2014

ISSN : 0974 - 7435 Volume 10 Issue 21

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 10(21), 2014 [13453-13458]

Study on E-commerce development scale measuring methods based on sample survey

Dachuan Zheng[1], Lei Huang[2], Zhongyan Lin[3]*

[1,2]Straits institute, Minjiang University, Fuzhou, (CHINA)

[3]Minjiang University, Fuzhou, (CHINA)

E-mail: lzy@mju.edu.cn, 80885338@qq.com; ginahuang615@gmail.com

ABSTRACT

In recent years, China's e-commerce trading volume has witnessed fast growth. In 2013, the transaction turnover on China's online retail market was RMB 185 million and surpassed U.S. to become the world's biggest e-commerce power. How to measure the development scale of e-commerce has become an important research field in the current industry. While analyzing China's e-commerce statistics status quo, the paper explored the measuring methods for the e-commerce development scale based on a sample survey, adopted some statistical modeling methods such as simple estimation method, ratio estimation method, index method and regression model method, conducted repeated tests and studies on the measuring methods of e-commerce scale of the entire society, and provided thoughts on designing a reasonable measuring methods system for e-commerce development scale.

KEYWORDS

E-commerce; Sample survey; Statistical model.

FOREWORD

Electronic Commerce is generally understood as commercial activities conducted via computer network. Hereby "commercial activities" include trade, service, economic matters and administrative affairs, etc. As a kind of new market operation method, e-commerce refers to economic activity methods guided by informationized market, networking and marketized information. The appearance of e-commerce is helpful to promoting the economic integration of different countries and regions and enhancing resource allocation efficiency. In particular, enterprises in the industrial manufacturing industry and modern service industry can improve their operation efficiency and benefit level, reduce manpower and time consumption and enhance service efficiency through e-commerce. According to the statistical report on China's internet development status, in 2013, the proportions of online purchase and online sales by Chinese enterprises were respectively 23.5% and 26.8%. The number of enterprises which carried out marketing activities on the internet was 20.9%. There's a great gap among different industries in terms of e-commerce application popularization. The e-commerce application is more common in the manufacturing and wholesale & retail industry. However, different from other economic activities, e-commerce, in dynamic development, shows frequent changes and appears in different industries and different types of operations. How to conduct accurate and feasible statistics and quantitative assessment on e-commerce has become one of the important research fields for e-commerce development.

ANALYSIS OF CHINA'S E-COMMERCE STATISTICS STATUS

According to the current framework of the National Bureau of Statistics of China, e-commerce enterprises are divided into three kinds of subjects:

First, e-commerce enterprises above designated size have been included in conventional statistical statements and submit data to the statistical department on a monthly basis;

Second, online sellers below designated size and registered in the industry and commerce bureau are listed in the directory of the statistical department. Sample survey is conducted by quarter to retrieve data and predict the overall situation;

Third, subjects are online sellers without business license and are not observed yet. Such kind of subjects is extensive with frequent changes and hard to be grasped by the statistical department. Collection of data on such subjects requires the big data of e-commerce enterprises.

Compared with the prosperous development of the e-commerce development, China's e-commerce statistics work is apparently backward. So far, an explicit and standard e-commerce statistical method and index system has not yet been formed. The traditional e-commerce statistical survey is quite limited. In the meantime, e-commerce statistical work falls short of macro-policy instruction and national unified supervision. The e-commerce surveys made by some non-official organizations, folk organizations or some professional websites are not statutory as official statistics, and their contents and scope are extremely confined, thus leading to systematic bias in the outcome.

MEASURING METHODS OF E-COMMERCE DEVELOPMENT SCALE BASED ON SAMPLE SURVEY

There are many forms of e-commerce activities and complicated operation categories. What's more, the number of participants is huge and a great proportion of participants haven't been registered through official channels. It makes it impossible for the e-commerce statistical departments to obtain complete information just as other ordinary business activities. As a result, during the e-commerce statistical work, sample statistical survey has become the only feasible data collection method at present. Unlike the common measuring methods for macro-economic development, the measuring for e-commerce development takes on its unique characteristics.

Usually certain overall parameters are concerned in the actual sample survey such as gross and average value. Generally speaking, two ways of deduction for these goals are as follows: first, deduction based on traditional sample design which randomly selects the sample probability of samples from the limited total as the deduction basis according to certain sample design; second, deduction based on models, that is, suppose limited total is a random realization of a certain super-population and the estimator is worked out on account of the super-population.

