Empirical Analysis on the Influencing Factors of Shanghai Residents' Tourism Consumption Level

Xixi Hu

School of Management, Shanghai University, Shanghai China

Abstract. In recent years, with the economic level of our country and the continuous development of people's consumption concept, residents' tourism consumption has become a popular trend of consumption. This paper is based on the tourism consumption related data of Shanghai residents in 1996-2015, selecting four main influencing factors of Shanghai third industry gross product, consumer price index, Engel coefficient and traffic cost. Through the establishment of econometric model, the paper did Multiple linear regression analysis of OLS using R software, getting the estimation of the corresponding parameters. Then for the model, the paper did Multiple collinear test, heteroscedasticity and autocorrelation test and correction. Finally, the conclusion is obtained according to the revised model. By studying the characteristics of tourism consumption level and related data, we find that the third industry gross output and Engel coefficient of Shanghai have a significant impact on the consumption level of Shanghai residents. In order to make some contributions to the economic development of our country, we put forward some countermeasures and suggestions on the development of Shanghai's tourism.

1 Introduction

Since the reform and opening up, with the rapid development of China's economy and the continuous improvement of people's living standards, people pay more and more attention to leisure and entertainment, thus, tourism has become a popular consumer trend. With the improvement of people's awareness of leisure and entertainment, tourism consumption is also improving. Especially in some first-and second-tier cities in China, tourism has become a part of residents' lives, and tourism consumption has become an important part of people's consumption expenditure. Therefore, it is of great practical significance to analyze the influencing factors of Shanghai residents tourism consumption for promoting the development of tourism industry.

By consulting the relevant literature in recent years, it is found that the current domestic scholars' research on tourism consumption is basically concentrated on the following aspects: from the perspective of consumption demand to study the impact of residents' income level, consumption preference and other factors on the level of tourism consumption, or from the perspective of supply to study China's transportation conditions, festivals. The influence of holiday time and other factors on tourism consumption level [1]. Scholars Jiang Ronghua and Zhou Jiuhhe used the grey relational analysis method to compare the calculation results between 1994-2000 and 2001-2008. The results show that the main factors affecting domestic tourism income are cities and towns. Per capita household disposable income, annual average wage of workers and per capita gross domestic product [2]. Scholar Li Xuewei, in the "Analysis of Influencing Factors of Domestic Urban Residents Tourism Consumption", analyzes the per capita disposable income of urban residents, per capita GDP of urban residents and virtual variables of leisure time. From the angle of equality, this paper studies the impact of tourism consumption on urban residents, and draws a relevant conclusion: the per capita tourism consumption of urban residents in China does not account for a high proportion of per capita disposable income, tourism consumption tendency is not obvious, it is necessary to strengthen residents' awareness of tourism consumption expenditure [3]. From the above, we can see that as tourism becomes more and more popular leisure mode, the factors affecting residents' tourism consumption are increasingly concerned by all walks of life.

The innovation of this paper is that, for the first time, we choose the GDP of tertiary industry and Engel's coefficient as the main influencing factors to study the impact of Shanghai's GDP of tertiary industry, consumer price index, Engel's coefficient and transportation cost on the tourism consumption of Shanghai residents.

2 Theoretical model

2.1 Selection of influencing factors

X1: Development level of service industry. First of all, the tourism consumption level of residents will be
affected by the development level of service industry. The gross domestic product of the tertiary industry determines the development level of the service industry. Generally speaking, the higher the development level of the service industry, the more recreational activities of the residents, and the smaller the tourism consumption.

X2: Consumer price index. Secondly, from the demand theory analysis, the commodity price will affect the consumption demand, generally speaking, the higher the product price, the lower the tourism demand, the less tourism consumption.

X3: Engel coefficient. Moreover, residents’ consumption habits and consumption consciousness will also affect the level of tourism consumption. With the continuous improvement of living standards, the focus of residents consumption has been shifting from material consumption to enjoyable consumption. Tourism is a typical representative of enjoying consumption. When residents strengthen their consumption awareness of tourism consumption, the level of tourism consumption will increase.

