Summer 6-30-2018

The Evaluation of E-commerce Efficiency in China using DEA-Tobit model: evidence from Taobao data

Zongwei Li
School of Management, Shanghai University of Science and Engineering, Shanghai,201620, China

Dongqing Luan
School of Management, Shanghai University of Science and Engineering, Shanghai,201620, China

Yanhui Zhang
School of business, East China University of Science and Technology, Shanghai,200237, China

Peng Guo
School of business, East China University of Science and Technology, Shanghai,200237, China

Follow this and additional works at: http://aisel.aisnet.org/whiceb2018

The Evaluation of E-commerce Efficiency in China using DEA-Tobit

model: evidence from Taobao data

Zongwei Li[1*], *Dongqing Luan*[1], *Yanhui Zhang*[2], *Peng Guo*[2]

[1]School of Management, Shanghai University of Science and Engineering, Shanghai,201620, China

[2]School of business, East China University of Science and Technology, Shanghai,200237, China

Abstract: Using the analytical framework of DEA-Tobit, this paper investigates the efficiency of e-commerce in China's provinces based on the cross-section data of 31 provinces in China and the data of e-commerce service providers from Taobao's open platform. The data envelopment analysis (DEA) is used to calculate the technical efficiency and scale efficiency. Furthermore the paper gives an empirical test on the relationship between the scale efficiency and influencing factors by using the censored Tobit model. The results show there are significant regional differences in the efficiency of e-commerce services in provinces of China, and the Real GDP per capita, the seller number on e-commerce platform, the retail sales and wholesale are important reasons for the different efficiency in each province of China. This study provides a domain-specific, integrative approach in evaluating the E-commerce development combining macro data from National Bureau of Statistics of China and micro data from taobao.com.

Keywords: E-commerce , DEA, Tobit, Efficiency Evaluation, Measurement

1. INTRODUCTION

For the past years, the rapid development of e-commerce in China has become the new engine of China's economic growth. Data from China e-Business Research Centre shows Chinese e-commerce transactions amounted to 13.35 trillion Yuan, an increase of 27.1% on year-over-year basis by the mid of 2017. Nevertheless, the e-commerce development of China is also facing some problems, such as development level of e-commerce in China's provinces is unbalanced, and the regional differences are huge[1] [2] .It is valuable to measure the development of e-commerce through the lens of integrative approach with an emphasis on empirical data. How should the regional efficiency of e-commerce development be measured? What factors determine the efficiency of e-commerce services? Measuring and assessing the efficiency of e-commerce efficiency is a difficult process and often involves methodological issues.

The main purpose of this paper is to investigative the efficiency of e-commerce services in China by applying the DEA-Tobit model. In particular, the paper concentrates on investigating: e-commerce efficiency measurement of 31 provinces in Chinese mainland combined with macroeconomic indicators and e-commerce platform indicators; Tobit regression analysis the determinants of e-commerce services efficiency.

The rest of the paper is arranged as follows. Section 2 shows the most relevant results in the literature and briefly outlines the related research using this method. Section 3 presents the DEA-Tobit structure design methodology and explains the data and variables. We provide the demonstration results for DEA-Tobit applied to 31 provinces in China. The paper offers our experimental results of DEA-Tobit model applied in China in section 4. Section 5 concludes. The results may serve as a reference for measuring the regional efficiency of e-commerce and e-commerce development promotion.

2. LITERATURE REVIEW

E-commerce can be regarded as the use of the Internet to conduct business transactions. In the context of

---

E-commerce efficiency, the literatures focus on measurement framework and index system design, measurement methods and index acquisition and so on.

2.1 E-commerce development measurement

The NSF Workshop presented e-commerce development measurement index system firstly, including the E-Readiness, E-Intensity and E-Impact. OECD pointed out the index system of e-commerce development measurement including three categories: the readiness, the application and the impact of e-commerce. APEC put forward nine kinds of index system for the measurement of e-commerce level. But the measuring method involved too many indexes to collect data. The China Internet Research Center e-commerce issued index system with 9 major categories of 32 indicators, including trading, efficiency, security, IT infrastructure, network economy, human capital, development potential, policy Environmental, user satisfaction. Similarly it was not easy to demonstrate the application of the model in reality. The calculation method of e-commerce development is mainly the use of information index model (RITE model) now, including four elements of information to conduct a comprehensive evaluation.