The differences between the two go like this: the sample design method is random deduction method without parameters and doesn't have any distribution assumptions for the total goals; the model deduction method supposes that the total comes from a certain distribution and is a kind of parameter deduction method. The model parameters can be estimated through samples.

At present, four methods can be applied in the e-commerce development measuring field, including simple estimation method, ratio estimation method, index method and regression model method.

Simple Estimation Method

Simple estimation method is the estimator obtained after simple expansion, that is, the commonly-known distance estimation.

Mean value estimation is a kind of typical simple estimation method. In mean value estimation, the certain index value of N units which constitute the total is $y_1,y_2,?,y_N$. In order to get the total gross $\sum_1^N y$, the random sample is $x_1,x_2,?,x_n$. The sample mean value $\overline{x} = \frac{1}{n}\sum_1^n x$ is regarded as the unbiased estimation of $\overline{y}$ to acquire the mean value of the total y. therefore, the total gross is $\sum_1^N y = N\overline{y}$.

In e-commerce statistics, we can use simple estimation method to estimate and calculate the e-commerce development scale of different regions and industries via the data collected on the "platform". Further, the average level of e-commerce sales revenue of surveyed enterprises can be obtained to serve as the average e-commerce sales level of the entire society. Hence, the gross e-commerce sales of the entire society=average e-commerce sales level of the entire society × the number of industrial and commercial enterprises of the entire society.

Ratio Estimation Method

Suppose the survey variable y_i is the to-be-described index value and the variable x_i is the auxiliary variable of the index value. Both of them are random variables. A stable ratio relationship exists between y_i and x_i, that is, $y_i = R \times x_i$. R is the ratio coefficient. Further, $R = y_i / x_i$. The total ratio formula is $R = \sum_1^N y_i \Big/ \sum_1^N x_i$. The sample ratio formula is

$$r = \sum_1^n y_i \Big/ \sum_1^n x_i .$$

In the ratio estimation method, in order to obtain the gross $\sum_1^N y$ of a certain index of the total, the sample ratio r is deemed as the unbiased estimation of the total ratio relationship and the ratio value of the total index is R. therefore, the total index value is $\sum_1^N y = R\sum_1^N x$.

The ratio estimation method is adopted to infer the e-commerce transaction value of the entire society based on the sample survey data. In other words, the proportion of e-commerce transaction value in the sales revenue r should be calculated first from the survey sample, which will further be treated as the unbiased estimation of R, proportion of the e-commerce transaction value of the entire society in the sales revenue. Afterwards, in accordance with the ratio relationship R and the total sales revenue of the entire society's commodities and services, the e-commerce transaction value of the entire society will be further inferred. The detailed formula is as follows:

$$\hat{R} = r = \frac{\text{E-commerce transaction value of the selected sample}}{\text{Sales revenue of commodities and services of the selected sample}}$$

$$\text{E}-\text{commerce sales revenue of the entire society} = S \times \hat{R}$$

(S means sales revenue of commodities and services of the entire society)-

The method of ratio estimation method can help reckon the e-commerce transaction value of the entire society from the perspective of industrial summary. The "platform" has collected e-commerce sales data of different industries. Therefore, we can calculate the corresponding ratio index r_q according to the transaction data of different industries. There into, q (q=1,2,…, p), which indicates all kinds of industries of the entire society. Then, the e-commerce transaction value of all industries of the entire society can be further calculated based on each ratio relationship and commodities and services sales

revenue of separated industry of the entire society. In the end, the total sales revenue of the entire society can be acquired after summary. The detailed formula is as follows:

$$\hat{R}_q = r_q = \frac{\text{Ecommerce transaction value of industry q in the selected sample}}{\text{Commodities and services sales revenue of industry q in the selected sample}}$$

$$\text{The e}-\text{commerce sales revenue of industry q of the entire society} = T \times \hat{R}_q$$

(T means total ties and services sales revenue of industry q of the entire society)

$$\text{E}-\text{commerce sales revenue of the entire society} = \sum T \times \hat{R}_q$$

