Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 15, No. 1, 2024

Analysis of Spatial Effects of New Energy Vehicles Based on Semi-parametric Spatial Durbin Model

Wanqing Wu, Yunquan Song*, Rui Guo, Jiaxin Liu

College of Science, China University of Petroleum (East China), Qingdao, Shandong, 266580, China
* **Corresponding author**: Yunquan Song (Email: statistics99@163.com)

Abstract: The global scientific and technological revolution and industrial change are developing rapidly, the integration of automotive technology with energy, transportation and information and communication fields is accelerating, and electrification, internet connectivity and intelligence have become the main trends. Therefore, studying the impact of digital development on the growth of new energy vehicle industry and consumer purchase intention is of great theoretical and practical significance for the digital transformation of automobile companies. In this paper, based on the provincial panel data of new energy vehicles in China from 2016 to 2023, we constructed the index system of digital economy and economic development, used the entropy method to calculate the level of development, and conducted spatial autocorrelation tests using global and local Moran indices. Considering the non-linear relationship between variables, a semi-parametric spatial Durbin model is established using local polynomial estimation to explore the impact of digital economy on new energy automobile industry. The results of the study show that the development of digital economy promotes the development of new energy automobile industry to a certain extent, but its development level may have a negative impact if it is too low or too high.

Keywords: Digital Economy; New Energy Vehicles; Entropy Method; Semi-parametric Spatial Dubin Model.

1. Introduction

The digital economy, as a strategically emerging industry, is thriving. In 2022, China's digital economy ranked first in the world, accounting for 41.5% of GDP. It is projected to continue growing, reaching 56.1 trillion yuan in 2023, thereby becoming a vital driver of economic growth. The extensive application of the digital economy has also had a significant impact across various fields, including the new energy automobile industry.

According to the report from the 20th Party Congress, the development of new energy vehicles is crucial for promoting a clean and low-carbon transformation in the transportation sector. Research both domestically and internationally has primarily focused on several key aspects. Li Miaoran (2020) assessed the relevant policies affecting the new energy automobile industry and their impact on regional economies and the environment, offering suggestions to enhance technological innovation and market penetration. Xiong Zhifei (2023) employed panel data and a spatial econometric model to discover that investment in the new energy automobile industry has a positive spillover effect on economic growth and innovation activities in neighboring regions. Wang(2019) [4] examined the geographic agglomeration effect of the new energy automobile industry using GIS data, illustrating its development trends and impact on the local economy.

This paper will explore how the integration of the digital economy and the new energy automobile industry can foster growth. By constructing a spatial Durbin model, it seeks to unveil the relationship between the digital economy and the sales of new energy vehicles, providing a theoretical foundation for the formulation of related policies.

2. Theoretical Mechanism

Economic development has significantly heightened the demand for new energy vehicles. As economic levels rise, consumers' purchasing power increases, thereby stimulating the consumption of new energy vehicles. In this paper, the entropy value method is employed to construct an evaluation index system for economic development, and nine representative indicators have been selected, as illustrated in Table 1.

Table 1. Indicators for Evaluating the Level of Economic Development

Primary indicator	Secondary indicator	Tertiary indicator
Level of economic development	Economic scale	Gross domestic product
		Value added of primary industry
		Value added of secondary industry
		Value added of tertiary industry
	Economic development	Per capita gross domestic product
		Total fixed asset investment
	Economic prosperity	Total import and export volume
		Local fiscal revenue
		Local fiscal expenditure

Population growth has directly driven the demand for new energy vehicles, as an expanding population base increases transportation need. Additionally, changes in industrial structure have influenced consumption patterns, with a decline in the proportion of energy-consuming industries and a rise in the share of tertiary industries, particularly high-tech sectors, altering energy consumption patterns. In this context, consumers are increasingly favoring new energy vehicles for their low energy consumption and environmentally friendly characteristics.

