Check for updates

Study on the Influence of Family Factors on the Number of Household E-bikes

Ruolin Jiang*, Lu Ma[a]

Department of Transportation Engineering, Beijing Jiaotong University, Beijing, China
**17120817@bjtu.edu.cn (Corresponding author)*
[a]lma@bjtu.edu.cn

Abstract

In recent years, e-bike, as a simple, economical and convenient travel tool, has become one of the main modes of transportation for short-distance travel. Based on the resident trip survey data of Zhoukou City, Henan Province, a data set of influencing factors with families as the unit has been constructed. In order to find the influence of family factors on e-bike ownership, a multinomial logistic regression model has been established. The results show that the total population of the family, the monthly household income and the family housing area are positively correlated with the number of e-bikes, while the number of private cars and of bicycles is negatively correlated with the number of e-bikes. Through the analysis and summary of the characteristics of different families, relevant suggestions on e-bike sales management and traffic planning are put forward.

Keywords-*multinomial logistic regression; e-bike; family factors*

1. Introduction

China is the largest country in the production and sales of e-bikes in the world. According to the data released by the Ministry of Automotive Industry Information, e-bike was only lower than bicycle in the national vehicle manufacturing volume from 2015 to 2020, ranking second, becoming a large-scale manufacturing industry in China's vehicle manufacturing field. According to the statistics of China Bicycle Association, the annual sales volume of e-bikes has exceeded 30 million, and the social ownership has been close to 300 million [1].

With the gradual enhancement of people's awareness of low-carbon environmental protection and the increasing congestion during rush hours, e-bike is recognized by consumers because of its small size, convenience, economy and energy saving. There has been some research about e-bike in recent years, but their scope and direction are different.

Astegiano P., Tampère C.M.J. and Beckx C. [2] have shown that 27% of Belgian residents preferred to own an e-bike, a bicycle and a car at the same time. Among them, e-bike was usually related to commuting, showing a trend of replacing bicycle. Sun Q., Feng T., Kemperman A. and Spahn A. [3] have found that in the Netherlands, people around 50 years old and retirement age (60-69 years old) owning e-bikes were more likely to change their original travel mode of private cars. Compared with areas with higher urbanization level, the shift from private cars to e-bikes was more significant in rural areas.

Taking Chengdu and Shanghai as research areas, Liu Y. [4] has found that most of e-bike users were in the lower middle-income level, and the substitution trend of e-bike for bicycle and bus was obvious. Taking Zhongshan metropolitan area as research area, Zhang Y., Li Y., Yang X.G., Liu Q.X. and Li C.Y. [5] have proposed that family attributes and measures were the main factors that affect household e-bikes. The attributes of building environment and other competitive modes of transportation also had a certain impact on it. Taking Nanning as research area, Hu Y.C., Chen J., Zou X.J. and Chen Z.W. [6] have shown that under the conditions of serious road congestion and poor comfort of public transport services, lower middle-income residents would prefer to choose e-bikes for commuting.

Most of the above research focused on the travel characteristics of e-bikes from the perspective of travel mode choice and external conditions, while there were few on the family ownership of e-bikes. In related research, the selection and classification of explanatory variables were not detailed enough, and the influence of

© The Author(s) 2023
G. Vilas Bhau et al. (Eds.): MSEA 2022, ACSR 101, pp. 1350–1356, 2023.
https://doi.org/10.2991/978-94-6463-042-8_196

family factors on e-bike ownership was not fully considered. In order to help formulate a reasonable management policy of e-bike and improve the supporting facilities for non-motorized travel, it's necessary to make statistics from the family level and study the change characteristics of household e-bikes.

2. DATA

2.1 Data Sources and Analysis

The data comes from the sampling of resident trip survey in Zhoukou City in 2017. The number of actual samples is 3017 households, with a total population of 9,099 people and an average household size of 3.02.

Among the daily trips of residents in downtown Zhoukou, the proportion of non-motorized travel (including walking, bicycle and e-bike) reaches 74.6%. Due to the low frequency of bus departure and the long waiting time at bus stop, bicycle (including e-bike) + walking is still the main travel mode for most residents. The proportion of e-bike travel has reached 51.7% of all modes.

