12-2015

THE FACTORS DETERMINING THE QUANTITY OF TAXIES - AN EMPIRICAL ANALYSIS OF CITIES IN CHINA

Yunzi Zhu
Clemson University, yunziz@g.clemson.edu

THE FACTORS DETERMINING THE QUANTITY OF TAXIES

—AN EMPIRICAL ANALYSIS OF CITIES IN CHINA

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillments
of the Requirements for the Degree
Master of Arts
Professional Communication

---

by
Yunzi Zhu
December 2015

Accepted by:
Dr. Chungsang Lam, Committee Chair
Dr. Kevin Tsui
Dr. Jaqueline Oliveira

ABSTRACT

The inefficient allocation of transport capacity in the taxi industry leads to the imbalance between the demand and supply of taxies. This paper tries to investigate the influencing factors on the quantity of taxies from the perspective of the demand. It is based on the empirical analysis and assumes that the quantity of buses (Bus), average wage per citizens (AW), population density (PD), road density (RD), the availability of subways (SW) and the e-hailing application (EH) co-determine the quantity of taxies. Under the assumption, I collected the data in 287 cities from 2010 to 2013 and put them into the fixed effect model to examine the impacts of factors mentioned above. The empirical result shows that Bus, AW and PD have positive and significant influences on the quantity of taxies at different levels. Comparatively, CR negatively impacts it and the other two play little role. This result may offer some reference for government regulation on the taxi industry.

TABLE OF CONTENTS

Page

TITLE PAGE ... i

ABSTRACT ... ii

LIST OF TABLES ... iii

LIST OF FIGURES ... iv

SECTION

I. Introduction ... 1

II. Literature ... 2

III. Methodology ... 5

An Empirical Model ... 6

Data ... 10

IV. Empirical Result and Analysis ... 15

Hausman Test ... 15

The Fixed Effect Model ... 15

V. Conclusion ... 20

APPENDICES ... 21

REFERENCES ... 26

LIST OF TABLES

Table Page

4.1 Table.3 Correlation Analysis of the Four-year Sample 13

4.2 Table.4 Correlation Analysis of Samples across Four years.............................. 14

4.3 Table.5 the Result of Hausman Test 16

4.4 Table.6 the Regression with Fixed Effects (2010-2013) 19

I. Introduction

Taxi is an important component of the public transportation system. Since taxies provide convenient and comfortable services, the demand for them has been increasing. Though taxi effectively relieves the burden on the public transportation, it also exacerbates the traffic congestion. Some cities in China are facing the excess supply of taxies. There are too many idle taxies on the streets, which cause the waste of road resources. In other cities, comparatively, the taxi industry cannot satisfy the increasing travel demand. Passengers are difficult to hail taxies especially during the rush hour. To adjust the imbalance between the demand and supply, local governments regulate the taxi industry. But relevant policies, especially limiting the quantity of taxies by offering licenses, have little effect. The imbalance of the demand and supply of taxies results from an inefficient allocation of transport capacity in the taxi industry. Before solving the problem of inefficiency, we first need to examine factors which can impact the demand and supply of taxies.

II. Literature Review

The taxi industry has fully developed in developed countries. Since 1980s, some scholars began to focus on the equilibrium problem in the taxi market, and developed many methods to examine determinants of demand and supply in the taxi industry. Also, the transport capacity was one of their concerns.

The first person who proposed an aggregated model of taxi market was George W. Douglas (1972). This aggregated model displayed a functional relationship between variables such as the total number of taxies in a city, passengers' waiting time, and the number of vacant taxies on the streets. The latter two directly decided the optimal quantity of taxies that a city should keep. Douglas also described features of the regulated taxi market. But he only took individual factors into account, ignoring those aggregated factors such as urban development, GDP growth and living standards, which could greatly influence the taxi demand.

The expansion and upgrading of the road network play a little role in reducing congestion. Anthony Downs (1962) proposed that congestion could not be solved by extending the length of road or increasing area of road. New demand for travel would be induced after the expansion of the road network. As a result, demand for vehicles would exceed the supply.

S.C. Wong, Hai Yang and K.I.Wong (1999) built a network model on the basis of the impact of congestion and the elasticity of demand for transport capacity. They took the relationship between taxi and bus into account, investigating the substitution effect of bus for taxi.

