YADI WANG, Ph.D.[1]
E-mail: 18113050@bjtu.edu.cn
XIAOCHUN LU, Ph.D.[1,2]
(Corresponding author)
E-mail: xclu@bjtu.edu.cn
YANG WANG, M.D.[1]
E-mail: 16120588@bjtu.edu.cn
[1] School of Economics and Management,
  Beijing Jiaotong University
  No.3, Shangyuan Village, Haidian District, Beijing, China
[2] Beijing Logistics Informatics Research Base
  No.3, Shangyuan Village, Haidian District, Beijing, China

EVALUATING THE MARKET SHARE OF CHINA RAILWAY EXPRESS LINES – A DATA-DRIVEN APPROACH

ABSTRACT

Fuelled by the Belt and Road Initiative, Eurasia railway transport has gained rapid traction. However, China Railway Express is in the development period, information about China Railway Express lines is in chaos, and it is difficult to appraise the market situation. This paper focuses on providing an approach to estimate the market share of every China Railway Express line. In this paper, the crawled data from the website are applied to estimate the customer demands for China Railway Express services in different areas of China. The government subsidy is a factor that cannot be ignored, but its level is unclear. Thus, a dummy regression model was established to estimate the subsidy. The regression result is in line with the data released by the Guangdong Provincial Government in 2017. To identify customers' preferred choices for particular lines, a multi-objective optimization model has been built. With the crawled demands data and this optimization model, the current and future market share of China Railway Express can be assessed. If the subsidy is cancelled, some China Railway Express lines will lose their market, and only three lines have a bright future.

KEY WORDS

China Railway Express; economic choice; international transit hub;

1. INTRODUCTION

The Belt and Road Initiative has greatly accelerated the Chinese economy due to economic co-operation across Europe, Central Asia, and China. Logistics across Europe and China is vital and has resulted in the development of the China Railway Express (CR Express) [1], which was inaugurated in Beijing on May 26, 2017, and has since been used by many companies, as shown in *Figure 1*. The Chinese government announced that, by December 2017, a total of 5,000 trains had traversed China and Europe [2]. However, there is chaos under the cover of prosperity. Since the CR Express is a general term of use, these railway lines are not centralized, and therefore, accessing statistics on the progress of this programme has proven difficult, and this difficulty has resulted in the lack of data on the market situation.

Research has been conducted on the current situation of the CR Express to clarify the situation. Lin [3] pointed out that the CR Express achieved impressive results in the past few years. Wang [4] analysed the factors influencing the development of CR Express and put forward suggestions on developing CR Express. Some researchers confirmed that the government subsidy resulted in an entanglement in CR Express lines [5]. Bi [6] and Zhang [7] suggested that for the CR Express to improve efficiency, it could eliminate subsidies and instead use economies-of-scale, though they did not say how much subsidy is provided by the government. Xu [8] noted that the government subsidies are still an important support for the normal operation of China-Europe trains. The key to the removal of subsidies in the future lies in the comprehensive benefits that China-Europe trains provide, but no quantitative analysis has been conducted. A survey in 2016 showed that the CR Express subsidy could be worth
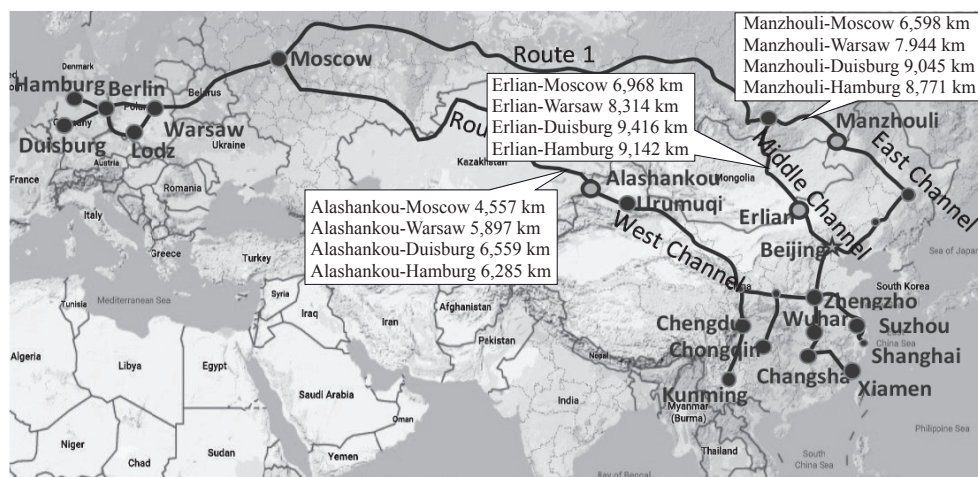
Figure 1 – CR Express Routes

7,000 USD/TEU, but these data have a significant difference compared with the subsidy published by the Guangdong government, which is 2,173 USD/TEU.