The data of family factors include population, income and expenditure, housing and vehicle information. Population information includes the family total population, the population over 6 years old and the number of temporary residents. Income and expenditure information includes the monthly household income and the proportion of transportation costs in monthly household expense. Housing information includes the family housing area, housing type and the year the house was built. Vehicle information includes the number of private cars, bicycles, e-bikes, motorcycles and elderly mobility scooters. According to the above information as the dividing standard, the number of families with different e-bike ownership was statistically analyzed.

2.1.1 Population information:

In the families with only one person, 50% of them do not have e-bikes. With the increase of family total population, the proportion of families without e-bikes gradually decreases, while the proportion of families with two or three e-bikes gradually increases. There is no obvious change trend between the number of family temporary residents, the population over 6 years old and the number of household e-bikes, so these two factors are not considered in subsequent model.

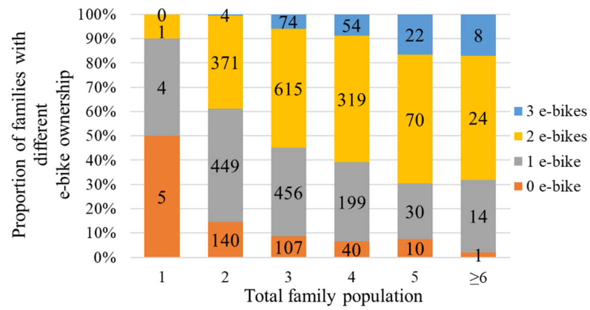

Figure 1. The number and proportion of household e-bikes in different family population

2.1.2 Income and expenditure information:

About 23% of the families with a monthly income of less than 1500 yuan don't have e-bikes, while the proportion of those with a monthly income of more than 1500 yuan doesn't have e-bikes remains around 10%. That is, the proportion of low-income families with e-bikes is significantly lower than that of high-income families. Except for the families whose monthly income is less than 1500 yuan, the proportion of families with two e-bikes is significantly more than that with one. This shows that with the progress of e-bike industry and the rapid development of national economy, families with better income tend to buy more e-bikes when the basic travel demands are met. With the increase of the proportion of transportation costs, the proportion of families without e-bike is generally decreases.

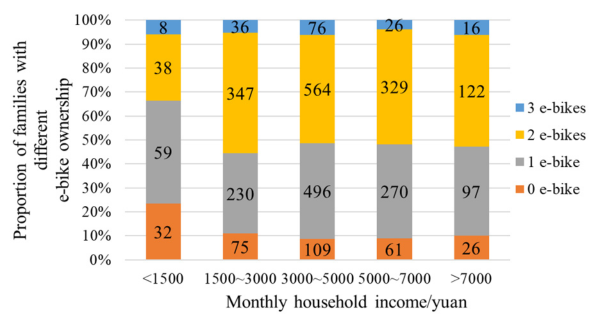

Figure 2. The number and proportion of household e-bikes in different monthly household income

2.1.3 Housing information:

With the increase of housing area, the proportion of families with two or more e-bikes has increased significantly. When the housing area is more than 150 m^2, the proportion of families with two or more e-bikes is more than 50%. With the change of housing type and building age, the change trend of the number of household e-bike is not obvious, so these two factors are not considered in subsequent model.

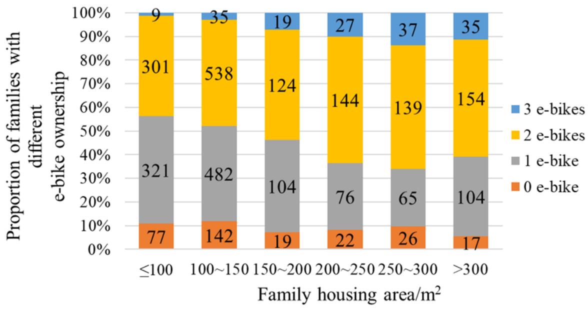

Figure 3. The number and proportion of household e-bikes in different family housing area

2.1.4 *Vehicle information:*

Due to the small number of motorcycles and elderly mobility scooters, more than 80% of the family don't have them, which is easy to affect the accuracy of the model results. Therefore, these two modes of transportation are not considered in subsequent model, only private cars and bicycles remain.

