

# *An analysis of the determinants of China's domestic catering industry revenue—Based on the data of 31 provinces in China from 2006 to 2020*

Peihao Song<sup>1,\*</sup>, Jingxiang Ruan<sup>1</sup>, Jiahui Xu<sup>2</sup>, Aowei Jiang<sup>3</sup>

<sup>1</sup>*School of Mathematics and Physics, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu, 215123, China*

<sup>2</sup>*School of Information Science and Engineering, East China University of Science and Technology, Shanghai, 200237, China*

<sup>3</sup>*School of Management, Shanghai University, Shanghai, 200444, China*

*\*Corresponding author*

**Keywords:** Revenue of catering industry; multiple linear regression; econometric models

**Abstract:** The catering industry in China has witnessed remarkable revenue growth over the years, underscoring its significance within the domestic economy. However, understanding the factors that drive the industry's performance is crucial for its sustained growth and competitiveness. The Stata software is employed for conducting multivariate regression analysis to examine and refine the parameters. Furthermore, recommendations are put forth to enhance the turnover of the catering industry.

## 1. Introduction

The catering industry is considered a crucial component of the domestic economy. According to data from the National Bureau of Statistics (2021), its revenue increased from 4864.01 billion in 2015 to 6557.38 billion in 2019.

The performance of the catering industry is influenced by both external and internal factors. An article on a website mentions that the total population of a province has an impact on the income of the catering industry. Moreover, further information obtained from the Internet illustrates that GDP is the most significant factor affecting the catering industry. Additionally, using a multiple linear regression model, He indicates a positive correlation between the number of catering enterprises and catering industry revenue (2022). Furthermore, Liu point out that the volume of freight transport may affect the revenue to some extent[1]. From Gong's perspective, agriculture and the development of the catering industry are closely linked[2]. The existing literature highlights the impact of factors such as the provincial total population, gross domestic product (GDP), the number of catering enterprises, volume of freight transport, and the total output value of agriculture, forestry, and fisheries. However, a deeper understanding of the relationships between these factors and the catering industry's performance is required to formulate effective strategies and policies for its sustained growth.

## 2. Empirical design and justification

### 2.1. Variable description

The explained variable in this report is catering service industry revenue (CIR). It is selected as the indicator of the domestic catering industry's income. Number of urban population (PTP), provincial gross domestic product (PGDP), number of catering enterprises (NOCE), the volume of freight transport (VOFT), the total output value of agriculture, forestry, and fisheries (TOVAFF) across 31 provinces, municipalities and autonomous regions (Hong Kong, Macao and Taiwan excluded) are selected as the explanatory variables in this report. The definition of variable and measurement of units are shown in Table 1.

Table 1: Definition of Variables

Variable	Definition	Unit
Catering service industry revenue (CIR)	All currency earned by providing catering services of 31 provinces between 2006 and 2020	Yuan in one million
Provincial total population (PTP)	The population of 31 provinces from 2006 to 2020	In one hundred million
Provincial Gross Domestic Product (PGDP)	The gross domestic product of 31 provinces from 2006 to 2020	Yuan
Number of catering enterprises (NOCE)	The number of enterprises providing catering services in the country of 31 provinces from 2006 to 2020	In one
Volume of freight transport (VOFT)	Total volume of freight carried by all routes (land, water and air...) of 31 provinces between 2006 and 2020	Tons in ten thousand
Total output value of agriculture, forestry and fisheries (TOVAFF)	Total volume of all products of agriculture, forestry, livestock and fisheries expressed in monetary terms in the 31 provinces between 2006 and 2020	Yuan in one hundred million

The brief explanation of selected variables is stated below:

#### A. catering service industry revenue (CIR)

Catering service industry revenue represents all currency received by trading food and beverage with related services. Vigorous promotion of the catering industry plays a vital role in improving people's living standards and increasing employment opportunities (Liu, 2021). This report selects the catering service industry revenue of 31 provinces of China over 15 years as the explained variable.

#### B. provincial total population (PTP)

Increased population density is driving the growth of the catering industry. The larger the population base, the greater the demand for the catering industry, and simultaneously the greater the labor force. Liu (2011) points out that the number of populations may affect the revenue to some extent [3].

#### C. provincial gross domestic product (PGDP)

A higher level of GDP means a higher standard of living, which will make consumers consume more in the tertiary industry. It is also demonstrated that GDP is the most significant factor affecting the income of the catering industry.

#### D. number of catering enterprises (NOCE)

In order to increase the catering service industry revenue, the most direct way is to increase the number of catering enterprises. He (2022) uses a multiple linear regression model to indicate that

both consumption levels of residents and the number of catering enterprises have positive correlations with catering industry revenue [4].

