Study on Influence Factors of Freight Ton-kilometers—Based on Data Analysis of Quantile Regression

Keming Wang
School of Management,Sun Yat-sen University Guangzhou, China
E-mail: wangkm@mail.sysu.edu.cn

Yezi Mo, Xueying Chen Yilin Wei and Gang Wang
School of business,Guangdong University Of Foreign Study, Guangzhou city, China

***Abstract*—Freight ton-kilometers is the product of shipment quantity multiplied by shipment distance, which, compared with the shipment quantity, can better reflect the development as well as the scale of logistics in a region. The paper is a study of the interrelations among the six influence factors of 21 cities in Guangdong province, namely, freight ton-kilometers and cities GDP, fiscal budget, gross retail sales of social consumption, cities household savings, cities total import volume and cities total export volume. It can be seen that the degree of influence displays differently in accordance to both high-level and low-level quantile models via Quantile Regression Analysis, otherwise via the conventional OLS (Ordinary Least Square) it can't achieve the same. The study of freight ton-kilometers related to the six influence factors mentioned above helps to acquire not only the knowledge of the structure and situation of the cargo market, but that of regional economic development in order to make a strategic decision and policy for the economic development in a better fashion.**

***Index Terms*—Freight Ton-kilometers, Influencing Factors, Impact Degree, Quantile Regression, Guangdong Province**

I. INTRODUCTION

In recent years, China's logistics industry has developed rapidly, with diverse logistics enterprises progressively springing up. However, compared with the advanced logistics technology and excellent logistics service of foreign developed countries, the service and efficiency of that in China keep rather defective. In terms of the proportion of logistics costs to the country's GDP, China's near 17% of GDP is higher than the world average of about 5%. Also, the logistics cost in China seems quite high compared with the 8% of the US logistics cost and 11% of Japan, which restricts the economic development of China (Nie Zhengyan, 2015).

This paper targets 21 cities of Guangdong, a major economy in China. Considering freight ton-kilometers as a dependent to study the influence factors of logistics development, the author hopes to further explore how the six factors (Fan Yuejiao, 2014) influence freight ton-kilometers in different regions with kinds of logistics scale so that the local decision makers can better formulate economic development strategy and investment budget of logistics construction(Cui Shudan, 2008.).

II. METHODS

The domestic researchers generally take the method of OLS (ordinary least square) to analyze freight ton-kilometers and its influence factors (Liang Xin, Wang Qianfeng, 2010), but OLS can merely explain the influence the independent variable exerts on the dependent variables in conditional distribution(in other words, in normal distribution) rather than comprehensive influence on varied dependent variables. To study the extreme case in skewed distribution, this paper adopts the Quantile Regression Model put forward by Roger Koenker and Gilbert Bassett in 1978 to analyze the freight ton-kilometers together with its six influence factors. Contrast to the traditional mean-oriented analysis, the quantile regression is able to explain the different marginal effects with different quantiles, hence to avoid the limitation of mean marginal effects.

Generally speaking, the application of quantile regression method abroad is more frequent and earlier than that of China (Frederick, Wiseman, Sangit, Chatterjee, 2010). However, in recent years an increasing number of domestic scholars have been tending to apply the method in various fields including finance, economics, medicine, etc. (Liang Lizhen, 2008), and now the method begins to be introduced into the logistics industry to study freight ton-kilometers.

III. EMPIRICAL RESEARCH

A. Sample Data

The data are collected from the official statistics report *Guangdong Statistical Almanac* in Guangdong statistical information network, with the incomplete or missing data of city and year excluded. The paper finally selects the data of 21 prefecture-level cities in Guangdong Province during 2001-2015 as the empirical data. 7 categories are classified, and they separately are freight ton-kilometers, GDP, fiscal budget, gross retail sales of

social consumption, household savings, total import volume and total export volume, all in city level. Finally the author has concluded 2205 sums of valid data.

Table 1 is the descriptive statistics of the sample data in this paper, with the mean value, standard deviation, minimum and maximum values respectively calculated.