X4: Transportation cost. Finally, because of the exotic nature of tourism, the transportation cost is an important aspect of the impact of tourism costs, generally speaking.

2.2 Model establishment

This paper will use the per capita tourism expenditure of Shanghai to represent the level of tourism consumption, and the gross domestic product of Shanghai's tertiary industry to represent the level of development of Shanghai's service industry. The statistical index of Shanghai residents consumption habits is expressed by Engel's coefficient, which refers to the proportion of food expenditure in total consumption expenditure. The smaller a family's income, the greater the proportion of household income or total expenditure on food purchases. As household income increases, household income or expenditure on food purchases will decline. Engel coefficient is an important index to measure the degree of family wealth. The smaller Engel coefficient indicates that the higher the living standard of the household, the more inclined the consumption habits and consumption consciousness of the residents to spiritual consumption. The data of per capita traffic cost in Shanghai represent traffic consumption. Because of the rich variety and different prices of tourism products in China, there is no uniform statistical index to reflect the price of tourism products, so this paper takes the per capita tourism expenditure of Shanghai residents as the explanatory variable Y, and explains the GDP of Shanghai's tertiary industry, consumer price index of Shanghai residents, Engel coefficient and transportation fee respectively. Variables X1, X2, X3 and X4 were analyzed and multiple regression models were used to analyze them.

3 Empirical analysis

3.1 Collection and sorting of sample data

This paper selects the time series data of 1996-2015 downloaded from the official websites of Shanghai Statistical Bureau and China Statistical Bureau. Based on the above data, this paper builds multiple regression models:

\[ Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4 + \epsilon_i \]

Among them, \( b_0 \) is a constant term; \( b_1, b_2, b_3, b_4 \) are the parameters of explanatory variables of \( X_1, X_2, X_3, X_4; \) \( \epsilon_i \) is a random disturbance term.

<table>
<thead>
<tr>
<th>Year</th>
<th>Y</th>
<th>X1</th>
<th>X3</th>
<th>X4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>676</td>
<td>1292.11</td>
<td>50.70%</td>
<td>106.1</td>
</tr>
<tr>
<td>1997</td>
<td>682</td>
<td>1592.74</td>
<td>51.70%</td>
<td>115.6</td>
</tr>
<tr>
<td>1998</td>
<td>687</td>
<td>1855.36</td>
<td>50.60%</td>
<td>103.7</td>
</tr>
<tr>
<td>1999</td>
<td>825</td>
<td>2129.60</td>
<td>45.20%</td>
<td>112.4</td>
</tr>
<tr>
<td>2000</td>
<td>887</td>
<td>2486.86</td>
<td>44.50%</td>
<td>110.9</td>
</tr>
<tr>
<td>2001</td>
<td>934</td>
<td>2728.94</td>
<td>43.40%</td>
<td>98.1</td>
</tr>
<tr>
<td>2002</td>
<td>1046</td>
<td>3038.90</td>
<td>39.40%</td>
<td>96.9</td>
</tr>
<tr>
<td>2003</td>
<td>1104</td>
<td>3404.19</td>
<td>37.20%</td>
<td>96.3</td>
</tr>
<tr>
<td>2004</td>
<td>1263</td>
<td>4097.26</td>
<td>36.40%</td>
<td>96.5</td>
</tr>
<tr>
<td>2005</td>
<td>1377</td>
<td>4776.20</td>
<td>35.90%</td>
<td>97.5</td>
</tr>
<tr>
<td>2006</td>
<td>1476</td>
<td>5508.48</td>
<td>35.60%</td>
<td>97.3</td>
</tr>
<tr>
<td>2007</td>
<td>1726</td>
<td>6821.11</td>
<td>35.50%</td>
<td>96.9</td>
</tr>
<tr>
<td>2008</td>
<td>1940</td>
<td>7872.23</td>
<td>36.60%</td>
<td>97.5</td>
</tr>
<tr>
<td>2009</td>
<td>2099</td>
<td>8930.85</td>
<td>35.00%</td>
<td>97.5</td>
</tr>
<tr>
<td>2010</td>
<td>2320</td>
<td>9833.51</td>
<td>33.50%</td>
<td>97.4</td>
</tr>
<tr>
<td>2011</td>
<td>2510</td>
<td>11142.86</td>
<td>35.50%</td>
<td>100.2</td>
</tr>
<tr>
<td>2012</td>
<td>2625</td>
<td>12199.15</td>
<td>36.80%</td>
<td>100.8</td>
</tr>
<tr>
<td>2013</td>
<td>2816</td>
<td>13785.45</td>
<td>31.20%</td>
<td>100.4</td>
</tr>
<tr>
<td>2014</td>
<td>3052</td>
<td>15275.73</td>
<td>31.00%</td>
<td>100.1</td>
</tr>
<tr>
<td>2015</td>
<td>3182</td>
<td>17022.63</td>
<td>30.60%</td>
<td>97.6</td>
</tr>
</tbody>
</table>

