

# The Empirical Analysis on Influencing Factors of Chinese E-commerce

<sup>1</sup>Xinting Xu, <sup>2</sup>Fengyu Wu, <sup>3</sup>Chengli Song, <sup>4</sup>Hongjian Yin

<sup>1</sup> College of Management, University of Shanghai For Science and Technology, Shanghai, China

<sup>2</sup> College of Management, University of Shanghai For Science and Technology, Shanghai, China

<sup>3</sup> Medical equipment, University of Shanghai For Science and Technology, Shanghai, China

<sup>4</sup> Medical equipment, University of Shanghai For Science and Technology, Shanghai, China

Email: [wyddbmdm@126.com](mailto:wyddbmdm@126.com)

**Abstract** –Chinese e-commerce develops quickly and the e-commerce volume of trade grows fast with the size of the market increasing dramatically. Based on the data of Internet users, online shoppers, domain names and e-commerce enterprises from year 2001 to 2010, this paper studies the relations between these factors and e-commerce volume of trade. It uses Eviews software to conduct empirical analysis. The conclusion is China should improve the popularity of Internet and encourage e-commerce construction of enterprises to increase the e-commerce volume of trade and promote the development e-commerce in China.

**Keywords** –e-commerce volume of trade; Preparation; Internet users; online shoppers; domain names; e-commerce enterprises; regression function

## 1. Introduction

Along with the rapid development of information network technology, e-commerce in China has achieved great development in recent years. It has been developing rapidly since 2006. The size of the market increasing dramatically and it is in a rapid growth period with the volume of trade increasing. By the year 2010, the volume of trade in the e-commerce market in China (including B2B and B2C, C2C) has amounted to 4.5 trillion yuan. The development of the electronic commerce is affected by many factors and many of the factors have been studied. Based on the data of Internet users, online shoppers, domain names and e-commerce enterprises from year 2001 to 2010, this paper uses Eviews to conduct empirical analysis to study the relations between these factors and e-commerce volume of trade.

## 2. Variation and determination of the model

In order to validate the factors that influence e-commerce, the paper uses the data from [2]-[5] as samples to conduct model validation of multi-element linear regression analysis.

### 2.1. Preliminary linear function model set

This paper takes the e-commerce volume of trade of our country as the dependent variable, the amounts of Internet users, online shoppers, domain names and e-commerce enterprises which can measure the development level of e-commerce as independent variables and establish multivariate linear regression model:

$$Y_i = C + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \varepsilon \quad (1)$$

In this equation,  $C$  is constant,  $Y_i$  is e-commerce volume of trade,  $X_1$  is the amount of Internet users,  $X_2$  is the amount of online shoppers,  $X_3$  is the amount of domain names and  $X_4$  is the amount of e-commerce enterprises.  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  and  $\beta_4$  are the coefficients of  $X_{1i}$ ,  $X_{2i}$ ,  $X_{3i}$  and  $X_{4i}$  respectively and  $\varepsilon$  is random variables. Based the data collected above, the multiple linear regression model is established with the Eviews software and the estimate result is shown in Figure 1.

Method: Least Squares  
 Date: 01/04/12 Time: 00:29  
 Sample: 2001 2010  
 Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5156.168	1041.410	-4.951141	0.0043
X1	0.655062	0.601035	1.089890	0.3255
X2	-0.960602	1.554128	-0.618097	0.5636
X3	6.127334	4.031235	1.519965	0.1890
X4	1.273945	0.265666	4.795283	0.0049

R-squared	0.997086	Mean dependent var	16150.00
Adjusted R-squared	0.994754	S.D. dependent var	16261.22
S.E. of regression	1177.771	Akaike info criterion	17.28749
Sum squared resid	6935727.	Schwarz criterion	17.43878
Log likelihood	-81.43744	Hannan-Quinn criter.	17.12152
F-statistic	427.6605	Durbin-Watson stat	3.280974
Prob(F-statistic)	0.000002		

Figure 1. Regression analysis output

2.1.1. Fitting test of the mode

$R^2 = 0.997086$  and adjusted  $R^2 = 0.994754$  shows that the model fits very well. The coefficient of  $X_2$  is minus which means the relationship of e-commerce volume of trade and the amount of online shoppers is minus and it is do not accord with economic meaning, so delete the independent variable  $X_2$ .

2.1.2. Significant test- F-statistic

Calculated  $F = 427.6605$ ,  $n = 10$ ,  $k = 4$ . Given the level of significance  $\alpha = 0.05$ , check the F distribution table and

get the critical value  $F_{0.05}(4, 5) = 5.19$ . Obviously, F is greater than  $F_{0.05}$ . So reject the null hypothesis, we can think that the linear relationship of the model is significant in the probability level of 95% is significant. That means the linear relationships of e-commerce volume of trade with the amounts of Internet users, online shoppers, domain names and e-commerce enterprises are significant.