In order to predict the aggregated demand for travel, Ningning Chen, Weijia Xu(2006), and Hao Ding(2008) used urban population and the size of urban areas to built a mathematical formulation. Combining this predicted demand with a city's transport capacity, they estimated the optimal quantity of taxies that city should supply. The urban population and the size of urban areas are the main factors determining aggregated demand for travel in this study.

Fang He and Max Shen (2015) developed a spatial equilibrium model to examine the equilibrium in the regulated taxi industry. They particularly investigated the impact of e-hailing application. To extend this model, the price elasticity of demand for taxies was also taken into account. The empirical result showed that e-hailing application could reduce passengers' average waiting time. The increasing demand for taxies resulted from the use of this application. We consider using e-hailing application may positively impact the quantity of taxies in cities.

Base on previous research, this paper tries to examine factors implicitly or explicitly impacting the demand for taxies. They are: (1) other public vehicles such as bus and subway, (2) road density (3) population density (4) average wage and (5) e-hailing application for taxies.

III. Methodology

In order to investigate the relationship between the taxi supply and factors of interest mentioned above, we build an empirical model. The general representation of the estimated model is as following:

$$Taxi = f(Bus; AW; PD; RD; SW; EH) \quad (1)$$

Among:

Taxi: the average quantity regulated by a city government per year. It is calculated by the quantity of taxies per ten thousand citizens.

Bus: the average number of buses per year; and assumed as a substitution for taxies. It is calculated by the quantity of buses per ten thousand citizens.

AW: per capita annual income in a specific district (unit: yuan per year). The more annual income earned, the more can be paid for taxies services. We assume that AW impacts the quantity of taxies positively.

PD: population density, the ratio of population to the total land area in a specific district (unit: population per square kilometer). High population density is followed by high travel volume, which reflects a high demand for vehicles, including taxies.

RD: road density, the ratio of the length of total road network to land area in a specific district (unit: kilometer of road per square kilometer of land). Road density is an index measuring volume of road network. High road density means a road network is capable of carrying a large number of vehicles, which can have a positive influence on the supply of taxies.

SW: the availability of subways, another substitution for taxies, a dummy variable. If a city owns subways, SW=1. Conversely, SW=0.

EH: the availability of e-hailing application, dummy variable. If e-hailing application is used to call taxies, EH=1. Conversely, EH=0. E-hailing application lowers passengers' waiting times and increases taxies' utilization. That is to say, it may have a positive influence on the demand for taxies.

3.1 An Empirical Model

3.1.1 Form of the Empirical Model

Panel data model becomes an important econometric method in recent years. It spatially extends data along the time series, adding degrees of freedom as well as reducing collinearity between explanatory variables. Meanwhile, this model can analyze not only the impacts of various cross sections in the same period, but also the impact of one particular cross section in different time. Therefore, panel data model has a superior advantage than common time-series model or cross-section model. Theoretically, the general form of linear panel data model can be shown as:

$$y_{it} = \alpha_{it} + \beta_{it} X_{it} + u_{it}$$

In this form, y_{it} is response variable. Specifically, α_{it} represents some time-invariant factors, which are sometimes hard to be observed or quantified, such as patterns of individual consumption or social institutions of a country. They are often defined as individual effects such that include variations in individuals. The coefficient β_{it} consists of a matrix of estimated parameters and X_{it} refers to a matrix of explanatory variables impacting the cross sections of response variable. The last term u_{it}, which is on the right-hand side of the equation, is the random error. It includes omitted factors which impact the response variables as cross section and time vary.

In addition, the letters i and t respectively refer to the ith cross section and the tth year in the panel data.

According to the research of interest, we need to set an original regression model. In order to find whether impacts of these variables are stable and consistent, we add them gradually into the model. Specifically, we compare each model to the previous one, checking whether coefficients of existing variables significantly change after adding new variables. If there is an obvious change in the coefficient of some existing variable, we should pay attention to it. Table 6 provides the result in detail.

