Analysis of Factors Affecting Domestic Tourism Income Based on Multiple Linear Regression

Shoudu Xie[a], Fei Zhou[b] and Xuanfeng Xu[c]

School of Shanghai University, Shanghai 201800, China

[a]2452095767@qq.com, [b]Zhouhuihui2014@163.com, [c]xuxuanfeng123@126.com

Keywords: Multiple linear regression; Multicollinearity; Autocorrelation; Tourism Industry

Abstract. This paper analyzes the impact of domestic tourism revenue influencing factors on domestic tourism revenue based on multiple linear regression models. First, perform the stationarity test of each variable, the unit root test, and then find that they are all second-order single-order sequences of the same order. Then the covariate test is carried out on the variables, and the relationship between domestic tourism income and each explanatory variable is obtained, and compared with the actual data. Finally, the direction of improvement of this program can be analyzed, and relevant policy recommendations are proposed for the development of the tourism industry.

Introduction

Background and significance of the topic. Tourism income is an important component of China's domestic income. It reflects the scale and level of tourism in a country and is the main indicator of tourism statistics, which affects China's economic status and economic decision-making in the tourism industry. Therefore, this paper studies the influencing factors of domestic tourism income, and then adopts appropriate measures to promote the growth of domestic tourism income.

Literature review. With the rapid development of China's tourism industry, the relationship between tourism and economic growth has increasingly attracted the attention of researchers. Shen Yujia[1] explored the issue of tourism-led economic development from a qualitative analysis of the theoretical level. Yan Min[2] used the input-output analysis method to study the relationship between tourism and economic development. Zhang Min's[3] Gnarger-causal test on the relationship between tourism and national economic development concluded that there is a two-way causal relationship between tourism and national economic development. Wu Guoxin[4] directly used the simple data analysis such as regression analysis to explore the pulling effect of tourism development on China's economic growth. Wang Weiguo[5] and Xu Yong[6] et al. analyzed the tourism marginal propensity, the narrow Chinese tourism revenue multiplier and the generalized Chinese tourism income multiplier from the tourism multiplier. Many scholars have studied the tourism revenue multiplier.

Qualitative analysis and multiplier analysis are based on a priori theoretical models and concepts. Therefore, there is a lack of theoretical basis for the role of tourism in economic growth, without considering the differences in data and the test of data stability[9]. Therefore, this paper uses the existing data to analyze the data stationarity test and cointegration test with as few a priori assumptions as possible, and studies domestic tourism income by mapping, tabulation, equation fitting, and calculation of feature quantity. The influencing factors, and then propose the corresponding recommended measures.

Theoretical support. The regression method is a widely used and theoretically powerful modeling method. In the regression analysis, the simple and commonly used model is the multiple linear regression model as follows:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon \quad (1)$$

Where X_1, X_2, X_3, X_4 is the independent variable, Y is the dependent variable, β_0 , β_1 , β_2 ,

  DOI: 10.25236/icess.2019.200

β_3 , β_4 is the regression coefficient, and ε is the random error.

In the regression analysis, the regression coefficient is generally estimated by the ordinary least squares method. If the data used is a non-stationary time series, the inferences derived from these statistics may be incorrect, so regression may be a meaningless false regression. Cointegration theory can solve this problem well, because there is a long-term equilibrium relationship between variables, that is, the cointegration relationship, which can eliminate the situation of false regression.

In this paper, a combination of regression analysis and cointegration theory is used to apply the cointegration theory to find the stable equilibrium relationship between variables, eliminate false regression, and establish a multiple linear regression model.

Data sources. The data used this time comes from the China Statistical Yearbook. The data is from 1994 to 2014. A multivariate regression model was established to analyze the impact of domestic tourism revenue, and finally put forward our suggestion.

The domestic travel rate of urban residents is the total number of domestic tourists of urban residents divided by the total number of urban residents in that year. The domestic travel rate of rural residents is the total number of domestic tourists in rural areas divided by the total rural population in that year. Using GDP, domestic tourists, urban residents' domestic travel rate and rural residents' domestic travel rate as independent variables, domestic tourism income as a dependent variable, using data from 1994 to 2014, using multiple linear regression analysis of each independent variable The effect on the dependent variable.

The Establishment and Testing of the Model

Stationarity test - unit root. Firstly, the unit root test of Y domestic tourism income can be concluded that Y domestic tourism income is a change of trend with time and has intercept items.