The ratio estimation method can assist in the deduction of e-commerce transaction value of the entire society from the perspective of dynamic development. The "platform" has gathered e-commerce sales data of different years. Suppose the e-commerce transaction value of the selected sample is in line with the growth speed of e-commerce transaction value of the entire society. We can use historical data and ratio estimation method to work out the current total e-commerce transaction value of the entire society. The detailed formula is as follows:

$$\hat{R}_q = r_q = \frac{\text{E}-\text{commerce transaction value}_t - \text{E}-\text{commerce transaction value}_{t-1}}{\text{E}-\text{commerce transaction value}_{t-1}}$$

$$\text{The e}-\text{commerce sales revenue of the entire society} = (\hat{R}_q + 1) \times E$$

(E means e-commerce transaction value of the entire society$_t$)

Index Method

The study on e-commerce level measuring originated from the 1990s. Domestic and international studies on e-commerce level measuring mainly focus on the information measuring. Some scholars equated the information level to the e-commerce level. Information index is usually adopted for e-commerce level measuring. CII proposed the e-commerce general index and indicator system (2001) and measured the total e-commerce index of China. The e-commerce index of CII is relatively suitable for China's current statistical system and accepted and approved by many Chinese scholars. Studies on e-commerce key statistical indexes in recent years have emphasized qualitative and quantitative analysis based on CII indexes.

CII e-commerce indexes provide referable framework for the e-commerce development index system. On this basis, index systems which are suitable for each region can be confirmed according to their own characteristics. After setting the e-commerce development index and indicator system, the comprehensive scoring analysis method can be applied to the total e-commerce indexes.

The analytic hierarchy process method is an important method of determining the weight W_{ij} of each index. The comparative matrix can be constructed by hierarchy in accordance with the established hierarchical structure model, questionnaire survey and expert scoring. Afterwards, it will be ranked by the influence degree for the superior hierarchical index. Nine-level scaling is adopted and the degree shows a progressive increase from 1 to 9. At the same time, the consistency ratio of the matrix can be calculated to test the matrix consistence. Based on the comparative matrix, we can work out the integrated weights of each hierarchical index for the development index.

Last but not least, the total score of the e-commerce development index can be obtained according to the formula below:

$$ECI = \sum_{1}^{n} (\sum_{1}^{m} P_{ij} W_{ij}) * W_i$$

Thereinto, ECI represents the score of the e-commerce development index; n is the first-level index number of the e-commerce index system; m refers to the number of second-level indexes of the index system; P_{ij} is the value of the j standardized second-level index of the i first-level index; W_{ij} is the weight of the j second-level index of the i first-level index.

Regression Model Method

In economics and social science field, regression model analysis is an extensively applied quantitative analysis method. The method describes the causal relationship between economic matters in the form of regression model, examines

the influence rules between different variables and help people precisely grasp the influence degree of other independent variables over dependent variables, thus predicting the future trend of the dependent variables in a scientific way.

The steps of the regression analysis are as follows:

The dependent variable (y) and independent variable (x) of the model should be determined. The regression analysis is used to analyze how dependent variables change as independent variables change. Hence, the first step of the analysis is to confirm the contents of independent variables and dependent variables. Regression model is a model which describes the relationship between y and x. while x conditions are given, the expected mean value of y is forecasted through the model.

The model form shall be considered. According to the function fitting, the mathematical relationship between dependent variables and independent variables can be decided via observation of scatter diagram to choose the right mathematical function for describing the regression relationship. If a linear relationship exists between independent variables and dependent variables, then the linear regression analysis will be adopted. If it's non-linear relationship, the scatter diagram shall be employed to judge what kind of non-linear relationship it is so as to set up the appropriate non-linear regression model.

The regression model shall be established for parameter estimation. Pursuant to the collected sample data and the determined regression model above, and under certain statistical principles and set confidence coefficient, parameters of the model shall be estimated to get the regression equation. The most common estimation methods are least square method and maximum likelihood estimation.

Parameter test can be conducted on the regression equation. The regression equation is concluded from the sample data. Whether its results truly reflect the statistical relationship of the total of a matter and whether the regression equation can be used for forecast shall be further tested in a statistical sense.

The estimated regression equation can be utilized for forecast. An important purpose of the econometric model is to make forecast with the estimated regression model and conduct advanced judgment of the future development trend. The forecast is divided into point forecast and interval forecast. Point forecast is to import a group of specific value of independent variables to the equation and further calculate the point forecast value of dependent variables. The interval forecast not only obtains the estimation value of the dependent variables but also provides the general scope of the estimated value.