3.2 Parameter estimation

According to the time series data from 1996 to 2015 and the initial regression model, the parameters of the model are estimated by using R software.

The regression model obtained is as follows:

1.342e-05

\[ Y=-252.302-41.442X1+32.324X2-157.154X3+5.731X4+\epsilon_i \]

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Coefficient</th>
<th>t test</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>-41.442</td>
<td>2.477</td>
<td>0.0256</td>
</tr>
<tr>
<td>X2</td>
<td>32.324</td>
<td>0.904</td>
<td>0.3801</td>
</tr>
<tr>
<td>X3</td>
<td>-137.154</td>
<td>8.467</td>
<td>4.25e-07</td>
</tr>
<tr>
<td>X4</td>
<td>5.731</td>
<td>0.286</td>
<td>0.7786</td>
</tr>
<tr>
<td>Constant term</td>
<td>-252.302</td>
<td>0.661</td>
<td>0.5158</td>
</tr>
</tbody>
</table>

R²=0.8741

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Coefficient</th>
<th>t test</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
<td>26.04</td>
<td>1.34e2-05</td>
<td>1.34e2-06</td>
</tr>
</tbody>
</table>

The results show that the parameters of X1 and X3 are negative, indicating that although the tertiary industry can promote the development of tourism to a certain extent, but the development of other leisure services on the restriction of tourism consumption is greater; the lower
Engel coefficient, the higher people's tourism consumption. The parameters of X2 and X4 are positive, which are conform to the economic significance and the actual theoretical analysis.

Goodness of fit: $R^2=0.8741$, adjusted $R^2=0.8406$, indicating that the goodness of fit of the model can also explain the variables of Shanghai’s tertiary industry gross domestic product, consumer price index, Engel coefficient and transportation costs for Shanghai’s per capita tourism consumption level as high as 84.06%.

T-test: When the saliency condition is 0.05, the significance of the corresponding explanatory variable can be judged by comparing the P value with our preset 0.05 (the original hypothesis of our test is whether the coefficient is significant 0, $P < 0.05$ rejects the original hypothesis, that is, the corresponding variable is not significant 0), we can see that both X1 and X3 can be considered. The T test of x2 and X4 is not significant, indicating that there may be multiple collinearity problems.

F-test: On the significance level of 0.05, F-statistic is often used to judge the significance of the equation as a whole. P value is 1.342e-06, obviously less than 0.05. We can think that the equation at the level of $P = 0.05$ or through the significance test, that is, all variables are connected. The combination has a significant impact on the model[4].

