exploring Travel Mode Preference of External Trips for Smart City Transportation Planning: Sejong, Korea. *Sustainability, 14*(2), 630.

Essam, A., & Sadi, A. (2013). Factors affecting mode choice of work trips in developing cities—Gaza as a case study. Journal of Transportation *Technologies*, 2013.

Gebre, G., & Quezon, E. T. (2021). Modeling public transport users' trip production in Hawassa city, Ethiopia. *Journal of Civil Engineering, Science and Technology, 12*(2), 75-90.

Harbering, M., & Schlüter, J. (2020). Determinants of transport mode choice in metropolitan areas the case of the metropolitan area of the Valley of Mexico. *Journal of Transport Geography, 87*, 102766.

Hasnine, M. S., & Nurul Habib, K. (2021). Tour-based mode choice modelling as the core of an activity-based travel demand modelling framework: a review of state-of-the-art. *Transport Reviews, 41*(1), 5-26.

Hensher, D. A., & Johnson, L. W. (2018). *Applied discrete-choice modelling*. Routledge.

Khan, O. A., Kruger, J., & Trivedi, T. (2007). *Developing passenger mode choice models for Brisbane to reflect observed travel behaviour from the South East Queensland Travel Survey*. In 30th Australasian Transport Research Forum.

Koppelman, F. S., & Bhat, C. (2006). A self instructing course in mode choice modeling: multinomial and nested logit models.

Lewis, F., Butler, A., & Gilbert, L. (2011). A unified approach to model selection using the likelihood ratio test. *Methods in ecology and evolution, 2*(2), 155-162.
Lim, E. C., & Alum, J. (1995). Construction productivity: issues encountered by contractors in Singapore. *International journal of project management, 13*(1), 51-58.
Limtanakool, N., Dijst, M., & Schwanen, T. (2006). The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium-and longer-distance trips. *Journal of transport geography, 14*(5), 327-341.
Mayo, F. L., Maglasang, R. S., Moridpour, S., & Taboada, E. B. (2022). Impact of transport policies to commuter safety in urban cities of a developing country: A sustainability and system perspective. *Case studies on transport policy, 10*(4), 2138-2152.
Mengistu, E.H. & Teklu, B. (2015). *Logit model for school and work trip mode choice on railway route corridor.* MSc thesis, Addis Ababa University, Ethiopia.
Müller, S., Tscharaktschiew, S., & Haase, K. (2008). Travel-to-school mode choice modelling and patterns of school choice in urban areas. *Journal of transport Geography, 16*(5), 342-357.
Nair, G. S., Dias, F. F., Ruiz-Juri, N., Kuhr, J., & Bhat, C. R. (2018). Travel Modeling in an Era of Connected and Automated Transportation Systems: An Investigation in the Dallas-Fort Worth Area (No. D-STOP/2016/143).
Olsson, A. L. L. (2003). *Factors that influence choice of travel mode in major urban areas*. The Attractiveness of Park & Ride.
Ortúzar, J.D. and Willumsen, L.G. (2011) *Modelling Transport*. 4th Edition, Wiley, Hoboken.
Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology, 49*(12), 1373-1379.
Richards, M. J., & Zill, J. C. (2019, October). Modelling mode choice with machine learning algorithms. In Australasian Transport Research Forum (ATRF), 41st, 2019, Canberra, ACT, Australia.
Rodrigue, J. P. (2020). *The geography of transport systems*. 5th ed., Langara College, pp. 83-321.
Saxena, N., Rashidi, T. H., & Auld, J. (2019). Studying the tastes effecting mode choice behavior of travelers under transit service disruptions. *Travel behaviour and society, 17*, 86-95.
Sekhar, C. (2014). Mode Choice Analysis: The Data, the Models and Future Ahead. *International Journal for Traffic & Transport Engineering, 4*(3).
Shang, B., & Zhang, X. N. (2013). *Study of travel mode choice behavior based on nested Logit model*. In Applied Mechanics and Materials, 253, 1345-1350.
Šinko, S., Rupnik, B., Prah, K., & Kramberger, T. (2021). Spatial modelling of the transport mode choice: Application on the Vienna transport network. *Transport, 36*(5), 386-394.
Suaa, A. J. Q., Chuaa, H. N., Khoob, H. L., Lowa, Y. C., Leea, A. S. H., & Ismailc, M. A. (2022). User Mode Choice Behavior in Public Transportation: A Systematic Literature Review. *Jurnal Kejuruteraan, 34*(1), 11-28.
Takahashi, T. (2019). Transportation mode choice and spatial structure of a city.
Tskeris, T., & Tsekeris, C. (2011). Demand forecasting in transport: Overview and modeling advances. *Economic research-Ekonomska istraživanja, 24*(1), 82-94.
Tuan, V. A. (2015). Mode choice behavior and modal shift to public transport in developing countries-the case of Hanoi city. *Journal of the Eastern Asia Society for Transportation Studies, 11*, 473-487.
Tushara, T., Rajalakshmi, P., & Bino, I. K. (2013). Mode choice modelling for work trips in Calicut City. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*.
Tyrinopoulos, Y., & Antoniou, C. (2013). Factors affecting modal choice in urban mobility. *European Transport Research Review, 5*(1), 27-39.