Research on Influencing Factors of Coal Price Based on Multiple Regression Model

Qing Lv

Anhui University of Finance and Economics, Bengbu, Anhui, 233030

ABSTRACT. *The article is based on a series of data related to the coal industry from May 2019 to April 2020, Firstly, the grey correlation model is used to screen out the factors with higher correlation degree with coal price, and then the multiple regression model is used to explore the influence degree and direction of each factor. Finally, it is concluded that rail traffic volume, ppi, fob and coal demand have influences on coal price, And put forward relevant Suggestions according to the results of the model, so as to promote the stable development of coal industry.*

KEYWORDS: *Coal Price, influencing factors, gray correlation, multiple regression*

1. Introduction

In recent years, China's energy mix has changed, with coal consumption gradually decreasing and people's consumption of natural gas increasing. However, coal is still one of China's important energy sources, and coal output has been increasing year by year. Data shows that China's coal output has increased to 3.68 billion tons in 2018. And the country continues to build coal railway transport ports, coal transport capacity has been greatly improved. The price of coal is the weathervane of the development of coal industry. The change of coal price has an influence on the development of coal industry and the operation of national economy. At present, due to the impact of the epidemic, the demand for coal remains sluggish and the imbalance between coal supply and demand has led to a sharp drop in coal prices. How to develop coal enterprises under the background of energy structure reform and epidemic situation and how to support the coal economy by the government to prevent a large number of coal enterprises from recession or even bankruptcy is worth studying.

At present, there are many researches on coal price in China, including the following aspects: In terms of influencing factors of coal price, Zhu Jimao [1] (2017) analyzed the factors influencing the coal price from four aspects, namely coal supply and demand, coal's own value, substitution effect of imported coal and national regulation policies, using comparative demonstration and other methods. The results show that the change of coal price is mainly affected by coal supply and demand. In the long run, coal prices are likely to stabilize between 500 and 550 a tonne. Tang Yijun and Su Xu [2] (2018) explored the relationship between coal production cost,

coal consumption, macroeconomic prosperity index and other factors and coal price by using the gray model, and finally concluded that coal production cost had the greatest positive impact on coal price. Zhang Tonggong and Zhao Dezhi [3] (2018) combined qualitative and quantitative methods and used the stepped-up regression model to study the influencing factors of coal price. The results showed that factors such as coal output, coal import and export, and coal industry downstream product output all had an impact on coal price fluctuation. In the coal price fluctuation and its formation mechanism, Feng Caicai [4] (2017) researched coal price fluctuation characteristics, from the four aspects of seasonality, trend, periodicity and randomness of data, the seasonal fluctuation of coal price is obvious. Zhang Dongyang [5] (2018) analyzed the supply factors, demand factors and comprehensive factors affecting coal price, He combined with the system dynamics model, so the coal price pricing system is established and the causal relationship diagram of coal price is analyzed. At present, most of the literatures firstly screen the influencing factors through qualitative analysis and then conduct quantitative analysis, which may omit some influencing factors. In this paper, the empirical analysis is carried out through the grey correlation model, and the possible influencing factors are accurately screened through the ranking of correlation degree data. Then, the multiple regression model is used to further explore the factors affecting the change of coal price.

2. Model principle

(1) Grey relational model

For the factors between two systems, the scale of the degree of correlation which varies with time or with different objects is called the correlation degree. In the process of system development, if the trend of change of the two factors is consistent, that is, the degree of synchronous change is high, then the degree of correlation between the two factors is high. On the contrary, the correlation between the two is low. Therefore, grey correlation analysis method, also known as "grey correlation degree", is a method to measure the degree of correlation among factors according to the degree of similarity or difference of development trend among factors. This paper uses the grey correlation model to screen the factors affecting the coal price through the correlation degree, and selects the factors with higher correlation degree to conduct multiple regression analysis. The main steps are as follows:

Step 1: Determine the comparison sequence and reference sequence between systems. In order to study the correlation between each factor and the coal price, the coal price positioning reference series is presented in this paper as x_0, the business index of macro-economic, commodity price index, coal supply and coal demand and other 25 data series are used as the comparative series. They are expressed as x_1, x_2, ..., x_{25}.

