Proceedings of ICEMGD 2026 Symposium: The Role of Blue Economy in Promoting Human Sustainable Development
DOI: 10.54254/2754-1169/2026.GT34199

Driving Factors of New Energy Vehicle Sales in the Stock Market: An Empirical Analysis from the User Perspective

Conrad Yuan

School of Mathematics and Statistics, Central South University, Changsha, China
17705935293@163.com

Abstract. As China's new energy vehicle market enters the stock competition phase, consumers' purchase decision logic has shifted from fundamental product competitiveness to multi-dimensional comprehensive experiences. However, most existing studies use contemporaneously matched sales and user-side data, failing to address the reverse causality endogeneity that "rising sales drive higher market exposure", and lacking systematic empirical verification of consumers' heterogeneous perception of segmented vehicle quality dimensions. Taking 65 on-sale NEV passenger models from Autohome in March 2026 as the sample, this paper adopts a time-mismatch design to eliminate reverse causality endogeneity fundamentally. It constructs a baseline OLS model with Newey-West HAC robust standard errors to examine the impacts of user behavior, product competitiveness, and vehicle quality on monthly model sales, and conducts subsample heterogeneity analysis by energy type. The results show that: First, user attention is the absolute core driver of NEV sales, with a far greater influence than product and quality indicators. Second, Manufacturer's Suggested Retail Price and CLTC driving range are core decision thresholds, with price exerting a significantly stronger impact than range. Third, consumers have significant structural differences in quality perception: only faults in high-frequency scenarios have a significant negative impact on sales. Fourth, user attention and price sensitivity have a stronger impact on battery electric vehicle sales, while driving range only significantly affects plug-in hybrid electric vehicle sales. This paper supplements microcausal evidence of NEV sales drivers from the user perspective, and provides empirical support for automakers in product planning, quality control and marketing strategy formulation.

Keywords: new energy vehicle, sales, driving experience, policy incentives

1. Introduction

As China's new energy vehicle (NEV) industry enters a new phase of stock competition, market penetration continues to rise, and model supply becomes increasingly diversified. Against this backdrop, identifying core drivers of consumers' purchase decisions has become critical for automakers. Existing studies on NEV sales drivers are mainly divided into supply-side and demand-side streams. Most supply-side studies focus on pricing and driving range, while a small number of demand-side studies involve model attention and word-of-mouth. However, two core limitations remain: first, most studies use contemporaneous data, which cannot avoid the bidirectional causality endogeneity between sales and user attention; second, existing research mostly focuses on overall quality indicators, lacking a systematic analysis of consumers' sensitivity to faults in different vehicle subsystems.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2026.GT34199

Taking 65 on-sale NEV models from Autohome's March 2026 sales ranking as the sample, this paper adopts a time-mismatch design to eliminate reverse causality interference, and uses a Heteroscedasticity and Autocorrelation Consistent (HAC) robust standard error Ordinary Least Squares (OLS) model to explore the impacts of product competitiveness, vehicle quality and user behavior on model sales. This research not only supplements causal evidence of NEV sales drivers from the user perspective, but also provides a reference for automakers to optimize product layout and user operation efficiency.

2. Theoretical overview and research hypotheses

2.1. Literature review

Existing studies on the influencing factors of new energy vehicle (NEV) sales are mainly divided into two streams: supply side and demand side. Supply-side research has formed a broad academic consensus that the Manufacturer's Suggested Retail Price (MSRP) has a significant negative inhibitory effect on sales, while driving range plays a positive driving role [1, 2]. Demand-side studies confirm that user-side information, including model attention, word-of-mouth and quality evaluation, significantly affects consumers' purchase decisions and NEV sales fluctuations [3, 4].

However, two core research gaps remain. First, most studies use contemporaneous sales and user-side data for empirical analysis, failing to effectively address the bidirectional causality endogeneity that "rising sales drive higher model exposure and word-of-mouth spread", leaving the rigor of causal identification to be improved. Second, existing research mostly focuses on the impact of overall quality indicators on sales, lacking a systematic analysis of consumers' heterogeneous sensitivity to faults in different vehicle subsystems.

