

# Allocation of Ride-hailing Resources in the "Internet +" Era: Taking Xuzhou City as an Example

Guodong Hu, Yihua Zhang\*

school of Business Administration, JiMei University, Xiamen, China

\*corresponding author

**Keywords:** Resource allocation; Diverse linear regression; Principal component analysis; Grey correlation analysis; Subsidy scheme

**Abstract:** Based on the problem of “difficulty in taxis” in major cities, taking Xuzhou City as an example, the paper analyzes the resource allocation of the network car, the effect of the existing subsidy program and proposes an improvement plan to reconfigure the vehicles in each district. Using python crawler to collect data on the supply and demand of Xuzhou City's network vehicles, and using the principal component analysis method to empirically analyze the factors affecting the network allocation of vehicles, using linear regression prediction method to predict the network of vehicles. The demand, using the gray correlation analysis method, further reveals that when the network car subsidy fails to effectively promote the resource allocation of the vehicle, and proposes a plan to improve and optimize the existing subsidy program, in order to promote the network car resource allocation significant improvement in rationality.

## 1. Introduction

According to the relevant report data released by the research institute, the ride-hailing industry in 2014 started a climax. Due to the relevant subsidy program by the ride-hailing software company, the general public began to gradually use the ride-hailing vehicle as a travel tool. It has grown rapidly, with a simultaneous increase of nearly six times. However, in recent years, the amount of subsidies has gradually weakened, and the growth has tended to slow down or even decrease slightly. In addition, the number of users on each platform has significantly decreased since the introduction of the New Deal on Internet Cars. By the end of 2017, the total number of ride-hailing users across the country reached 336 million [1]. With the gradual improvement of the living standards of citizens, the quantity and quality of travel requirements are constantly increasing. Many citizens tend to use the Internet for taxis because of the congestion of passengers in buses, the slow speed of travel, and the inability to reach their destinations directly. As an important choice for travel, the phenomenon of “difficult to take a taxi” is also becoming increasingly serious, especially in some large cities with large populations [2]. From 2012 to 2016, the total number of registered users of taxis has continued to rise. In 2012, the number of registered users was less than 5 million, and the number of registered users increased significantly in 2013. This period is the period of publicity and use of taxi software. Many citizens felt the convenience brought by the taxi software and gradually started using it. The emergence of such a taxi software platform can realize the information exchange between the customer and the owner of the online taxi, and then the owner of the online taxi can accurately and quickly respond to the corresponding information on the mobile phone, greatly adjusting the vehicle's resource allocation [3]. However, this model still cannot completely alleviate the situation of “difficult taxi”. This paper analyzes and analyzes the resource allocation of ride-hailing cars in Xuzhou City to analogize cities in the country.

## 2. Research Review

Scholars at home and abroad have begun to explore the hot issue of "difficult travel" for residents. Among them, one of the important transportation specifics has been attracting widespread

attention.

Most countries have issued a series of related regulations to solve this problem. In Singapore, for example, every taxi company has an app. A taxi app has become popular in Singapore. Most taxi software companies have their own taxi platforms to attract consumers. Customers only need to pay a part of the deposit to enjoy the taxi service [4]. In 2018, Uber conducted a pilot cooperation with the local government in a small city in Japan to dock passengers' taxi software and ride-hailing on mobile phones, and car owners and companies will not charge a docking fee to achieve the effect of carpooling. Make better use of online car resources. In addition, the taxi software company did not specify clearly in terms of auto insurance, and some disputes may occur in future development [5].

Domestic scholars have explored the problem of supply and demand for online ride-hailing cars and have achieved fruitful results. Li Xiaochun (2019) researched the relationship between the number of peak air-travel network rides and passenger demand during a certain period of time in Changsha, and found that the matching degree in some time periods is less than 0.5, which can indicate that the supply of online-travel cars at these times is Seriously inadequate, it is difficult for passengers to successfully reach the car [6]. Cui Yuzhao, Cao Yongzhao, and Wang Mengxuan (2017) used factor analysis to explore the supply-demand matching situation of Internet-linked taxis in 11 cities, and concluded that the detection indicators in Tangshan, Shenzhen, and Xiamen did not reach the average level, and the distribution of Internet-linked taxis Unbalanced, the matching degree of the city's network of ride-hailing is not directly related to the degree of urban development [7]. Fu Anrui and Dai Hanhui (2017) evaluated the subsidy scheme for ride-hailing based on game theory and found that the existing subsidy scheme can only solve the problem of short-distance taxis to a certain extent [8].