The general representation of the regression model is:

$$
\begin{aligned}
LNTaxi_{it} &= \beta_0 + \beta_1 * LNBus_{it} + \beta_2 * LNAW_{it} + \beta_3 * LNPD_{it} \\
&+ \beta_4 * LNPD_{(i-1)t} + \beta_5 * LNPD_{(i-1)t} + \beta_6 * LNRD_{it} + \beta_7 * CR_{it} + \beta_8 * EH_{it} \\
&+ \sum_{t=2012}^{2013} Year_t + \sum_{i=1}^{286} City_i + u_{it}
\end{aligned}
$$

Except for the six explanatory variables we mentioned above, some other variables need to be explained. Previous year's population density and road density may impact the next year's supply of taxies. That is to say, the quantity of taxies needs time to respond to the impact of population density and road density. So, two variables lagged by one year, i.e., $LNPD_{(i-1)t}$ and $LNRD_{(i-1)t}$ are put in the model. In addition, in order to investigate individual effects and time effects, we add dummy variables such as Year 2012, Year 2013 and $City_i$. Using those variables lagged by one year results in a lost data on Year 2010. Therefore, there are two year dummies and 286 city dummies in the model. Term u_{it} is the random error.

3.1.2 The Choice between Fixed Effect Model and Random Effect Model

The general form of the linear panel data model mentioned above is classified as the fixed effect model and the random effect model. The main difference between these two models is the way in which dealing with individual effects. The fixed effect model treats them as time-invariant fixed factors. Comparatively, the latter treats individual effects as random factors. Specifically, the fixed effect model assumes that individual variations are correlated with explanatory variables. That is to say, factors of interest are assumed to be influenced by obvious individual differentials. Conversely, the random effect model assumes that individual variations are uncorrelated with explanatory variables. In our study, for example, we can consider regulations on different taxi industries, different public transport systems, and travel patterns of individuals as a series of individual effects. Moreover, in order to improve estimated result, the fixed effect model applies Least Square with Dummy Variable (LSDV) to estimation, while the random effect model uses Estimated Generalized Least Squares (EGLS) to solve the problem arising from the random error term.

Greene (2003) used Hausman Test to decide which model can be applied to estimation. Basically, under the null hypothesis that individual effects are uncorrelated with explanatory variables, both estimates $\overline{\theta}_i$ from LSDV and $\theta_i \sim$ from EGLS are consistent, but $\overline{\theta}_i$ become inefficient due to the loss of degrees of freedom. Therefore, the difference between $\overline{\theta}_i$ and $\theta_i \sim$ will be small or close to zero. Under the alternative hypothesis that individual effects are correlated with explanatory variables, $\overline{\theta}_i$ are still consistent. However, $\theta_i \sim$ are not consistent any more. Hence, the difference between $\overline{\theta}_i$ and $\theta_i \sim$ will be large. Conclusively, Hausman Test transformed the null hypothesis as no difference between estimates of fixed effect model and of random effect model. If we fail to reject the null hypothesis, we should choose random effect model. Conversely, if we reject the null hypothesis, fixed effect model should be used instead. And we can conduct Hausman Test by using STATA 12.0.

3.2 Data

The sample period of this study is from year 2010 to 2013, which includes the time when e-hailing applications for taxies were built. The quantity of taxies (Taxi), the quantity of buses (Bus), average wage (AW), population density (PD), road density (RD), city subways (SW) and e-hailing application (EH) are collected from "China Urban Construction Statistical Year Book" and "China City Statistical Year Book" during the period from 2010 to 2013. Statistical data were gathered from 287 different municipal districts in all 34 provinces in China. Since the municipal district is the main part of a city, which is characterized as having high population density, concentrated floating population, high proportion of urban residents and developed economy. Public transportation serves the most of population that is concentrated in this region.

Also, the logarithms are taken of all variables except the two dummy variables. Comparing the box plot of original variables to that of log transformed variables (see Figure 1 and 2 in the Appendices), we find that in the original data there are many extreme values which may impact the stability of the estimated result. Taking logarithms of all these variables can eliminate this negative impact to some degree. Both descriptive statistics relative to all variables in the four-year sample and in samples from 2010 to 2013 are shown below (see Tables 1 and 2 in Appendices).