2.2 E-commerce development index system

According to prior research in this stream, most of the index systems include economic and financial resources, trading volume, e-commerce system user, IT infrastructure and logistic & delivery system. We further discuss some of these studies to apply specific indicators that will help us to measure the level of development of e-commerce efficiency.

2.2.1 Economic and Financial Resources

It is generally assumed that Gross National Product (GNP) is an indicator of national wealth or drivers[3]. Wealth has been shown to be the strongest individual factor driving e-commerce sales. Economic and financial resources appear to play a direct role as drivers of e-commerce. Thus, GDP per capita as a key factor is mentioned strong correlation with the measures of e-commerce development[4].

2.2.2 Trading volume

Trading volume is significantly relative to e-commerce transaction, which depends on the nature of the good and offline trading volume[5]. And E-commerce will also have a notable impact on trade in services. Liu and Wang(2009) utilize the transaction volume(retail and wholesale) and proportion of e-commerce as key index to demonstrate the e-commerce level[6]. As a result of the rapid development of e-commerce, a lot of manufacturers have opted to choose the online and offline distribution channels at same time. Although the multi-channel system enables manufacturers to offer their customers products through synchronized channels, it is still a common phenomenon that the online channel is controlled by the traditional retailer or the supplier.

2.2.3 E-commerce system user

The number of Internet users in the world reaches 3.88 billion worldwide by June 2017 and is growing. In terms of e-commerce, the primary users are customers(buyers) or suppliers(sellers).Buyers and sellers could establish a genuine market price in e-commerce system[7]. From a business ecology perspective, e-commerce business ecosystem includes five populations as bellows: Leader Population(e-commerce platform), key species, supportive species, related Population and parasitic species[8]. Besides customers or suppliers, ISV(independent software vendor) as user in e-commerce system may develop some e-commerce relative system based on the e-commerce giant open platform[9].

2.2.4 IT infrastructure

Kraemer et al.(2003) and Lawrence et al.(2010) point out that information infrastructure is an important facilitator of e-commerce development[10] [11].In the diffusion of B2C e-commerce, the information infrastructure

is also considered as an important booster. It is not easy to shop online for people without reasonably quick Internet access. Thus, available IT infrastructure, such as ICT underlying framework, electronic data interchange (EDI) and open systems, is a prerequisite for e-commerce growth[12].

2.2.5 Logistic & delivery system

Logistics capability is critical for e-commerce progress[13].Logistics service quality can exert an influence on online shopping experiences, loyalty, and satisfaction[14].It is critical to support the complexity capabilities and diversity of supply chains for e-commerce collaboration technologies[15].By using SEM model, Y Nurdani et al.(2016) shows that express delivery's service quality has positive influence on both customer satisfaction and trust of online shopping [16].In the Brazilian scenario, de Oliveira et al.(2017) evaluates the applicability of different solutions to urban e-commerce deliveries[17]. Although automatic delivery stations have the potential to reduce delivery problems, it still relies on the transportation system to complete delivery.

2.3 E-commerce development evaluation approach

The index system of network economic measurement established by U.S. government and the Ministry of Commerce can help to understand the classification of information technology measurement. OECD, APEC, Harvard University and the Nordic countries basically adopt qualitative method. The OECD's research on knowledge economy brings the concept of "knowledge-based economy", which not only opens the way for the theory of knowledge economy but also sets up framework on the index system of knowledge economy. Various studies are conducted from different aspects of the network economy, the information, e-commerce, and digital issues in Japan respectively. Based on e-commerce firms, H. Joseph Wen et al.(2003) utilize a DEA model to evaluate the relative efficiency[18].In terms of the empirical research, Liu and Wang(2009) use AHP model to determine the weight of indicators and give the method of this index of e-commerce on city level[19].