2.2. Core theoretical foundations

This research is developed based on three core theories: (1) Consumer Purchase Decision Theory [5], which explains how model attention and word-of-mouth affect consumers' purchase evaluation; (2) Anchoring Effect Theory [6], which illustrates that vehicle sales rankings act as a core decision anchor for consumers; and (3) Perceived Quality Theory [7], which reveals that vehicle fault indicators change consumers' quality perception and purchase intention. These theories form the logical basis for the research hypotheses below.

2.3. Core research hypotheses

Based on the above literature review and theoretical foundations, this paper proposes research hypotheses from four dimensions:

First, sales anchoring effect dimension. Higher monthly sales rankings signal stronger market recognition and act as a core decision anchor for consumers [6], thus boosting sales in the subsequent month. Accordingly, this study proposes:

H1: The higher the monthly sales ranking of a new energy vehicle (NEV) model, the higher its sales volume in the subsequent month.

Second, user behavior: Higher model attention indicates stronger purchase intention, while more fault evaluation participants mean wider negative information spread. Thus:

H2a: The higher the daily attention of a vehicle series on the ranking page, the higher the sales volume of the NEV model in the subsequent month.

H2b: The higher the number of users participating in fault evaluation, the lower the sales volume of the NEV model in the subsequent month.

DOI: 10.54254/2754-1169/2026.GT34199

Third, core product competitiveness dimension. Longer China Light-Duty Vehicle Test Cycle (CLTC) driving range alleviates consumers' range anxiety and enhances product competitiveness, while higher MSRP increases purchase costs and suppresses purchase intention. Accordingly, we propose:

H3a: The longer the official CLTC comprehensive driving range of a vehicle, the higher the sales volume of the NEV model in the subsequent month.

H3b: The higher the Manufacturer's Suggested Retail Price (MSRP) of a model, the lower the sales volume of the NEV model in the subsequent month.

Fourth, word-of-mouth and quality dimension. Higher comprehensive word-of-mouth indicates stronger consumer recognition [8], while higher Problems per Hundred Vehicles (PHH) and subsystem fault proportions mean worse quality stability and lower perceived quality. Accordingly, we propose:

H4a: The higher the Problems per Hundred Vehicles (PHH) of a vehicle, the lower the sales volume of the NEV model in the subsequent month.

H4b: The higher the proportion of body exterior faults on the word-of-mouth page, the lower the sales volume of the NEV model in the subsequent month.

H4c: The higher the proportion of driving process faults on the word-of-mouth page, the lower the sales volume of the NEV model in the subsequent month.

H4d: The higher the proportion of function operation faults on the word-of-mouth page, the lower the sales volume of the NEV model in the subsequent month.

H4e: The higher the proportion of electronic equipment faults on the word-of-mouth page, the lower the sales volume of the NEV model in the subsequent month.

H4f: The higher the proportion of air conditioning system faults on the word-of-mouth page, the lower the sales volume of the NEV model in the subsequent month.

H4g: The higher the proportion of vehicle interior faults on the word-of-mouth page, the lower the sales volume of the NEV model in the subsequent month.

H4h: The higher the proportion of power system faults on the word-of-mouth page, the lower the sales volume of the NEV model in the subsequent month.

3. Research design and data preparation

3.1. Data sources and experimental design

All research data are collected from Autohome, China's leading automotive vertical platform, covering four dimensions: vehicle sales, product parameters, user word-of-mouth, and quality evaluation.

To address the core reverse causality endogeneity in empirical research, this paper adopts a time-mismatch design: all core explanatory and control variables are collected at the end of February 2026 (before the official release of March 2026 NEV sales data), while the dependent variable (monthly model sales) is collected after the release of the March sales ranking. This design fundamentally cuts off the reverse impact of sales on user-side variables and improves the rigor of causal identification.