Among the families without private cars, less than 10% don't have e-bikes, and about 60% own two and three e-bikes; among the families with two private cars, about 20% don't have e-bikes, and only about 3% own 3 e-bikes. It can be seen that with the increase of the number of private cars, the proportion of families with e-bike gradually decreases. More than 90% of the families without bicycles have at least one e-bike. With the increase of the number of bicycles, this proportion is gradually decreasing. Private cars, bicycles and e-bikes can be substituted for each other within a certain travel range, so it's necessary to consider these two factors.

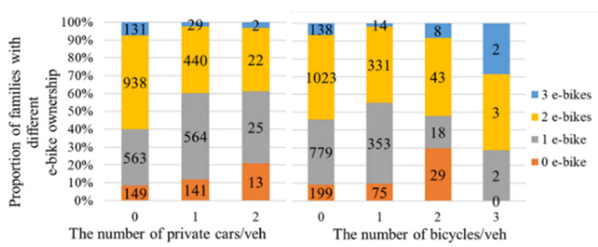

Figure 4. The number and proportion of household e-bikes in different number of private cars and of bicycles

2.2 *Variable Selection*

Through the above analysis, the explanatory variables of the model were preliminarily determined, including family total population, monthly household income, the proportion of transportation costs in monthly household expense, family housing area, the number of private cars and the number of bicycles. The explanatory variables and explained variable were defined and assigned respectively in the following tables.

TABLE 1. DEFINITION AND ASSIGNMENT OF EXPLANATORY VARIABLES

Definition	Type	Assignment
X1: Family total population	Discrete	1,2,3,4,5,6,7
X2: Monthly household income	Ordinal categorical	<1500=0,1500~3000=1, 3000~5000=2,5000~7000=3, >7000=4
X3: Proportion of transportation costs in monthly household expense	Ordinal categorical	<5%=0,5%~10%=1, 10%~15%=2,15%~20%=3, 20%~30%=4, ⩾30%=5
X4: Family housing area	Continuous	0~900m^2
X5: Number of private cars	Discrete	0,1,2
X6: Number of private bicycles	Discrete	0,1,2,3

TABLE 2. DEFINITION AND ASSIGNMENT OF EXPLAINED VARIABLE

Definition	Category	Assignment
Y: Number of household e-bikes	Y1	0
	Y2	1
	Y3	2
	Y4	3

Spearman rank correlation test was used to test the numerical explanatory variables: family total population, family housing area, number of private cars and number of bicycles. The results were shown in the table below. The correlation of these four explanatory variables was significant at the level of 0.01 (bilateral), which meant they passed the correlation test.

TABLE 3. CORRELATION TEST

Spearman Rho	Family total population	Family housing area	Private cars	Bicycles
Correlation coefficient	.224**	.147**	-.200**	-.088**
Sig. (bilateral)	.000	.000	.000	.000
Number	3017	3017	3017	3017

TABLE 4. ONE-WAY ANOVA RESULTS OF MONTHLY HOUSEHOLD INCOME AND PROPORTION OF TRANSPORTATION COSTS

E-bike		Quadratic sum	df	Mean square	F	Sig.
Income	Intraclass	14.233	4	3.558	6.427	.000
	Interclass	1667.479	3012	.554		
	Total	1681.712	3016			
Proportion	Intraclass	2.375	5	.475	.851	.513
	Interclass	1679.338	3011	.558		
	Total	1681.712	3016			

One-way ANOVA was used to test the categorical explanatory variables: monthly household income and the proportion of transportation costs in monthly household expense. Taking the significance level as 0.05, the probability p was less than 0.05, which indicated that different monthly household income had a significant impact on the number of household e-bikes. And it can be preliminarily judged that there is a certain correlation between the them, while the proportion of transportation costs doesn't pass the test.

To sum up, family total population, monthly household income, family housing area, number of private cars and number of bicycles were taken as the relevant explanatory variables.

