Financial Engineering and Risk Management (2022)
Clausius Scientific Press, Canada

DOI: 10.23977/ferm.2022.050316
ISSN 2523-2576 Vol. 5 Num. 3

Empirical Analysis of Influencing Factors of Energy Consumption in Shandong Province based on Stepwise Regression and Path Analysis

Wang Xinyue

School of Statistics, Jilin University of Finance and Economics, Changchun, Jilin, 130117, China

Keywords: Energy consumption in Shandong Province, Influencing factors, Stepwise regression, Path analysis method

Abstract: In the context of energy consumption or urgent economic development, it is necessary to explore the factors that affect the energy consumption of a country. Shandong Province is a major energy consuming Province in China. This paper takes its data from 2000 to 2020 as an example. The index system is established from the three dimensions of economy, society and energy. Firstly, the stepwise regression method is used to screen the indicators, and four significant impact indicators are obtained: output value of secondary industry, per capita consumption expenditure of urban residents, energy consumption per unit GDP and output value of tertiary industry. After establishing the multiple regression model, according to the path analysis method, the indirect influence between the indexes is obtained, and the models pass the test. Finally, it puts forward the policy suggestions that Shandong Province needs to further optimize the industrial structure, improve energy efficiency and enhance consumers' awareness of saving consumption.

1. Introduction

Energy is an important material basis for China's economic and social development. In 2020, China's energy use has surpassed the United States to become the largest energy consuming country. Therefore, under the situation of resource shortage, it is very necessary to deeply explore the factors affecting energy consumption [1]. Energy consumption has always been a key research direction in China's academic circles. The research on this issue plays a key role in the development of China's low-carbon economy and the rational utilization of energy. As a major energy consumption province in China, Shandong Province has more value in its research, which can improve the overall energy consumption of the country. The research object of this paper is the influencing factors of energy consumption in Shandong Province. On the issue of energy consumption, some scholars have studied the impact of energy factor prices on energy consumption from 2000 to 2020 [2], but due to the single index, the factors affecting energy consumption cannot be fully considered. Or some scholars have studied the dynamic relationship between total industrial output value and energy consumption and concluded that total industrial output value is the Granger reason for raw coal consumption, but it lacks in-depth exploration of the impact relationship [3]. At the same time, it is also found that some scholars only have a simple written description of the problem of energy consumption, but lack the

corresponding empirical support. Therefore, this paper will analyze the problem of energy consumption in Shandong Province more scientifically and perfectly, introduce the latest data of energy consumption in Shandong Province, use the stepwise regression method to scree the index system, establish a multiple regression model, and use the path analysis method to refine the influencing factors of the interaction of various indicators on energy consumption in Shandong Province, so as to seek the way to reduce energy consumption in Shandong Province.

2. Establishment of Index System and Data Source of Influencing Factors of Energy Consumption in Shandong Province

By comprehensively comparing relevant literature and considering the actual situation of Shandong Province, this paper establishes an index system from the three dimensions of economy, society and energy [3,4,5,6,7]. In the economic dimension, economic development has a strong dependence on energy consumption, and production and life are inseparable from energy consumption; The development of the secondary industry dominated by industry and the tertiary industry dominated by service industry and manufacturing industry is inseparable from energy consumption. Therefore, four indicators are selected under this dimension: GDP, output value of the secondary industry, output value of the tertiary industry and the number of industrial enterprises. The social dimension reflects the effect on energy consumption based on the improvement of people's living standards and quality. Therefore, five indicators are selected under this dimension: urban residents' disposable income, urban residents' per capita consumption expenditure, provincial people's per capita consumption expenditure, civilian car ownership and urbanization rate. In the energy dimension, the consumption of different types of fuels is the direct effect and embodiment of energy consumption. In this dimension, four indicators are selected: the proportion of coal fuel, the proportion of oil fuel, the average daily coal consumption and the energy consumption per unit of GDP. According to the availability and real-time of data, the data of this study are from the statistical yearbook of Shandong Province from 2000 to 2020.