3.2. Sample screening and data cleaning rules

The initial sample includes all on-sale NEV models listed in Autohome's March 2026 NEV sales ranking. We apply two core screening rules to ensure sample validity and data integrity: (1) Exclude models with serious missing values in core variables (sales, core product parameters, quality evaluation data); (2) Exclude discontinued, unlaunched or undelivered non-on-sale models, only retaining normally mass-produced on-sale passenger vehicles.

After processing, 65 valid samples are finally obtained, covering NEVs with different price levels, body types, brand positioning and energy types with good market representativeness.

DOI: 10.54254/2754-1169/2026.GT34199

3.3. Feature selection

We select three types of dummy variables as control variables: vehicle type (sedan, Sport Utility Vehicle, and Multi-Purpose Vehicle (MPV)), brand positioning (luxury brand, joint-venture brand, and independent brand), and energy type (battery electric vehicle (BEV) and plug-in hybrid electric vehicle (PHEV)).

Core explanatory variables are divided into three dimensions: (1) User behavior dimension: Concern (daily average attention of the vehicle series on the ranking page), and Participants (number of valid users participating in model fault evaluation); (2) Core product competitiveness dimension: Range (official CLTC comprehensive driving range of the model, unit: km), and MSRP (average manufacturer's suggested retail price of the model, unit: 10,000 RMB); (3) Vehicle quality dimension: PHH (Problems per Hundred Vehicles of the model), and fault proportions of body exterior, driving process, function operation, electronic equipment, air conditioning system, vehicle interior, and power system.

4. Empirical results and analysis

Based on the research framework and sample data constructed above, this chapter clarifies the overall distribution characteristics of the sample through descriptive statistics, conducts empirical tests with the OLS regression model with HAC robust standard errors as the baseline model, systematically analyzes the effects of various dimensional factors on the sales of new energy vehicles (NEVs), verifies the research hypotheses proposed above, and identifies the importance ranking of core influencing factors.

4.1. Descriptive statistics and model specification

4.1.1. Descriptive statistics

Descriptive statistics of core variables (Table 1) show that the sample covers mainstream NEV models in China with good market representativeness.

Table 1. Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Median	Min	Max
ln (Sales)	65	8.6079	0.8067	7.0825	8.3680	10.5923
MSRP	65	20.5448	10.7963	4.5800	17.9300	53.4800
Range	65	486.8923	219.6359	120.0000	530.0000	835.0000
Score	65	4.5185	0.0638	4.3000	4.5300	4.6500
PHH	65	123.1077	52.2650	25.0000	123.0000	262.0000
Body	65	0.1378	0.0061	0.0000	0.1300	0.4100
Driving	65	0.2690	0.0889	0.1000	0.2500	0.5400
Function	65	0.1218	0.0575	0.0000	0.1200	0.2300
Device	65	0.1695	0.0726	0.0000	0.1700	0.3900
AirConditioning	65	0.0844	0.0601	0.0000	0.0800	0.4000
Interior	65	0.1160	0.0511	0.0300	0.1000	0.2500
PowerSystem	65	0.0438	0.0332	0.0000	0.0400	0.1600
Participants	65	98.0000	62.3159	31.0000	77.0000	296.0000
Concern	65	54607	30815	21228	47855	159066

DOI: 10.54254/2754-1169/2026.GT34199

4.1.2. Multicollinearity test

This paper conducts a multicollinearity test using the Variance Inflation Factor (VIF). The initial test shows that the comprehensive word-of-mouth score (Score) has a VIF value far exceeding the critical threshold of 10, with severe multicollinearity, so this variable is excluded from the baseline model (Table 2).

Table 2. Multicollinearity test

Variable	VIF
MSRP	4.65
Range	7.37
Score	75.90
PHH	4.08
Body	5.42
Driving	8.96
Function	4.66
Device	5.35
AirConditioning	3.03
Interior	4.08
PowerSystem	2.29
Participants	3.46
Concern	4.42

A re-test after exclusion shows that the VIF values of all remaining variables are below 10, indicating no severe multicollinearity. The high condition number in regression results stems from numerical scale differences between variables, not multicollinearity, and will not interfere with the regression results (Table 3).