Eliminate the independent variable  $X_2$ , and then conduct regression, get Figure 2.

Dependent Variable: Y  
 Method: Least Squares  
 Date: 01/05/12 Time: 00:31  
 Sample: 2001 2010  
 Included observations: 10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5633.887	661.0629	-8.522467	0.0001
X1	0.296025	0.146195	2.024862	0.0893
X3	8.425915	1.473760	5.717292	0.0012
X4	1.251893	0.249335	5.020937	0.0024

R-squared	0.996863	Mean dependent var	16150.00
Adjusted R-squared	0.995294	S.D. dependent var	16261.22
S.E. of regression	1115.473	Akaike info criterion	17.16112
Sum squared resid	7465677.	Schwarz criterion	17.28215
Log likelihood	-81.80559	Hannan-Quinn criter.	17.02834
F-statistic	635.5430	Durbin-Watson stat	2.787835
Prob(F-statistic)	0.000000		

Figure2. Regression analysis output

2.1. Preliminary linear function model set

2.2.1. Statistics test of the regression model

- Fitting test of the model

$R^2 = 0.996863$  and adjusted  $R^2 = 0.995294$ . It means that the model fits very well.

- Model significant test- F-statistic

Calculated  $F=635.5430$ ,  $n=10$ ,  $k=3$ . Given the level of significance  $\alpha= 0.05$ , check the F distribution table and get the critical value  $F_{0.05} (3, 6) = 4.76$ . Obviously, F is greater than F. So reject the null hypothesis, we can think that the linear relationship of the model is significant in the probability level of 95% is significantly. It means the linear relationships of e-commerce volume of trade with the amounts of Internet users, domain names, and e-commerce enterprises are significant.

- Significant test of the explanatory variables -t statistic

All the t statistics value are calculated with Eviews software,  $t_0=8.522467$ ,  $t_1 = 2.024862$ ,  $t_2 = 5.717292$ ,  $t_3 =$

$5.020937$ . Given significant level  $\alpha= 0.05$ , the degree of freedom is 6 in the t distribution table. We get  $t = 1.943$ , apparently,  $|t_0|$ ,  $|t_1|$ ,  $|t_2|$ ,  $|t_3|$  are greater than t. So the null hypothesis is rejected and we can think that causal relationships between the three explain variables and the explained variables are significant.

### 2.2.2. The autocorrelation inspection

Here we use Durbin-Watson (DW) test to test the autocorrelation of the model. When  $\alpha= 0.05$ ,  $n = 10$ ,  $k = 2$ , by checking the table, we can get that (dL, du) is (0.525, 2.016). From Figure 3, it is known that  $t DW = 2.787835$ . So  $4-du < DW < 4-dL$ , we can't judge whether there is the autocorrelation. And residual diagram analysis is conducted (Figure 3).

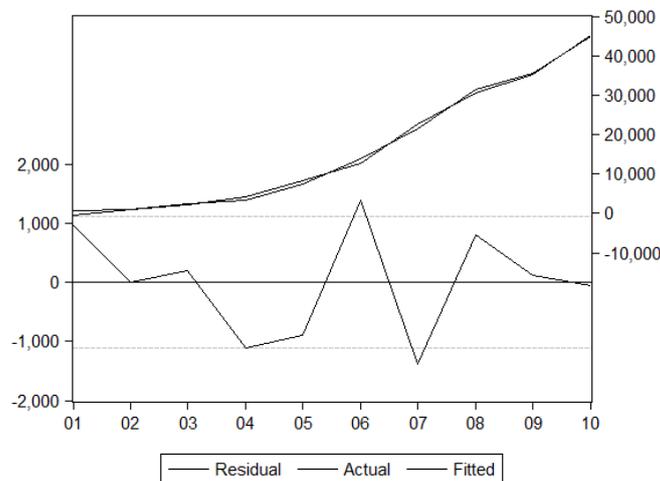


Figure3. Durbin-Watson Test

From the residual distribution plot, we can see there are periodic fluctuations in the residual distribution,

showing that there may be autocorrelation. And partial correlation analysis is conducted. (Figure 4)

Date: 01/14/12 Time: 14:41  
 Sample: 2001 2010  
 Included observations: 10

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
█	█	1	-0.455	-0.455	2.7579	0.097
█	█	2	0.084	-0.155	2.8641	0.239
█	█	3	0.032	0.008	2.8816	0.410
█	█	4	-0.293	-0.335	4.5935	0.332
█	█	5	0.187	-0.141	5.4361	0.365
█	█	6	-0.164	-0.219	6.2456	0.396
█	█	7	0.101	-0.110	6.6576	0.465
█	█	8	0.014	-0.131	6.6694	0.573
█	█	9	-0.007	-0.069	6.6759	0.671

Figure4. Residual distribution plot

AC in the figure represents the autocorrelation coefficients in different periods and PAC represents the partial correlation coefficients. On left side of the graphic, the dotted lines represent 0.5 and -0.5. If the histogram of the partial correlation coefficient S phase is more than the dashed part it means that the partial

correlation coefficient  $|\rho t-S| > 0.5$ , there is S order autocorrelation. From the PAC in the figure we can see, there is no histogram of a period that is more than the dotted lines, so there are no autocorrelation.

