

Research on Chinese Import Trade Based on Multiple Linear Regression Model

Biao Xu

School of North China Power University, Baoding 071000, China.

18632016554@163.com

Keywords: Import; multiple linear regression model; OLS; econometric test.

Abstract. This paper uses OLS parameter estimation method and GDP, per capita disposable income, export, foreign investment and exchange rate fluctuations as the five factors for our multiple linear regression analysis, and the economic significance, statistical significance and econometric test are made. The conclusions are that: (1) the main factors that affect the import trade are GDP, urban residents per capita disposable income, total exports, influence of foreign investment on import trade, and fluctuations in the exchange rate, which has a negative influence on import trade. The two factors: GDP and the impact of foreign investment, have a serious overlap relationship and causes the col-linearity problem. The influence of exchange rate on import trade, can display by urban residents per capita disposable income. (2) According to the results of the test for heteroscedasticity can be concluded that the effect of 2007-2014 data of the model don't have heteroscedasticity. (3) The autocorrelation test shows that, the import trade by a year ago has majorly impact on the nowadays, total imports often with national fundamentals associated closely. First three quarter of imported trade resulted in an increase in the import trade increased. In front of the fourth quarter of imports often have a negative impact.

1. Introduction

Since joining the WTO, China's foreign trade import and export scale rapid growth. In 2001 Chinese import and export volume is 2661 billion, has grown to \$4.16 trillion in 2013. At the same time, trade surplus is also expanding, the trade surplus in 2008 reached amazingly 298.13 billion, in 2013 still high up 259.75 billion dollars and is 10.2 times in 2001. The imbalance of international payments' balance has become the outstanding problems in Chinese current economic operation.

Qing Shan Ni, Zeng Fan *The experience analysis of Chinese import trade relationship between continuous time and its influencing factors* said, the factors like specification language, GDP's products trade volume and the initial transactions, have a positive effect on Chinese import and export ongoing relationship, negative effect factors such as distance of Chinese import trade of a continuing relationship.

In this paper, the author based on the method of econometric researching strategy, using OLS parameter estimation method, to analyze five factors, occupying multiple linear regression model, and proceed economic significance, statistical significance and econometric test, ultimately concluded.

2. Model Establishment and Parameter Estimation

Through the economic analysis shows that GDP, urban residents per capita disposable income, total exports, foreign investment, exchange rate fluctuations are the main factors influencing the total import (Y), X_1 represents GDP, X_2 represents the per capita disposable income of urban residents, with the X_3 said total exports, X_4 said foreign-invested enterprises export volume, with X_5 represents exchange rate fluctuations. Establishing the five members linear regression model.

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + \mu \quad \mu \sim N(0, \sigma^2)$$

Here μ contains some other factors which affect the total imports, random factors and so on.

Table 1 shows the Chinese 1979-2014 years imports (billion U. S. dollars), GDP (million RMB), urban residents per capita disposable income (RMB), exports (billion U. S. dollars), actually utilized

foreign investment amount (billion U. S. dollars), exchange rate fluctuations (direct quotation method under the yuan against the dollar exchange rate, here expressed by the arithmetic average between the opening price begin of season and the closing price end of season).

Table 1 Chinese 1979-2014 years imports

Particular year	Import (Y)	GDP(X ₁)	Disposable income (X ₂)	Export (X ₃)	Foreign investment (X ₄)	Exchange (X ₅)	
2007	First Quarter	2056.3	56686.6	3934.9	2521.4	158.9	777.05
	Second Quarter	2285.5	64648.9	3117.1	2947.8	160.0	767.185
	Third Quarter	2583.4	68789.5	3294.2	3314.7	153.3	755.8
	Forth Quarter	2637.4	77894.3	3439.6	3397.6	275.5	740.745
2008	First Quarter	2645.4	68778.4	4385.6	3059.7	274.1	715.825
	Second Quarter	3025.7	78537.6	3679.3	3604.5	249.7	693.31
	Third Quarter	3246.4	81619.5	3799.9	4079.8	219.9	685.135
	Forth Quarter	2401.6	87816.3	3915.9	3544.7	180.2	683.38
.....							
2013	First Quarter	4653.7	128083.5	7202.9	5089.2	299.1	622.055
	Second Quarter	4784.3	143031.8	6043.7	5441.4	320.8	617.42
	Third Quarter	5006.9	150719.8	6598.6	5622.4	266.3	612.75
	Forth Quarter	5048.0	166183.6	6621.9	5953.5	289.8	608.775
2014	First Quarter	4747.5	138738.0	7911.9	4913.3	315.5	613.455
	Second Quarter	4848.6	155201.0	6607.8	5708.0	317.8	620.6
	Third Quarter	5069.4	163467.0	7177.3	6350.4	240.3	617.075
	Forth Quarter	4965.5	178732.8	7146.9	6460.4	322.1	617