There are dozens of cities for CR Express companies to choose as a hub; the hub is applied as a dry port to undertake the functions of transhipment and cargo reloading. So, some researchers study the "Belt and Road" problem from the railway company perspective, which suggests the need for the development of a logistics centre. Cullinane and Wilmsmeier [9] stress the significance of dry ports, claiming that dry ports will be applied more and more in the future. Martins de Sá et al. [10] focus on the hub line location problem and put forward an exact algorithm based on a Benders decomposition of a strong path-based formulation. Ishfaq and Sox [11, 12] address the strategic planning for an interregional, hub-based, intermodal logistics network operated by a logistics service provider and use taboo search meta-heuristic to solve this problem. Roso [13] and Hanaoka and Regmi [14] study the dry port problem from the economic and environmental perspective, and discuss factors that influence the location of a dry port in Asia. Zheng et al. [15] design environment-friendly distribution of inland nodes using a bi-objective inland nodes location model. Yang and Chen [16] explore global logistics hub and build the assessment criteria, then they compare three logistic hubs in Northeast Asia using the criteria. As for the Western countries, Limbourg and Jourquin [17] considered the p-hub problem in the European countries and draw the conclusion that the optimal locations for the European transfer terminals are embedded in a hub-and-spoke network. The researcher Dong suggested that

Xi'an inland port needs to be an international transit centre to actualize the CR Express programme [18, 19]. Another researcher, Wang et al. [20], stated that the CR Express can only reduce its running cost and improve its efficiency by building a hub. Jiang et al. [21] performed a study to analyse the suitable location for a hub using a mixed integer model where they suggested the hub should be located in Urumchi. After studying four CR Express lines, a team of two researchers, Fu et al. [22], stated that building a transit centre in Zhenghou would reduce the logistics cost by 15% per container.

Although some studies have examined the marketing strategy of CR Express, few papers have provided a data analysis. For instance, Chen and Shi [23] suggested an integrated CR Express brand by establishing an industrial strategic union. Ye et al. [24] also proposed with a Less-Than-Container Load (LCL) logistics plan to solve the problem of few routes, low frequency, high logistics cost and an empty return trip rate.

To analyse the market statistics on the client's demand is vital; however, no one in China has described these demands. Interestingly, these demands have been described with web crawler technology using recruitment data published by CR Express freight forwarders, which is shown in *Figure 2*. This information can help analyse the business volume in different regions. To analyse the CR Express market, the inevitable need for a government subsidy becomes clear. This paper provides a regression model to estimate the subsidy, and its result is close to data provided in 2017 by the Guangdong local government. In contrast to some papers, it was found that the subsidy has ebbed with the introduction of the CR Express, and this result is used in

Figure 2 – Map of released recruitment information related to the CR Express business

order to analyse the market situation. This paper also provides a portfolio for market selection, which allows a client to choose the best CR Express line. This model can be used in different regions to calculate the market situation of the CR Express.

2. CUSTOMER DEMAND DISTRIBUTION ESTIMATION

To evaluate the market situation of the CR Express, gathering information on the clients' demands is vital, although this has proven difficult. Despite the China Railway Company being liable for the CR Express logistics, it has never published each line's statistics. Freight forwarders have their own customer data, but it is impossible to access the data. Recently, many companies have invested in the CR Express business. These companies hire employees through a recruitment website, and this recruitment information could be used to analyse the current level of customer demand.

Between February 2018 and July 2018, 370 recruitment samples were obtained after downloading 543 records from a Chinese website, (http://www.51job.com), as shown in *Table 1* and *Figure 2*. *Table 1* shows that the highest recruitment is concentrated in Guangdong, Shanghai, and Sichuan, and it is in these regions that customers' demand is high. We will later use the percentage of recruitment as the distribution of customers' demands.

3. DUMMY REGRESSION MODEL USED FOR SUBSIDY ANALYSIS

Understanding each CR Express line subsidy is vital, since price defects might be caused by government subsidies. In addition to the Guangdong local government, which announced through its finance department that the subsidy for the CR Express was 87,217,750 CNY (about 12.83 million USD) no other local government had done so. These data are not enough to evaluate the average subsidy of CR Express.

We obtained from a CR Express freight forwarder the logistics cost and the distance to destination to evaluate the subsidy as listed in *Table 2*. The prices in *Table 2* are classified into two classes; Lines 1-11 are class 1, and they have a subsidy; and Lines 12-15 are classified as class 0, and they do not have a subsidy; therefore, variables 1 and 0 are used as the CR Express classes. It is assumed that there is a regression model as follows:

$$y_{price} = \beta_0 \cdot d_{class} + \beta_1 \cdot x_{distance} \tag{1}$$

where y_{price} is a dependent variable representing the freight logistics cost from China to Europe, and $x_{distance}$ is an independent variable representing

Table 1 – Recruitment information statistics

Province ID	Province/City	Recruitment released	%	Province ID	Province/City	Recruitment released	%
1	Guangdong	126	34.05	11	Tianjin	9	2.43
2	Shanghai	55	14.86	12	Henan	7	1.89
3	Sichuan	40	10.81	13	Shandong	7	1.89
4	Chongqing	25	6.76	14	Guangxi	5	1.35
5	Beijing	18	4.86	15	Fujian	3	0.81
6	Zhejiang	18	4.86	16	Gansu	3	0.81
7	Jiangsu	15	4.05	17	Xinjiang	3	0.81
8	Hunan	11	2.97	18	Liaoning	2	0.54
9	Hubei	10	2.70	19	Yunnan	2	0.54
10	Shaanxi	10	2.70	20	An'hui	1	0.27

distance, while d_{class} is a dummy variable, where d_{class}=1, meaning its price has a subsidy, and it has no subsidy when the variable is d_{class}=0.