In summary, there are very slight changes in the distributions of all these variables across four years except the distribution of average wage (AW). There is an upward trend in the distribution of average wage (AW). Combining results displayed in Table 2, we can find that the values of standard deviation of all these variables are also stable across the four years. Therefore, we may consider, except for the variable average wage (AW), that the impact of time effect on these variables may be slight. And it needs to be tested later.

It is also worth to note the coefficient of variation (CV). CV is similar to standard deviation, and it reflects the degree of dispersion of a variable. But standard deviation cannot be used to compare the degrees of dispersion of a variable in different time periods or that of different variables in the same condition, due to different mean values and magnitudes of variables. The CV value eliminates the impacts of different mean values and their magnitudes. Tables 1 and 2 provide information on comparisons of variations within group. We find that average wage (AW) has the smallest CV, which means the average wage differentials across cities are the smallest, compared to other factors in question. Based on this result, we may consider individual effects have a relatively slight impact on average wage. Comparatively, there are the largest variations in the quantity of buses (Bus). The individual effects may have relatively strong impact on the supply of buses

3.2.1 Correlation Analysis

A Pearson correlation coefficient is used to investigate the direction and strength of the linear association between the response variable and each explanatory variable. If there is a strong linear association between them, we can tell that the explanatory variable may have some influence on the response variable and should be added into the regression model to test the accurate relationship. The results of correlation analysis of the four-year sample and that of samples across four years are shown as Tables 3 and 4.

From Table 3, comparing each explanatory variable's correlation coefficient, it is easy to find the quantity of buses has the strongest positive association with the quantity of taxies at significant level of 0.01. That is to say, the quantity of buses may have some positive impact on the supply of taxies. Though other variables such as average wage and population density are significantly correlated with the response variable, their linear associations are weak. We consider that their impacts on the quantity of taxies may be weak. In addition, the road density is almost uncorrelated with the quantity of taxies.

Table.3 Correlation Analysis of the Four-year Sample

	LNTaxi	LNBus	LNAW	LNPD	LNRD
LNTaxi	1				
LNBus	0.6312***	1			
LNAW	0.1969***	0.3892***	1		
LNPD	0.1203***	0.3427***	0.0909***	1	
LNRD	0.0406	0.0149	-0.0160	-0.0852***	1

*** p<0.01, ** p<0.05, * p<0.1

After conducting correlation analysis in different year groups, we find strong positive correlations between the explanatory variables and the response variables across years. The result is shown in Table 4. In detail, in all the year groups, the quantity of buses has a positive correlation with the quantity of taxies. So dose average wage Also, the population density has displayed a positive correlation with the response variable in year 2010, 2011 and 2013. However, the road density is uncorrelated with the quantity of taxies in any year group.

Table.4 Correlation Analysis of Samples across Four years

	LNTaxi	LNBus	LNAW	LNPD	LNRD
2010 Year					
LNTaxi	1				
LNBus	0.6181***	1			
LNAW	0.3048***	0.4359***	1		
LNPD	0.1480**	0.3769***	0.1303**	1	
LNRD	0.086	-0.0012	-0.0457	-0.0105	1
2011 Year					
LNTaxi	1				
LNBus	0.6313***	1			
LNAW	0.2586***	0.4424***	1		
LNPD	0.1399**	0.3415***	0.1559***	1	
LNRD	0.0225	0.0259	0.0189	-0.1403**	1
2012 Year					
LNTaxi	1				
LNBus	0.6253***	1			
LNAW	0.2386***	0.4332***	1		
LNPD	0.0864	0.3298***	0.1306**	1	
LNRD	-0.04	-0.0003	0.0396	-0.1538***	1
2013 Year					
LNTaxi	1				
LNBus	0.6572***	1			
LNAW	0.2042***	0.4220***	1		
LNPD	0.1064*	0.3316***	0.0762	1	
LNRD	0.0652	0.0594	0.0214	-0.0948	1

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

IV. Empirical Result and Analysis

4.1 Hausman Test

The empirical result is shown in Table 5. The value of chi-square statistic is 56.20, and its p-value is less than significant level of 0.01. Based on this result, we can reject the null hypothesis, i.e., reject the assumption that individual effects are uncorrelated with other explanatory variables. Therefore, the fixed effect model is proper for regression.