Table 3. Multicollinearity test

Variable	VIF
MSRP	4.22
Range	7.07
PHH	4.03
Body	4.30
Driving	2.84
Function	3.46
Device	2.51
AirConditioning	2.38
Interior	1.85
PowerSystem	1.15
Participants	3.46
Concern	4.36

DOI: 10.54254/2754-1169/2026.GT34199

4.1.3. Baseline model specification

This paper takes the natural logarithm of monthly sales of new energy vehicles as the dependent variable and constructs the following baseline OLS linear regression model:

$$\begin{aligned}\ln\left(\mathrm{Sales}_{i,t+1}\right) &= \beta_0 + \beta_1 \mathrm{MSRPs}_{i,t} + \beta_2 \mathrm{Range}_{i,t} + \beta_3 \mathrm{PHH}_{i,t} + \beta_4 \mathrm{Body}_{i,t} + \beta_5 \mathrm{Drving}_{i,t} \\ &+ \beta_6 \mathrm{Function}_{i,t} + \beta_7 \mathrm{AirConditioning}_{i,t} + \beta_8 \mathrm{Inteior}_{i,t} + \beta_9 \mathrm{PowerSystem}_{i,t} \\ &+ \beta_{10} \mathrm{Participants}_{i,t} + \beta_{11} \mathrm{Concern}_{i,t} + \beta_{12} \mathrm{Energy}_{i} + \beta_{13} \mathrm{CarType}_{i} + \beta_{14} \mathrm{Brand}_{i} + \varepsilon_{i}\end{aligned}$$

Where $\ln(\cdot)$ denotes the natural logarithm of the corresponding variable, β_0 is the constant term, β_1 - β_{11} are the regression coefficients of core explanatory variables, and ε_i is the random disturbance term. Energy, $\mathrm{CarType}$ and Brand are control variables. Detailed definitions of all variables are shown in Section 3.3.

4.2. Baseline regression results analysis

Considering that cross-sectional data are prone to heteroscedasticity and serial correlation problems, the standard error estimation of ordinary OLS may be biased [9]. Therefore, this paper adopts the Newey-West Heteroscedasticity and Autocorrelation Consistent (HAC) robust standard errors to correct the model, so as to obtain more robust coefficient estimation results and significance test conclusions, and ensure the rigor of empirical analysis. The data are standardized by Z-score, and the regression results are shown in Table 4.

Table 4. Baseline regression results

Variable	Next-period Sales (ln)
const	5.551e-16*** (0.072)
MSRP	-0.5125*** (0.086)
Range	0.2032*** (0.073)
PHH	-0.0686 (0.083)
BodyExteriorFaultRate	-0.1121 (0.106)
DrivingFaultRate	-0.1480 (0.151)
FunctionFaultRate	-0.0843 (0.111)
ElectromagneticDeviceFaultRate	-0.3003** (0.150)
AirConditioningFaultRate	-0.1993** (0.097)
InteriorFaultRate	-0.0828 (0.106)
PowerSystemFaultRate	-0.1503 (0.121)
Participants	0.0017 (0.067)
Concern	0.7791*** (0.096)
Energy	Controlled
CarType	Controlled
Brand	Controlled
N	65

DOI: 10.54254/2754-1169/2026.GT34199

Table 4. (continued)

R^2	0.827
adj. R^2	0.667
F	76.21

4.2.1. Overall model fit analysis

According to the overall model test results, the F-statistic is 30.92 with a corresponding p-value of 1.94e-19, which is highly significant at the 1% statistical level, indicating that the model is overall valid with significant joint explanatory power of explanatory variables. The adjusted R^2 of the model is 0.679, meaning the model can explain 67.9% of the variation in the logarithm of NEV sales in the subsequent month, which is at a favorable level for cross-sectional empirical research on automobile sales.

4.2.2. Coefficient interpretation by dimension and research hypothesis verification

Holding other variables constant, this study interprets the regression coefficients and verifies the research hypotheses in three dimensions.

(1) User Behavior Dimension: The regression coefficient of Concern is significantly positive at the 1% statistical level, which is the core driving variable of sales, verifying Hypothesis H2a. The coefficient of participants fails to pass the significance test, so Hypothesis H2b is not verified, indicating consumers pay more attention to the structural distribution of faults rather than the scale of evaluation participants.