From the regression model (*Equation 1*), coefficient β_0 is the CR Express subsidy while β_1 is the unit price per kilometre; therefore, the subsidy value of the CR Express can be obtained, which is an important factor in China's market. The data from *Table 2* are applied to the dummy linear regression of the model (*Equation 1*), but the results show collinearity between variable $x_{distance}$ and d_{class} which are eliminated using a ridge regression model.

A ridge regression model is used to reduce the regression errors:

$$\frac{1}{n}\| Y - X\beta \|^2 + \alpha \| \beta \|^2 \qquad (2)$$

where n is the number of independent variables, which equals 15; Y is the vector form of y_{price}; X is the vector form of d_{class} and $x_{distance}$; β is the vector form of β_0 and β_1; α is the hyper-parameter.

The objective is to find the value of the hyper-parameter α, and there are the ridge score and MSE curve under different α values as shown in *Figure 4*.

From *Figure 4*, if α<0.2, then MSE is lower, and the score is higher. Therefore, we set α=0.1, and we obtain the coefficient of the model (*Equation 1*) as: β_1=0.4882, β_0=-1,719.18. By cross-validation, the model MSE=83,476.85, RMSE=288.92, R^2= 0.93 is obtained.
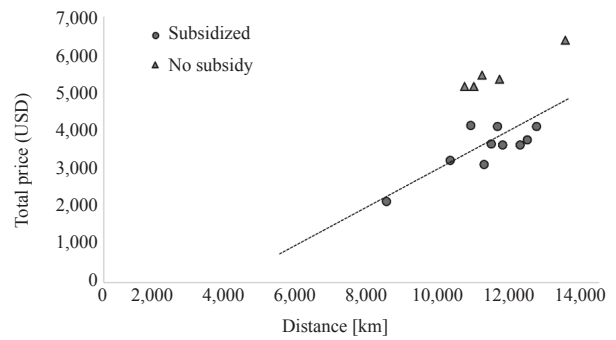

Figure 3 – The scatter plot of total price and distance of the CR Express lines
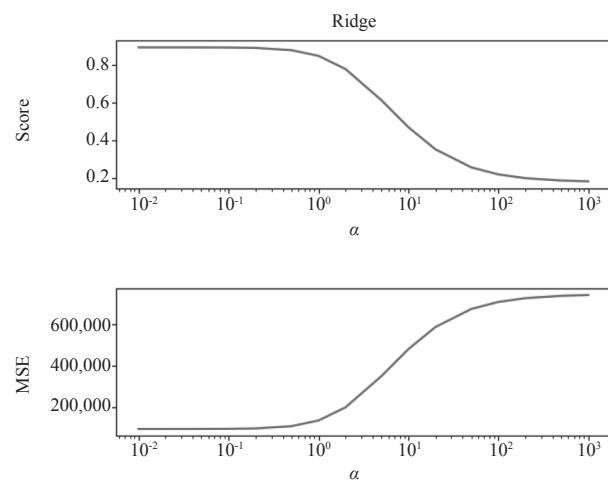
Figure 4 – Score and MSE curves of the ridge regression with different alpha values

Table 2 – Price of CR Express [25]

Line ID	Name	Original city (China)	Border port (China)	Destination city	CR Express price [a] y_{price}	Distance [b] [km] $x_{distance}$	Class d_{class}
1	Haou	Haerbin	Manzhouli	Hamburg	3,191	9,820	1
2	Hexinou	Hefei	Alashankou	Hamburg	4,118	10,403	1
3	Lanxinou	Lanzhou	Alashankou	Hamburg	2,103	8,027	1
4	Rongou	Chengdu	Alashankou	Lodz	3,088	9,826	1
5	Shenmanou	Shenyang	Manzhouli	Hamburg	3,632	11,000	1
6	Xiangou	Changsha	Erlian	Hamburg	3,603	11,305	1
7	Xiangou	Changsha	Manzhouli	Hamburg	3,603	11,808	1
8	Xiarongou	Xiamen	Alashankou	Lodz	3,719	12,000	1
9	Yuxinou	Chongqing	Erlian	Duisburg	4,066	12,268	1
10	Yuxinou	Chongqing	Alashankou	Duisburg	4,066	11,179	1
11	Sumanou	Suzhou	Manzhouli	Warszawa	5,353	11,200	0
12	Hanou	Wuhan	Manzhouli	Hamburg	5,456	10,700	0
13	Zhengou	Zhengzhou	Erlian	Hamburg	5,147	10,484	0
14	Zhengou	Zhengzhou	Alashankou	Hamburg	5,147	10,214	0
15	Yixinou	Yiwu	Alashankou	Madrid	6,382	13,052	0

[a] *The above shipping prices are the 40 ft container market prices (including subsidy), unit: USD/TEU; the exchange rate of USD and CNY is 1 USD=6.8 CNY;* [b] *The data of distance are from the official website of each line.*

From the regression result, β_0 is the subsidy value, which is 1,719.19 USD/TEU.