4.2 The Fixed Effect Model

After applying xtreg command in STATA 12.0, we derive the final regression model:

$$
\begin{aligned}
LNTaxi = & -1.805 + 0.389 * LNBus + 0.303 * LNAW + 0.114 * LNPD \\
& -0.0235 * 1.LNPD + 0.036 * 1.LNRD - 0.00515 * LNRD \\
& -0.0705 * CR + 0.0315 * EH \\
& -0.0623 * Year2012 - 0.0961 * Year2013
\end{aligned}
$$

Estimates of the explanatory variables' coefficients are displayed in Table 6.

Initially, the joint significance test rejects the null hypothesis that all coefficients of explanatory variables are equal to zero. It tells us that this regression model is still useful to measure the impacts of explanatory variables after taking logarithms of them. Specifically, from model (1) to model (5), the quantity of buses has a positive and significant influence on the quantity of taxies. So does average wage. At the significant level of 0.01, keeping other explanatory variables constant, there will be a 0.398% increase in the quantity of taxies per ten thousand people as a 1% increase in the quantity of buses per ten thousand people. That is to say, when a city develops its public transportation by increasing the supply of buses, the supply of taxies will also increase. This result contradicts our expectation that buses play a role in partially substituting for taxies.

Table.5 The Results Of Hausman Test

Variable	Coefficient		(b-B)	sqrt(diag(V_b-V_B))	$\square^2$	Prob>chi2
	(b)	(B)				
	fixed	random	Difference	S.E.		
LNBus	0.3887	0.4944	-0.1057	0.0196		
LNAW	0.3027	0.2859	0.0168	0.0376		
LNPD						
--.	0.1139	0.0572	0.0568	0.0228		
L1.	-0.0235	-0.0491	0.0256	0.0111		
LNRD						
--.	-0.0051	-0.0042	-0.0009	0.0022	56.20	0
L1.	0.0360	0.0467	-0.0106	0.0028		
SW	-0.0705	-0.0364	-0.0340	0.0456		
EH	0.0315	0.0296	0.0019	.		
Year						
2012	-0.0623	-0.0643	0.0021	0.0046		
2013	-0.0961	-0.1024	0.0063	0.0087		

Also, small- and medium-sized cities are a large proportion of the sample. The public transportation system in these cities may be at an early stage of development. Therefore, the passengers' demand for buses greatly exceeds the supply. Under the circumstance of a limited transport capacity, increasing the supply of buses still cannot meet passengers' needs, the supply of taxies as a supplementary of buses may increase to satisfy the excess demand for travel.

There will be a 0.303% increase in the quantity of taxies per ten thousand people as a 1% increase in average wage. The rise of average wage allows people to increase their spending on transportation. It gives people more latitude in choosing their preferred transportation means.

At the significant level of 0.10, when there is a 1% increase in population density, the quantity of taxies per ten thousand people will increase by 0.114%. The Increasing population density refers to a rise in the population in a specific region and subsequently induces aggregated travel demand to go up. Therefore, the demand for taxies will experience an increase. Though the result shows a positive impact of road density on the quantity of taxies, it is not significant. Combining the values of road density's CV in four years, we can find the road density is relatively stable across all cities. We may then think that the construction of road network takes time and there exists a lag effect on the response of the taxi industry to the change in road density. Due to the limited number of years observed, the data cannot reflect the full information of cities analyzed. In addition, two variables lagged for one year have small impacts on the quantity of taxies. This result may infer that the adjustment of the taxi supply, according to the size of population and volume of road density, will take more than one year.

Moreover, if a city provides subways, the quantity of taxies will decrease by 0.0705%. Subways are more time-saving, convenient than other vehicles, especially during the rush hour. Passengers may prefer subways to taxies, since there is no congestion. E-hailing application for taxies impacts the quantity of taxies slightly. This technology has just emerged recently and become popular in some medium- and large- sized cities with high demand for travel. Therefore, as for most cities in the sample, it is hard to tell whether e-hailing application can reduce passengers' waiting time, improve taxi utilization and stimulate demand for taxies. More data on this part should be collected.