Step 2: standardize the data. Then, use the formula to calculate the correlation coefficient (δ_{0i}), the relevant data of the coal industry are standardized:

$$\delta_{0i} = \frac{\Delta(min) + \rho\Delta(max)}{\Delta_{oi}(k) + \rho\Delta(max)}$$

Step 3: since the correlation coefficient compares the correlation degree value of the reference sequence and the comparison sequence at different moments, the number is more than one, so the correlation degree r is calculated on the basis of the correlation coefficient:

$$r = \frac{1}{N}\sum_{k=1}^{N}\delta(k)$$

In the above mentioned r is the grey correlation between coal price and various factors.

(2) Multiple regression model

Regression analysis model is generally used to analyze the statistical quantitative relationship and change rules between things. It is a quantitative analysis method, which can finally show the quantitative relationship in the form of regression equation, so as to explore the degree of influence of independent variables on dependent variables [6].Regression models are widely used. In this paper, multiple regression models are used to explore the linear relationship between the coal price of the explained variable and the explanatory variables, such as coal demand. Its mathematical model is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \varepsilon$$

In the above equation, y represents the price of coal, and x_p represents the factors influencing the change of coal price obtained after screening by the grey correlation model. Multiple regression model is used to explore the influence direction and degree of each influencing factor on the change of coal price.

3. Empirical analysis

(1)Correlation analysis based on grey correlation model

The daily price data of Qinhuangdao port steam coal on May 1, 2019, solstice and April 30, 2020, as well as the data of 25 independent variables preliminarily selected, are imported into the grey correlation model and solved by MATLAB software. The specific results are shown in the table1.

As can be seen from the above table, the highest grey correlation with coal price is the coal export price index, its value is 0.95234, this may be related to the coal price index as a reference for coal pricing. The lowest grey correlation with coal prices is the quantity of steam coal exports, its value is 0.75817,Generally speaking, the grey correlation degree is greater than 0.9, indicating that the change trend between the reference series and the comparison series is consistent, and the correlation is strong, reaching a strong degree of correlation, Therefore, 14 variables such as international coal price, real Estate Climate Index, railway traffic volume,

commodity price index and business index of macro-economic are selected in order from large to small. These 14 variables are taken as independent variables and Qinhuangdao port steam coal price as dependent variables for multiple linear regression.

Table 1 Correlation degree solution

Index	Correlation degree	Rank
Business index of macro-economic	0.91595	8
Commodity Price Index	0.91694	7
Coal supply	0.90616	9
Coal demand	0.90398	14
Real Estate Climate Index	0.92224	3
Rail capacity	0.92185	5
Industrial added value increased year on year	0.87249	20
Raw coal: output	0.90579	11
Coal: Sales volume	0.87676	17
Coal: Quantity imported	0.90438	12
Coal: amount of import	0.83785	23
Price index: Coal: export price	0.95234	1
Price index: Coal: FOB	0.94755	2
Quantity of steam coal imported	0.8636	22
Quantity of steam coal export	0.75817	25
Futures settlement price: Crude Oil:TOCOM	0.87233	21
Market price: liquefied natural gas	0.87789	16
Market price: LPG	0.88096	15
Key power generation enterprises consume coal	0.90423	13
Steam coal: amount of import	0.87351	19
Steam coal: export amount	0.87547	18
Coal stockpiles	0.80506	24
The national Consumer Price Index	0.90586	10
ppi	0.9222	4
Total Retail Sales of Consumer Goods	0.91694	6

(2) Analysis of influencing factors of coal price under multiple linear regression

In this paper, a stepwise regression strategy is adopted by SPSS to analyze the influencing factors of coal price, and significance test, goodness of fit test and multiple linearity diagnosis are conducted. The 14 variables obtained by grey correlation analysis are introduced into the multiple linear regression model. After

four stepwise regression, the goodness of fit of the final model reached 91.8%.The specific results are shown in the table below:

Table 2   model summary

Model	R	R^2	Adjusted R^2	error
1	0.884	0.781	0.778	48.9824
2	0.923	0.852	0.847	40.596
3	0.954	0.91	0.906	32.7923
4	0.958	0.918	0.913	30.6232