#### E. Volume of freight transport (VOFT)

A fast and convenient transportation network is an extremely necessary condition for the catering industry. This ensures the supply source of the catering industry. Quality goods transport will boost the catering industry [5]. It is easy to infer that freight volumes can be used as a criterion for assessing freight.

#### F. Total output value of agriculture, forestry and fisheries (TOVAFF)

Catering is an industry in which a caterer provides food, beverages, and services to a client. The source of food and beverages is agriculture, forestry and fisheries. Agricultural development could affect the supply of raw materials for the catering industry. Sun also claims that the primary sector (agriculture, forestry and fisheries included) has a pivotal influence on the tertiary sector (catering included) [6]. It is feasible to use money as a measure of the total volume of all products from agriculture, forestry and fisheries (National Bureau of Statistics China).

## 2.2. Empirical econometric model

All the variables are in the form of the natural logarithm and the influence of measurement units is omitted so as to test the elasticity influence of independent variables on the explanatory variable. Moreover, to minimize the bias resulting from unpredictable variables that vary with time or cross region, we select the entity and time fixed effects regression model. Therefore, based on the former analysis, a panel regression model is

$$\ln CIR_{kt} = \beta_1 \ln PTP_{kt} + \beta_2 \ln NOCE_{kt} + \beta_3 \ln PGDP_{kt} + \beta_4 \ln TOVAFF_{kt} + \beta_5 \ln VOFT_{kt} + \mu_{kt}$$

Where k represents provinces (k=1,2...31) and t represents time (t=2006,2007....2019).  $\mu_{kt}$  represents the error term. The specific explanations are in Table 1.

## 2.3. Sample and data

The observational data are collected from the National Bureau of Statistics for authority purposes. Considering the lack of relevant data from the previous years, we select the annual data from 2006 to 2020 across 31 provinces, autonomous districts and municipalities (Taiwan, Hong Kong, and Macao excluded). All the samples mentioned are utilized for regression analysis and the descriptive statistics of each determinant are displayed in Table 2.

Table 2: Descriptive Statistics of Variables

Regressor	Obs	Min	Median	Max	Mean	SD
CIR	465	130.000	5380.000	94600.000	12155.054	16489.176
PTP	465	285.000	3823.000	12624.000	4388.127	2826.026
NOCE	465	0.000	423.000	4393.000	718.848	733.265
VOFT	465	348.000	103833.000	434298.000	120655.100	90082.964
PGDP	465	285.900	13587.800	111151.600	19441.421	18788.242
TOVAFF	465	71.400	2442.730	10190.580	2845.520	2252.192

Judging from Table 2, a relatively big gap can be seen between the max and min of CIR and it also features a comparatively high standard deviation. Therefore, the CIR has a large range of volatility. The reason behind this may be that China is in the midst of rapid development and the Reform and Opening up have given China a great impetus, which has led to the rapid development of the catering industry. Additionally, a positively skewed distribution is shown, indicating that there is a noticeable difference among different provinces with regard to the catering industry. Considering VOFT, it

varies considerably, implying that the development of transportation in recent years has had a significant impact on the overall industry in China. Moreover, the rapid development of transportation has promoted the exchange of food in different regions and the spread of advanced catering concepts.

Table 3: Correlations

<b>Variables</b>	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>	<b>(5)</b>	<b>(6)</b>
(1) lnCIR	1.000	-	-	-	-	-
(2) lnPTP	0.414*	1.000	-	-	-	-
(3) lnNOCE	0.728*	0.756*	1.000	-	-	-
(4) lnVOFT	0.382*	0.860*	0.783*	1.000	-	-
(5) lnPGDP	0.694*	0.834*	0.903*	0.847*	1.000	-
(6) lnTOVAFF	0.139*	0.138*	0.104	0.182*	0.256*	1.000
* Shows significance at the .01 level						

Table 4: Variance inflation factor

-	<b>VIF</b>	<b>1/VIF</b>
lnPGDP	9.367	0.107
lnNOCE	6.061	0.165
lnVOFT	4.887	0.205
lnPTP	4.57	0.219
lnTOVAF	1.21	0.827
MeanVIF	5.219	0.827

Table 3 displays the pairwise correlation. Concluding from Table 3 and Table 4, a positive correlation can be found between the explained variable and the NOECE appears to have the highest correlation with CIR. Furthermore, in order to further test whether there is multicollinearity in the model, the variance inflation factors are tested, and it has been demonstrated in Table 4 that all VIFs are less than 10. Consequently, multicollinearity in the model can be omitted.