Accelerated urbanization and improved infrastructure have made it easier to use new energy vehicles, while the expansion of sales channels has facilitated access for more consumers to purchase these vehicles. Technological advancements have enhanced vehicle performance and range while also reducing production costs, further stimulating market demand. In the digital economy, digital platforms enable users to easily locate charging stations, alleviating charging anxiety and promoting the adoption of new energy vehicles.

the impact mechanism of the digital economy on new energy vehicle sales. Drawing on the research of Li Xiaomin (2024), we selected four secondary indicators—digital infrastructure, digital industry, digital innovation capability, and digital inclusive finance—to construct an indicator system. We then employed the entropy method to calculate the values of these indicators. Based on this, we created a quartile chart to reflect the regional distribution characteristics of the development of the digital economy in China, as shown in Fig. 1.

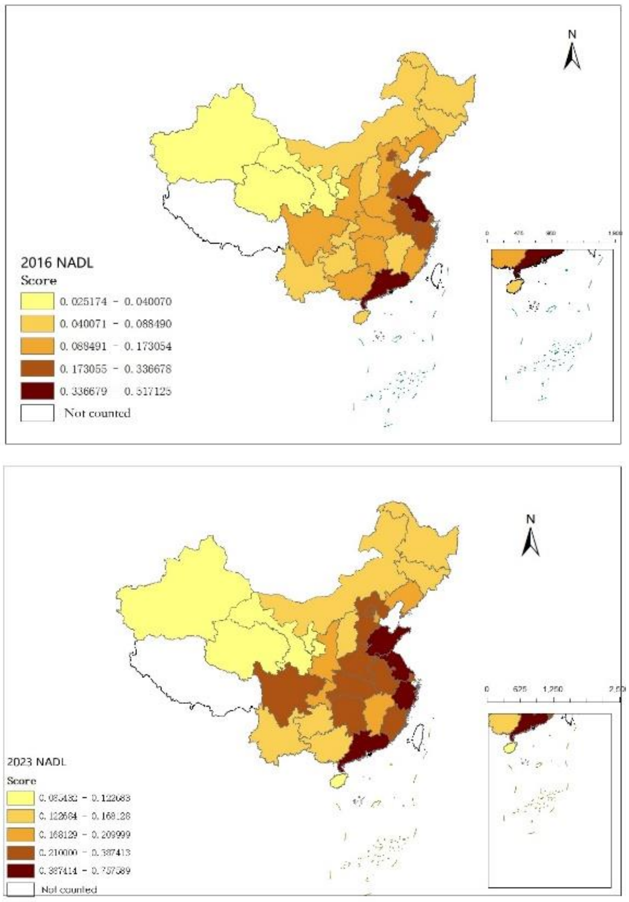

Figure 1. Digitalization Index Quartile Chart

The digital economy index and the economic development level index serve as the primary explanatory variables. Additionally, the control variables mentioned above are incorporated into the model to construct a spatial Durbin model. This approach allows for a thorough analysis of the impact of digital economy development levels on the sales of new energy vehicles both regionally and nationally, with the goal of providing a theoretical foundation for policy formulation. The specific variable selection is presented in Table 2.

Table 2. Variable Table

Variable table	Name	Meaning
Dependent variable	New energy vehicle sales (NEVS)	Number of new energy vehicles insured
Independent variable	Digital economy development level (NADL)	Digital economy development index
	Economic development level (EDL)	Economic development index
Control variables	Urbanization level (UR)	Percentage of urban population in total population
	Total population (POP)	Resident population
	Industry structure (RS)	Proportion of secondary and tertiary industries
	Charging pile ownership (CPO)	Number of charging piles in each province

3. Methods

3.1. Spatial Autocorrelation Analysis

When constructing the model for new energy vehicle sales, it is essential to assess the spatial correlation of digitalization across provinces and regions. If a significant spatial effect is identified, it should be included in the model; if not, model parameters can be estimated using ordinary least squares (OLS). The analysis of spatial autocorrelation involves three main processes: establishing a spatial weight matrix, conducting global spatial autocorrelation analysis, and performing local spatial autocorrelation analysis.

Typically, a binary symmetric spatial weight matrix W is defined to represent the proximity of spatial regions at individual locations in the following form:

$$\begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{pmatrix}, \tag{1}$$

where w_{ij} denotes the proximity between regions i and j.