Existing issues related to the allocation of ride-hailing resources provide the basis for the study of this article, but there are still shortcomings: there are not many impact indicators selected in the existing research, and there may be important indicators that have not been considered to make the results different. It only analyzes the relationship between the supply and demand of online ride-hailing cars and the existing subsidy incentives. There is no targeted improvement or the effect is not obvious.

### **3. Construction of Data Source and Indicator System**

#### **3.1. Data collection and cleaning**

##### **3.1.1. Data collection**

The quantitative index data used in this article is collected from the data of Xuzhou's 2010-2018 online rides may affect the index data, some of which are from the summary of previous years, the annual report of the Xuzhou Statistics Bureau, and the other part are from the software of Didi Taxi Database for python web crawlers.

In the end, we obtained 3128 pieces of data such as per capita GDP, the number of buses, the urban population, and the number of tourists.

##### **3.1.2. Data cleaning**

Data cleaning is to correct wrong information, delete incorrect data, and keep data synchronized. Because the database data is extracted from several business systems, it is bound to cause some conflicting data, and of course there is no way to guarantee the accuracy of the data. This is commonly referred to as "dirty data." By cleaning the 3128 "dirty data" obtained initially, identifying and processing invalid values, filtering the data that does not meet the requirements, and finally obtaining 2596 valid data.

#### **3.2. Construction of indicator system**

##### **3.2.1 Supply and demand of online car hailing**

Considering the situation of different time and space, the selection of data is used to distinguish the relevant influencing factors of the city's online car rental in different years in different regions,

and the following indicators are analyzed according to the factors. Because the question was raised to analyze the degree of "supply-demand matching", and the supply is the number of city car rides, which can be obtained through the statistical yearbook of each city. It also analyzes the related influencing factors of the demand for online ride-hailing cars. There are many factors that affect the demand for online car rental, and each index factor has its own meaning. Optimize the indicators to get the following table:

According to the above method, because the regional size of Xuzhou is constant every year, this factor is excluded in all indicators, and the relevant data is finally obtained as shown in Table 1 and 2:

Table 1 Resource allocation index of online car hailing in Xuzhou.

Region	Indices Years	GDP per capita (Yuan)	Number of buses (Car)	Urban population (Million)	Per capita consumption (Yuan)	Bus operating mileage (10,000 / km)	Per capita income (yuan)
Xu Zhou	2010	45031	6271	838	7895	8328	13132
	2011	49676	6489	844	9487	8582	14225
	2012	53230	6525	851	10553	7495	16035
	2013	57985	6588	857	13113	8057	17015
	2014	59327	6630	860	14023	8498	19080
	2015	62285	6711	866	16982	9079	20219
	2016	65898	6778	871	17760	9831	22168
	2017	72005	6798	876	19350	9989	24535
	2018	77084	6886	879	22337	10153	33178

Table 2 Resource allocation index of online car hailing in Xuzhou.

Region	Indices Years	Bus passenger traffic (10,000 / person)	Investment amount of online car (Million)	Urban road area (Square kilometers)	Total tourism income (Billion yuan)	Net number (Car)	Number of tourists (Ten thousand people)
Xu Zhou	2010	236578	154	3443	323	6000	2689
	2011	114587	181	3087	482	6646	3590
	2012	112542	212	3632	753	6646	4715
	2013	225245	248	3941	90	6650	4560
	2014	373272	300	2247	144	6746	6669
	2015	382452	389	2355	425	6851	8390
	2016	324537	374	2258	864	6851	1404
	2017	418585	407	2271	301	6851	3134
	2018	405875	426	2295	666	7046	5189

### 3.2.1 Current subsidy for online car hailing

Through the analysis of various influencing factors, it is found that there are many indicators that affect the needs of network-reserved cars. Therefore, the reasonable configuration of vehicles is also related to many of the above indicators. Therefore, four factors with higher significance are selected here. As the basis for analysis, there are: per capita income, kilometers traveled by Internet-linked vehicles, total number of passengers carried by Internet-linked vehicles per month, and per capita consumption capacity of citizens.