### 3. Model selection

To account for the reasons why we select the model, we demonstrate it from two aspects: realistic significance and statistical analysis.

From the economic point of view, the model may involve two dominant features. Considering entity effects, we introduce a linear trend in our model and assume some unpredictable factors, such as natural hazards, consumption concepts, and religious beliefs, constant over time [3].

Secondly, the existing component may give us a deeper insight into the model and its implication in real life. Some factors act as important indicators of the development of China's catering industry. For instance, the number of catering enterprises reflects the fierceness of competition in the industry to some extent. He (2022) declared that since the reform and opening up, the catering industry, acting as the key component of the tertiary industry, has witnessed astonishing business expansion. However, according to the annual report of the catering industry (2021), China's catering industry only contributes to approximately 4.7% of the GDP and 10% of the tertiary industry [7]. Additionally, the tertiary industry only accounts for about 50% of GDP, which is merely about 70% of that of the developed country, indicating that the economy is underdeveloped and still needs to be adjusted. During the past decades, with expanding the number of catering enterprises, the structure of the industry is likely to improve.

In order to further determine the validity of the model, we use the stata to test which model should be used, random effect models or fixed effect models.

(1) Hausman test

Hausman test is used to build the panel data regression model, and the results are displayed in Table 5 to illustrate which panel data model is adequate for the subsequent analysis. The testing outcomes show that the p-value is 0, in which case the individual effect null hypothesis at 1% can be rejected. Consequently, we choose the fixed effect model.

(2) Sequence correlation test

The sequence-related test is used to test whether there is a sequence correlation. The results are significant, hence the null hypothesis which is there is no sequence correlation should be rejected.

(3) Heteroskedasticity test

Next is the heterogeneous variance test, the null hypothesis is homoskedasticity, the results are significant, and the p-value is 0 in which case we are able to reject the null hypothesis. Thus, the sample is heteroskedastic at a 5% significant level, according to the heteroskedasticity test's results. The traditional Hausman test will no longer be effective since there is heteroskedasticity exists in our model. Thus, the Robust Hausman test is necessary to illustrate which model should be chosen.

(4) Robust Hausman Test

Based on the test result, the p-value is less than 0.001. Therefore, the null hypothesis should be strongly rejected, which means the fixed effect model should be selected.

Table 5: Hausman test

Variables	(1)	(2)
	FE	RE
PTPln	1.527*** (0.247)	0.103 (0.152)
NOCEln	0.425*** (0.035)	0.431*** (0.037)
VOFTln	0.121** (0.054)	0.022 (0.056)
lnPGDP	0.533*** (0.067)	0.612*** (0.065)
lnTOVAFF	-0.301*** (0.066)	-0.212*** (0.062)
Constant	-10.472*** (1.913)	0.768 (1.176)
Observations	464	464
R-squared	0.816	-
Number of city	31	31
Hausman	64.06	-
p-value	0	-
Standard errors in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

#### 4. Regression results and analysis

The CIR panel regression results (Table 6) also demonstrate that the PTP, NOCE, PGDP and TOVAFF regression coefficients are extremely significant, however, the VOFT is not statistically significant. The estimation's regression results are shown in Table 6. PTP has the greatest impact on CIR, followed by NOCE, TOVAFF and PGDP, all had a positive 5% significance test effect on CIR. The 95% confidence interval (-0.1579065, 0.2142748) contains 0, indicating that there may not be a

significant association between CIR and VOFT. Theoretically, all computed coefficients have the anticipated signs, with the exception that VOFT is not strongly statistically significant. To assess the overall significance of the regressors, the F-test is introduced.

Table 6 shows that the R square is 0.856 which is close to 1. That means the model has an excellent fitting effect. The estimated coefficient of explanatory variables is interpreted as the elasticity of explained variable with respect to explanatory variables. When all other factors are held constant, the elasticity of PTP equals 1.343, which means that for every 1% increase in PTP, CIR will grow by 1.343%. Although PGDP and VOFT have a big impact, their marginal effects on CIR growth are unquestionably less than those of PTP. CIR will increase by 1.06% and 0.0282%, respectively, for every 1% increase in PGDP and VOFT. Moreover, the elasticity of TOVAFF is -0.306, which means for every 1% increase in TOVAFF, CIR will decrease by 0.306%.