The regression result was confirmed by an announcement by the CR Express Subsidy Project Plan in 2017 by the Guangdong local government, which stated that the total subsidy was 87.22 million CNY (approximately 12.83 million USD). The Chinese government estimated that approximately 144 CR Express trains traversed Europe from Guangdong Province, with each train transporting 41 containers [26]. Therefore, the subsidy of each container is approximately 2,173.00 USD/TEU, which has a difference of 454 USD/TEU from our regression result, translating to a 21% error. The regression result is reliable, since it highlights the whole subsidy.

4. LINES CHOICE MODEL OF CR EXPRESS

The customer's market choice is necessary to evaluate the market share. As so many CR Express lines are operational, it is difficult to estimate which line will be popular. Therefore, a multi-objective optimization model has been developed to try to find which line will be chosen by a customer, so that the CR Express lines market share can be evaluated.

Model assumption

Assume that freight originates from City A in China via railway station B in China to city C in Europe. Therefore, the container transits along route A to B and finally C by the CR Express, as shown in *Figure 5*. The following are the assumptions for the model:
1) Only the capital cities of the provinces in China are considered, while other small cities are not considered;
2) The unit price of a container per kilometre for road transport is the same, but it is different for the CR Express lines.
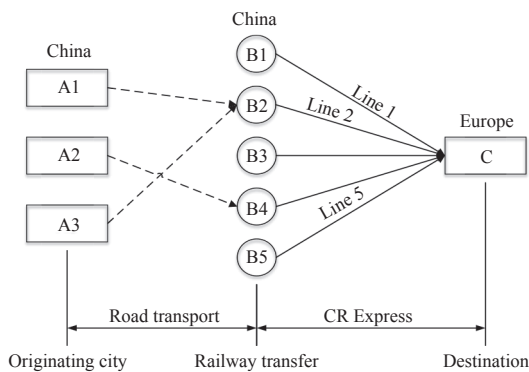

Figure 5 – Transport route diagram

Optimization modelling

From the customer's perspective, the goal is to receive the freight in the shortest time possible at a lower cost. The model is obtained as *Equation 3*. In *Equation 3*, $f(i,j,k)$, represents a customer's choice. The freight transits from the original city (i) by a trailer to a railway station (j) and then to the destination (k). From *Equation 3*, c_r is the trucking cost of TEU/km, while d_{ij} is the distance between the original city (i) and the railway station (j); c_{jk} is the total cost from the railway station (j) to the destination city (k) in Europe; d_{jk} is the distance between railway station (j) and destination (k) and finally y_{ij} is the binary variable.

$$\begin{cases} f(i,j,k) = \arg\min \sum_{j=1}^{m} y_{ij}\left[d_{ij} \cdot c_r + d_{jk} c_{jk} + \alpha \cdot w_j\right] \\ \sum_{j=1}^{m} y_{ij} = 1, \ y = 0 \text{ or } 1 \\ d_{ij} > 0, \ rfc_{jk} > 0, \ c_r > 0, \ c_{jk} > 0, \ w_j > 0, \ \alpha > 0 \\ i = 1, 2, \dots, n; \ j = 1, 2, \dots, m; \ k = 1, 2, \dots, u; \end{cases} \quad (3)$$

In this model, coefficient α represents the time cost per day that can be derived from *Table 2*, and w_j represents the average waiting time for a train from the railway station (j).

By means of stochastic theory, w_j can be calculated if it is assumed that the train schedule is a process. It is assumed that $\{N(t), t \geq 0\}$ is the process of train (N) to arrive at time (t), which is a continuous cycle. As a result, random variable X_n is noted as the interval time between the $(n-1)$-th and the train (n). Let us assume $E(X)$ is the mean of X, $Var(X)$ is its variance, $W(t)$ is the waiting time for the next train and $E(W)$ is its mean. $W(t)$ is also referred to as the excess time.

From renewal process theory [27], $E(W)$ is obtained, that is:

$$E(W) = \lim_{x \to \infty} \frac{\int_0^x W(t)dt}{x} = \frac{E(X^2)}{2E(X)} = \\ = \frac{E(X)}{2}\left[1 + \frac{Var(X)}{E(X)^2}\right] \quad (4)$$

Using *Equation 4*, the container waiting time (w_j) at railway station (j) can be calculated.

5. OPTIMIZATION MODEL SOLUTION

Before solving *Equation 3*, some variables or parameter values need to be identified; their values are w_j, α, c_{jk}, c_r, d_{jk}, d_{ij}.

The w_j is one of the variables in *Equation 3*. This variable is the mean waiting time of a container at a railway station. By the train timetable of every CR Express line that is listed in *Table 3*, w_j can be calculated according to *Equation 4*.

For example, in *Table 3*, the freight train timetable for Rongou (Line ID=4) is Tuesday, Wednesday (a.m.), Wednesday (p.m.), Thursday, Friday, Saturday, Sunday. Random variable X in *Equation 3* (the interval time between trains) is [2, 0.875, 0.25, 0.875, 1, 1, 1] day(s). The mean interval time $E(X)=1.0$ is also obtained as well as the variance of interval time $Var(X)=0.26$. By *Equation 3*, the mean waiting time $E(W)=0.63$ day can be obtained. The mean waiting time of all CR Express lines can be calculated in this way. The calculation result is listed in *Table 3*.