Conclusion of the Four E-commerce Development Measuring Methods

The simple estimation method and ratio estimation method are simple and easy to implement. It doesn't require so much for the sample quantity and easy to carry out in actual work. However, both of them assume there's a constant quantitative relation, consistent mean value or fixed ratio relationship between the sample and total data. Nonetheless, as a matter of fact, it's hard to realize, which affects the estimation accuracy of the two methods.

The index method is the mainstream method for e-commerce development measuring at present. Owing to the complete index system and scientific quantitative analysis, the method can precisely predict the e-commerce development level to some extent. However, the precondition for doing so is to establish objective expert scoring. In the actual practices, the outcome of expert scoring is usually affected by individual factors of experts and therefore, it's greatly subjective and has an impact on the accuracy of the development indexes.

The regression model method, as a scientific quantitative analysis method, has been widely applied to varied fields especially economic field, thus proving its reasonability and operability. It can conduct detailed analysis of economic issues from several perspectives and provide reasonable forecast. Yet, the method is limited as well. The regression model is quite demanding for the sample data. Too little sample data or the singular value will affect the estimation and forecast of the model. At the same time, the data distribution is also extremely strict in the model. If the practical data doesn't conform to the distribution assumption, the obtained estimation results will deviate and further affect the forecast ability.

The four methods mentioned above have both merits and demerit so they can't be just randomly ranked. In actual operations, the four methods can be simultaneously adopted to compare different results and thus acquire more scientific forecast outcome.

The Experimental Simulation of the Software Operating Environment

In order to better conduct the study on e-commerce statistical measuring methods, the researcher conducted an experimental simulation of the software operating environment for the previous model. The overall design thought of the program goes like this: analyzing the data in each sheet of the document by reading a.xls document which meets with the data analysis requirements; working out the corresponding curve graph after reading the value in fixed ranks; storing the calculated value of the model in the document of "value analysis. xls".

The program requires fixed titles for each sheet in the read.xls document and data in each sheet must be listed in the designated ranks. Otherwise, the program cannot correctly work out the expected outcome. Excel.exe is an executable program generated by matlab2012b. It's easy to use. Users just need to directly click the Choose Document button and after the corresponding xls document is selected, the program can automatically display the extracted data from the sheet and the finally drawn graphs.

CONCLUSION

To sum up, the study on the e-commerce statistical measuring methods must be laid on complete data samples and data warehouses. However, at present, e-commerce involves data of the entire society and industry as well as great coverage, which makes e-commerce statistics a gap field in the present statistics. The researcher's pre-stage study and industrial survey shows that the sample statistical survey can be regarded as the e-commerce statistical data collection method for in-depth studies; in the meantime, it's suggested that the four statistical methods including simple estimation method, ratio estimation method, index method and regression model method can be applied in experimental environment to conduct repeated tests and studies on the e-commerce scale measuring methods of the entire society. Besides, the e-commerce development features and experience shall be taken into account to design a reasonable e-commerce scale measuring method system for the entire society, timely grasp the dynamic trends of the regional economy and e-commerce development, and provide reference for government macro-control and enterprises' operation decisions.

BIBLIOGRAPHY

[1] D.Gefen, D.W.Straub; The relative importance of perceived ease of use in IS adoption: a study of e-commerce adoption[J], Journal of the Association for Information Systems, **1(1)**, 8 **(2000)**.

[2] Y.Zhang, Y.Fang, K.K.Wei, et al.; Repurchase intention in B2C e-commerce—A relationship quality perspective[J], Information & Management, **48(6)**, 192-200 **(2011)**.

[3] S.H.Chiu; Interactive Recommend System Regarding Time Factor-Women and Children Merchandise E-Commerce Platform[J], **(2013)**.

[4] ZhongyanLin; On Cobweb model of the Information Service Industry from the Perspective of Industrial Chain, Journal of Henan University of Science & Technology, **1,** 76-79 **(2009)**.

[5] ZhongyanLin; Improving Path for ICT Industry of Fujian in the Post-ECFA Era.Journal of Fuzhou University, **5**, 32-36 **(2012)**.

[6] China Internet Network Information Center; The 33rd Statistical Report on Internet Development in China, **(January 2014)**.

[7] Jiang Shu; Government Statistics in the Era of Big Data. The World of Survey and Research. **(April 2014).**

[8] Qiu Xiaohua, Yang Jingying; Research Report on E-commerce Statistics and Survey Solutions, China Information Almanac, **(2006).**