3. METHODOLOGY AND DATA

3.1 Research Design

In this paper, we combine the DEA and Tobit regressing model to appraise the e-commerce efficiency of 31 provinces in China. In the first stage, the indicators of e-commerce efficiency are defined. And then five group factors are selected to construct input indicators, respectively Economy level, Trading volume, User, IT infrastructure and Transportation. The relative e-commerce efficiency is then assessed by DEA model. Finally, we use Tobit regression model to examine the effect of inputs on the e-commerce efficiency score calculated.
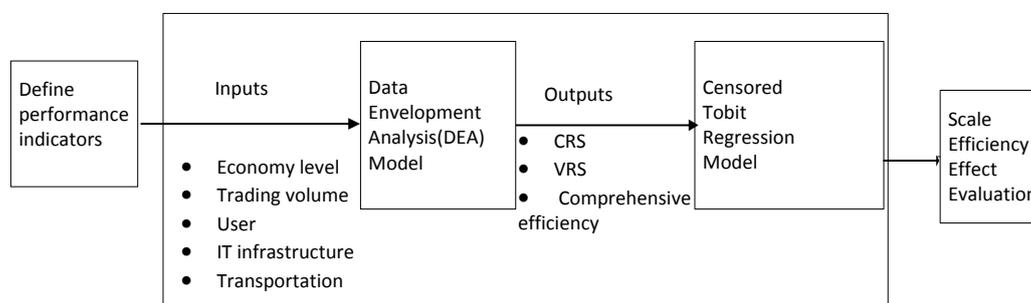

Figure 1. Diagram of methodological steps

3.2 Methodology

By using data from National Bureau of Statistics of China online database and Taobao's open platform, We apply DEA-Tobit analysis framework to evaluate the e-commerce development relative efficiency in China. The

relative efficiencies can been estimated by using DEA model ,which is a non-parametric approach combining various input/output. In the particular case of our research, there are 6 inputs and 2 outputs for 31 DMUs, which represent provinces in China. The model is expressed mathematically as follows:

$$\begin{cases} \underset{\theta,\lambda}{M\,in}\,\theta \\ -y_i + Y\lambda \geq 0 \\ \theta x_i - X\lambda \geq 0 \\ N_1\lambda \leq 1 \\ \lambda \geq 0 \end{cases} \tag{1}$$

In order to further understand the impact factors of input-output efficiency, this paper carries on the Tobit regression analysis to the comprehensive efficiency. However, the efficiency index obtained by the DEA model is between 0 and 1, and if the ordinary least squares (OLS) method is adopted, the estimated values of the parameters will be biased and inconsistent. In this paper, the censored Tobit regression model limited dependent variable is used in the regression analysis, which was proposed by Tobin[20]:

$$Z_i = \begin{cases} \beta^T X_i + \varepsilon_i, \beta^T X^m + \varepsilon_i > 0 \\ 0, \beta^T X^m + \varepsilon_i \leq 0 \end{cases} \tag{2}$$

3.3 Data and variables

Statistical data for 31 provinces in China are extracted from the National Bureau of Statistics of China online database and Taobao's open platform. Taobao is the world's largest C2C e-commerce platform based in China, started in May 2003.According to Alibaba, its online trading volume exceeded the milestone of ¥3 trillion in 2016, accounting for 10 percent of the overall retail volume in China. We selected two output variables: Number of Sellers and number of ISV (Independent Software Vendors) on Taobao's open platform and six input variables: Real GDP per capita, Total Retail Sales of Consumer Goods, Total Sales Value of Wholesale and Retail Trades, Number of Internet Users, Number of Port of Internet, and Freight Traffic.

According to Taobao's open platform, the number of sellers represents the number of active e-commerce shops providing transaction on taobao.com. The number of sellers is considered one of the most direct indicators of the e-commerce systems. Data from Alibaba group show that there are more than 8.5 million active sellers distributed in different provinces in mainland China in Taobao platform.

Number of ISV measures the number of Independent Software Vendors providing third-party services on Taobao's open platform, such as e-shop decoration services, marketing services, and data services etc. Number of ISV is treated as output variable because it represents E-commerce services capabilities. At present, the service market has gathered tens of thousands of providers to supply services for Taobao and Tmall's sellers, the transaction size has reached billions in 2016.

As input variables, Real GDP per capita, Total Retail Sales of Consumer Goods, Total Sales Value of Wholesale and Retail Trades represent Chinese provincial economic level and trading volume. The degree of development of e-commerce is positively correlated with the level of regional economic development.