3. Empirical Analysis on Influencing Factors of Energy Consumption in Shandong Province Based on Stepwise Regression

3.1 Research ideas and research methods

In the process of studying the influencing factors of energy consumption in Shandong Province, the main problem at this stage is that there are too many indicators involved, and the previously established index system is too theoretical, which may not have a good interpretation effect in the actual modeling process, so it is necessary to screen the indicators. In the process of index screening, this paper adopts the method of stepwise regression, which automatically identifies the indexes that have a significant impact on the total energy consumption through the system. Its essence is still regression analysis, but on its basis, it automatically removes the indicators that are not statistically significant, dynamically introduces an indicator into the model every time, and then immediately tests the significance of the model. If the model is no longer significant after the introduction of indicators, the indicators shall be eliminated and repeated for many times until a series of indicators with the highest significance level of the model are selected. This method is easy to operate and can obtain more concise and significant important variables. Then, multiple regression is carried out according to the selected indicators to explore the functional relationship of the impact of each indicator on energy consumption in Shandong Province. Multiple regression analysis method refers to the statistical method of establishing the linear or nonlinear mathematical model relationship between multiple variables and analyzing the sample data. After establishing the multiple regression model,

the influence of each index on the explained variables can be obtained.

3.2 Result analysis

3.2.1 Screening results of index system based on stepwise regression method

With the help of SPSS software, all indicators in the initially established index system are added to the model fitting process. After running, the following results are obtained. It can be seen that model 4 has the highest goodness of fit when introducing four indicators: output value of secondary industry, per capita consumption expenditure of urban residents, energy consumption per unit GDP and output value of tertiary industry, and finally reaches 0.992. At present, retaining these four indicators has the best effect on the analysis of the influencing factors of energy consumption in Shandong Province. The model is no longer lack of important variables. These four indicators can be used to explain the impact of energy consumption in Shandong Province. Specifically analyzing these indicators, the four indicators respectively reflect the economic, social and energy dimensions of the indicators; The increase of output value of secondary and tertiary industries in Shandong Province is the main incentive mechanism for the continuous improvement of economic benefits in Shandong Province in recent years. At this stage, the pace of industrial development in Shandong Province is rapid. A system with energy, chemical industry and metallurgy as the pillar industries has been formed, but these pillar industries are high energy consumption industries. The tertiary industry in Shandong Province is dominated by the industries of production and living services and circulation services. In recent years, residential consumption and food, tobacco and alcohol expenditure account for a large proportion of its urban per capita consumption expenditure, which also reflects the increase of residents' consumption of the products produced by the secondary and tertiary industries. The energy consumption per unit GDP is calculated by dividing the total energy consumption of various industries by the total production value of the region in that year. This index can reflect the development degree of regional low-carbon economy. Due to its rising trend, it can be seen that the low-carbon development of Shandong Province still needs to be vigorously promoted.

Table 1 Stepwise Regression Model Summary Table

Model	R	R side	R square after adjustment	Error in standard estimation
1	.981a	.962	.960	2234.891
2	.991b	.983	.981	1530.760
3	.994c	.988	.985	1346.338
4	.997d	.993	.992	1019.781

a) Predictor variable: (constant), output value of secondary industry
b) Predictor variable:(constant), output value of secondary industry, per capita consumption expenditure of urban residents
c) Predictor variable:(constant), output value of secondary industry, per capita consumption expenditure of urban residents, energy consumption per unit GDP (energy consumption intensity)
d) Predictor variable:(constant), output value of secondary industry, per capita consumption expenditure of urban residents, energy consumption per unit GDP (energy consumption intensity), output value of tertiary industry

3.2.2 Model Results of Influencing Factors of Energy Consumption in Shandong Province

Based on the screening of the above indicators, four indicators are introduced: output value of secondary industry, output value of tertiary industry, per capita consumption expenditure of urban residents and energy consumption per unit of GDP. The multiple regression model is established in Stata, and the regression results are shown in Table 2 below. The significance level of variables passed the test at the significance level of 0.01. However, in order to make the model more robust and reliable,

the next test is needed.