Table 6: Regression result

Variables	(1)
	CIRln
lnPTP	1.343*** (0.357)
lnNOCE	0.371*** (0.0631)
lnVOFT	0.0282 (0.0911)
lnPGDP	1.064*** (0.286)
lnTOVAFF	-0.306** (0.142)
Constant	-12.29*** (3.187)
Observations	464
Number of city	31
R-squared	0.856
*** p<0.01, ** p<0.05, * p<0.1	

From an economic standpoint, higher provincial GDP indicates higher purchasing power as well as a higher living standard, implying that people will devote more money to recreational activities, consequently creating the necessary cash flow for the industry. Besides, with a larger provincial population, the consumer group will be expanded, which in turn taps the potential of the catering market and stimulate consumption needs, thus increasing revenue to a large extent. Also, concluding the regression results, the provincial population turns out to have the greatest influence on CIR. Furthermore, the increase in the number of catering enterprises reflects the expanding scale of the catering industry and lead to an increasingly competitive marketing environment, providing more employment position as well. As a result, both the productive force and quality of labor are improved. Meanwhile, offered more choices, people are likely to make a consumption, therefore increasing revenue and improving the structure of a tertiary industry. However, the marginal effects of PGDP and NOCE turn out to be weaker than that of PTP. Furthermore, the total output value of agriculture, forestry and fisheries shows a negative relationship with CIR.

Ultimately, the volume of freight transportation appears to have little influence on the growth of CIR. A possible explanation may be that with the rapid development of China's transportation industry over the past decades, the VOFT has reached a comparatively high level and there is relatively little space to progress.

## 5. Conclusion and Discussion

This report explores the influence of five determinants, provincial total population, provincial gross domestic product, the number of catering enterprises, volume of freight transport and the total output value of agriculture, forestry and fisheries respectively through a panel of 31 provinces and municipalities in China from 2006 to 2020.

Firstly, the fixed-effect model is preferred over the random-effect model because the Hausman test demonstrates the rejection of the null hypothesis. Secondly, the Sequence correlation test shows that the results are significant, hence the null hypothesis which is there is no sequence correlation should be rejected. Thirdly, the Heteroskedasticity test shows that the traditional Hausman test will no longer be effective since there is heteroskedasticity exists in our model. Finally, we use Robust Hausman Test which helps us to choose the fixed effects model. Based on this, a specific multiple non-linear regression model has been established. The regression results have indicated that, at a 5% significance level, there is strong evidence that growth in the Provincial total population and Provincial Gross Domestic Product, an increase in the Number of catering enterprises and a Volume of freight transport do help boost the Catering service industry revenue as expected. Among these, the Provincial total population is the most influential. However, the Volume of freight transport does not significantly affect the Catering service industry revenue at a 5% significance level.

From an economic perspective, higher provincial GDP signifies greater purchasing power and a higher standard of living, which leads to increased consumer spending on leisure activities and generates the necessary cash flow for the catering industry. Additionally, a larger provincial population expands the consumer base and stimulates consumption demand, resulting in substantial revenue growth. Consequently, it is recommended that the government continue to invest in supporting the development of catering enterprises and implement relevant regulations to improve market standards.

However, it is important to acknowledge the limitations of this study. The dataset used in this report may not be entirely comprehensive, as only statistics related to the selected influential factors and 31 provinces were utilized. Some earlier years' data were not available, and the dataset itself may have limitations. Additionally, one observation in the sample exhibited high variance inflation factor (VIF), indicating a potential issue of multicollinearity. Furthermore, there might be omitted variable biases behind the selected factors, as other unaccounted variables could also play a role in the macroeconomic landscape.

In conclusion, this report offers valuable insights and recommendations for the development of the catering industry in China by conducting an economic analysis of the determinants of industry revenue. However, readers should consider the limitations mentioned in this study and take into account other factors and data sources when applying the findings in practical situations.

## References

- [1] Liu Xiaojuan, Yin Yixiao, Zhao Shanshan. *Analysis on the development relationship between the Transportation Industry and the tertiary Industry* [J]. 2010.
- [2] Gong S. *What does the recovery of the catering industry mean for the 'three rural areas'*. 2022.
- [3] Liu M. *Research on China's Catering Industry Development Level and its Influential Factors*. 2011.
- [4] He C. *An econometric analysis of turnover in China's restaurant industry*. *Industrial Innovation*. 2022, 22(14), pp. 117-119.
- [5] Fitch Solutions Group Limited. *China Logistics & Freight Transport Report-Q4 2022*. 2022. (Accessed: 4 October 2022).
- [6] Sun M. *Econometric Analysis on the Relationship between the Primary Industry and the Tertiary Industry in Shandong Province*. *Journal of Heze University*, 2007, 29(2), pp. 19-23. doi: 10.16393/j.cnki.37-1436/z.2007.02.038.
- [7] Liu B. *The Effect of Informal Economy on Employment in China—An Input-Output Analysis*. *Statistical Research*. 2021, 38(2), pp. 88-98.