In *Equation 3*, coefficient α is the waiting cost of a container in a railway station; it is an opportunity cost to evaluate the customers' waiting loss. In this paper, the railway freight price is divided in *Table 2* by the train running time in *Table 3*. Then, an average value of transport cost per day can be obtained. Let this value be α. Finally, $\alpha=281$ USD/day is obtained.

The c_{jk} in *Equation 3* is the railway freight unit price. This variable can be directly calculated by *Table 2*.

To analyse the CR Express market without subsidy support, the unit price that does not contain a subsidy needs to be determined. The subsidy value was previously obtained by using dummy regression. This value is 1,719 USD. Besides, the CR Express Lines 1 to Line 10 in *Table 2* are supported by subsidies. Therefore, if 1,719 USD is added to these lines freight price, a new unit price with no subsidy can be calculated. All the calculation results are listed in *Table 4*, and it can be found from the table that if the subsidy is excluded, the railway freight unit price is approximately 0.5 USD/km. This value is close to the market price.

In this paper, the road transport unit price for a container is $c_r=10.5$ RMB/km (1.54 USD/km). This price is quoted from the freight forwarder.

The customer demand was previously discussed, and in this paper, *Table 1* reflects the distribution of customer demand. Therefore, the data of variable d_{ij} are the distance from origin city (i) in *Table 1* to transfer railway station (j) in *Table 2*. In this paper, the origin city only refers to the capital city of a province. For ease of description, the transfer railway station name is used instead of the CR Express line ID.

The d_{jk} is the distance from the transfer railway station (China) to the destination city (Europe). In *Table 2*, the destination city of the CR Express line is different. Most frequent destination city is Hamburg, but Rongou's (Line ID=4) destination is Lodz, and its transport volume is much higher than of the other lines. To solve *Equation 3*, Warsaw is set as the study destination city. The distance to Warsaw of the CR Express line is shown in *Table 5*.

Some CR Express lines are subsidized by local governments. These lines can operate at lower prices. The model is solved and an attempt is made to determine the market share with the subsidy of CR Express.

*Table 3 – Train timetable of CR Express lines**

Line ID	Line name	Timetable	Waiting time mean (days)	Train running time (days)
1	Hamanou	1 train/week	3.50	15
2	Hexinou	2 trains/week	2.39	15
3	Lanxinou	2 trains/month	7.50	13
4	Rongou	7 trains/week	0.63	14
5	Shenmanou	1 train/week	3.50	14
6	Xiangou	3 trains/week	1.67	18
7	Xiangou	1 train/week	3.50	18
8	Xiarongou	1 train/week	3.50	15
9	Yuxinou	3 trains/week	2.31	17
10	Yuxinou	7 trains/week	0.50	14
11	Sumanou	1 train/week	3.50	14
12	Hanou	3 trains/week	1.67	16
13,14	Zhengou	7 trains/week	0.85	13
15	Yixinou	1 train/week	3.50	18

* *Data are from the China Railway Cooperator official freight service web site: http://hyfw.95306.cn/gateway/gjly/DzswGlyd/action/HomeIndex*

Table 4 – Railway freight unit price per kilometre

Line ID	Line name	Unit price c_{jk} [a]	Unit price c'_{jk} [b]	Line ID	Line name	Unit price c_{jk} [a]	Unit price c'_{jk} [b]
1	Haou	0.32	0.50	9	Yuxinou	0.33	0.47
2	Hexinou	0.40	0.56	10	Yuxinou	0.36	0.52
3	Lanxinou	0.26	0.48	11	Sumanou[*]	0.48	0.48
4	Rongou	0.29	0.45	12	Hanou[*]	0.51	0.51
5	Shenmanou	0.33	0.49	13	Zhengou[*]	0.49	0.49
6	Xiangou	0.32	0.47	14	Zhengou[*]	0.50	0.50
7	Xiangou	0.31	0.45	15	Yixinou[*]	0.49	0.49
8	Xiarongou	0.31	0.45				

[*] *These CR Express lines are not subsidized. The unit prices of these lines are unchanged.*
[a] *The unit price contains subsidy. Unit: USD/km*
[b] *The unit price does not contain subsidy. Unit: USD/km*

Table 5 – Distance to Warsaw

Line ID	CR Express line name	Distance to Warsaw [km]	Line ID	CR Express line name	Distance to Warsaw [km]
1	Haou	8,993	9	Yuxinou	11,166
2	Hexinou	10,015	10	Yuxinou	10,517
3	Lanxinou	7,639	11	Sumanou	11,200
4	Rongou	9,705	12	Hanou	9,873
5	Shenmanou	10,173	13	Zhengou	9,656
6	Xiangou	10,477	14	Zhengou	9,826
7	Xiangou	10,981	15	Yixinou	11,457
8	Xiarongou	11,879			

The model solution is listed in *Table 6*. This result is the customer's best choice under current situation. The CR Express Lines 11, 12, 13, 14, 15, which do not receive any subsidy are excluded from *Table 6*.