Number of Internet Users and Number of Port of Internet are included in the model as input variables. Number of Internet Users represents the degree of access to e-commerce applications, and Number of Port of Internet is handled as an important indicator of information infrastructure. Both of the variables represent information technology facilities and Applications.

In terms of Logistics & delivery capability, Freight Traffic is also included in the model as input variable as

transportation infrastructure. The rapid growth of e-commerce in the United States benefits from the development of the logistics infrastructure and the rapid delivery system. Similar system is boosting in China and supports application of e-commerce, for example cainiao use big data and intelligence to promote the smart logistics upgrade. The variable description see table1.

Table 1. Variable description

New Variable	Variable
lnGDP	the Real GDP per capita
lnretail	Total Retail Sales of Consumer Goods
lnwholesale	Total Sales Value of Wholesale and Retail Trades
lnnetuser	Number of Internet Users
Lnfreight	Freight Traffic
Lnnetport	Number of Port of Internet
lnISV	Number of ISV
lnseller	Number of Sellers

4. RESULT AND DISCUSSION

We first present the descriptive statistics for the input and output variables and then discuss the results of the DEA model (see table 2).

Table 2. Descriptive statistics for input and output variables (the natural logarithm adopted)

statistics	lnGDP	lnretail	lnwholesale	lnnetuser	lnfreight	lnnetport	lnISV	lnseller
Max	11.51	9.155	11.04	8.853	12.89	8.109	10.31	14.44
Min	10.04	4.539	5.195	4.745	7.523	3.786	3.951	7.997
Average	10.68	7.606	9.056	7.311	11.40	6.722	7.438	11.62
SD	0.402	1.134	1.275	0.873	1.090	0.974	1.502	1.526
Coeff.of Variation	0.0377	0.149	0.141	0.119	0.0956	0.145	0.202	0.131

On the whole, the coefficient of variation is between 3.77% and 20.2%, and the coefficient of variation of lnISV is the highest(20.2%), thus the sample for the variables is homogeneous.DEA Model result shows as figure 2. The model CRS is obtained with average performance score of 0.961, VRS score of 0.975,then the scale efficiency is calculated as the ratio between efficiency scores in CRS and VRS models. Ten provinces out of 31 are scale efficient: Beijing, Tianjin, Ningxia, Tibet, Jiangxi, Fujian, Zhejiang, Shanghai, Guangdong, and Guangxi. Interestingly Ningxia and Tibet are usually regarded as underdeveloped provinces in western China, Therefore, the development level of E-commerce is also relatively backward. Seven provinces out of 31 are decreasing return to scale: Neimenggu, Xinjiang, Jilin, Qinghai, Shandong, Liaoning, and Jiangsu. This means that an increase in input will result in less output increase. Therefore, for these provinces, to enhance the service efficiency of e-commerce is critical, rather than blindly expand the scale of e-commerce. The remaining 14 provinces are increasing return to scale. In conclusion, these provinces need to increase investment in e-commerce and improve service efficiency.
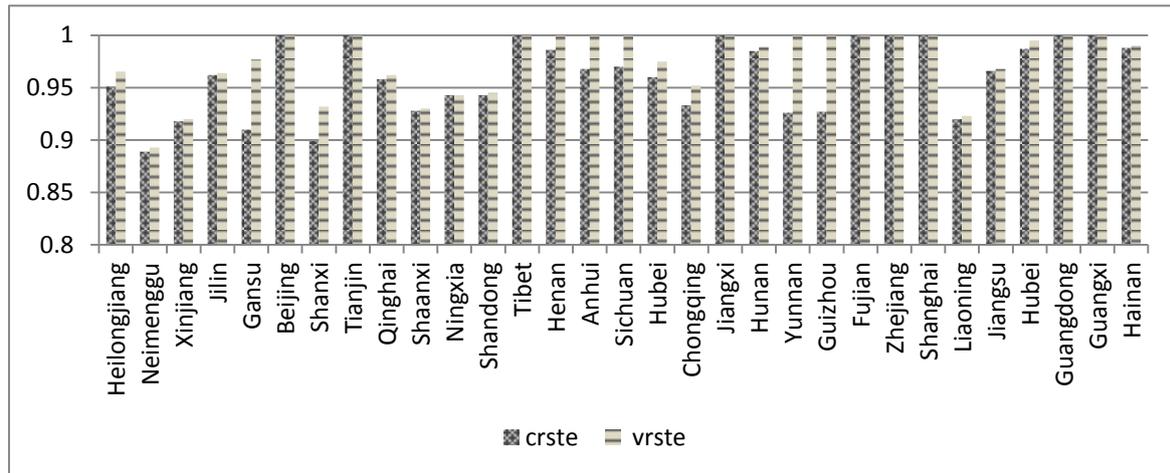
Figure 2. Efficiency scores for DEA model