Since the subsidies for ride-hailing vehicles are dynamically changing, the analysis of relevant indicators is performed on a monthly basis. The indicator data is as follows in table 3:

In order to consider whether the above indicators have influences, whether the problem of

"difficult to ride a car" can be effectively solved, but the degree of priority impact cannot be judged, so each indicator is processed by gray correlation analysis to facilitate later judgment.

Table 3 Changes in subsidies.

months	Monthly average number (Times)	rides every month (Km)	Average monthly consumption(Yuan)	Average monthly income (Yuan)	Internet ride subsidies (yuan)
2	5138	23585	2164	3128	8
3	6435	45635	2365	4104	10
4	5375	47258	2915	3853	12
5	4625	23658	2389	3883	5
6	6485	46857	2358	3864	7
7	5782	36474	2206	4013	5
8	5433	33945	2265	4165	2

## 4. Empirical Analysis

### 4.1. Index selection

Use SPSS statistical software to perform principal component analysis on it, remove the indicators with relatively small correlation, and then perform another principal component analysis to obtain its principal component. The obtained principal component is predicted by the multiple linear regression prediction method for each year of the city The demand of online car rental, and finally, analyze the deviation between demand and supply.

After processing the data according to the spss software, it is found that there are significant relationships: the city's resident population, the number of tourists, the number of kilometers traveled by the bus, the monthly passenger capacity of the bus, per capita consumption, per capita income, and total tourism profit. There is a significant difference between the two indicators of road area in Xuzhou and the total investment of ride-hailing vehicles in Xuzhou and other indicators. Therefore, in order to ensure the accuracy of subsequent prediction values, these two indicators are deleted, and the principal component analysis is performed on the processed indicator data again. As a result, it can be found that the sum of the variance principal component extraction tables of the first three eigenvalues has exceeded the overall 91.652. The gradual decrease in the contribution of the variance from the fourth principal component is negligible. Therefore, only the first three principal components are extracted as analysis data.

### 4.2. Demand forecast

After dealing with the relevant index values that affect the car Hailing demand of Xuzhou network, the regression analysis of the model with SPSS tool is shown in Table 4:

Table 4 Forecast result table.

Years	main ingredient 1	main ingredient 2	main ingredient 3	Number of Internet Cars	Forecast	Deviation
2011	20660.1658	5125.2318	-27749.7616	6600	6607	-0.106%
2012	19690.1353	5530.9466	-24338.5664	6664	6662	-0.030%
2013	20538.4534	5646.2324	-25199.4528	6664	6646	0.270%
2014	24291.7653	6602.6432	-30848.6432	6650	6654	-0.061%
2015	26851.4623	7290.6423	-34193.5643	6746	6757	-0.163%
2016	28278.6456	7647.8984	-35613.3135	6848	6767	1.226%
2017	29846.4832	7908.9474	-37103.8411	6851	6878	0.394%
2018	31228.7587	7585.9761	-39448.4184	6851	6757	1.372%

From the analysis results in Table 4, it can be obtained that there is not much difference between the current number of Xuzhou city's network of ride-hailing vehicles and the predicted demand, and the value of the deviation is also low, which is less than 1.5%. In terms of quantity, it can satisfy the daily travel of citizens, as well as other cities like Xuzhou. Therefore, it is speculated that the cause of the "difficult taxi" problem in Xuzhou City is not because there are not enough Internet-based taxis, but because of the unreasonable allocation of Internet-based taxi resources. From the "Data Analysis of Xuzhou City 2017 Online Carpooling Software Data Analysis" statistics, it can be concluded that the empty driving rate of Xuzhou City's online car rental in 2017 was extremely high, reaching 20% -60%. During the more serious period, it was more than normal when it was idle status. However, after the taxi software company has adopted corresponding subsidies, the allocation of ride-hailing resources has been improved, and the no-load phenomenon has also been alleviated. However, the "difficult taxi" phenomenon has not been fundamentally resolved. The following will take the supply and demand analysis and the grey correlation analysis method to explore whether the relevant subsidy is effective and how effective it is.

