

# Research on Influencing Factors of the Willingness to Buy New Energy Vehicle Based on Binary Logistic Regression

Yijia Tian<sup>1</sup>, Zhouying Fan<sup>2</sup>

1.Ningbo Culture Square Commercial Management Co., Ltd Ningbo City, Zhejiang Province, 315000

2.Shaoxing Pingding Knitting Co., Ltd Zhejiang Province, Shaoxing City, 312000

**Abstract:** The binary logistic regression method was used to research and analyze the influence of 10 factors in the questionnaire data, such as gender, age and corporate image, on consumers' willingness to buy new energy vehicles. Through the binary logistic regression model, it is proved that age, educational background, occupation, annual income, corporate image, environmental awareness, government policies and product quality and performance can have an impact on consumers' willingness to buy. According to Wald value, environmental awareness has the greatest impact on the willingness to buy new energy vehicles. The results show that binary logistic regression can predict the probability of an event. In order to increase consumers' willingness to buy new energy vehicles, we must start from population variables, corporate image, environmental awareness, government policies and product quality and performance.

**Key words:** new energy vehicles; willingness to buy; binary logistic regression

## 1 Introduction

With the global change and the rapid development of science and technology, the automobile industry has also been developed. The ensuing environmental pollution and energy shortage also remind people of the importance of energy saving and the necessity of developing and applying new energy vehicles. China has also followed the trend of the times, made greater efforts to develop the new energy automobile industry, expanded the market scale, and increased the output and sales volume of new energy vehicles, which is now among the best in the world. In order to improve the enthusiasm of consumers to buy new energy vehicles. China's automobile enterprises have taken various measures, but they are still far from the target, and the attitude of consumers is unclear. Therefore, it is still necessary to explore the influencing factors of consumers' willingness to buy in order to achieve the purpose of increasing output and sales volume.

By combing foreign literature, it is found that foreign scholars have carried out many studies on the influencing factors of consumers' willingness to buy new energy vehicles, and the research results are quite rich and wonderful. Hanna L analyzed the cost of purchasing a car in the United States. In most cities, the price and rapid depreciation rate of purchasing new energy vehicles exceed the saved fuel cost, and the cost disadvantage greatly reduces customers' willingness to buy new energy vehicles. Kai Chen found that three different dimensions of customer perceived value, price factor, function quality and service quality had different impacts on the willingness to buy. Hardman et al., analyzed the reasons why British consumers bought new energy vehicles, and found that government financial subsidies would positively affect consumers' willingness to buy.

Domestic scholars, Sheng Lijun et al., put forward that when consumers bought new energy vehicles, they were most concerned about the performance of the vehicles. Feng Leidong et al., put forward a product innovation strategy and environmental protection incentive policy that combined incentive subsidies with punitive taxes. Liu Liu found that license policy had the greatest impact on the willingness to buy, followed by price policy, and environmental protection policy has the smallest impact with the stepwise regression analysis.

The above domestic and foreign scholars have studied the influencing factors of consumers' willingness to buy new energy vehicles from the aspects of fuel price, customer perceived value, environmental value and government financial subsidies. This paper focuses on five index systems: corporate image, product quality and performance, environmental awareness, involvement in products and government policies, and studies the influencing factors of consumers' willingness to buy new energy vehicles. By using the binary logistic regression method, this paper builds a binary logistic regression model based on the influencing factor index of the willingness to buy new energy vehicles and whether to buy new energy vehicles in the future, accurately predicts consumers' willingness to buy new energy vehicles, and obtains the important influencing factors of consumers' willingness to buy new energy vehicles.

## 2 Research methods

### 2.1 Binary logistic regression method

Binary logistic regression is used to solve the problem of binary classification. The output of classification model is discrete, the frequency of each classification in classified variable is not less than zero, and the output of regression model is continuous. When using binary logistic regression method, the classified samples of sample size and dependent variable should be sufficient. This paper shows that whether consumers will buy new energy vehicles in the future is a binary variable, so the binary logistic regression method is adopted. Logistic regression model can be expressed as:

$$(1) p = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m)}$$

Then, the probability of not happening is:

$$(2) 1 - p = \frac{1}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_m x_m)}$$

The output is regarded as probability P, that is, the probability of an event that happens, and 1-p is the probability of an event that does not happen. You can also refer to the formula for details:

$$(3) \text{Logit}(p) = \ln\left(\frac{p}{1-p}\right)$$

The Internet questionnaire platform used in this study is Sojump. The questionnaire design adopts scale technology and the questionnaire is designed by building a theoretical model. The questionnaire was distributed to consumers in Zhejiang Province. All the 400 questionnaires collected are used in this study.