Using Eviews and occupying the OLS parameter estimation, the results shown that:

$$\hat{a} = 1923.699 \quad \hat{b}_1 = -0.0109 \quad \hat{b}_2 = 0.1620 \quad \hat{b}_3 = 0.8424 \quad \hat{b}_4 = 2.7067 \quad \hat{b}_5 = -3.2867$$

$$\hat{Y} = 1923.699 - 0.0109x_1 + 0.1620x_2 + 0.8424x_3 + 2.7067x_4 - 3.2867x_5$$

$$SE = (1926.6) \quad (0.006) \quad (10.064) \quad (0.135) \quad (1.160) \quad (2.2788)$$

$$t = (0.999) \quad (-1.825) \quad (2.512) \quad (6.231) \quad (2.333) \quad (-1.442)$$

$$R^2 = 0.9642 \quad \bar{R}^2 = 0.9573 \quad Df = 32 \quad DW = 0.7872 \quad F = 139.8694$$

$$t_{\alpha}(n-k-1) = t_{\alpha}(26) = 1.706$$

3. Model Checking

3.1 Econometric Significance Test.

Multicollinearity test

Although the overall degree of model is favorable, the independent variable X5 's t-statistic is not significant, so there may be multicollinearity (Fig. 1).

Pearson							
	Y	X1	X2	X3	X4	X5	
Pearson	Y	1.000	.945	.855	.961	.733	-.892
	X1	.945	1.000	.855	.965	.687	-.904
	X2	.855	.855	1.000	.788	.681	-.873
	X3	.961	.965	.788	1.000	.641	-.846
	X4	.733	.687	.681	.641	1.000	-.737
	X5	-.892	-.904	-.873	-.846	-.737	1.000
Sig.	Y	.	.000	.000	.000	.000	.000
	X1	.000	.	.000	.000	.000	.000
	X2	.000	.000	.	.000	.000	.000
	X3	.000	.000	.000	.	.000	.000
	X4	.000	.000	.000	.000	.	.000
	X5	.000	.000	.000	.000	.000	.
N	Y	32	32	32	32	32	32
	X1	32	32	32	32	32	32
	X2	32	32	32	32	32	32
	X3	32	32	32	32	32	32
	X4	32	32	32	32	32	32
	X5	32	32	32	32	32	32

Fig. 1 Correlation Matrix

From the correlation matrix, we can see that the correlations between variables are higher, so there should be multicollinearity.

Multicollinearity's correction - stepwise regression

A: Calculate regression equations for each independent variable, namely, the basic regression equation:

$$Y = 468.4473 + 0.029199X_1 \quad R_1^2 = 0.892421 \quad t_1 = 15.78366$$

$$Y = 213.8784 + 0.672854X_2 \quad R_2^2 = 0.731706 \quad t_2 = 9.045300$$

$$Y = -53.7961 + 0.870689X_3 \quad R_3^2 = 0.923235 \quad t_3 = 18.99480$$

$$Y = -236.962 + 15.35994X_4 \quad R_4^2 = 0.536916 \quad t_4 = 5.897719$$

$$Y = 16869.99 - 19.74794X_5 \quad R_5^2 = 0.795908 \quad t_5 = -10.81631$$

Comprehensively comparing the regression results, the addition of R^2 for X_1 、 X_5 , although both have increased, the parameter of the T statistics are not significant, so the final measurement model is as follow:

$$\hat{a} = -656.7266 \quad \hat{b}_2 = 0.1523 \quad \hat{b}_3 = 0.6506 \quad \hat{b}_4 = 2.9524$$

$$\hat{Y} = -656.7266 + 0.1523x_2 + 0.6506x_3 + 2.9524x_4$$

$$SE = (213.572) \quad (0.0529) \quad (0.0582) \quad (1.1299)$$

$$t = (-3.0750) \quad (2.8785) \quad (11.188) \quad (2.6130)$$

$$R^2 = 0.9587 \quad \bar{R}^2 = 0.9542 \quad Df = 32 \quad DW = 0.9844 \quad F = 216.5156$$

Heteroscedasticity test

For Goldfeld-Quandt test. First, proposing the original hypothesis H_0 : μ_i is homoscedasticity; the alternative hypothesis H_1 : μ_i is heteroskedasticity. Then, since X is increasing sequence, deleting eight center observation values, the rest of the data are divided into two sub samples, one sample contains X lower values, others contains the values of X are higher. calculating the F statistic of Goldfeld-Quandt

$$F = \frac{RSS_2 / \frac{n-c}{2} - k - 1}{RSS_1 / \frac{n-c}{2} - k - 1} = \frac{RSS_2}{RSS_1} = \frac{113164.4}{320987.7} = 0.3526$$

the degrees of freedom for RSS_2 and RSS_1

$$\frac{n-c}{2} - k - 1 = 12 - 3 - 1 = 8$$

Under the 5% level of significance, finding the critical value of F distribution, the first degree of freedom is 9, second degree of freedom is 9, the critical value of F distribution threshold is 3.44. $F = 0.3526 < 3.44$, rejecting the assuming of alternative hypothesis H_1 : μ_i is heteroskedasticity, accepting the original hypothesis H_0 : μ_i is homoscedasticity and proving that there is no difference in the variance.