### 4.3. Analysis of current subsidy policies

The processed data is programmed into the following code through MATLAB software, and a weight table of the impact of each factor on the "difficult to drive" problem is output in table5:

Table 5 Weighting table

Per capita income	Spending power	Kilometers traveled	Impact of the subsidy scheme
0.7780	0.7911	0.8010	0.6066

By analyzing and ranking the weights of relevant influencing factors from Table 5, we can conclude that the per capita income of the citizens is the most influential, followed by the consumption capacity, then the number of kilometers traveled by the network, and the impact of subsidies. Therefore, online car subsidies can not effectively promote the resource allocation of vehicles, but in a sense, certain subsidies can stimulate car owners' desire to take orders, thus playing some roles.

## 5. Conclusions and Recommendations

### 5.1. Conclusions

#### 5.1.1. The real reason of "difficult taxi"

In response to the problem of "difficult to ride a car", various factors and related data affecting the passenger demand of Xuzhou City's online car rental were collected. After conducting a principal component analysis of the data, it was found that the two factors of the investment amount of urban online car rental and the area of urban roads were related to There are significant differences between other factors. After excluding these two factors, the principal factors analysis was performed again for the remaining factors. The first three principal components were used to make a multiple linear regression prediction of the number of ride-hailing cars, and they were compared with the actual ride-hailing cars in Xuzhou. It was concluded that Xuzhou's online car rental can meet normal demand, and the cause of "difficult taxis" is caused by vehicle owners' unwillingness to take orders from remote areas and traffic peaks.

#### 5.1.2. Poor results from existing taxi schemes

After learning about the existing subsidy schemes for online ride-hailing and the configuration of online ride-hailing in each district in Xuzhou City, the grey correlation was analyzed. The relevant weights are ranked as follows: per capita income> per capita consumption> average monthly passenger mileage of Internet-linked vehicles> subsidies for Internet-linked vehicles. Therefore, the existing related subsidy schemes can have a certain effect on the relief of the "difficult to ride a taxi", but they are not ideal. There are certain problems in the resource allocation of various districts

in Xuzhou City.

## 5.2. Recommendations

There is a problem with the fare subsidy schemes that can be implemented through the above analysis. That is, the subsidy for ride-hailing drivers is fixed. In this way, the same amount of subsidy policy for each order will cause a serious network. The owner of the car is singled out.

Most ride-hailing drivers are more willing to go to the central area. It is difficult for drivers in some remote areas to return after receiving orders. So even with the existing subsidy policy, most drivers still do n't like to go to some remote areas, preferring to drive empty. Based on this, a dynamic subsidy scheme for passengers related to the distance from the driver and the remoteness of the destination is established.

In addition to the remote destination of passengers, another more serious phenomenon is the high empty rate of vehicles during the rush hour, which is still one of the important reasons leading to the unreasonable allocation of network-reserved car resources. Therefore, consider separately increasing the subsidy amount during the peak period, and implement two different subsidy schemes for the peak and off-peak periods to stimulate the online taxi driver to receive orders during the peak period.

### 5.2.1 Dynamic subsidy scheme related to remoteness of demand

Therefore, according to the distance from the place where customers want to go to the city center, corresponding subsidies are set up. Set the minimum amount of subsidy as  $a$ , the area near the mobile phone display of a driver is  $I$ , the distance from the target customer's taxi destination to the city center is  $l_i$ , the average speed of the vehicle is  $v$ , and the time required to reach the target

location is  $t_i$ ,  $t_i = \frac{l_i}{v}$ .