If the CR Express lines were all unsubsidized, the market share would be different. By *Table 4*, and the regression result of Section 3, the model without subsidy can be solved. The unit price of container c'_{jk} in *Table 4* needs just to be used to obtain the optimization solution, and the result is shown in *Table 7*.

6. CR EXPRESS MARKET SHARE ANALYSIS

After completing the previous work, the market share of CR Express can be evaluated. *Equations 5-7* are calculation formulas.

$$customer\_choice_{ij} =$$
$$= (recru\_pcnt_i) \cdot (y_{ij}) \cdot (total\_customers) \tag{5}$$

$$demand_j = \sum_{i=1} customer\_choice_{ij} \tag{6}$$

$$market\_share_j = \frac{demand_j}{total\_customers} \tag{7}$$

In *Equations 5-7*, i is province (or city) ID listed in *Table 1*. The $recur\_pcnt_i$ is the percentage of recruitment released by province (i), data of $recur\_pcnt_i$ are listed in *Table 1*.

Variable y_{ij} is the binary variable obtained from *Equation 3*. If y_{ij}=1, it means the customer in province (i) will choose CR Express line (j) as the best solution; the data of y_{ij} are listed in *Table 6*. For example, in *Table 6*, the first row data means: the customer in Province 3 will choose CR Express Line 4, so here y_{34}=1. Thus, the variable $customer\_choice_{ij}$ means the number of customers in province (i) who will choose CR Express line (j).

In this way the demand of CR Express line (j) can be estimated with variable $demand_j$ in *Equation 6* and also its market share can be evaluated by *Equation 7*.

Here, a data example is given: Suppose there are 1,000 customers (*total_customers*=1,000) in China who will transport their cargoes to Europe by CR Express. According to *Equation 5*, we look up in *Table 1* to get variable $recur\_pcnt_i$. Assume that we look up Guangdong for the recruitment percentage (Province ID=1) in *Table 1* and get $recur\_pcnt_1$=34.5%.

Table 6 – CR Express line choice of different provinces in China (with subsidy)

Customer's province	Province ID	CR Express line selected by customer	Line ID	Total cost
Sichuan	3	Rongou	4	2,952.73
Chongqing	4	Rongou	4	3,455.15
Shaanxi	10	Rongou	4	4,128.34
Henan	12	Rongou	4	4,846.42
Shandong	13	Rongou	4	5,505.88
Guangxi	14	Rongou	4	4,857.99
Yunnan	19	Rongou	4	4,258.42
Xinjiang	17	Rongou	4	1,972.52
Guangdong	1	Xiangou	6	4,833.77
Zhejiang	6	Xiangou	6	5,163.96
Hunan	8	Xiangou	6	3,807.44
Hubei	9	Xiangou	6	4,343.70
Beijing	5	Shenmanou	5	5,401.61
Tianjin	11	Shenmanou	5	5,377.44
Liaoning	18	Shenmanou	5	4,342.44
Shanghai	2	Hexinou	2	5,355.16
Jiangsu	7	Hexinou	2	4,899.93
An'hui	20	Hexinou	2	4,636.80
Gansu	16	Lanxinou	3	4,108.85
Xinjiang	17	Lanxinou	3	7,106.34
Fujian	15	Xiarongou	8	5,064.43

Table 7 – CR Express line choice of different provinces in China (without subsidy)

Customer's province	Province ID	CR Express line selected by customer	Line ID	Cost (USD)
Shanghai	2	Zhengou	13	6,432.88
Beijing	5	Zhengou	13	6,045.75
Zhejiang	6	Zhengou	13	6,408.01
Jiangsu	7	Zhengou	13	6,002.81
Tianjin	11	Zhengou	13	6,063.22
Henan	12	Zhengou	13	4,979.79
Shandong	13	Zhengou	13	5,679.34
An'hui	20	Zhengou	13	5,853.42
Hubei	9	Hanou	12	5,502.64
Guangdong	1	Xiangou	6	6,426.86
Hunan	8	Xiangou	6	5,400.54
Fujian	15	Xiangou	6	6,730.97
Liaoning	18	Shenmanou	5	5,932.20
Sichuan	3	Rongou	4	4,497.44
Chongqing	4	Rongou/Yuxinou	4/10	4,999.86/ 5,582.92
Shaanxi	10	Rongou	4	5,673.05
Guangxi	14	Rongou	4	6,402.70
Yunnan	19	Rongou	4	5,803.13
Gansu	16	Lanxinou/Rongou	3/4	5,744.76/ 6,026.49
Xinjiang	17	Lanxinou/Rongou	3/4	8,742.25/ 9,014.93

Next, by *Equation 3*, it can be found that the customers in Guangdong (i=1) choose CR Express Line 6 (j=6) and can get the lowest cost, that is: y_{16}=1 is the optimization result of *Equation 3*. This result is listed in the 9th row in *Table 4*. So, by *Equation 5*, the *customer_choice*$_{16}$=345 is known. *Table 6* shows that all the customers from Zhejiang (Province 6), Hunan (Province 8) and Hubei (Province 9) will also choose Line 6 as their best choice (y_{66}=1, y_{86}=1, y_{96}=1). In *Table 1*, we know *recur_pcnt*$_6$=4.86%, *recur_pcnt*$_8$=2.97%, *recur_pcnt*$_9$=2.70%, and then by *Equation 6 demand*$_6$=445 is obtained.