Considering that this paper examines the impact of digitization of new energy vehicles on sales, we selected spatial panel data for 30 provinces in China from 2016 to 2023. Therefore, constructing a geographic distance weight matrix is more advantageous than an economic distance matrix, as it better reflects spatial differences. The basic expression form of the matrix is:

$$w_{ij} = \begin{cases} \frac{1}{d_{ij}^{n}}, i \neq j \\ 0, \quad i = j \end{cases}. \tag{2}$$

Typically, n is set to 1. The Euclidean distance is used to calculate the inverse distance between the centers of two entities.

Moran's I measure of spatial correlation, which has clear boundaries and is capable of determining positive and negative correlations, was chosen for this study.

$$\text{Moran's I} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i-\bar{x})(x_j-\bar{x})}{S^2\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}\ , \tag{3}$$

where $s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$ is the sample variance, w_{ij} is a spatial weight matrix element that represents the distance between region i and region j, and $\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}$ is the sum of all spatial weights. If the weight matrix rows are normalised, $\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} = n$. At this point, the Moran's index can be expressed as:

$$I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}. \tag{4}$$

The global Moran's index is calculated using Stata software, and the results indicate that the Moran's I values for digital economic development, economic development, and new energy vehicle sales are all significantly positive. This suggests that the spatial distribution of these variables exhibits a positive and stable correlation. Partial results are presented in Table 3.

Table 3. Global Moran's Index

Year	Digital economy	Economic development level	New energy vehicle sales
2016	0.067**	0.038**	0.047**
2017	0.058**	0.039**	0.062**
2018	0.067**	0.031**	0.031**
2019	0.055**	0.043**	0.019**
2020	0.061**	0.044**	0.049**
2021	0.052**	0.044**	0.071**
2022	0.054**	0.048**	0.077**
2023	0.055**	0.048**	0.063**

A scatterplot of the Moran's index is created to visualize the level of spatial agglomeration of new energy vehicle sales in each province from 2016 to 2023. To enhance aesthetics, numbers are used in place of province names in the graph to represent the Moran's index for 2016 and 2023, as shown in Fig. 2.

When examining spatial agglomeration in the vicinity of a region, we use the local Moran's index and the results are shown in Table 4.

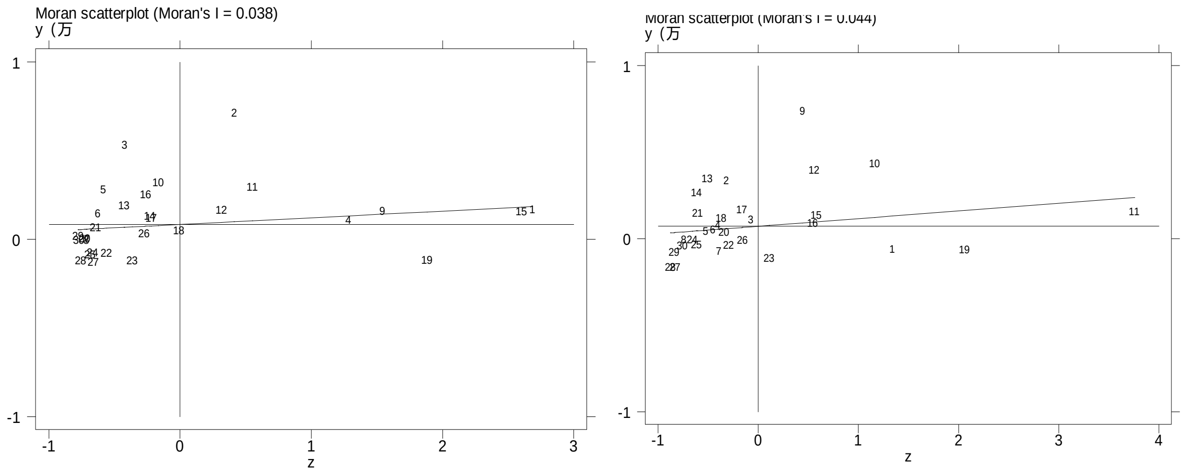

Figure 2. Scatterplot of the Moran's Index for 2016 and 2023

Table 4. Local Moran's Index

	Local Moran's index						
	Beijing	Tianjin	Shanghai	Shandong	Guangdong	Zhejiang	Qinghai
2016	0.399**	0.294**	0.217**	0.365**	0.271**	0.154*	0.113
2020	0.018	0.036	0.549**	0.098*	0.223**	0.492**	0.188
2023	0.115	0.103	0.469**	0.070	0.183*	0.516**	0.171

$$I_i = \frac{(x_i-\bar{x})}{s^2}\sum_{j=1}^{n} w_{ij}\left(x_j-\bar{x}\right). \tag{5}$$