3. METHODS

3.1 Multinomial Logistic Regression Model

The basic theory of multinomial logistic regression is similar to that of binomial logistic regression. Its purpose is to reflect the comparison between each category and the reference category of the explained variable. Suppose that there are J categories of the explained variables, the last one (i.e., the J-th category) is selected as the reference category. Then other categories are compared with the reference one, and a total of J-1 Logit models are established.

$$\text{Logit } P(y=1|x) = \ln\left(\frac{P(y=1|x)}{P(y=J|x)}\right) = \alpha_1 + \sum_{k=1}^{K}\beta_{1k}x_k = V_1$$
$$\text{Logit } P(y=2|x) = \ln\left(\frac{P(y=2|x)}{P(y=J|x)}\right) = \alpha_2 + \sum_{k=1}^{K}\beta_{2k}x_k = V_2$$
$$\cdots$$
$$\text{Logit } P(y=(J-1)|x) = \ln\left(\frac{P(y=(J-1)|x)}{P(y=J|x)}\right) = \alpha_{J-1} + \sum_{k=1}^{K}\beta_{(J-1)k}x_k = V_{J-1} \quad (1)$$

This is generalized logit model. α_j and β_{jk} are undetermined parameters. α_j is the constant term in each logistic model, and β_{jk} is the relative coefficient of category j relative to the explanatory variable x_k. V_j is the relative effect of category j on the reference category J. According to the sum of the probabilities of all categories of the explained variable is 1, the probability of the j-th category can be obtained.

$$P(y=j|x) = \frac{\exp\left(\alpha_j + \sum_{k=1}^{K}\beta_{jk}x_k\right)}{1+\sum_{j=1}^{J-1}\exp\left(\alpha_j + \sum_{k=1}^{K}\beta_{jk}x_k\right)} \qquad j=1,2\ldots\ldots,(J-1) \quad (2)$$

The above is multinomial logistic regression model [7]. The model is used to explain the changes of the number of household e-bikes under different family factors.

3.2 Model Establishment and Test

IBM SPSS statistics 24 was used to establish multinomial logistic regression model. Among the explanatory variables, monthly household income was an ordinal categorical variable, so the dummy variables were set before model establishment. The reference category was set as the first category, that is, the number of household e-bikes was 0 as the reference.

TABLE 5. MODEL FITTING INFORMATION

Model	AIC	BIC	-2*LL*	Chi-square	df	Sig.
Only intercept	4808.852	4826.888	4802.852			
Final	4308.927	4471.252	4254.927	547.924	24	.000

The significance value was less than 0.01, which showed that the multinomial logistic model was better than the model with only intercept.

TABLE 6. GOODNESS OF FIT

	Chi-square	df	Sig.
Pearson	5611.200	4122	.000
Deviation	3494.775	4122	1.000

The test value of Pearson chi-square statistic was less than 0.01, which indicated that the model fitted the original data well. The significance of deviation was 1.000, which indicated that the fitting of multinomial logistic model to the original data was comparatively accurate.

TABLE 7. LIKELIHOOD RATIO TEST

Explanatory variable	-2*LL*	Likelihood ratio test		
		Chi-square	df	Sig.
Intercept	4481.275	226.347	3	.000
Family total population	4455.398	200.470	3	.000
Family housing area	4284.420	29.493	3	.000
Number of private cars	4523.206	268.279	3	.000
Number of bicycles	4312.494	57.567	3	.000
Income1	4293.421	38.493	3	.000
Income2	4318.183	63.256	3	.000
Income3	4322.922	67.995	3	.000
Income4	4307.285	52.358	3	.000

Seeing that the above -2*LL* were relatively large and the probability p of chi-square test was less than the significance level, all explanatory variables had significant impact on the explained variable. The model establishment was reasonable.

4. RESULTS & DISCUSSION

4.1 Parameter Estimation Results

The parameters of the model were estimated and results were shown in the table 8. From left to right, each column in the table is the estimated value of regression coefficient, the standard error of regression coefficient, the observed value of Wald statistic, the degree of freedom, the probability *p* corresponding to the observed value of Wald statistic and odds ratio (the abbreviation is OR).

Taking the significance level as 0.05, compared the category that the number of e-bikes is 1 with that the number is 0, family total population, number of private cars, number of bicycles and monthly household income were significant factors, while the family housing area was not significant. Compared the category that the number of e-bikes is 2 and 3 respectively with that the number is 0, all explanatory variables were significant factors. Therefore, the model has certain statistical significance.