Table 3 reports the Tobit regression results for scale efficiency. According to Table 3, four variables besides constant—lnseller, lnGDP, lnretail and lnwholesale—are statistically significant at the 5% level, while others show non-significance, both at the 5% and 1% levels.（1）The Real GDP per capita and Number of Sellers have a significant impact on the scale efficiency, and the influence coefficient is positive, indicating that the greater the Real GDP per capita and the greater the number of sellers can lead to the greater the scale efficiency value. This shows that increasing province's economic level and improving the number of e-commerce sellers will promote the efficiency of e-commerce development and enhance the competitiveness of e-commerce services.（2）There is a negative correlation Total Retail Sales of Consumer Goods ,Total Sales Value of Wholesale and Retail Trades with scale efficiency, which indicates that the development of e-commerce would be damaged if a province overdevelop the offline trade, because channel conflict may occur when the supplier enters the online channel.（3）The Number of Internet Users and Number of Port of Interne have correlation with scale efficiency, but they are not significant, which indicate that the network infrastructure is no longer a bottleneck for the development of electronic commerce; there is a negative correlation of freight volume and comprehensive efficiency, but not significantly either.

Table 3. Tobit regression on efficiency related to inputs

Variable	Coef.	Std. Err.	t	P>t	[95%Conf.	Interval]
lnseller	0.013874	0.005614	2.47	0.021**	0.002287	0.025461
lnGDP	0.068679	0.010328	6.65	0**	0.047364	0.089994
lnretail	-0.02463	0.008206	-3	0.006**	-0.04156	-0.00769
lnwholesale	-0.01847	0.005422	-3.41	0.002**	-0.02966	-0.00728
lnnetuser	0.020161	0.019262	1.05	0.306	-0.01959	0.059917
Lnfreight	-0.00131	0.004768	-0.27	0.787	-0.01115	0.008536
Lnnetport	0.000974	0.015165	0.06	0.949	-0.03033	0.032273
constant	0.307809	0.108742	2.83	0.009	0.083377	0.53224

** indicates 5% level of significance

Thus, based on the Tobit regression model, the DEA efficiency score of e-commerce development could be

improved through two main variables: lnseller(Number of Sellers), lnGDP (the Real GDP per capita) and declined by offline trade volume. Hence, economic development level , e-commerce participation have a direct effect on increasing the efficiency score of e-commerce for regional development.

5. CONCLUSIONS

Due to the data from Taobao Open Platform and the National Bureau of Statistics of China, this paper presents a DEA model consisting of 6 inputs and 2 outputs to measure the efficiency of e-commerce services across provinces in China.

The contributions of this research are that we have proposed a "broader picture" of the efficiency of e-commerce initiatives by combining macroeconomic statistics with the micro-data from e-commerce platforms. The results of our study found some interesting phenomena, the efficiency level of e-commerce development in economically developing provinces is not necessarily lower than developed provinces, such as Tibet and Ningxia. In addition, it empirically demonstrated using censored Tobit model that some factors which affect the scale efficiency, such as the number of sellers on the platform and the level of economic development in the region.

Although this study provides meaningful implications for efficiency of e-commerce evaluation, it has some limitations and thus has further research issues. On one hand, if the representative indexes should be discussed and the long time sequence selection should be chosen in the future research, the results may be more accurate. Although the results show that the factors of economic development level, e-commerce participation have a direct effect on increasing the regional e-commerce efficiency e-commerce,. Recognizing that other factors may also play an important role, such as economic, political environment and tax policies. On the other hand, dynamic methods method can be used, for example Malmquist method. The research results of China's economically developed and developing provinces need to find more ways to allocate resources for their e-commerce development.