2.2 Test of binary logistic regression equation

The test of binary logistic regression equation includes: significance test of regression equation, significance test of regression coefficient and goodness-of-fit test of regression equation. Hosmer Lemeshow goodness-of-fit test can show the goodness-of-fit between fitted values and observed values, that is, the difference between fitted values and observed values.

3 Result analysis

The results show :Chi-square value is 275.313, and p value is 0.000, which passes the significance test with significance level of 1%. Cox & Snell R2 is 0.498, and Nagelkerke R2 is 0.917, which is between 0 and 1 and is close to 1. Therefore, it shows that the goodness-of-fit of the equation is high, that is, the degree of explanation of the model to the original data is ideal and the model has specific statistical significance.

The results of the Hosmer Lemeshow goodness-of-fit test shows:Chi-square value is 9.551, and p value is 0.215 and greater than 0.05. The original hypothesis should be accepted, indicating that there is no significant difference between the predicted value and the true value, that is, the prediction goodness-of-fit of the model is good.

The initial steps of binary logistic regression analysis shows that the accuracy rate of inconsistency between the choice and prediction of 53 consumers who will actually buy new energy vehicles in the future is 86.8%. the confusion matrix of the current model shows that among the 53 consumers who will not buy new energy vehicles in the future, the model correctly identifies 50 consumers and misidentifies 3 consumers, with an accuracy rate of 94.3%. Compared with the initial steps of binary logistic regression analysis, it is found that the prediction accuracy rate of not purchasing new energy vehicles increases. Similarly, the initial steps of binary logistic regression analysis shows that the accuracy rate of consistency between actual choice and prediction of 347 consumers is 100%. the confusion matrix of the current model shows that among the 347 consumers who will actually buy new energy vehicles in the future, the model correctly identifies 345 consumers and misidentifies 2 consumers, with an accuracy rate of 99.4%. After analysis and calculation, the total prediction accuracy rates of the models and the confusion matrix of the current model are 86.8% and 98.8%, respectively, that is, the overall prediction accuracy rates of the model increase, which shows that the prediction results of the models are very ideal.

The ROC curve constructed by prediction probability and real value shows that:it is found that the prediction probability has certain value when the willingness to buy is on the upper left side of the reference line. The area under it is 0.843 and the significance is 0.000, which shows that the model has a good prediction effect for target variables. According to its sensitivity and specificity, the highest Uden index is 0.558, and its best cut-off point is 0.862.

Table 1 shows the process of variable screening and the regression coefficient test results of each explanatory variable. The coefficient of X2(1) in age is 0.117 and the p value is 0.045, which is less than 0.05. It has passed the significance test with a significance level of 5%. The OR value is 1.124. Therefore, on the basis of keeping other variables unchanged, the probability of the willingness to buy new energy vehicles for samples with age of X2(1) will increase by 12.4% compared with samples with other ages.

Table1 Variables in the Equation

Variable	B	Significance	Exp(B)	Variable	B	Significance	Exp(B)
Sex X1(1)	0.368	0.314	1.445	X4(4)	-2.648	0.004	0.771
Age		0.367		X4(5)	-1.179	0.242	0.308
X2(1)	0.117	0.045	1.124	X4(3)	-0.485	0.517	0.616
X2(2)	0.411	0.002	1.309	Annual income		0.687	
X2(3)	-1.316	0.087	0.268	X5(1)	0.149	0.788	1.161
X2(4)	-21.261	1.000	0.000	X5(2)	0.651	0.324	1.918
Educational background		0.106		X5(3)	0.971	0.264	2.640
X3(1)	-0.683	0.343	0.505	X5(4)	0.180	0.025	1.198
X3(2)	1.368	0.049	1.255	Corporate image	0.662	0.024	1.638
X3(3)	1.791	0.022	1.367	Environmental awareness	0.873	0.000	1.846
Occupation		0.023		Involvement in products	0.483	0.076	1.621
X4(1)	-0.702	0.516	0.496	Government policies	0.257	0.023	1.293
X4(2)	-0.927	0.337	0.396	Product quality and performance	0.509	0.012	1.401
X4(3)	-0.485	0.517	0.616	Constant	-3.595	0.015	0.027