Autocorrelation test

Hypothesizing the model is first-order autocorrelation, putting forward the original hypothesis H_0 : $\rho = 0, \mu$ has no first-order autocorrelation; the alternative hypothesis H_1 : $\rho \neq 0, \mu$ has first-order autocorrelation.

Because of $DW = 0.9844$, querying DW statistics table, in the upper and lower bounds of 0.1, the linear equation which number of variables is 3 and the sample size is 32 has the $D_{l1} = 1.039$ and $D_u = 1.428$, so the $0 < DW < D_{l1}$, rejecting the original hypothesis H_0 : $\rho = 0, \mu$ has no first-order autocorrelation; accepting the alternative hypothesis H_1 : $\rho = 0, \mu$ with first-order autocorrelation.

Self related elimination-Estimation of AR (p) model and residual test

Establish the following regression model:

$$y_t = a + b_2x_{2t} + b_3x_{3t} + b_4x_{4t} + \varepsilon_t$$

$$\varepsilon_t = \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-2} + \phi_3\varepsilon_{t-3} + \phi_4\varepsilon_{t-4} + \phi_5\varepsilon_{t-5} + \mu_t$$

Using Eviews , the results showed that:

$$y_t = 0.198662x_{2t} + 0.617121x_{3t} + \varepsilon_t$$

$$t = (12.93977) (27.15470)$$

$$\varepsilon_t = 1.362729\varepsilon_{t-1} - 1.157318\varepsilon_{t-2} + 1.017790\varepsilon_{t-3} - 0.466968\varepsilon_{t-4} + \mu_t$$

$$t = (6.815961) (-4.218116) (3.629775) (-2.287816)$$

$$R^2 = 0.980564 \quad DW = 1.766255$$

3.2 Economic Significance Test.

According to the definition of economics, increase of urban residents per capita disposable income will bring the increase of domestic consumption level and inevitably bring the increasing of such import demand. Corresponding increase in export led to the country's economic growth and import enlarged. So our conclusions meeting the economic significance, \hat{b}_2, \hat{b}_3 is theoretically greater than zero.

3.3 Statistical Significance Test.

Coefficient of determination

$$R^2 = 0.980564 \quad \bar{R}^2 = 0.976147$$

Per capita disposable income of urban resident, the total amount of exports are two factors that affect the total amount of imports can explain 98.06% of change in the total amount of imports.

Variables significance test

$$t = (12.93977) (27.15470)$$

Because the $t_2=12.93977, t_3=27.15470$, we can see \hat{b}_2, \hat{b}_3 obviously not zero.

Test of significance of equation

Because the F statistic =114.1320, we can see the equation is obvious

4. Conclusion

According to the above test results, we can draw the conclusion that:

(1) In the factors that affect the import trade, GDP, urban residents per capita disposable income, total exports, influence of foreign investment have positive influence for import trade, fluctuations of exchange rate on import trade has negative influence. Urban residents per capita disposable income, GDP appear serious overlap, cause analysis of the problem serious multicollinearity problem; the influence of exchange rate on import trade, based on urban residents per capita disposable income and total factor the influence of import displayed.

(2) According to the results of the test for heteroscedasticity can determine 2007-2014 data to model the effect with the variance assumption, from 2007 to 2009, due to the economic crisis, such as force majeure, urban residents per capita disposable income increase effect of import trade is negative, the residents affected by the impact of the global economic crisis, reducing the demand for imports and increase in the demand for domestic goods. Exports and foreign investment to import trade produced positive effect. 2012-2014 years, chinese economy out of the morass of economic crisis, per capita disposable income of urban residents, the total export have positive influence for import trade. Residents's confidence on the domestic economy contributed to the purchase of imported goods booming again. Foreign investment to import trade influence is negative. This is mainly due to China's macro economy has entered a period of adjustment, foreigners treated chinese economy cautious, many investors quited. But new manufacturers with innovation, and increasing demand for imports, thus making foreign investment reduction has inversely brought the increase of import trade.

(3) The autocorrelation test shows that, the import trade by a year ago has majorly impact on the nowadays, total imports often associating with national fundamentals closely. First three quarter of imported trade resulted in an increase in the import trade increased. In front of the fourth quarter of imports often have a negative impact.

References

- [1]. Zhang Baofa, Economic Meteorology, 2006, Fifth Edition.
- [2]. Fan Huanhuan, Zhang Lingyun, EViews statistic analysis and application, 2009.
- [3]. Yang Gongpu, Xia Da, Gong Yangjun, Industrial Economics Total, 2008, Third Edition.