ACKNOWLEDGEMENT

REFERENCES

[1] Yu Jinguo，Wang Li - hua&Li Na. (2010).E- Shops Spatial Distribution Rule:A Case Study of Taobao Net, Economic Geography, 30,1248-1253(in Chinese).

[2] Tang Yinghan. (2015).A Spatial Difference Study on China's E-Commerce Development Level, Economic Geography, 35, 9-14 (in Chinese).

[3] Dewan, S., & Kraemer, K. L. (2000). Information technology and productivity: evidence from country-level data. Management Science, 46(4), 548-562.

[4] Oxley, J. E., & Yeung, B. (2001). E-commerce readiness: Institutional environment and international competitiveness. Journal of International Business Studies, 32(4), 705-723.

[5] Kaplan, S., & Sawhney, M. (2000). E-hubs: the new B2B marketplaces. Harvard business review, 78(3), 97-106.

[6] Liu, Y., & Wang, W. (2009). The reconstruction and empirical analysis of regional e-commerce development index. Science of Science and Management of S. & T, 30(7), 144-147.

[7] Cai, G. G., Zhang, Z. G., & Zhang, M. (2009). Game theoretical perspectives on dual-channel supply chain competition with price discounts and pricing schemes. International Journal of Production Economics, 117(1), 80-96.

[8] Huang, L., Hu, G., & Lu, X. (2009). E-business ecosystem and its evolutionary path: the case of the Alibaba group in china. Pacific Asia Journal of the Association for Information Systems, 1(4).

[9] Chen, J., Wang, C., Liu, F., & Wang, Y. (2017). Research and Implementation of a Software Online Testing Platform Model Based on Cloud Computing. In Advanced Cloud and Big Data (CBD), 2017 Fifth Internataional Conference, 87-93.

[10] Gibbs, J., Kraemer, K. L., & Dedrick, J. (2003). Environment and policy factors shaping global e-commerce diffusion: A cross-country comparison. The information society, 19(1), 5-18.

[11] Lawrence, J. E., & Tar, U. A. (2010). Barriers to e-commerce in developing countries. Information, society and justice journal, 3(1), 23-35.

[12] Iacovou, C. L., Benbasat, I., & Dexter, A. S. (1995). Electronic data interchange and small organizations: Adoption and impact of technology. MIS quarterly, 465-485.

[13] Joong-Kun Cho, J., Ozment, J., & Sink, H. (2008). Logistics capability, logistics outsourcing and firm performance in an e-commerce market. International journal of physical distribution & logistics management, 38(5), 336-359.

[14] Ramanathan, R. (2010). The moderating roles of risk and efficiency on the relationship between logistics performance and customer loyalty in e-commerce. Transportation Research Part E: Logistics and Transportation Review, 46(6), 950-962.

[15] Tan, P. S., Lee, S. S. G., & Goh, A. E. S. (2012). Multi-criteria decision techniques for context-aware B2B collaboration in supply chains. Decision Support Systems, 52(4), 779-789.

[16] Nurdani, Y., & Sandhyaduhita, P. I. (2016). Impact of express delivery service quality towards repurchase intention by B2C and C2C: A case of Indonesia. In Advanced Computer Science and Information Systems (ICACSIS), 2016 International Conference , 221-227.

[17] de Oliveira, L. K., Morganti, E., Dablanc, L., & de Oliveira, R. L. M. (2017). Analysis of the potential demand of automated delivery stations for e-commerce deliveries in Belo Horizonte, Brazil. Research in Transportation Economics.

[18] Wen, H. J., Lim, B., & Lisa Huang, H. (2003). Measuring e-commerce efficiency: a data envelopment analysis (DEA) approach. Industrial Management & Data Systems, 103(9), 703-710.

[19] Liu Yue, Wang Wenqing. (2009).The Reconstruction and Empirical Analysis of Regional E－Commerce Development Index. Science of Science and Management of S.& T, 30.7:144-147(in Chinese).

[20] Tobin, J. (1958). Estimation of relationships for limited dependent variables. Econometrica: journal of the Econometric Society, 24-36.