Analysis of the Factors Affecting China's CPI Based on Multiple Linear Regression Model

Xinlei Li and Shibo Xin *

Beijing Technology and Business University, Beijing, China; lixinlei9404@163.com
*Correspondence: 1210539474@qq.com

Abstract: As an important indicator reflecting macroeconomic operation, CPI (consumer price index) is closely related to the life of the people. Accurately identifying the economic factors affecting the fluctuation of CPI is of great significance for the country to formulate macroeconomic policies and promote economic growth. This paper selects seven factors including the social commodity retail price index and the industrial product ex-factory price index that affect CPI from 2000 to 2017 to construct a multiple linear regression model. The most important ones are obtained through parameter estimation, multicollinearity and correction, and heteroscedasticity test. The two influencing factors are the social commodity retail price index and the per capita GDP (Gross Domestic Produc) index. At the same time, it is found that the multiple linear regression model has a high degree of prediction of CPI, which can provide reference value for the country to formulate macro policies related to the consumer price index.

Keywords: CPI, Multivariate linear regression model; the Factors Affecting China's CPI

1. Introduction

CPI is the abbreviation of consumer price index and an important index reflecting macroeconomic operation. The fluctuation of CPI not only affects the life and consumption of urban and rural residents, but also affects the formulation of the government's monetary and fiscal policies and consumer price policies.

In recent years, domestic and foreign scholars have done a lot of research on the influencing factors of CPI. Cao Wei, Luo Hao and Deng Shengjun [1] studies show that there is a negative correlation between nominal effective exchange rate of RMB and domestic price level. Yao Mingming [2] concluded that there is a long-term co-integration equilibrium relationship between broad money supply and CPI. Chen Biqiong and He Yan [3] found a long-term co-integration relationship between import and export trade, foreign exchange reserves and prices. Feng Chaojun [4] found that the retail price of social commodities was positively correlated with CPI, while Kyrtsou and Labys found that PPI changes in the United States would cause CPI changes to some extent; Akcay found that PPI in Finland and France had a one-way causal relationship with CPI while CPI in Germany was negatively correlated with CPI [5]. There is a two-way causal relationship between PPI (Producer Price Index) and CPI. The research results of Wang Xiuqing and Qian Xiaoping [6] show that the impact of the rising prices of agricultural products on the overall national price level is gradually declining. Liu Wei [7] concluded that there is a long-term equilibrium relationship between grain retail price and CPI, and the impact of grain price fluctuation on CPI is significant.

To sum up, domestic and foreign scholars have carried out in-depth analysis on the influencing factors of CPI in many aspects. Multivariate regression analysis has been used to analyze the influencing factors of CPI, such as interest rate, money supply, factory price index of industrial products, retail price of social commodities, factory price of industrial products, total import and export and investment in fixed assets. The panel data analysis, VAR model and hypothesis test are fully studied. Based on this, this paper, on the basis of previous research results, includes as many factors as possible affecting CPI changes.

2. Date Acquisition and Model Construction

2.1. Data Acquisition

On the basis of Feng Chaojun's [4] research, eight indicators, such as Retail Price Index of Commodities, PPI, Fixed Asset Investment Price Index, Price Index of Agricultural Means of Production, Industrial Producer Purchase Price Index, Total import and export index, Per capita GDP index and CPI, are selected. The data are derived from the statistical yearbooks and bulletins of the National Bureau of Statistics over the years. The data of economic indicators from 2000 to 2017 are collected, and the indicators are processed on a fixed basis.

2.2. Model Construction

A multivariate regression model was established based on the selected eight economic indicators:

$$Y=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_4X_4+\beta_5X_5+\beta_6X_6+\beta_7X_7+\varepsilon \quad (1)$$

β_i is the coefficient of each influencing factor. X_i affects various economic indicators of CPI, which is Retail Price Index of Commodities, PPI, Fixed Asset Investment Price Index, Price Index of Agricultural Means of Production, Industrial Producer Purchase Price Index,

Total import and export index, Per capita GDP index and CPI. $\varepsilon \sim (0, \sigma^2)$ indicates random error.

The CPI consumer price index is selected as the dependent variable Y. Using stata14.0 as an analytical tool, multiple regression analysis was conducted on the influencing factors of the consumer price index from 2000 to 2015, and the consumer price index for 2016-2017 was predicted to test the validity of the model.

2.3. Descriptive Statistics

The descriptive statistical analysis of the eight indicators from 2000 to 2015 is shown in Table 1.

Table 1. Descriptive statistics.

Variable	mean	standard deviation	minimum	maximum
y	510.43	65.98	433.5	615.2
x1	390.43	38.15	346.7	445.6
x2	348.28	36.69	292.6	400.2
x3	238.23	30.29	198.6	276.6
x4	510.39	129.63	349.9	677.5
x5	305.16	56.10	222.7	379.3
x6	398.55	205.23	100.0	672.8
x7	1237.0	512.55	575.7	2110.9

The results in the table show that the data selected in this paper are within the normal range, and there are no abnormal values.

2.4. Multivariate Linear Regression Analysis

This paper uses Stata software to estimate parameters. The results are shown in table 2.