Finally, using *Equation 8*, the market share of Line 6 is evaluated at 44.5%, which is shown in *Table 8*. By this analogy, the market share of each line can be calculated and the results are shown in *Table 8*.

In *Table 8*, the optimization solution of a customer in Gansu Province (ID 16) or Xinjiang Province (ID 17) is to choose CR Express Line 3 (Lanxinou). However, this interval time of Line 3 is 2 weeks (it is listed in *Table 3*). Therefore, most often the customer will choose the second optimization solution, CR Express Line 4 (Rongou).

Finally, using this method, the CR Express market share in China without subsidy is also obtained, and it is listed in *Table 9*.

From *Table 9*, it is known that the CR Express Lines 6, 13, and 4 will occupy a large share of the market. The CR Express Line 10 will occupy a smaller share of the market if the government subsidy is lost. However, current Line 10 has a great impact on CR Express because it has created a great deal of publicity and it is one of the earliest operating CR Express lines. However, *Table 8* and *Table 9* show that the market prospects of this line are pessimistic. Geographically, this line is too close to CR Express Line 4, and they have become competitors. The total price of CR Express Line 10 is 1,000 USD higher than that of Line 4, so it cannot compete with Line 4. Except for CR Express Lines 4, 6, and 13, other lines will have to compete for smaller market share, similar to Line 10.

7. CONCLUSION

As a carrier for the implementation of the "One Belt One Road" strategy, the international container transport channels have gained more attention. As the representative, the CR Express has garnered support nationwide. However, because of the lack of operational data, the market share of the CR Express lines is unclear to the public. In this paper, three problems have been solved.

The first problem is how to estimate the customer demand of CR Express. Currently, it is not possible to obtain customer demand data through public channels. In this paper, the CR Express job information issued on a recruitment website was crawled, and in this way, the demands estimation problem was solved. Of course, this demand estimation is not accurate, but it provides a method of gaining preliminary data. From the data analysis results, it was noticed that three provinces have the highest recruitment, and it can be considered that the customers' demand in these regions is much higher.

The second problem is to estimate the government subsidy of CR Express. The officially released data on the CR Express subsidy are relatively limited. In this paper, a dummy regression model was built to estimate the subsidy, and the obtained value is 1,719.19 USD/TEU. This result is reasonable and it is basically consistent with the data released by the Guangdong Province in 2017, so this value could be used in our later market analysis.

The third problem is to estimate the market share of different CR Express lines. It is difficult to estimate the market share when you cannot get the survey data. In this paper, an optimization model was built trying to help different customers to choose their best CR Express lines. In this way, it will be known which line will be more popular on the market. Combining the recruitment data and the opti-

Table 8 – CR Express market share in China under the current situation

Line ID	CR Express line name	Market share [%]
6	Xiangou	44.58
4	Rongou	25.94
2	Hexinou	19.18
5	Shenmanou	7.83
3	Lanxinou	1.62
8	Xiarongou	0.81

Table 9 – CR Express market share in China without subsidies

Line ID	CR Express line name	Market share [%]
6	Xiangou	37.83
13	Zhengou	35.11
4	Rongou	20.28
10	Yuxinou	3.38
12	Hanou	2.70
5	Shenmanou	0.54
3	Lanxinou	0.12

mization model, the market share of CR Express can be calculated. This analysis shows that only three CR Express lines occupy a large market share among dozens of lines, and most CR Express lines may face this market dilemma.

In our model, only the departure stations of each line were considered. In the future, due to data limitation, this optimization model can be improved by collecting detailed data about the shipping stations of each CR Express line in the hope of improving our model.

ACKNOWLEDGEMENT

王亚迪，博士[1]
18113050@bjtu.edu.cn
鲁晓春，博士*[1,2]
xclu@bjtu.edu.cn
王杨，硕士[1]
16120588@bjtu.edu.cn
[1] 经济与管理学院
  北京市海淀区上园村3号，中国，北京
[2] 北京物流信息化研究基地
  北京市海淀区上园村3号，中国，北京

基于实证分析及客户选择模型的中欧班列国内市场份额研究

摘要：

在"一带一路"的影响下，中欧班列运输随之迅速发展。然而，中欧班列正处在发展期，很多班列的运营数据、信息难以获取，对中欧班列的市场状况更加难以评估，基于实证分析与运筹学优化选择模型相结合的方法，对中欧班列各条线路的市场份额进行了研究。首先通过爬虫在网页上获取了中欧班列的相关数据，以此来评估国内不同地区客户对中欧班列的运输需求。因中欧班列价格受政府补贴影响较大，因此构建了虚拟变量回归模型估计补贴额度，回归结果与2017年广东省政府公布的补贴数据基本吻合。另外，以客户的视角，建立了客户对班列的最优选择模型，进一步利用爬取的数据，得到目前中欧班列的市场份额分布估计。最后，利用回归结果，假设政府补贴全部取消后，估计了中欧班列市场份额的分布。

关键词：中欧班列；经济选择；国际转运中心

REFERENCES

[1] China Railway Corporation. *Development Plan for Railway express (2016-2020)*. National Development and Reform Commission, China.