3.3 Model test and modification

3.3.1 Unit root test and correction are made for residual sequences generated by regression

In order to prevent pseudo regression, this paper carries out DF unit root test on the residual sequence of the model[5], $t=0.003642$, less than 0.05, at this time reject the original hypothesis, there is no unit root, through the test does not need to carry out unit root correction.

3.3.2 Multiple-collinearity test and modification.

Multiple collinearity tests were conducted by R software. The results are as follows:

<table>
<thead>
<tr>
<th>Table 3. Multiple collinear regression results.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>X1</td>
</tr>
<tr>
<td>X2</td>
</tr>
<tr>
<td>X3</td>
</tr>
<tr>
<td>X4</td>
</tr>
</tbody>
</table>

By observing the correlation coefficients between interpreted variables and interpreted variables and between interpreted variables and interpreted variables, we can find that the coefficients between X1 and X4, X4 and X3 are greater than 0.4, which can judge the existence of multiple collinearity in the model. The model revision was done step by step in R, and the results are as follows:

![Table 4. The first Multiple collinearity correction.](image)

<table>
<thead>
<tr>
<th>Table 5. The second Multiple collinearity correction.</th>
</tr>
</thead>
<tbody>
<tr>
<td>START: AIC=235.26</td>
</tr>
<tr>
<td>$y=x_1+x_2+x_3+x_4$</td>
</tr>
<tr>
<td>RSS</td>
</tr>
<tr>
<td>-X1</td>
</tr>
<tr>
<td>-X2</td>
</tr>
<tr>
<td>-X3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6. The third Multiple collinearity correction.</th>
</tr>
</thead>
<tbody>
<tr>
<td>START: AIC=234.37</td>
</tr>
<tr>
<td>$y=x_1+x_3$</td>
</tr>
<tr>
<td>RSS</td>
</tr>
<tr>
<td>-X1</td>
</tr>
<tr>
<td>-X2</td>
</tr>
</tbody>
</table>

The result shows that when X1, X2, X3, X4 are used as the coefficient of regression equation, the value of AIC is 237.16. The AIC value of X4 regression equation is 235.26, and the AIC value of X2 equation and X4 regression equation is 234.37. Since removing X2 and X4 can minimize AIC, R automatically removes X2 and X4; removing X2 and X4 increases AIC values, and the stepwise regression analysis terminates, resulting in the current optimal regression equation.

The equation is regressed again. The results of the regression analysis show that the parameters of X1, X3 have changed, but the trend has not changed much, which is consistent with the economic significance and actual theoretical analysis. By using step function to eliminate the variables X2 and X4, from the above correlation analysis, we can get the correlation between the interpreted variables and the interpreted variables. It is known that the correlation between X2 and X4 and the interpreted variables is low.

The revised model is: $y=-3683.33-47.81X1-155.10X3$

The goodness of fit of the modified model is: $R^2=0.8662$, adjusted $R^2=0.8505$. It shows that the goodness of fit of the model can be used now, and the explanatory variables, such as the development level of service industry and the Engel coefficient of residents, can explain more than 85% of the per capita tourism consumption level of Shanghai residents. At the significant level of 0.05, the F-test and t-test of the revised model passed the test at this time, indicating that the equation is significant on the whole, and the explanatory variables of Shanghai's tertiary industry GDP and Engel's coefficient have obvious influence on the explanatory variables of per capita tourism expenditure.

<table>
<thead>
<tr>
<th>Table 7. The second OLS regression results.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanatory variable</td>
</tr>
<tr>
<td>X1</td>
</tr>
<tr>
<td>X3</td>
</tr>
<tr>
<td>Constant term</td>
</tr>
<tr>
<td>$R^2=0.8662$</td>
</tr>
<tr>
<td>F-statistic</td>
</tr>
</tbody>
</table>
3.3.3 Heteroscedasticity test and modification.