Table 2 Results of multiple regression model

Independent variable	Regression Coefficient	Standard deviation	T value	P value	95% significance level confidence interval	
Output value of secondary industry	.926	.084	10.930	0.000**	.746	1.105
Output value of the tertiary industry	-.013	.003	-3.690	0.002**	-.021	-.006
Per capita urban residents Consumption expenditure	.605	.085	7.100	0.000**	.424	.785
Energy consumption per unit GDP	8839.085	1992.897	4.440	0.000**	4614.332	13063.840
Constant	6082.006	2594.631	-2.340	0.032*	11582.380	-581.634

3.2.3 Model Verification

It is very necessary to test the fitted model, including the existence test of its multicollinearity, autocorrelation and heteroscedasticity. Multicollinearity refers to the high degree of correlation between explanatory variables. The existence of multicollinearity will make the economic significance of parameter estimators unreasonable and the significance test function of variables will lose its significance. This paper uses the variance expansion factor test method and runs through Stata software, The test results are shown in Table 3. The average value of VIF is 7.40, less than 10, which passes the test.

Table 3 Table of multicollinearity test results

Independent variable	VIF	1/VIF
Output value of secondary industry	14.220	0.070
Per capita consumption expenditure of urban residents	7.810	0.128
Energy consumption intensity per unit GDP	6.080	0.164
Output value of the tertiary industry	1.490	0.672
VIF average	7.400	

The second econometric test after the establishment of multiple regression model is autocorrelation test, which refers to the expectations of random interference terms exist correlation between values, which can also be called sequence autocorrelation. If there is autocorrelation in the model, the parameter estimator will be ineffective and the significance test of variables will be meaningless. Here, we need to carry out BG test on the model. The p value is 0.7803, which is significantly greater than 0.05, so there is no sequence autocorrelation. At the same time, through the Stata graphical test method, as shown in Figure 1, the lag order changes from order 1 to order 8, and the model can pass the test at the 95% confidence level, that is, all results fall in the shaded area, so it can also be considered that there is no sequence autocorrelation in the model.

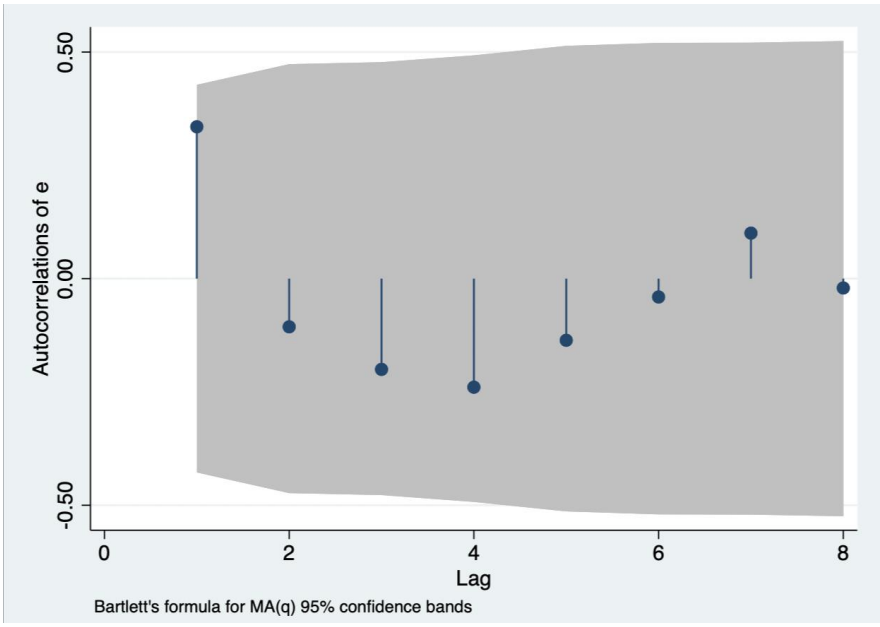

Figure 1 Figure of model autocorrelation test results

Then we need to carry out heteroscedasticity test on it. Heteroscedasticity refers to that the variance of random interference items in the model is no longer constant and different from each other. Its occurrence will still lead to the ineffectiveness of the parameter estimator of the model and the significance test of variables. Here, BP test method is adopted, and the p value is 0.5765, which is significantly greater than 0.05, which passes the heteroscedasticity test. The three tests show that the multiple regression model established by the four indicators is statistically reliable, so it has the function of further explanation.