TABLE Ⅰ
DESCRIPTIVE STATISTICS OF THE SAMPLE DATA

Variables in city level（Unit:Billion）	Mean Value	Standard Deviation	Minimum Value	Maximum Value
Freight ton-kilometers	29.564	91.138	0.298	839.658
GDP	191.817	297.818	9.728	1810.041
Fiscal budget	20.324	34.956	1.095	352.167
Gross retail sales of social consumption	71.194	108.103	3.395	798.796
Household savings	144.664	221.438	6.122	1360.238
Total import volume	12.898	30.919	0.046	231.773
Total export volume	18.104	42.757	0.093	305.702

TABLE Ⅱ
NUMERICAL STATISTIC FOR EACH QUANTILE

Main quantiles of freight ton-kilometers in city level	0.25	0.5	0.75
Quantile value(Billion)	3.052	5.999	16.271

In Table 2, given that freight ton-kilometers acts as a dependent variable, the author marks three quantiles level as 0.25, 0.5, 0.75 for the 315 data of freight ton-kilometers in city level for better observation.

B. Empirical Analysis

In the empirical analysis, freight ton-kilometers acts as the dependent variable (Y), with independent variables: GDP (X1), Fiscal budget (X2), Gross retail sales of social consumption (X3), Household savings (X4), Total import volume(X5) and Total export volume (X6) for quantile regression analysis. In order to study the difference of estimation coefficients of different models, the three quantiles level 0.25, 0.5, 0.75 are set.

First, adaptively sample the data 20 times by simultaneous quantile regression.

As shown in Table 3, in the vertical column of quantile level as 0.25, only X2 appears 1% P value, indicating that only fiscal budget (X2) has an obvious impact on freight ton-kilometers (Y) in the low-level quantile model (Y). Combined with the reality, it indicates that in a region with relatively small scaled freight ton-kilometers, the effect of fiscal expenditure on logistics industry is very significant and thus appropriately increasing the fiscal budget on logistics construction in this area can greatly improve freight ton-kilometers, further meet the logisticsss demands and accelerate the economic growth.

In the vertical column of quantile level as 0.5, total import volume (X5) shows 5% P value, which means that total import volume has a certain effect on freight ton-kilometers in the areas with moderate scaled freight ton-kilometers. Besides, fiscal budget (X2) and total export volume (X6) have 10% P value, indicating fairly small impacts.

Several independent variables affect freight ton-kilometers in the volumn of quantile level as 0.75, seen from table 3. In the areas with large-scaled freight ton-kilometers, the factors with 5% P value including household deposits (X4), total imports volume (X5) and total export volume (X6) exert a certain extent effects.

TABLE III
USING ALL KINDS OF REGRESSION MODELS TO INVESTIGATE HOW INDEPENDENT VARIABLE AFFECTS THE STATISTICAL VALUE OF FINAL TERM ASSESSMENT RESULTS 1

Model reps(20)	ols			Q0.25			Q0.5			Q0.75		
Var	Coef.	T	P	Coef.	T	P	Coef.	T	P	Coef.	T	P
X1	-0.129	-1.39		-0.060	-0.99		-0.110	-0.69		-0.269	-1.17	
X2	-0.059	-0.26		0.615	4.31	***	0.523	1.83	*	0.307	0.62	
X3	1.301	9.15	***	0.162	1.43		0.330	1.03		0.709	1.02	
X4	-0.063	-0.74		0.012	0.28		0.111	1.1		0.341	2.34	**
X5	2.024	3.75	***	0.647	1.16		1.859	2.28	**	3.541	2.28	**
X6;	-1.767	-4.09	***	-0.354	-0.58		-1.314	-1.73	*	-2.739	-2.17	**

Note: *means that it is significant under 10% P value;
**means that it is significant under 5% P value;
***means that it is significant under 1% P value.

Then, repeatedly sample the data 100 times by simultaneous quantile regression to achieve more accuracy with less variance, shown in table 4.

In the vertical column of quantile level as 0.25, GDP (X1) has a 10% P value while that of the fiscal budget (X2) continues to be low. What's more, there appears another variable of gross retail sales of consumption (X3), which is supposed to have a P value of 5%. In other words, gross retail sales of consumption in regions with small-scales freight ton-kilometers entails profound impact.

In the one as 0.5, fiscal budget (X2) is still in a stage of average P value, total import volume (X5) and total export volume(X6) is still in relatively high P value. Compared with the 20 times sampling, the P value of household savings is 10%, a medium significance level , whose effect on areas with moderate freight ton-kilometers works in terms of the increase of cargo demand.

For that as 0.75, the P value of household savings (X4) keeps unchanged at 5%. Whereas, the P values of total import volume (X5) and total export volume (X6) decrease from that in the 20 times sampling to 1% in the 100 times sampling. It shows that the increase of total import volume and total export volume will greatly promote freight ton-kilometers in large areas where freight ton-kilometer is already large.