Table.6 The Regression with Fixed Effects (2010-2013)

VARIABLES	(1) LNTaxi	(2) LNTaxi	(3) LNTaxi	(4) LNTaxi	(5) LNTaxi
LNBus	0.280*** (0.0624)	0.281*** (0.0625)	0.384*** (0.0796)	0.389*** (0.0787)	0.389*** (0.0793)
LNAW	0.165** (0.0713)	0.169** (0.0711)	0.298*** (0.0750)	0.300*** (0.0757)	0.303*** (0.0766)
LNPD	0.0117 (0.0255)	0.0122 (0.0255)	0.116* (0.0635)	0.115* (0.0606)	0.114* (0.0632)
LNRD	0.00357 (0.0318)	0.00337 (0.0318)	-0.00650 (0.0290)	-0.00468 (0.0302)	-0.00515 (0.0303)
SW		0.0196 (0.0499)	-0.0706** (0.0333)	-0.0649* (0.0336)	-0.0705** (0.0336)
EH		0.0272 (0.0284)	0.0314 (0.0291)	0.0314 (0.0290)	0.0315 (0.0290)
2011.Year	-0.0572** (0.0224)	-0.0584*** (0.0224)			
2012.Year	-0.0988*** (0.0315)	-0.101*** (0.0315)	-0.0617*** (0.0166)	-0.0625*** (0.0164)	-0.0623*** (0.0166)
2013.Year	-0.112*** (0.0420)	-0.118*** (0.0430)	-0.0952*** (0.0275)	-0.0961*** (0.0275)	-0.0961*** (0.0276)
1.LNPD			-0.0219 (0.0151)		-0.0235 (0.0152)
1.LNRD				0.0350 (0.0323)	0.0360 (0.0318)
Constant	0.514 (0.852)	0.466 (0.852)	-1.630 (1.047)	-1.928* (1.008)	-1.805* (1.058)
Fixed Effects	YES	YES	YES	YES	YES
Observations	1,102	1,102	826	835	825
R-squared	0.101	0.102	0.205	0.209	0.208
F	3.98***	3.17***	5.50***	5.07***	5.11***
Number of City	286	286	284	284	284

Note: robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

V. Conclusion

Using panel data model to test the impacts of variables of interest, we find that the quantity of buses, average wage, and population density are significantly and positively correlated with the quantity of taxies at different levels. These factors may induce demand for travel from different aspects. Comparatively, the availability of subways negatively affects the quantity of taxies, considering a partial substitution effect. However, because of the limited data, the sample cannot completely depict the fact of the majority of cities in China. This makes the impacts of some variables like road density and e-haling application for taxies insignificant. In conclusion, when conducting research on the demand and supply of taxies, these factors having obvious impacts should be taken into account.