Assuming that the randomness of the pick-up selection of the drivers of the network car Hailing

Assuming that the randomness of the pick-up selection of the drivers of the network car Hailing satisfies the distribution of the double index (gunbel), the logit selection probability model can be obtained, and the stimulation of picking up the pick-up orders in remote areas can be considered in reverse. The relevant dynamic subsidy formula is as follows:

$$A_i = a \left( 1 - \frac{\exp(-\frac{l_i}{v})}{\sum_{k \in I} -\frac{l_k}{v}} \right) \quad (1)$$

Through this subsidy scheme, it not only helps to alleviate the phenomenon that customers are difficult to take a taxi, but also helps to alleviate the phenomenon that the number of online car hailing in the central area is excessive and the supply exceeds the demand, so that the online car Hailing resources can be better allocated. However, the implementation area of this subsidy policy should remove the tourist attractions, stations, airports and other exceptions in remote areas.

### 5.2.2. Subsidies for peak periods

During the peak period, if there is a traffic jam, the actual cost is

$$R = (l_1 - l_2) w_1 + l_2 w_2 - p_t L - p_t t_j - f \quad (2)$$

Where  $p_t$  indicates the probability of traffic jam in peak period;

$l_1$  indicates the distance traveled by the vehicle;

$w_1$  Indicates the fuel cost when there is no traffic jam;  $L$  indicates the total distance;

$w_2$  Indicates fuel consumption in traffic jam;  $p_t$  indicates mileage cost;

$p_i$  Represents the waiting fee;  $f$  represents the subsidy amount.

In fact, in general, ride-hailing drivers generally only consider whether the extra cost can make up for the loss of income, so the above expression is simplified to:

$$R_d = l_2 w_2 - f \quad (3)$$

If the amount of subsidy  $f$  for online car hailing in peak period is enough, the required cost  $R_d$  will be negative. Under the stimulation of money, the enthusiasm of online car Hailing drivers can be mobilized, which indirectly and effectively alleviates the problem of taking a taxi.

In addition to the monetary subsidy schemes in the above two different situations, you can also consider using virtual points to solve the problem of taxi difficulties, so that those owners who are willing to receive tickets from remote areas and traffic jams will receive corresponding monetary subsidies while obtaining corresponding Virtual points. The higher the virtual points, the easier it is for owners to receive ordinary orders. This virtual points subsidy scheme can alleviate the pressure of excessive monetary subsidies on the company to a certain extent, and can also largely motivate drivers and increase their dependence on taxi software. As a result, the problem of "difficult taxi" can be solved better.

### Acknowledgements

The project was supported by the Fujian Social Science Fund General Project (2014B092), Fujian Science and Technology Plan Soft Science Project (B15013, B15013), Key Research Project of Science and Technology Think Tank of Fujian Science and Technology Association (A1610) and Major Decision Consulting Project of Fujian Provincial Government Development Research Center (2018JCZX04)

### References

- [1] Pan Liang, Cai Xiaowei, Zhang Wenjuan. Research on Taxi Resource Allocation in the "Internet +" Era. *Journal of Mudanjiang Teachers College: Natural Science Edition*, 2016 (3): 18-21.
- [2] Guan Jinping, Zhu Zheng. Study on the Spatio-temporal Characteristics and Causes of Taxi Driving Based on FCD. *Journal of Sun Yat-sen University: Natural Science Edition*, 2010, 49: 29-36.
- [3] Wang Shujia, Shen Zenghong, Gong Xiang. Measurement and comparison standards of urban taxi demand. *Special Economic Zone*, 2009 (7): 270-271.
- [4] Yang Haoxiong, Wei Bin. Study on the Competition Game between Online Car Booking and Taxi——Taking Platform Subsidies as the Background. *Beijing Social Science*, 2016, (5): 68-76.
- [5] Douglas, G.w. Price regulation and optional service standards: The taxicab industry. *Journal of Transport Economics and Policy* 1972, 20: 116-127.
- [6] Li Xiaochun. Research on Taxi Resource Allocation in the "Internet +" Era. *Soft Science*, 2019, 30 (7): 1-2.
- [7] Cui Yuzhao, Cao Yongzhao, Wang Mengxuan. Solving the problem of taxi resource allocation in the Internet + era based on factor analysis. Kunming: North China University of Technology 2017.
- [8] Fu Anrui, Dai Hanhui. Evaluation model of taxi subsidy scheme based on game theory. *Contemporary Economy*, 2017 (34): 2-3.