Table 2. Preliminary regression results.

y	coefficient	t-value	P- value
x1	0.816	3.270	0.011
x2	-0.452	-0.540	0.605
x3	-0.097	-0.220	0.833
x4	0.108	1.180	0.273
x5	0.222	0.310	0.766
x6	0.043	1.000	0.348
x7	0.036	2.100	0.069
C	187.085	2.300	0.051

Goodness of fit is 0.9993,F-value is 1549.55. According to the results of the table above, a preliminary multivariate linear regression model can be obtained.

$$Y = 187.085 + 0.816X_1 - 0.452X_2 - 0.968X_3 + 0.108X_4 + 0.222X_5 + 0.433X_6 + 0.363X_7 \quad (2)$$

However, according to the above table, most of the initial multivariate linear models are P > 0.1, not significant, and the goodness of fit is close to 100%, F value is as high as 1549. Therefore, this paper speculates that the model may have multiple collinearity problems.

2.5. Multiple Collinearity and Correction

Based on the classical judgment method, this paper considers whether there are multiple collinearities. Therefore, the variance expansion factor VIF is calculated by the variance expansion factor method. The results are shown in Table 3.

Table 3. VIF Results of Variance Expansion Factor.

Variable	VIF	1/VIF
x5	4085.280	0.000
x2	2371.790	0.000
x3	454.450	0.002
x4	352.680	0.003
x1	226.480	0.004
x6	198.640	0.005
x7	195.860	0.005
Mean VIF	1126.450	--

The results show that the VIF of each index is far greater than 6, among which the best fit is the variable X5, i.e. the industrial producer purchase price index, and the largest variance expansion factor VIF is 4085.28, which fully confirms the conjecture that there are multiple collinearities in the preliminary multivariate linear regression model. In order to solve this problem, this paper chooses the variable X5. Explain scalars for stepwise regression, and the results are shown in Table 4.

Table 4. Stepwise regression results.

Variable	X1	X2	X3	X4	X5	X6	X7	R^2
X7	--	--	--	--	--	--	0.1281	0.9904
X7,X1	0.8538	--	--	--	--	--	0.0656	0.9988
X7,X1,X2	0.8208	0.0315	--	--	--	--	0.0661	0.9988
X7,X1,X3	0.7783	--	0.1291	--	--	--	0.0638	0.9989
X7,X1,X4	0.0629	--	--	0.0427	--	--	0.0629	0.9989
X7,X1,X5	0.8241	--	--	--	0.0260	--	0.0653	0.9988
X7,X1,X6	0.8366	--	--	--	--	0.0175	0.0601	0.9989

The results show that consumer price index is positively correlated with commodity retail price index and per capita GDP. The model is as follows:

$$Y = 95.88408 + 0.8537747X_1 + 0656491X_7 \quad (3)$$

Subsequently, to test the selected data by White test, the test result P > 0.1, rejects the original hypothesis, there is no heteroscedasticity.

2.6. Analysis of Prediction Result of Regression Model

According to the multiple linear regression model taking the retail price of X1 social commodities and the per capita GDP index of X7 as independent variables, the predicted CPI values from 2010 to 2017 are obtained. The comparison between the predicted values and the real values is shown in Table 5.

Table 5. Comparison of predicted results.

year	CPI	predicted value	Error value

2010	536.10	540.05	3.95
2011	565.00	565.81	0.81
2012	579.70	580.84	1.14
2013	594.80	594.28	-0.52
2014	606.70	606.28	-0.42
2015	615.20	614.90	-0.30
2016	627.50	626.04	-1.46
2017	637.50	639.45	1.95

According to the fitting results, it can be seen that the fitting degree of the multiple linear regression model is better. At the same time, it shows that the retail price index and the per capita GDP index are the two most important factors in predicting CPI.

3. Conclusions

The model estimated in this paper is ideal and has a high degree of fitting. From the model, we can see that the level of consumer price is positively correlated with the retail price of social commodities and GDP per capita, but not with other variables. This coincides with its economic significance.

In the multivariate linear regression model established in this paper, the influence coefficient of commodity retail price index on consumer price index is 0.8538, and that of per capita GDP on consumer price index is 0.6565. Therefore, it is considered that the independent variable commodity retail price index and per capita GDP have greater influence on consumer price level.

Acknowledgement

This work was supported in part by a grant from Research on Dynamic Prediction Mechanism of Internet Public Opinion in the Background of Big Data (18YJZH203).

References

[1] Cao W., Luo H., Deng S. “Empirical Analysis of the Impact of RMB Exchange Rate Transfer on China's Price Level: 2005-2008”, World Economic Research, 2009, pp. 25-31+87-88.
[2] Yao M. “Research on the correlation between interest rate, exchange rate and CPI”, Liaoning University, 2012.
[3] Chen B., He Y. “Study on the dynamic transmission mechanism of China's foreign exchange reserves and CPI fluctuations”, Exploration of international economy and trade, 2009, pp. 69-74.
[4] Feng C. “Identification of influencing factors of CPI in China based on multiple regression analysis”, Statistics and decision-making, 2017, pp. 125-127.
[5] Jonathan W. “An empirical analysis of price transmission by stage of processing”, Monthly Labor Review, 2002.
[6] Wang X., Qian X. “The ripple effect of the rising prices of agricultural products in China from 1981 to 2000”, Rural Economy of China, 2004, pp. 12-15+61.
[7] Liu W. “An Empirical Study on the Dynamic Relationship between Grain Price and CPI in China: Based on Panel Data Model Analysis”, Shang, 2016, pp. 282-283.