### 2.2.3. The autocorrelation inspection

Heteroskedasticity Test: White				
F-statistic	21.66735	Prob. F(6,3)	0.0144	
Obs*R-squared	9.774443	Prob. Chi-Square(6)	0.1345	
Scaled explained SS	1.908740	Prob. Chi-Square(6)	0.9279	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 01/14/12 Time: 23:01				
Sample: 2001 2010				
Included observations: 10				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	23119660	1961174.	11.78868	0.0013
X1^2	-1.482197	0.430261	-3.444882	0.0411
X1*X3	20.60965	7.882157	2.614721	0.0794
X1*X4	5.387492	1.304928	4.128573	0.0258
X3^2	-159.2976	59.40687	-2.681468	0.0750
X3*X4	-22.02610	7.508722	-2.933403	0.0608
X4^2	-5.318358	1.138214	-4.672546	0.0185
R-squared	0.977444	Mean dependent var	9784073.	
Adjusted R-squared	0.932333	S.D. dependent var	9207522.	
S.E. of regression	2395143.	Akaike info criterion	32.41181	
Sum squared resid	1.72E+13	Schwarz criterion	32.62362	
Log likelihood	-155.0591	Hannan-Quinn criter.	32.17946	
F-statistic	21.66735	Durbin-Watson stat	3.457085	
Prob(F-statistic)	0.014423			

Figure 5. White test

Because  $nR^2 = 10 \times 0.977444 = 9.7444$ , take significant level  $d = 0.05$ ,  $\chi_{0.05}(6) = 12.5916 > nR^2$ , so we can't reject the virtual hypothesis that there is homoscedastic and there does not exist different variance.

To sum up, the modified function model has gone through the tests. Therefore, the sample regression function for China's e-commerce volume of trade is:  

$$Y_i = 5633.887 + 0.296025X_{1i} + 8.425915X_{3i} + 1.251893X_{4i}$$
(2)

### 3. Results and enlightenments

From the regression function we can see that China's e-commerce volume of trade is influenced by amounts of Internet users, online shoppers, domain names and e-commerce enterprises. So the Internet resources and the e-commerce construction of enterprises will influence the e-commerce volume of trade. In order to promote the development of the e-commerce in China, we should take some measures.

#### 3.1. Increase China's Internet resources

The informatization construction in our country started relatively later than foreign countries and lacks of mature and effective methods for guidance. Although in recent years, the basic resources in the network of our country is

increasing and the Internet resources is growing rapidly, but compared to the United States, Japan, South Korea and other developed countries, the popularization of the Internet in China is still at low level which can't create a good development environment for e-commerce. The government should actively support the infrastructure construction of e-commerce, increase expenditures in the information infrastructure and give the capital investment reasonable guidance and speed up the construction of the high speed information transmission network and Internet construction to improve the Internet speed and reduce the cost of going online. In addition, China should further the introduction and development of the hardware and software of the electronic information and narrow the gap of Internet between China and the developed countries.

#### 3.2. Support the e-commerce construction of the enterprises

At present, the enterprises which have informatization in our country take percentage of only 5% to 10% of the total number and the small and medium-sized enterprises take just 1.9%. These impede the process of developing e-commerce in our country. Regarding the domestic enterprise, the management modes and management ideas are backward which seriously affect the information construction of the enterprises. According to relevant survey data, the government can take direct investment or

the government gives most of the investment to support the e-commerce development projects in enterprises which are important but have investment difficulties. In short, the government and enterprises should adopt feasible measures to intensify the information construction investment and reinforce the enterprise internal information management and information integration source to promote the development of e-commerce in China.

## References

- [1] Panghao. *Econometrica*, Press of University of Southwest Finance and Economics, Beijing, 2007 pp. 63-157.
- [2] Caolei, E-commerce Market Data Monitoring Reports, in: IOEB, Hangzhou, 2011, pp. 54-61.
- [3] Caolei, E-commerce Market Data Monitoring Reports, in: IOEB, Hangzhou, 2010, pp. 23-45.
- [4] CNNIC, 2010 China Network Shopping Market Research Report, Beijing, 2011, pp. 21-30.
- [5] CNNIC, 2009 China Network Shopping Market Research Report, Beijing, 2010, pp. 36-47.

## Vitae

Xu Xinting, was born in 1988 in Jiangsu, China. She is studying for her Master degree now in College of Management, University of Shanghai for Science and Technology.

Wu Fengyu is the professor of College of Management, University of Shanghai for Science and Technology.