TABLE IV
USING ALL KINDS OF REGRESSION MODELS TO INVESTIGATE HOW INDEPENDENT VARIABLE AFFECTS THE STATISTICAL VALUE OF FINAL TERM ASSESSMENT RESULTS 2

Model reps(100)	ols			Q0.25			Q0.5			Q0.75		
Var	Coef.	T	P	Coef.	T	P	Coef.	T	P	Coef.	T	P
X1	-0.129	-1.39		-0.060	-1.68	*	-0.110	-0.99		-0.269	-1.41	
X2	-0.059	-0.26		0.615	4.16	***	0.523	1.85	*	0.307	0.71	
X3	1.301	9.15	***	0.162	2.37	**	0.330	1.62		0.709	1.38	
X4	-0.063	-0.74		0.012	0.29		0.111	1.68	*	0.341	1.97	**
X5	2.024	3.75	***	0.647	0.95		1.859	2.48	**	3.541	2.98	***
X6	-1.767	-4.09	***	-0.354	-0.54		-1.314	-2.16	**	-2.739	-2.91	***

Note: *means that it is significant under 10% P value;
**means that it is significant under 5% P value;
***means that it is significant under 1% P value.

It is demonstrated in Table 5 that the difference of household savings, total import volume and total export volume turns out to be distinct when comparing the data of the 0.75 (high level) quantile and the 0.25 (low lever) quantile. It is consistent with the estimated results in table 4. Among them, the coefficients of total import volume and export volume level also differ evidently.

Between the 0.5 quantile and the 0.25 quantile, only when comparing occurs the difference at an average P value, whilst there seen no change in the comparison of 0.75 quantile and 0.5 quantile.

To conclude the analysis, the differences of the later three variables is relatively obvious, which are mainly reflected in the low level quantile and the medium level quantile.

TABLE V
DIFFERENCES OF VARIOUS QUANTILE COEFFICIENTS AND ASSESSMENT RESULTS

Model	q0.75 0.25			q0.50 0.25			q0.75 0.50		
Var	Coefficient difference	F	P	Coefficient difference	F	P	Coefficient difference	F	P
X1	-0.209	1.09		-0.050	0.28		-0.159	1.28	
X2	-0.308	0.39		-0.092	0.11		-0.216	0.37	
X3	0.547	0.9		0.168	0.78		0.379	0.59	
X4	0.328	4.91	**	0.099	1.8		0.230	2.25	
X5	2.894	3.61	**	1.212	2.99	*	1.682	2.2	
X6	-2.385	3.45	**	-0.960	2.7	*	-1.425	2.61	