As shown in Fig. 2, new energy vehicle sales (NEVS) in 2016 exhibited a dispersed trend. Beijing, Shandong Province, Shanghai, and other regions were located in the first quadrant, indicating relatively high sales and a "high-high concentration" pattern. This can be attributed to the early-stage promotion of new energy vehicles, where first- and second-tier cities received strong policy support. In the second quadrant, provinces such as Hebei, Jiangsu, and Henan had lower sales, but their neighboring regions performed better. Most of the western regions fell within the third quadrant, while the fourth quadrant had almost no data points.

By 2023, Jiangsu Province, Zhejiang Province, and Shanghai remained in the first quadrant, continuing the "high-high agglomeration" pattern. This reflects the clustering effect of the new energy vehicle industry, particularly the rapid development in regions such as the Yangtze River Delta, the Greater Bay Area, and the Beijing-Tianjin-Hebei region.

Overall, the level of digital economy development and new energy vehicle sales across China's provinces show significant spatial autocorrelation, which meets the criteria for spatial econometric analysis.

3.2. Analysis of Spatial Regression Results

When constructing the model for new energy vehicle sales, it is essential to assess the spatial correlation of digitalization across provinces and regions. If a significant spatial effect is identified, it should be included in the model; if not, model parameters can be estimated using ordinary least squares (OLS). The analysis of spatial autocorrelation involves three main processes: establishing a spatial weight matrix, conducting global spatial autocorrelation analysis, and performing local spatial autocorrelation analysis.

The basic linear model considered in this paper takes the following form:

$$SNEV_{it} = \beta_0 + \beta_1 NADL_{it} + \beta_2 EDL_{it} + \beta_3 UR_{it} + \beta_4 POP_{it} + \beta_5 RS_{it} + \beta_8 CPO_t + \varepsilon_{it}. \quad (6)$$

When using panel data regression to deal with specific individuals, models such as Spatial Lag Model (SLM), Spatial Error Model (SEM), Spatial Durbin Model (SDM), etc., are usually considered to fit this data, and the forms of the above models are as follows:

$$\text{SLM:} \begin{cases} Y_{it} = \rho \sum_{j\neq i} w_{ij} Y_{jt} + X_{it}\beta + \varepsilon_{it} \\ \varepsilon_{it} \sim N(0, \sigma^2 I_n) \end{cases}, \quad (7)$$

$$\text{SEM:} \begin{cases} Y_{it} = X_{it}\beta + u_i \\ u_i = \lambda \sum_{j\neq i} w_{ij}\, u_i + \varepsilon_{it}, \\ \varepsilon_{it} \sim N(0, \sigma^2 I_n) \end{cases} \quad (8)$$

$$SDM\text{:} \begin{cases} Y_{it} = \rho \sum_{j\neq i} w_{ij} Y_{jt} + X_{it}\beta + \sum_{j\neq i} \gamma_i w_{ij} X_{it} + \varepsilon_{it} \\ \varepsilon_{it} \sim N(0, \sigma^2 I_n) \end{cases}, \quad (9)$$

where Y is a vector of $NT \times 1$, X is a matrix of $NT \times p$, N denotes the number of observed provincial domains, T denotes the time point of observation, p denotes the dimension of the independent variables, ρ and λ are the spatial autocorrelation coefficients of the corresponding models, γ_i is the coefficient of the spatial lag term of the explanatory variables. W is the spatial weight matrix, β is the explanatory variable, μ is the random error, and ϵ is the perturbation term of the random error.

The selection of models for empirical analysis requires additional testing. The following presents a test of the baseline regression model for new energy vehicle sales in 30 provinces across the country from 2016 to 2023, with the results displayed in Table 5.