Build the model according to Table 1 as follows:

$$\text{Logit } P = -3.595 + 0.117 * X_2(1) + 0.411 * X_2(2) + 1.368 * X_3(2) + 1.791 * X_3(3) - 2.648 * X_4(4) + 0.180 * X_5(4) + 0.662 * F_5 + 0.837 * F_3 + 0.257 * F_2 + 0.509 * F_1$$

## 4 Conclusions and discussions

The results show that X2(1) and X2(2) in age, X3(2) and X3(3) in educational background, X4(4) in occupation, X5(4) in annual income, corporate image, environmental awareness, government policies and product quality and performance all have significant impacts on consumers' willingness to buy new energy vehicles ( $p < 0.05$ ). Among them, X2(1) and X2(2) in age, X3(2) and X3(3) in educational background, X5(4) in annual income, corporate image, environmental awareness, government policies and product quality and performance have significant positive impacts on consumers' willingness to buy new energy vehicles. However, X4(4) in occupation has a negative impact on consumers' willingness to buy new energy vehicles.

Based on the influencing factors of consumers' willingness to buy new energy vehicles in this study, although the binary logistic regression model is relatively simple, resulting in weak expression ability of the model, making it unable to carry out a series of "advanced" operations such as feature crossing and feature screening, and difficult to fit the true distribution of data. However, binary logistic regression can be used to easily explain the importance of features according to different weights. Which factors will affect the results and the influence degree of each factor can be located quickly. It is convenient for parallelization in training, and only needs to weigh the features linearly in prediction, so the performance is better. The resource occupation is small, and only the features with large weights and the weights corresponding to the features need to be stored.

## References

- [1] Jiarui Du, Lequan Gao, Yan Liu, Bin Li, Zhongmin Yang. Research on influencing factors of purchasing intention of new energy vehicle consumers in Beijing, tianjin and hebei [J]. Value engineering, 2019, 38(19): 220-223.
- [2] Hanna L. Breetz. Do electric vehicles need subsidies? Ownership costs for conventional, hybrid and electric vehicles in 14 U.S. cities [J]. Energy Policy, 2018, 120(9): 238-249.
- [3] Chen K, Ren C, Gu R. Exploring Purchase Intentions of New Energy Vehicles: From the Production of Frugality and the Concept [J]. Journal of Cleaner Production, 2019, 230(9): 700-708.
- [4] Scott Hardman, Amrit Chandan, Eric Shiu, Robert Steinberger-Wilckens. Consumer attitudes to fuel cell vehicles post trial in the United Kingdom [J]. International Journal of Hydrogen Energy, 2016, 41(15): 6171-6179.
- [5] Li SHENGLi, Xie Tong. Research on influencing factors of consumers' purchasing intention of new energy vehicles [J]. China business theory, 2019, 11(21): 84-87.
- [6] Feng Leidong, GU Mengdi. Research on vertical product innovation strategy and consumption policy based on environmental quality preference [J]. Journal of jiaxing university, 2020, 32(02): 102-113.
- [7] Liu LIU. Research on influencing Factors of Consumers' Purchase intention of New energy vehicles in Shanghai [D]. Shanghai Jiaotong University, 2019.
- [8] BI Yuanyuan, ZHANG Shuaibing. Research on The Influencing Factors of Consumer Purchase Intention of New Energy Vehicle based on Binary Logical Model [J]. Journal of Harbin University, 201, 42(09): 42-45.
- [9] Wu Yilan. Binary logistic regression analysis of community correction for dangerous driving offenders [J]. Journal of Harbin university, 2019, 40(01): 58-61.
- [10] Yan Lulin. Multivariate linear regression analysis of college students' learning concentration using SPSS [D]. Lanzhou University, 2019.
- [11] Bai Jinfu, Zhuo Daoxi, Lu Hongyu. Research on the influencing factors of NEV purchase intention based on ELM [J]. Modern Business, 2021 (7): 3.
- [12] Li Wenting, Lou Hao Shuai, Wei Xinxin. Study on the influencing factors of neV purchase intention based on factor analysis [J]. Consumer Guide, 2018
- [13], Chen Xinru. Study on the influencing factors of automobile purchase intention based on multiple ordered Logistic regression [J]. Progress in Applied Mathematics, 2022, 11 (5): 10.
- [14], Ke Han. Research on the influencing factors of neV purchase intention based on the responsible environmental behavior model [D]. China University of Geosciences (Beijing), 2019.
- [15] Niu Liwei. Impact factors of new energy vehicles purchase intention and guiding policy research [D]. China University of Mining and Technology, 2015.