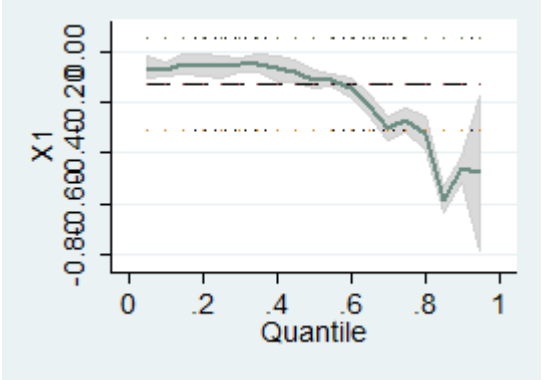

Table 1

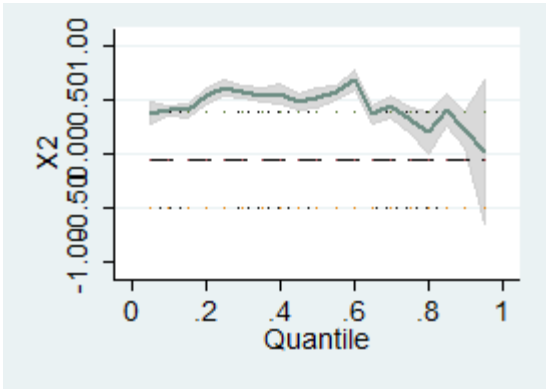

Table 2

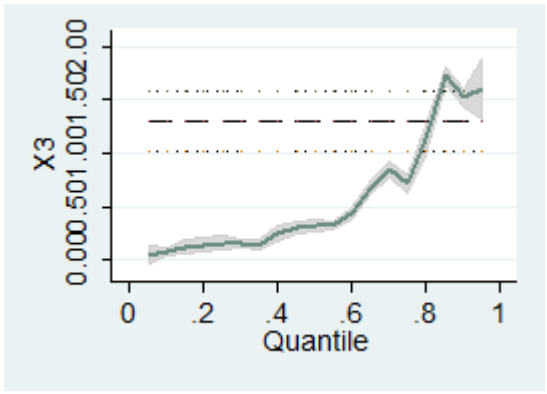

Table 3

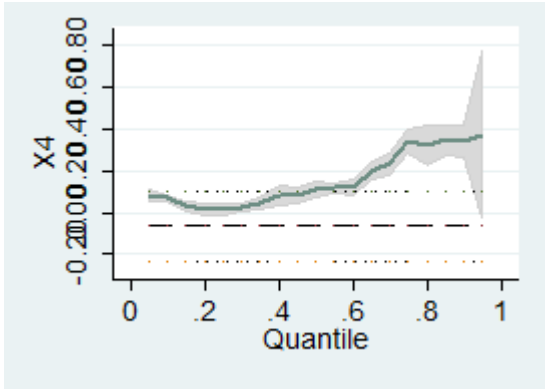

Table 4

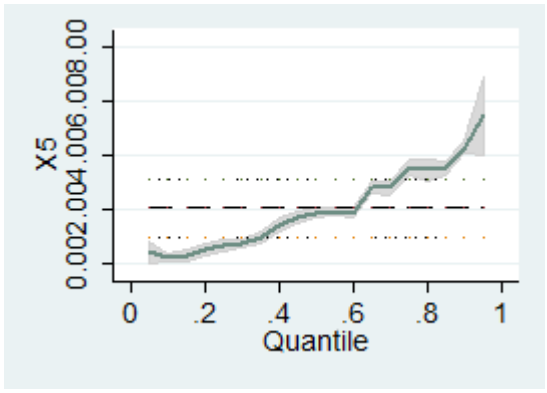

Table 5

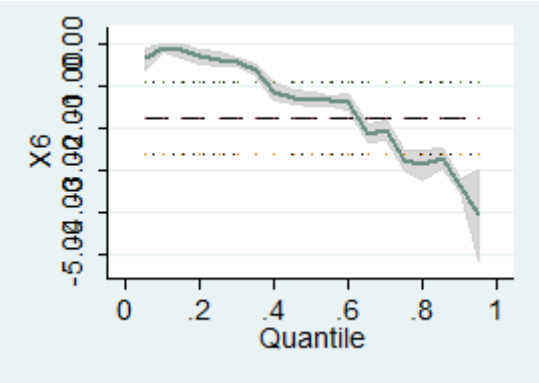

Table 6

Note: The shaded part means the 95% confidence interval of the quantile regression line. The ordinary least squares method is displayed by a horizontal coarse dashed line, with two horizontal fine dotted lines for the confidence intervals of the ordinary least squares (OLS) method.

The above six charts reveal the contrast between the quantile regression model and the OLS model, from which it is more intuitive to find the advantages of the former one. The curve means the quantile regression model, and the shadow section around means the 95% confidence intervals. The straight dotted line means the OLS model, with two finer dotted lines up and down its confidence intervals.

As shown in Chart 1, the curve of GDP (X1) in high level quantile has a rapid decreasing tendency, indicating that the effect of quantile regression on independent variable GDP is increasingly underestimated. At this point, OLS will overestimate the impact of GDP, resulting in deviations from results.

Chart 2 demonstrates the model of fiscal budget (X2), and it can be seen that the curve of quantile regression mostly distributes in OLS confidence interval except that of the quantile phase of 0.2-0.6 has exceeded the confidence interval of OLS, which indicates that in the OLS model, the influence of fiscal budget in the low level quantile is underestimated.

In Chart 3, the curve of quantile regression of gross retail sales of consumption (X3) is quite different from that of the OLS line. In the quantile phase 0.0-0.8, the OLS is much overrated. In the phase of quantile 0.8-1.0, the slope of the curve increases suddenly, and the overestimation of OLS becomes smaller and smaller.

Chart 4 shows that the curve of household savings (X4) is close to the OLS line in both the low level and medium level quantiles, indicating the similar estimated results of quantile regression and OLS at this stage. However, in the phase of high level quantile, the quantile regression curve gradually rises and even extends beyond the OLS confidence interval, which indicates that the influence degree of household savings is increasing when freight ton-kilometers is fairly large.