3.2.4 Model analysis

Through the above test, we can further explore the influencing factors of energy consumption in Shandong Province according to the results of multiple regression model.

$$y = -6082.006 + 0.926x_1 - 0.013x_2 + 0.604x_3 + 8839.085x_4 \quad (1)$$

Y represents the total energy consumption of Shandong Province, $x_1$refers to the output value of secondary industry, $x_2$represents the output value of tertiary industry, $x_3$means energy consumption per unit of GDP, $x_4$means the per capita consumption expenditure of urban residents. It can be concluded from the regression model that every unit increase in the output value of the secondary industry will increase the total energy consumption by an average of 0.926 units, which is mainly due to the huge role of the high-energy consumption industry in the secondary industry, including the typical high-energy consumption and high emission industries such as refining, coking and coal liquid fuel in Shandong Province. The coefficient before the output value of the tertiary industry is negative, which means that every increase of one unit will inhibit the increase of the total energy consumption by 0.013 units, but this does not mean that the tertiary industry will reduce energy consumption, but the role of the tertiary industry is relatively small compared with the impact of the secondary industry on energy consumption; It can also be understood that when the GDP is certain, the larger the proportion of tertiary industries, the smaller the energy consumption. As an important indicator of low-carbon development, energy consumption per unit of GDP will increase the total energy consumption by an average of 0.604 units. The increase of energy consumption per unit of GDP is also the impact of the increase of output value of secondary industry. Every increase of per capita

consumption expenditure of urban residents can increase the total energy consumption by 8839.085 units on average. This is because the increase of per capita consumption expenditure will lead to the increase of energy consumption in the process of product production.

4. Based on path analysis to explore the size and interaction of influencing factors of energy consumption in Shandong Province

4.1 Research ideas and methods

In the multiple regression analysis, the impact of each index on the total energy consumption has been obtained. However, in order to further study the impact of the interaction between each index on energy consumption, the research methods of standardized regression and path analysis are adopted to decompose the influencing factors into direct and indirect effects. This method is simple, scientific and reliable, It can more intuitively reflect the impact of the interaction between indicators on the explained variables, so as to make a more detailed analysis of energy consumption in Shandong Province. Path analysis refers to the decomposition of the coefficients in the multiple regression model into two parts [8]. One part is the direct effect of explanatory variables on the explained variables, which is called direct path coefficient, and the other part is the indirect impact of explanatory variables on the mass production of the explained variables through the action of other explanatory variables, which is called indirect path coefficient. The indirect path coefficient is equal to the correlation coefficient multiplied by the path coefficient, so the key of path analysis is to calculate the path coefficient.

4.2 Analysis of results

Because the dimensions of each index are different, it is necessary to standardize them. Then we can directly compare the influence coefficient of each index, and get the standardized regression equation from the results of Stata software operation.

$$y^* = 0.845{x_1}^* - 0.092{x_2}^* + 0.407{x_3}^* + 0.224{x_4}^* \quad (2)$$

The results show that increasing the proportion of the tertiary industry and reducing the energy consumption per unit GDP and the per capita consumption expenditure of urban residents can effectively reduce the total energy consumption. The coefficient in the standardized regression model is the direct path coefficient. According to the results of the regression equation, it can be seen that the output value of the secondary industry has the greatest impact, followed by the energy consumption per unit GDP and the per capita consumption expenditure of urban residents, and the output value of the tertiary industry has a relatively small impact; Therefore, it can optimize the secondary industry structure and reduce the energy consumption per unit of GDP. Through the correlation coefficient between variables, the indirect path coefficient is calculated, and the results are shown in Table 4 below. From the path analysis, it can be seen that among the four independent variables, the output value of the secondary industry has the greatest direct impact on the energy consumption y. through the analysis of each indirect path coefficient, it is found that the output value of the secondary industry has a greater impact on the total energy consumption by acting on the energy consumption per unit of GDP. For each additional unit of the role of the secondary industry acting on the energy consumption per unit of GDP, it will indirectly promote the energy consumption by an average of 0.279 unit, It shows that the output value of the secondary industry has obvious interaction with the unit GDP, and the secondary industry should be given priority in adjusting the energy structure. It is worth noting that the development of the tertiary industry will indirectly inhibit the energy consumption by an average of 0.029 units for every unit of increase in the per capita

consumption expenditure of urban residents. Each time the output value of the secondary industry is located in the tertiary industry, it will promote energy consumption by 0.025 units. Therefore, in order to reduce energy consumption, we should also consider the close relationship between the output value of the secondary industry and the output value of the tertiary industry.