It can be seen from the above table that the fourth model is the final multiple linear regression model with good fitting, Eliminate variables that will produce multicollinearity, and the final results are shown in the following table:

Table 3   Regression results

Index	coefficient	significance
Constant	-1200.285	0.000
Fob	1.815	0.000
ppi	12.926	0.000
Rail traffic volume	131.783	0.000
Coal demand	0.003	0.025

It can be seen from the above that the multiple regression model ultimately retains four variables, namely, the international coal price, PPI, the rail traffic volume and the coal demand, Take the significance level as 0.05,all the above variables passed the significance test. For every unit increase in the FOB coal price, the coal price increases by 1.815 units, for every unit increase in the PPT, the price of coal will increase by 12.926 units, for every unit increase in rail traffic volume, the coal price increases by 131.783 units, for every unit increase in coal demand, the price of coal goes up by 0.003 units. The change of rail capacity causes the biggest change of coal price, while the variable of coal demand has little influence on coal price.

The final equation is:

$$y = -1200.284 + 1.851fob + 12.926ppi + 131.783rt + 0.003cd$$

From the absolute value of coefficient, it can be seen that the influencing factors of Qinhuangdao port steam coal price are, in order from large to small, railway traffic volume, ppi, fob coal price and coal demand. The quantity of railway transportation has a great influence on the coal price and the quantity of coal inventory, Railway transportation has great advantages in the transportation of coal and other bulk commodities, and is the primary transportation mode of coal supply

and demand, the supply of coal is complementary to rail transport. Prices fluctuate around supply and demand.

4. Conclusion

(1) Conclusion

In this paper, gray correlation model and multiple regression model are used to combine the daily coal price data from May 2019 to April 2020, and MATLAB and SPSS software are used to explore the factors influencing the change of coal price and the degree of influence. Based on the empirical analysis, the following conclusions can be drawn:

First, the price of coal is affected by rail traffic volume, ppi, fob price and demand for coal. Among them, the rail traffic volume has the greatest influence on the coal price, which is related to the coal transportation mainly depends on the railway transportation mode，The quantity of iron wheel transport and the setting of ports have an influence on the coal price to a certain extent, and the setting of ports along the river and along the coast also promotes the circulation of coal. The second is the national producer price index, coal is a major raw material for industrial production. The ppi represents the fluctuation of ex-factory price of industrial products, which reflects the change of coal price.

Second, due to the impact of the epidemic, demand for coal, international coal prices and rail transport have all been impacted, showing a downward trend.The above four factors have a positive impact on the coal price, and the government can control the coal price by changing the above four factors.

(2) Suggestions

Based on the above conclusions, the following Suggestions are proposed: First, with regard to the balance between coal supply and demand, the epidemic situation has gradually improved. The government should make an overall plan for the resumption of work and production of coal enterprises to ensure the rational operation of coal enterprises and complete coal supply channels. At present, the coal market is out of order, policies should be introduced to promote the relative balance of coal supply and demand.

Second, in terms of coal transportation, it is necessary to build and maintain the operation of coal ports, increase coal transport capacity, optimize transport organization and link production and transport demand, and promote the normal circulation of coal. The government should maintain the stability of coal prices. Long-term decline in coal prices will lead to losses in the coal industry and affect the national energy structure.

References

[1] Zhu jimao. Analysis and forecast of influencing factors of coal price in China [J].

China coal, 2017, 43(11):20-24.

[2] Tang Yijun, Su Xu. Analysis of Influencing Factors of Coal price in China based on multidimensional Grey Model [J]. Journal of Liaoning University of Engineering and Technology (Social Science edition), 2012, 20(06):418-423.

[3] Zhang Tonggong, Zhao Dezhi. Empirical Research on influencing Factors of Coal price Fluctuation in China [J]. China Energy, 2008, 40(03):16-21.

[4] Feng Caicai. Based on the mechanism of dynamic and static perspective coal price fluctuation and the coal enterprise strategy research in China [D]. China mining university, 2017.

[5] Zhang Dongyang. Research on Coal Pricing Model and Design of Decision Support System in China [J]. China Coal, 2008, 44(10):113-118.

[6] Guo Yu. The impact of aging population on China's Labor Supply [J]. Economic Theory and Economic Management, 2013(11):49-58.