[2] Qin H-H, Qin S. Goods for China-Europe Block Train Based on Freight Value. *Railway Freight Transport.* 2016;34(8): 13-19.

[3] Lin B-Z. The Operational Situation and Suggestions for Sino-Euro Block Trains. *Port Economy.* 2017;4: 27-30.

[4] Wang Y-B. Discussion on Construction and Development Planning of China Railway Express. *Railway Transport and Economy*. 2017;39(1): 41-45.

[5] Li F-J. Countermeasures of optimizing China-Europe block trains under "the Belt and Road" Strategy. *Railway Transport and Economy*. 2016;38(5): 41-45.

[6] Bi G-T. The Development Situation of China-Europe Block Trains under "the Belt and Road" Strategy and Suggestions. *China Market*. 2017;22: 121-122.

[7] Zhang C, Jiang R. Development Status and Strategy Analysis of China-Europe Block Trains Based on SWOT. *China Market*. 2016;41: 148-149.

[8] Xu Y-M. Discussion on the Development Path of China-Europe Train from the Perspective of "One Belt and One Road" Strategy. *Southwest Finance*. 2015;10: 70-73.

[9] Cullinane K, Wilmsmeier G. The Contribution of the Dry Port Concept to the Extension of Port Life Cycles. In: Böse JW. (ed.) *Handbook of Terminal Planning*. Springer Science+Business Media; 2011.

[10] Martins de Sá E, Contreras I, Cordeau J-F, Saraiva de Camargo R, de Miranda G. The Hub Line Location Problem. *Transportation Science*. 2015;49(3): 500-518.

[11] Ishfaq R, Sox CR. Intermodal logistics: The Interplay of Financial, Operational and Service Issues. *Transportation Research Part E: Logistics and Transportation Review.* 2010;46(6): 926-949.

[12] Ishfaq R, Sox CR. Hub Location–allocation in Intermodal Logistic Networks. *European Journal of Operational Research*. 2010;210(2): 213-230.

[13] Roso V. Factors Influencing Implementation of A Dry Port. *International Journal of Physical Distribution & Logistics Management*. 2008;38(10): 782-798.

[14] Hanaoka S, Regmi MB. Promoting intermodal freight transport through the development of dry ports in Asia: An environmental perspective. *IATSS Research*. 2011;35(1): 16-23.

[15] Zheng C, Zhuang Q, Lü J. Research on bi-objective inland port location problem with the perspective of environmental constraints. *Journal of Systems Engineering*. 2014;29(5): 662-671.

[16] Yang Y-C, Chen S-L, Determinants of global logistics hub ports: Comparison of the port development policies of Taiwan, Korea, and Japan. *Transport Policy*. 2016;45: 179-189.

[17] Limbourg S, Jourquin B. Optimal rail-road container terminal locations on the European network. *Transportation Research Part E: Logistics and Transportation Review.* 2009;45(4): 551-563.

[18] Dong Q-L, Yang L, Chang X-H. Research on China-Europe Trains Integrated Operation Based on International Transit Hub Port Strategy Theory. *Science and Technology Management Research*. 2016;36(22): 230-236.

[19] Dong Q-L. The Advantage, Condition and Realization Way of International Transit Hub Strategy under the Background of the Belt and Road Initiative. *China Business and Market*. 2017;31(2): 46-54.

[20] Wang J-E, Jing Y, Wang C. Study on Better Organization of China-Europe Express Train. *Bulletin of Chinese Academy of Sciences*. 2017;32(4): 370-376.

[21] Jiang X-Y, Wang X-L, Zhu D-L. Research on the Problem of Site Selection of Logistics Infrastructure under the "One Belt and One Road" Development Strategy. *Shanghai Management Science*. 2015;20(7): 2694-2697.

[22] Fu X-P, Zhang Y-X, Wan H-H. Study on Economy Optimization of China—Europe Block Train Corridor in Central West of China. *Railway Transport and Economy*. 2017;39(4): 26-30.

[23] Chen R, Shi G-X. Reflections on creating "One Belt and One Road" China-Europe train brand. *Railway Transportation and Economy*. 2015;37(6): 71.

[24] Ye Y-C, Zhu D, Wang X. Study on Consolidation Pattern of Sino-European Rail Freight. *Shanghai Management Science*. 2014;36(6): 1-7.

[25] Li N-Y. The foundation of China Railway Express alliance under the background of "One belt and One Road". *Urban Mass Transit*. 2016;5: 73.

[26] Belt and Road Portal. *2017 China-Europe Express Report: 3271 Trains in 35 Cities Will Turn to "High Quality" in the Future*. Available at: https://www.yidaiyilu.gov.cn/xwzx/gnxw/42245.htm [Accessed 27th September 2019].

[27] Ross SM. *Introduction to Probability Models*. 11th ed. Boston: Academic Press; 2014; p. 425.