Chart 5 is the model of total import volume(X5), and it is easy to see that the quantile regression model is wholly in an increasing state. In the low level quantile, the estimated influence of total import volume by the quantile regression model is lower than the that by the OLS model. The significance level of the medium component continues to rise and manages to exceed that of the OLS model when reaching the high level quantile phase. At this point, the impact of OLS on total import volume is underestimated.

The last one, chart 6 is the model of total export volume (X6), and the trend of the quantile regression is opposite to that of imports, showing a decreasing tendency. In the low level quantile, the impact of total export volume estimated by the quantile regression is greater than that of OLS, which is confirmed that OLS underestimates its impact. The curve decreases gradually in the medium level quantile. In the high level pahse, the quantile regression estimates a smaller impact of the total export volume than that of the OLS, and it is confirmed that OLS overestimates its impact at this point.

IV. CONCLUSION AND SUGGESTION

Based on the data of Guangdong Bureau of Statistics for 2001-2015 years, this paper uses the quantile regression model to comparably analyze the influence in the low, medium and high level quantile on freight ton-kilometers and finally draws some conclusions and makes suggestions according to the results of empirical analysis.

In the low level quantile, fiscal budget and gross retail sales of social consumption have great influences on freight ton-kilometers. The paper assumes that it is because the regions with low freight ton-kilometers are incomplete in logistics infrastructure in the early stages of economic development with infrequent trade exchanges. If the government can increase the expenditure of infrastructure in these areas and formulate economic and trade policies to promote the local economy, it will also be of great help to the local logistics industry. As for the retail sales of social consumption, it is suggested to increase accompanied with household savings to stimulate the purchasing power of the society, or to stimulate social consumption by innovating the consumption patterns .

Total import as well as export volumes produce great influence on freight ton-kilometers in both medium and high level quantile. This paper holds the opinion that areas with large amount of imports shows great demand, and thus expands the demand for domestic logistics, and then freight ton-kilometers will also increase correspondingly. The factors that affect imports includes the rising exchange rate of renminbi, the increasing demand for cheaper imported goods and the increase of disposable incomes of household. A larger amount of exports means a larger quantity of goods shipped abroad, resulting in a greater demand for logistics. There are several ways to increase the export volume: increasing domestic production capacity to produce surplus products for export; declining the RMB exchange rate to grantee more profits for corporations; improving the product competitiveness to create world famous brands conducive to export.

In the high level quantile, the impact of household

savings sis till relatively significant. It shows that only in the regions with large freight ton-kilometers, the increase of household savings can serve to improve the demand for freight volume. Many factors affect the residents' savings deposits, such as improvement of economic level, expansion of business scale, increase of people's disposable income and the prosperity of the whole society, etc. In addition, a high interest rate of bank will promote household savings and vice versa, so monetary policy can have an effect on the residents' savings. Moreover, it is the same with the change of consumption concepts.

To sum up, the study of freight ton-kilometers plays an important role in the development of China's logistics industry, and how to improve the freight demand requires applicable approaches to shoot the target.

ACKNOWLEDGMENT

The research is sponsored by Guangdong Province Philosophy and Social Sciences Project (GD14CGL11), Ministry of Education Humanities and Social Sciences Project (15YJA630067), Guangdong Science and Technology Soft Science Project (2015A070704055), The National Natural Science Foundation of China (71673064), Guangdong Science and Technology Soft Science Project (2013B070206058, 2015A070704054).

REFERENCES

[1] Nie Zhengyan. "Analysis of the impact of logistics industry on regional economic growth in China"[J].*Industrial Technology Economy*, 77-82: 10th Issue of 2015.

[2] Fan Yuejiao. "Analysis of influencing factors degree of logistics service capability in China". *logisticsss Technology*, 193-196: 1st Issue,33nd Volume of 2015.

[3] Cui Shudan. "Assessment of interaction between economic growth and freight ton-kilometers in China". *Statistics and Strategies*, 75-77: 8th Issue of 2008.

[4] Liang Xin,Wang Qianfeng. "Analysis and forecast of air freight ton-kilometers in China based on SPSS". *Jiangsu Commercial Forum*, 45-47: 2010.

[5] Frederick Wiseman; Sangit Chatterjee. "Negotiating Salaries through quantile Regression". *Journal of Quantitative Analysis in Sports*, 1st Issue, 6th Volume of 2010.

[6] Liang Lizhen. "Study on the relationship between price and quantity in Chinese stock market by quantile regression". *Statistical Research*, 73-78: 12th Issue,25th Volume of 2008.