(2) Core Product Competitiveness Dimension: The coefficient of range is significantly positive at the 1% level, and the coefficient of MSRP is significantly negative at the 1% level, both consistent with theoretical expectations, verifying Hypotheses H3a and H3b. Specifically, driving range improvement can significantly alleviate consumers' range anxiety and boost sales, while higher pricing increases purchase costs and inhibits sales performance.

(3) Word-of-Mouth and Quality Dimension: The coefficients of PHH and all subsystem fault proportions are negative as expected, but with significant structural differences in significance. PHH fails to pass the 10% significance test, so Hypothesis H4a is not verified. Only the coefficients of electronic equipment and air conditioning system fault proportions are significantly negative at the 5% level, verifying Hypotheses H4e and H4f; the remaining subsystem fault variables are not significant, so the corresponding hypotheses are not verified. The core insight is that consumers are highly sensitive to faults in high-frequency daily usage scenarios, but not to low-frequency scenario faults and overall quality indicators.

4.3. Core influencing factors analysis

4.3.1. Importance ranking of core influencing factors

Based on the Z-score standardized regression results, we divide the sales influencing factors into three tiers: The first tier is the absolute core dominant factor, Concern, whose standardized coefficient is significantly positive at the 1% level, with an absolute value far higher than other variables, as the

DOI: 10.54254/2754-1169/2026.GT34199

primary prerequisite for sales conversion. The second tier is core decision threshold factors, including MSRP and electronic equipment fault proportion, both significant at the 5% level and above, which are the core hard indicators for consumers to screen models [10]. The third tier is supplementary decision factors, including range and air conditioning system fault proportion, which affect consumers' final decisions after the core thresholds are met.

4.3.2. Explanation for the causes of insignificant factors

The insignificant variables, including overall PHH, most subsystem fault proportions and participants, are mainly due to three reasons: First, the overall quality difference of mainstream mass-produced models is small, making it difficult for consumers to form perceived differences. Second, consumers cannot form clear expectations of low-frequency scenario faults at the purchase stage, so such faults will not affect their decisions. Third, consumers' sensitivity to negative quality information depends on fault scenarios rather than the scale of evaluation participants.

4.4. Heterogeneity test

To identify structural differentiation in consumer demands, the sample is divided into Battery Electric Vehicle (BEV) and Plug-in Hybrid Electric Vehicle (PHEV) groups for sub-sample regressions using the baseline model. The results are shown in Table 5.

Table 5. Baseline regression results

Variable	BEV	PHEV
Const	0.0330 (0.256)	-0.0525 (0.117)
MSRP	-0.5279*** (0.202)	0.4141*** (0.138)
Range	0.2209 (0.284)	0.2231*** (0.110)
PHH	-0.0286 (0.113)	0.0345 (0.108)
BodyExteriorFaultRate	-0.1614 (0.171)	0.0188 (0.147)
DrivingFaultRate	-0.0989 (0.287)	0.0045 (0.171)
FunctionFaultRate	-0.0608 (0.169)	0.0149 (0.180)
ElectromagneticDeviceFaultRate	-0.2923** (0.223)	0.2259 (0.145)
AirConditioningFaultRate	-0.2225** (0.172)	0.1232(0.126)
InteriorFaultRate	-0.1871 (0.206)	0.0277 (0.132)
PowerSystemFaultRate	-0.1968 (0.144)	0.0483(0.156)
Participants	0.1721 (0.115)	0.0755 (0.100)
Concern	0.7791*** (0.096)	0.5151*** (0.076)
Energy	Controlled	Controlled
CarType	Controlled	Controlled
Brand	Controlled	Controlled
N	26	39
R-squared	0.827	0.660
Adjusted R-squared	0.667	0.503
F-statistic	76.21	22.74

DOI: 10.54254/2754-1169/2026.GT34199

4.4.1. Verification of the common core conclusions from the full sample

User attention (Concern) remains a significant positive driver at the 1% level in both groups. However, its standardized coefficient in the BEV group (0.7791) is much higher than in the PHEV group (0.5151), indicating a stronger driving effect on BEV sales. This occurs because BEV market supply is highly concentrated with strong product homogeneity, making purchase decisions rely heavily on market exposure.