Table 4 Table of drift diameter coefficient

	Direct path coefficient	Urban residents Per capita consumption	Output value of the tertiary industry	Energy consumption per unit GDP	Output value of secondary industry
Urban residents Per capita consumption	0.224	-	0.012	-0.025	-0.599
Output value of the tertiary industry	-0.092	-0.029	-	-0.150	-0.229
Energy consumption per unit GDP	0.407	-0.0139	0.034	-	0.578
Output value of secondary industry	0.845	-0.159	0.025	0.279	-

5. Conclusions and policy recommendations

Based on the empirical research on the influencing factors of energy consumption in Shandong Province by stepwise regression, multiple regression and path analysis, this paper obtains ideal results. Through the analysis, the following conclusions and suggestions are obtained:

(1) In terms of economy, first of all, we should optimize the industrial structure, increase the proportion of tertiary industry, optimize the internal structure of secondary industry and change the mode of economic growth. The development mode of extensive economy can not really promote economic development. We should change it into an intensive mode of production, limit the number of industrial enterprises and give priority to the development of necessary industrial enterprises. In addition, the path analysis results show that the indirect effect of the output value of the tertiary industry on the output value of the secondary industry will inhibit energy consumption. Therefore, it can strengthen the connection between industrial technologies among departments and improve the relative inhibition of energy consumption.

(2) At the social level, the per capita disposable expenditure of urban residents affects energy consumption. The reason is that the products purchased by residents come from material production, which contains a lot of energy consumption. Therefore, residents can be encouraged to have a sense of conservation and advocate the recycling of goods when purchasing goods.

(3) In terms of energy, energy consumption intensity, i.e. energy consumption per unit of GDP, is also a key factor to increase energy consumption. Its value can reflect the use of energy. The larger the value, the lower the efficiency of energy use. Therefore, Shandong Province can improve energy use efficiency, and pay attention to avoiding waste of resources in the process of transportation and use.

References

[1] Wang Jinghui, Xu Ziyuan, Li Xinyu, Zhu Jiaming. Econometric Analysis of the Factors Influencing the Growth of Energy Consumption in Anhui Province [J]. Journal of Pingxiang University ,2018,35(06):20-24.(in Chinese)

[2] Han Songyan. Measurement of Energy Factor Price Distortion and Its Impact on China's Energy Consumption [J].

Technical Economy and Management Research, 2022(04):8-13. (in Chinese)
[3] Shi Xu, Li Wanbin. Research on the Dynamic Relationship between Energy Consumption and Industrial Development [J]. Statistics and decision making, 2022, 38(07):127-130.DOI:10.13546/j.cnki.tjyjc.2022.07.025. (in Chinese)
[4] Wang An, Gao Fuyi, Yu Jihai. Ideas and Countermeasures to Promote Energy Transformation in Shandong Province [J]. Research on financial development, 2016(08):83-85. DOI:10.19647/j.cnki.37-1462/f.2016.08.014. (in Chinese)
[5] He Lihua, Yang Pan, Meng Yanlin, Kong yuan. Potential contribution of energy structure optimization to low-carbon Shandong [J]. Chinese population ·Resources and Environment, 2015, 25 (06): 89-97. (in Chinese)
[6] Lei Mingyu, Cai Wenjia, Liu Wenling, Wang Can. The heterogeneity in energy consumption patterns and home appliance purchasing preferences across urban households in China[J]. Energy, 2022,253.
[7] Chen Jing, Chen Huanxin, Zeng Yuke. The energy consumption prediction method of chillers based on gradient boosting regression tree [J]. Refrigeration and Air-Conditioning,2020,11
[8] Du Jiaju, Chen Zhiwei. Use Path analysis by SPSS linear regression [J]. Bulletin of Biology,2010,45(02):4-6. (in Chinese)