Table 5. Local Moran's Index

Year	Test method	Test statistic (p-value)	Test result
SLM, SEM and SDM	LM-lag	89.380(0.000)	Both are significant, take the next test
	LM-error	148.835(0.000)	
	Robust LM-lag	27.662(0.000)	Both are significant, take the next test
	Robust LM-error	87.117(0.000)	
Degrade into SLM	LR	13.77(0.032)	No degradation
Degrade into SEM	LR	16.51(0.011)	No degradation
Degrade to SEM	Hausmann test	573.80(0.000)	Fixed effects
SDM fixed effects model selection	LR	128.82(0.000)	SDM with time fixed effects

Based on this, a spatial Durbin model with time fixed effects is necessary to control for the impact of time variation on new energy vehicle sales (NEVS), allowing for a greater focus on the influence of spatial factors on sales. This does not imply that time is irrelevant to sales; rather, it emphasizes that the time effect is "time-varying but not individual-varying." Consequently, a spatial Durbin model with fixed time effects can help eliminate temporal disturbances, improve the accurate assessment of the impact of spatial factors, and enhance the model's explanatory power.

While the time-fixed spatial Durbin model effectively captures the spatial spillover effects of the digitization level of new energy vehicles, further investigation is needed to determine whether its impact on sales remains consistent as the digitization level increases. Therefore, a more precise representation of the relationship between the digitization index and sales is required.

To accurately depict the relationship between the digitization index and sales volume, this paper includes a scatter plot (see Fig. 3), which illustrates a non-linear relationship between the two variables. As a result, the study employs a semi-parametric spatial Durbin model for analysis.

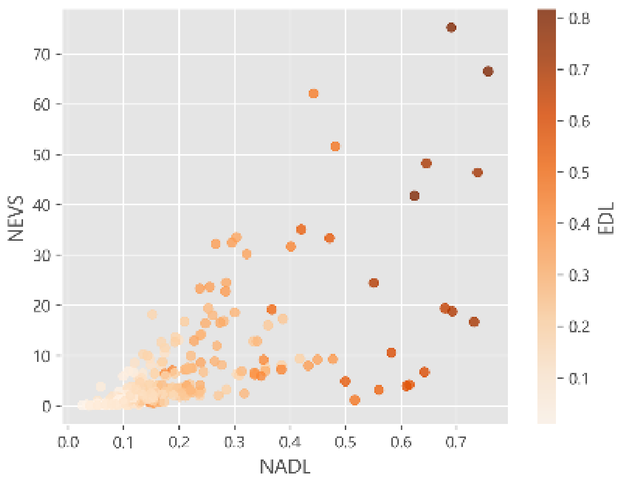

Figure 3. Scatterplot of the Digitalization Index and New Energy Vehicle Sales

The local polynomial estimation method is widely used in practice because it can estimate the regression function itself and its different order derivatives (depending on the smoothness of the regression function itself), and it also estimates a regression function better near the boundary of the support set. One can first understand the idea of local polynomial estimation by arbitrarily choosing an estimator $c = \widehat{m}(x)$ of $m(x)$ such that the sum of squares $\sum_{i=1}^{n} (y_i - c)^2$ is minimised, thus solving for $\widehat{m}(x) = \frac{1}{n}\sum_{i=1}^{n} y_i$. By the extreme value solution method, it can be solved that $\widehat{m}(x) = \frac{\sum_{i=1}^{n} w_i(x) y_i}{\sum_{i=1}^{n} w_i(x)}$, this happens to be kernel regression estimation. This means that the estimate can be improved by using a local polynomial rather than a local constant. For a value x in any field of x_0, $m(x)$ can be approximated by a Taylor expansion as follows:

$$m(x) \approx \sum_{i=0}^{p} \frac{m^i(x_0)}{i!}(x - x_0)^i \equiv \sum_{i=0}^{p} \xi_i (x - x_0)^i \quad (10)$$

Choose the $\hat{\xi}$ that minimises the local weighted sum of squares above to estimate ξ_i, i= $1,2,\ldots,p$.