In this paper, we use the BP method to test whether the new model has heteroscedasticity. Using R software, the results are as follows:

<table>
<thead>
<tr>
<th>Table 8. Results of heteroscedasticity test.</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP</td>
</tr>
<tr>
<td>df</td>
</tr>
<tr>
<td>P value</td>
</tr>
</tbody>
</table>

The p value of the concomitant introduction is 0.02292, which is less than 0.05 of the significance level, indicating that the original hypothesis of the same variance is not accepted, indicating the existence of heteroscedasticity. Therefore, it is necessary to correct the heteroscedasticity of the revised model.

Using R software to carry out weighted heteroscedasticity for heteroscedasticity correction, the revised model is:

\[ y = 3722.183 - 53.168X1 - 157.064X3 \]

At this time, the significance of X1 and X3 has been greatly improved, and goodness of fit \( R^2 = 0.9856 \), adjusted \( R^2 = 0.9839 \). It shows that the goodness of fit is very good at this time. The explanatory variables, such as the development level of service industry and the Engel coefficient of residents, can explain the per capita tourism consumption level of Shanghai residents as high as 98%. At the significant level of 0.05, the F-test and t-test of the revised model passed the test at this time, indicating that the equation is very significant on the whole at this time. The explanatory variables of Shanghai service industry development level, Engel coefficient on the explanatory variables per capita tourism expenditure is also obvious. The heteroscedasticity test is performed again, and the p value is 0.9704 at 0.05 at that time.

<table>
<thead>
<tr>
<th>Table 9. The third OLS regression results.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explanatory variable</td>
</tr>
<tr>
<td>X1</td>
</tr>
<tr>
<td>X3</td>
</tr>
<tr>
<td>Constant term</td>
</tr>
<tr>
<td>( R^2 = 0.9856 )</td>
</tr>
<tr>
<td>F-statistic</td>
</tr>
</tbody>
</table>

3.3.4 Autocorrelation test

In this paper, DW test is used to test in R. By analyzing the test results and looking up the DW distribution, we can see that the result Du < d < 4 - du is not rejecting the original hypothesis at this time, and the concomitant probability p value is 0.2084, which is higher than the significant level 0.05. This shows that the modified model has no autocorrelation, so it is not necessary to revise the autocorrelation.

4 Conclusions and recommendations

By studying the characteristics of tourism consumption level in China and relevant data, it is found that Shanghai's tertiary industry gross domestic product and Engel coefficient have a significant impact on the tourism consumption level of Shanghai residents. In order to improve the tourism consumption of Shanghai residents, we draw some recommendations from the above conclusions.

(1) Efforts should be made to improve the consumption structure of Shanghai residents.

The proportion of tourism consumption expenditure of Shanghai residents in leisure and entertainment is increased. Although our national standard of living has been greatly improved, the third industry is also developing. The service industry has made great progress, but the consumption structure of leisure and entertainment is still not very reasonable. Although there is no need to worry about food, clothing and transportation, but due to the development of the Internet, people's leisure and entertainment activities are increasingly inclined to stay at home, rather than go out. This also explains why the consumption of tourism will decrease with the development of the third industry. Therefore, the relevant government and units should take necessary measures to improve the consumption structure of residents, make it more reasonable and promote the development of tourism in China.

(2) Efforts should be made to enhance the awareness of tourism consumption of Shanghai residents.

Encourage people to "go out" for entertainment rather than relying solely on Internet or electronic media. Relevant departments can introduce some policies to guide residents tourism consumption, such as scenic spots can strengthen the publicity of scenic spots to attract people's interest in tourism, thus stimulating tourism consumption. It also provides people with more healthy entertainment choices.

(3) Strive to improve the price system of tourism products.

With the continuous development of China's economy, the consumer price index is also growing, and the rise in prices to a certain extent also has an impact on the tourism consumption of residents. Relevant departments should take measures to control the price of tourism products, such as reducing the admission tickets to scenic spots, so as to stimulate people to have the impulse to travel, so as to promote people to travel consumption.

References