4.2 Discussion

In the multinomial logistic regression model, if the estimated value of regression coefficient is positive, the value range of OR is (1, +∞), indicating that the explanatory variable is positively correlated with the explained variable. And the gr- eater the OR value is, the stronger the relationship is. On the contrary, if the estimated value of regression coefficient is negative, the value range of OR is (0,1), indicating that the explanatory variable is negatively correlated with the explained variable. And the smaller the OR value is, the stronger the relationship is [8]. The impacts of explanatory variables on the ownership of e-bikes were quantitatively analyzed as follows.

TABLE 8. PARAMETER ESTIMATION

E-bike		B	SE	Wald	Sig.	Exp(B)
1	Intercept	-.161	.323	.249	.618	
	Total population	.308	.092	11.257	.001*	1.361
	Housing area	.001	.001	1.520	.218	1.001
	Private cars	-.442	.139	10.057	.002*	.643
	Bicycles	-.367	.111	10.906	.001*	.693
	Income1	.554	.262	4.463	.035*	1.740
	Income2	1.084	.255	18.030	.000*	2.957
	Income3	1.136	.285	15.911	.000*	3.113
	Income4	.912	.349	6.834	.009*	2.490
2	Intercept	-2.043	.355	33.079	.000*	
	Total population	.764	.092	68.489	.000*	2.147
	Housing area	.002	.001	7.542	.006*	1.002
	Private cars	-1.518	.144	111.129	.000*	.219
	Bicycles	-.685	.114	35.925	.000*	.504
	Income1	1.633	.290	31.651	.000*	5.118
	Income2	2.146	.286	56.239	.000*	8.549
	Income3	2.426	.313	60.109	.000*	11.309
	Income4	2.300	.372	38.248	.000*	9.979
3	Intercept	-6.489	.617	110.539	.000*	
	Total population	1.353	.121	125.341	.000*	3.870
	Housing area	.004	.001	23.258	.000*	1.004
	Private cars	-2.623	.253	107.466	.000*	.073
	Bicycles	-1.287	.215	35.941	.000*	.276
	Income1	1.548	.501	9.547	.002*	4.704
	Income2	2.643	.490	29.099	.000*	14.059
	Income3	2.480	.531	21.793	.000*	11.936
	Income4	2.749	.597	21.192	.000*	15.630

4.2.1 Family total population:

When other variables are the same, for each increase in family population, the OR of owning one e-bike in the family becomes 1.361 times of owning zero, and the OR of owning two and three e-bikes becomes 2.147 and 3.870 times respectively. This shows that family total population has a positive correlation with the probability of owning e-bikes. The OR in the case of owning three e-bikes is the largest, which indicates that the positive correlation between family total population and the probability of owning three e-bikes is the strongest. It can be seen that the more the family population is, the higher the travel demand of residents is. So the need for e-bikes is also increased, that is, there is a greater possibility of owning more e-bikes.

4.2.2 Family housing area:

When other variables are the same, for every $1m^2$ increase in family housing area, the OR of owning two e-bikes in the family becomes 1.002 times of owning zero, and the OR of owning three e-bikes becomes 1.004 times. The regression coefficients of the two cases are positive, and the ORs are close to 1, indicating that the positive correlation degree of family housing area to the probability of owning e-bikes is small.

4.2.3 Number of private cars and of bicycles:

When other variables are the same, for each increase in the number of private cars or bicycles, the OR of owning one e-bike in the family becomes 0.643 or 0.693 times of owning zero, the OR of owning two e-bikes becomes 0.219 or 0.504 times, and the OR of owning three e-bikes becomes 0.073 or 0.276 times. This shows that the number of private cars and of bicycles have a negative correlation with the probability of owning e-bikes, and the negative correlation in the case of owning three e-bikes is the strongest. Horizontally, the negative correlation between the number of private cars and the probability of owning e-bikes is stronger than that between bicycle. It can be seen that families with more private cars and bicycles are less likely to own more e-bikes than those with fewer. Especially when a family owns a private car, the probability of purchasing e-bikes is significantly reduced.