One key task in nonparametric estimation is the choice of the window width \$h\$, and in this paper, we choose the missing cross validation method.

$$CV = \frac{1}{n}\sum_{i=1}^{n} (Y_i - \hat{r}_{(-i)}(X_i))^2, \quad (11)$$

here, $\hat{r}_{(-i)}(X_i)$ is obtained by excluding the ith data point (X_i, Y_i) from the sample data, leaving us with $n-1$ data points. We can select the optimal window width parameter h by minimizing the equation above.

$$\begin{aligned} NEVS_{it} &= \rho WNEVS_{it} + m(NADL_{it}) + \beta_2 EDL_{it} \\ &+\beta_3 UR_{it} + \beta_4 POP_{it} + \beta_5 RS_{it} \\ &+\beta_6 CPO_{it} + \sum_{i\neq j} \gamma_i w_{ij} X_{it} + \varepsilon_{it.} \end{aligned} \quad (12)$$

The results of the parameter estimation are shown in Table6.

Table 6. Parameter Estimation

Variable	Sales volume
ρ	0.124*(0.128)
EDL	1.310***(0.269)
UR	-0.062*(0.298)
POP	-1.329*(2.502)
RS	0.035*(0.040)
CPO	0.416***(0.081)
Observations	210
R-squared	0.339
Fixed time	YES
Province Fixed	NO

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6 shows that the coefficients for each variable are significant at the 1% significance level. The spatial autocorrelation coefficient is 0.124, which is also significant at the 1% level, indicating a positive spatial correlation: provinces with high sales of new energy vehicles tend to have higher sales in their neighboring provinces, and vice versa.

Fig. 4 presents the scatter plot of new energy vehicle sales (NEVS) estimates by digitization level (NADL), while Fig. 5 displays the estimation of the partial derivatives of the non-parametric component of the semi-parametric time-fixed-effects spatial Durbin model, specifically the estimation of the output coefficients by digitization level (NADL).

With the non-parametric digitization level (NADL), it is evident that the output coefficients of the digitization level are not invariant constants; rather, they vary with time and space. Furthermore, when the digitization level becomes excessively high, it can have a detrimental effect. Therefore, an increase in the level of digitization is not necessarily "the higher, the better." Instead, it requires a balance of various factors to achieve sustainable economic, social, and environmental development.