4.4.2. Significant heterogeneity in the core product competitiveness dimension

Core product competitiveness shows completely different characteristics between the two types. First, while MSRP is significantly negative in both groups, BEV consumers exhibit higher price sensitivity (coefficient -0.5279 vs. -0.4141) due to stricter budget constraints. Second, CLTC driving range (Range) only significantly promotes PHEV sales. Since mainstream BEVs generally exceed 500km, range anxiety is widely alleviated. Conversely, PHEVs' pure electric ranges vary widely, directly determining daily commuting costs and making consumers more sensitive to this metric.

4.4.3. Heterogeneity characteristics in the quality dimension

Subsystem fault proportions (electronic and air conditioning), which were significant in the full sample, lose statistical significance in both subsamples. This is because subsampling significantly improves the homogeneity of intelligent configurations and quality control standards within each group. Consequently, differences in fault indicators are reduced, making it difficult for consumers to form perceived differences.

5. Conclusion

This study empirically investigates the core drivers of new energy vehicle (NEV) sales using 65 on-sale models from Autohome in March 2026. The empirical results yield four main conclusions. First, user attention is the absolute core driver of NEV sales, outperforming product competitiveness and quality indicators across all segments. Second, product indicators serve as core screening thresholds; specifically, MSRP exerts a significantly stronger negative impact than the positive effect of CLTC driving range. Third, consumers exhibit structural differences in quality perception. Overall problem rates (PHH) do not significantly affect sales, whereas faults in high-frequency usage scenarios (electronic equipment and air conditioning) exert strong negative impacts. Fourth, significant heterogeneity exists between technical routes: BEV sales are more sensitive to user attention and price, while PHEV sales are uniquely driven by driving range. These findings offer targeted practical implications for automakers. Strategically, companies should prioritize expanding market exposure to drive purchase conversion. Product-wise, differentiated strategies are essential: BEV development should focus on price control and intelligent configurations, whereas PHEV planning must prioritize pure electric range. Furthermore, quality control resources should shift from lowering overall PHH to eradicating high-frequency functional faults. Future research could address current limitations by utilizing longitudinal panel data, expanding heterogeneity analysis to include price ranges and brand attributes, and integrating multi-platform datasets to further validate these findings.

References

[1] Liao, F., Molin, E., & van Wee, B. (2017). Consumer preferences for electric vehicles: a literature review. Transport Reviews, 37(3), 252-275.

DOI: 10.54254/2754-1169/2026.GT34199

[2] Li, W., Long, R., Chen, H., & Geng, J. (2017). A review of factors influencing consumer intentions to adopt battery electric vehicles. Renewable and Sustainable Energy Reviews, 78, 318-328.
[3] Ratchford, B. T., Lee, M. S., & Talukdar, D. (2003). The impact of the Internet on information search for automobiles. Journal of Marketing Research, 40(2), 193-209.
[4] Liu, Y., Li, X. (2023). Predicting the demand for new energy vehicles by incorporating perceived quality measures extracted from online reviews. Journal of Retailing and Consumer Services, 75, 103485.
[5] Engel, J. F., Blackwell, R. D., & Miniard, P. W. (1995). Consumer behavior (8th ed.). Dryden Press.
[6] Tversky, A., Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185(4157), 1124-1131.
[7] Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. Journal of Marketing, 52(3), 2-22.
[8] Chevalier, J. A., Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. Journal of Marketing Research, 43(3), 345-354.
[9] Wooldridge, J. M. (2015). Introductory econometrics: A modern approach (6th ed.). Cengage Learning.
[10] Han, L., Wang, S., Zhao, D., & Li, J. (2017). The intention to adopt electric vehicles: Driven by functional and non-functional values. Transportation Research Part A: Policy and Practice, 103, 185-197.