4.2.4 Monthly household income:

When other variables are the same, taking the families with monthly income less than 1500 yuan and no e-bike as the reference category, the OR of owning one e-bike in the family with monthly income of 1500-3000 yuan becomes 1.740 times of the reference category, and the OR of owning two and three e-bikes becomes 5.118 and 4.704 times respectively. When the monthly income is 3000-5000 yuan, 5000-7000 yuan and more than 7000 yuan, the ORs are also greater than 1. It can be seen that monthly household income is positively correlated with the probability of owning e-bikes. In the probability model of owning one and two e-bikes, OR becomes the largest when the monthly income is 5000-7000 yuan, which indicates that monthly income within this range has the strongest effect on the probability of owning e-bikes. Moreover, for the probability of owning three e-bikes, monthly income more than 7000 yuan has the strongest effect on it.

5. Conclusions

By analyzing the demand for e-bikes of families with different characteristics, relevant suggestions are put forward from the aspects of sales management and traffic planning.

- Family population is directly related to the travel situation. In families with a large population, the age distribution of family members is relatively scattered, and the travel demand is also different. Most of the office workers use e-bikes to commute, and the elderly mainly use e-bikes as a single-person daily transportation tool. In families with children, e-bikes are often used to carry people. In order to meet the requirements of different consumer groups for the performance and safety, e-bike sellers should design elaborate marketing strategies according to family population composition and travel habits.
- Family housing area and monthly income reflect the family's economic foundation. As a common non-motorized vehicle, parking and price of e-bike have become important factors that affect the decision-making of family car purchase. For residential districts with a large housing area, developers should plan supporting e-bike parking sites and charging facilities, and strengthen the management of irregular parking. Due to the high sensitivity of consumers to price, e-bike sellers can adjust the proportion of vehicle types on sale according to the income level of local families, and balance the market share of low-end, middle-end and high-end types. E-bike manufacturers can enrich the intelligent functions of e-bikes and enhance product competitiveness, so as to meet the expectations of families with higher price acceptance.
- Families that already own private cars or bicycles are less likely to buy e-bikes, especially the substitution of private cars for e-bikes. Starting from traffic planning, practitioners should give e-bike more sufficient road resources and traffic facilities to ensure the safety and efficiency of e-bike travel. Starting from the marketing approach, in order to increase car-free families' preference for e-bikes, sellers not only need to pay more attention to those families, but also should always maintain symmetry of information, trying to improve the disadvantage of e-bikes compared with cars and bicycles. Starting from the travel concept, the community should strengthen the publicity of non-motorized travel and advocate low-carbon and environmentally friendly travel modes.

References

[1] The Xinhua News Agency, 2020. The number of e-bikes in China is approaching 300 million. http://www.gov.cn/xinwen/2020-11/20/content_5562997.htm.

[2] Astegiano, P., Tampère, C.M.J., Beckx C. (2015) A Preliminary Analysis Over the Factors Related with the Possession of an Electric Bike. Transportation Research Procedia, 10: 393-402.

[3] Sun Q., Feng T., Kemperman A., Spahn A. (2020) Modal shift implications of e-bike use in the Netherlands: Moving towards sustainability? Transportation Research Part D: Transport and Environment, 78.

[4] Liu Y. (2011) Research on Travel Property of Urban Electric Bicycle. Traffic and Transportation, 02: 36-39.

[5] Zhang Y., Li Y., Yang X.G., Liu Q.X., Li C.Y. (2013) Built Environment and Household Electric Bike Ownership: Insights from Zhongshan Metropolitan Area, China. Transportation Research Record, 2387(1): 102-111.

[6] Hu Y.C., Chen J., Zou X.J., Chen Z.W. (2017) Research on Mode Choice of Electric Bicycle Travelers Based on Two-Step Cluster Method. Journal of Guangxi Normal University (Natural Science Edition), 35(03):22-29.

[7] Wang J.C., Guo Z.G. (2001) Logistic Regression Model—Method and Application. Higher Education Press, Beijing.

[8] Xue W. (2013) SPSS Statistical Analysis Method and Application. Publishing House of Electronics Industry, Beijing.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.