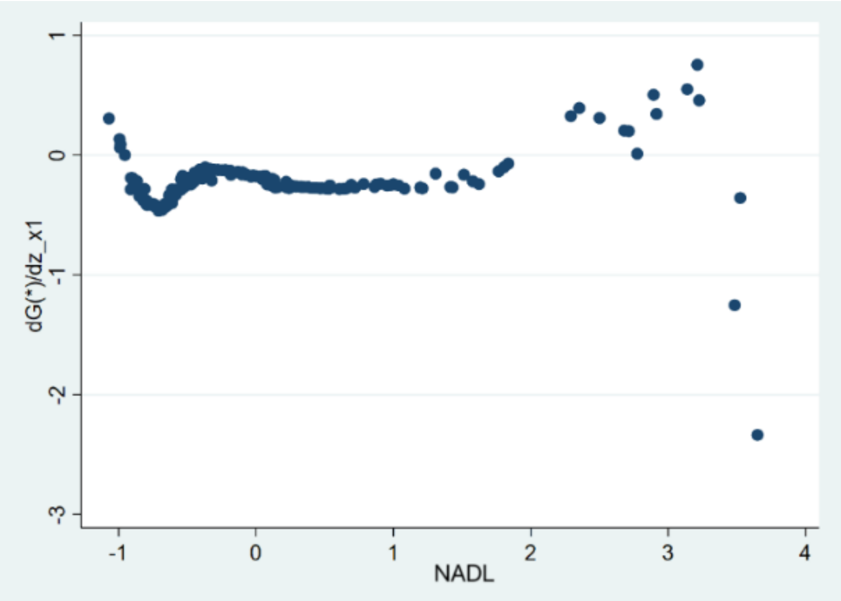

Figure 4. Scatterplot of NADL on SNEV Estimates

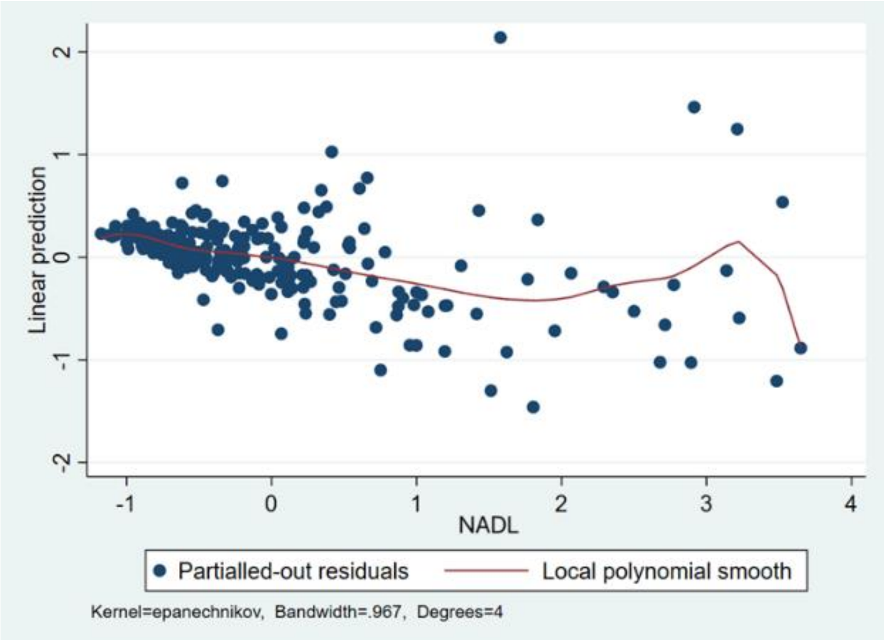

Figure 5. Scatterplot of Partial Derivatives of NADL on SNEV Estimates

4. Conclusion

The study finds that the digital economy significantly influences local new energy vehicle sales and contributes to growth in neighboring regions through spatial spillovers. The digital economy index has continued to rise from 2016 to 2023 and is concentrated in specific economic clusters. However, as digitization increases, new energy vehicle sales are negatively impacted, possibly due to the uncertainty surrounding the status of this new industry and consumer bias toward traditional vehicles, resulting in a saturated automotive market.

From 2016 to 2023, the digital economy index has increased and concentrated in specific economic clusters. Regional variations in economic development and digitization affect new energy vehicle sales (NEVS) differently. In developed provinces with high levels of digitization, sales have declined, whereas regions with lower digitization have seen growth.

To enhance the new energy vehicle industry through the digital economy, the following recommendations are proposed:

For car manufacturers: Focus on developing smart products to enhance user experience, adopt digital marketing strategies to effectively target customers, and collect user feedback for product optimization.

For the government: Increase subsidies for vehicle purchases and charging infrastructure, support research and development in digital technology, and strengthen digital infrastructure to improve user convenience.

References

[1] Getis A. Cliff, ad and ord, jk 1973: Spatial autocorrelation. london: Pion[J]. Progress in Human Geography, 1995, 19(2): 245-249.

[2] Anselin L. Spatial Econometrics: Methods and Models [M]. Dordrecht: Kluwer Academic,1988a: 7-133.

[3] Lesage J P, Parent O. Bayesian Model Averaging for Spatial Econometric Models[J]. Geographical Analysis, 2007, 39(3): 241–267

[4] Wang, Qunwei, et al. "Spatial agglomeration effect of new energy vehicle industry in China." Journal of Cleaner Production 229 (2019): 1025-1036.

[5] Anselin L. Spatial Econometrics: Methods and Models [M]. Dordrecht: Kluwer Academic,1988a: 7-133.1.

[6] Lee L F, Yu J. Estimation of spatial autoregressive panel data models with fixed effects[J].Journal of Econometrics, 2010, 154(2):165-185.

[7] Du J, Sun X, Cao R, et al. Statistical inference for partially linear additive spatial autoregressive models [J]. SpatialStatistics, 2018:S2211675318300198.

[8] Guo, Jiaming, et al. "Spatial spillover effects of new energy vehicle industry in China." Energy Policy 109 (2017): 512-521.

[9] Wang, Qunwei, et al. "Spatial agglomeration effect of new energy vehicle industry in China." Journal of Cleaner Production 229 (2019): 1025-1036.

[10] Liu, Ming, et al. "Evaluation on the Influence of China's Policies on the Development of New Energy Automobile Industry." Economic Review 3 (2016): 111-122.