Appendices

Figure.1 the Box Plot of Variables

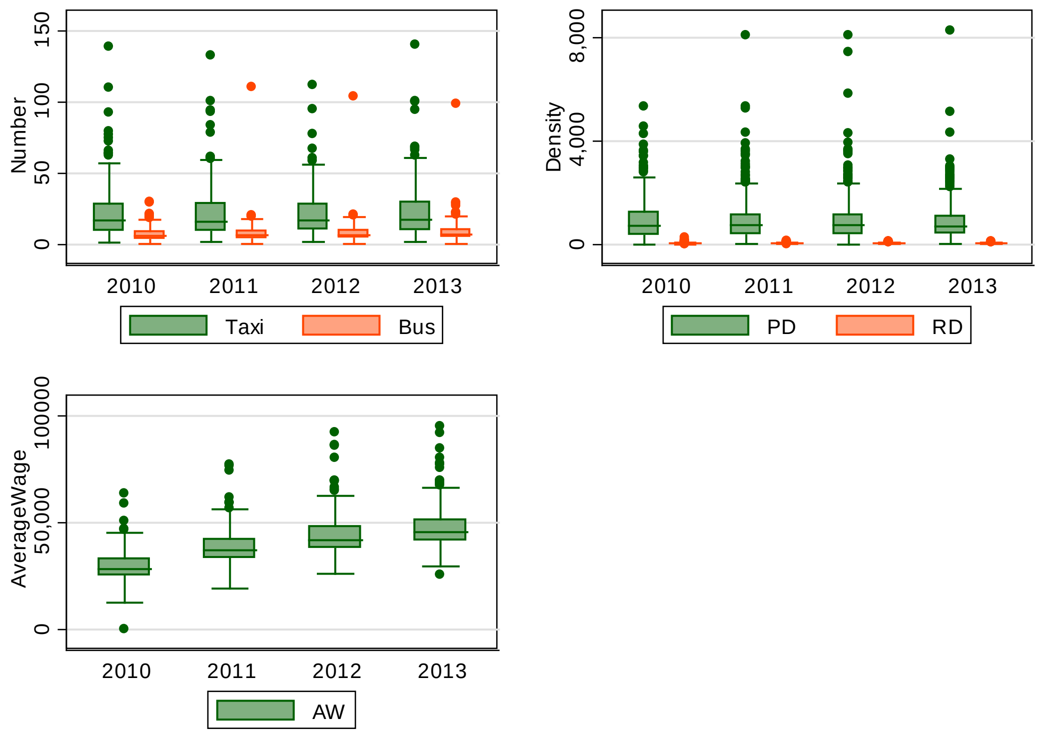

Figure.2 the Box Plot of Log Transformed Variables

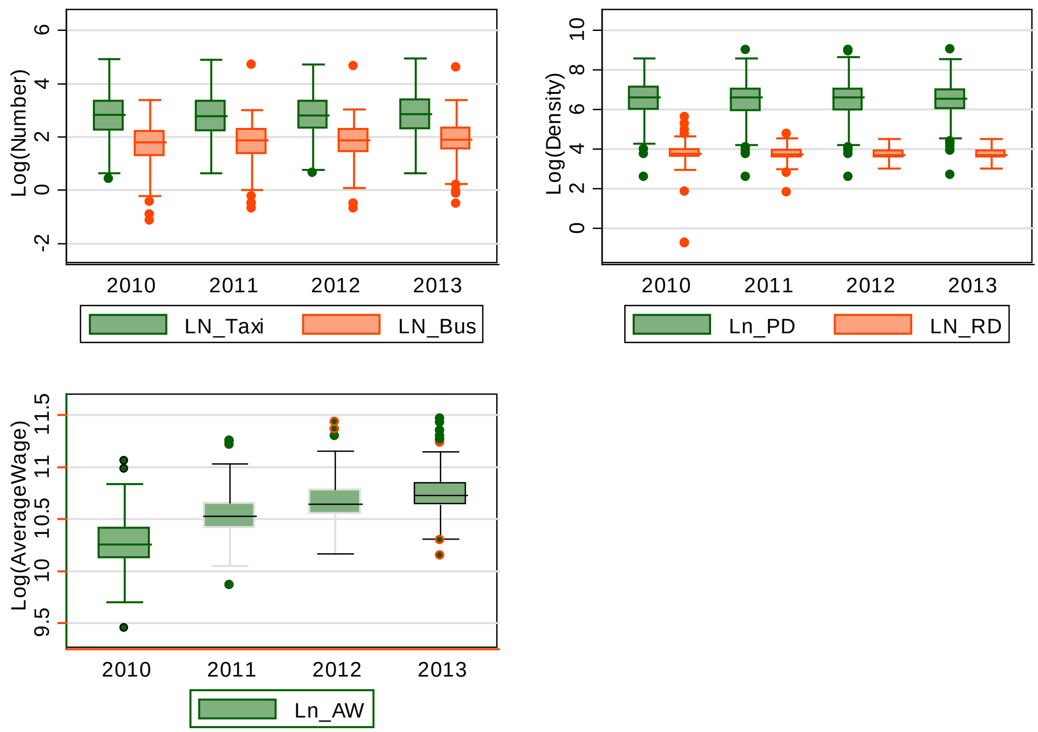

Table.1 Descriptive Statistics of the Four-year Sample

Varible	N	Mean	SD	Min	Max
Taxi	1138	21.77	17.53	1.50	140.24
LNTaxi	1138	2.78	0.81	0.41	4.94
Bus	1135	7,68	6.80	0.32	110.52
LNBus	1135	1.80	0.72	-1.14	4.71
AW	1121	39874.22	11200.88	0.00	95029.65
LNAW	1120	10.56	0.28	9.45	11.46
PD	1142	964.56	926.55	0.00	8248.04
LNPD	1131	6.49	0.94	2.57	9.02
RD	1144	44.75	16.58	0.46	265.64
LNRD	1144	3.74	0.37	-0.78	5.58

Table.2 Descriptive Statistics of Samples across Four Years

Periods	Varibles	N	Mean	SD	CV	Min	Max
2010	Taxi	285	21.93	18.19	0.83	1.50	138.86
	LNTaxi	285	2.78	0.82	0.29	0.41	4.93
	Bus	284	6.98	4.65	0.67	0.32	29.86
	LNBus	284	1.71	0.74	0.43	-1.14	3.40
	AW	279	29580.61	7293.49	0.25	0.00	63545.69
	LNAW	278	10.27	0.23	0.02	9.45	11.06
	PD	286	966.52	881.08	0.91	0.00	5324.12
	LNPD	276	6.52	0.96	0.15	2.59	8.58
	RD	286	47.07	22.53	0.48	0.46	265.64
	LNRD	286	3.76	0.52	0.14	-0.78	5.58
2011	Taxi	283	21.94	18.19	0.83	1.50	138.86
	LNTaxi	283	2.78	0.82	0.29	0.65	4.89
	Bus	283	7.59	7.49	0.99	0.50	110.52
	LNBus	283	1.78	0.72	0.40	-0.69	4.71
	AW	284	38586.59	8633.90	0.22	19266.57	77144.95
	LNAW	284	10.54	0.21	0.02	9.87	11.25
	PD	285	984.12	962.71	0.98	13.13	8063.40
	LNPD	285	6.48	0.97	0.15	2.57	9.00
	RD	286	44.48	14.57	0.33	6.03	117.11
	LNRD	286	3.74	0.33	0.09	1.80	4.76
2012	Taxi	284	21.18	15.79	0.75	1.91	112.05
	LNTaxi	284	2.78	0.78	0.28	0.65	4.72
	Bus	283	7,86	7.26	0.92	0.50	103.77
	LNBus	283	1.83	0.70	0.38	-0.69	4.64
	AW	284	43894.45	9776.62	0.22	25981.17	92357.49
	LNAW	276	10.67	0.21	0.02	10.17	11.43
	PD	285	982.80	1003.29	1.02	0.00	8074.84
	LNPD	285	6.49	0.96	0.15	2.57	9.00
	RD	286	43.93	13.94	0.32	20.23	91.65
	LNRD	286	3.74	0.30	0.08	3.01	4.52

(Continued) Table.2 Descriptive Statistics of Samples across Four Years

2013	Taxi	286	22.02	17.89	0.81	1.89	140.24
	LNTaxi	286	2.79	0.81	0.29	0.64	4.94
	Bus	285	8.29	7.36	0.89	0.59	98.53
	LNBus	285	1.88	0.70	0.37	-0.53	4.59
	AW	282	47420.37	9945.28	0.21	25545.88	95029.65
	LNAW	282	10.75	0.19	0.02	10.15	11.46
	PD	285	924.74	854.85	0.92	14.57	8248.04
	LNPD	285	6.48	0.89	0.14	2.68	9.02
	RD	286	43.53	13.44	0.31	20.51	89.85
	LNRD	286	3.73	0.29	0.08	3.02	4.50

References

[1] Douglas G. W. Price Regulation and Optimal Service Standards: The Taxicab Industry, *Journal of Transport Economics and Policy*, 1972, 6(2): 116-127.

[2] Downs A. The Law of Peak-Hour Expressway Congestion, *Eno Transportation Foundation*, 1962, 16(3): 393-409

[3]Wong K.I., et al. Modeling Urban Taxi Services in Congested Road net works with elastic demand, *Transportation Research Part B*, 2001, 35:819-842

[4]Chen N., et al. Taxi Planning in Urban Traffic Management, *Mathematics in Practice and Theory*, 2006, 136(17):115-116

[5]Ding H. Model of Taxi Planning in Urban Traffic Management, *Mathematics in Practice and Theory*, 2008, 38(6):94-101

[6]He F., et al. Modeling Taxi Services with Smartphone-based E-hailing Applications, *Transportation Research Part C*, 2015, 58:93-106