Next, do a 0th-order differential test to check whether the original sequence is stable. From the results test, the *Mackinnon* critical values of the unit roots are -0.4532598, -3.673616, -30277364, and the t-statistic value of 0.825975 is greater than the corresponding critical value at the three significance levels of 1%, 5%, and 10%, indicating The domestic tourism income sequence has a unit root and is a non-stationary sequence.

Then continue to make first-order difference and second-order difference for Y domestic tourism revenue. The sequence of the first-order difference of Y domestic tourism income still has a unit root, and after the second-order difference, it is seen from the result test that the *Mackinnon* critical value of the unit root is at the three significance levels of 1%, 5%, and 10%. They are -4.616209, -3.710482, - 3.297799,t- statistic value - 4.128870, a critical value corresponding to a level of significance greater than 1%, and a corresponding critical value of less than 5% and 10% significance level. We believe that the level of significance is below 5%. There is no unit root in the domestic tourism income sequence, which is a stationary sequence. The domestic tourism income of Y is second-order and monotonous, that is Y~I(2).

Similarly, the same unit root test is performed on other independent variables, which are second-order single-integral sequences, namely

$X_1 \sim \mathrm{I}(2)$, $X_2 \sim \mathrm{I}(2)$, $X_3 \sim \mathrm{I}(2)$, $X_4 \sim \mathrm{I}(2)$.

Due to the limitation of the layout, the above inspection process will not be shown here.

Cointegration test. From the above stationarity test, it can be seen that the five variables all have the same order, that is, the second-order single-sequence, so the regression equation can be established to perform the cointegration test. According to the cointegration test theory, variables with the same single integer order are likely to be cointegrated, but whether the variables are cointegrated or not must be co-integrated. This article uses the Engle-grange two-step method: the first step is to use the least squares method for cointegration regression. The second step is to test whether there is a unit root in the residual of the cointegration regression.

Table 1 Multiple returns of Y domestic tourism income

Dependent Variable: Y
Method: Least Squares
Date: 01/07/16 Time: 15:31
Sample: 1994 2014
Included observations: 21

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-992.5769	760.4919	-1.305177	0.2103
X1	0.005473	0.006613	0.827625	0.4201
X2	21.31499	2.940968	7.247610	0.0000
X3	-68164.91	22555.35	-3.022117	0.0081
X4	-147705.4	25131.55	-5.877290	0.0000

R-squared	0.998826	Mean dependent var	8547.681
Adjusted R-squared	0.998533	S.D. dependent var	8729.934
S.E. of regression	334.3610	Akaike info criterion	14.66658
Sum squared resid	1788756.	Schwarz criterion	14.91527
Log likelihood	-148.9991	Hannan-Quinn criter.	14.72055
F-statistic	3404.480	Durbin-Watson stat	1.559977
Prob(F-statistic)	0.000000		

With the domestic tourism income Yt, GDP X_{1t}, domestic tourists X_{2t}, urban residents' domestic travel rate X_{3t}, rural residents' domestic travel rate X_{4t} four variables, and the unit residual test of the residuals.

Table 2 Stationarity test of regression residuals

Null Hypothesis: E1 has a unit root
Exogenous: None
Lag Length: 0 (Fixed)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.484305	0.0014
Test critical values:	1% level	-2.685718	
	5% level	-1.959071	
	10% level	-1.607456	

*MacKinnon (1996) one-sided p-values.

From the above table 6, it can be seen that the regression residuals are at the three significant levels of 1%, 5%, and 10%, and the *Mackinnon* threshold values per unit root are -2.685718, -1.99591, -1.707456, and the t-statistic value is -3.784305. A threshold value corresponding to less than 1%, 5%, and 10% significance levels indicates that the residual sequence does not have a unit root and is a stationary sequence. This indicates that there is a cointegration relationship between domestic tourism income Y_t and GDP X_{1t}, domestic tourists X_{2t}, domestic travel rates of urban residents X_{3t} and domestic travel rates of rural residents X_{4t}. The cointegration regression

equation is as follows:

$$\hat{Y}_t = -992.5769 + 0.005473X_{1t} + 21.31499X_{2t} - 68164.91X_{3t} - 147705.4X_{4t} \quad (2)$$

Conclusions and Policy Recommendations

Conclusion. As seen from the regression process above, it shows that the relationship between variables X_1, X_2, X_3, X_4 and domestic tourism income is significant. However, it can be seen from the regression model that the correlation between X_3, X_4 and domestic tourism income is negative, which is difficult to match with the actual situation.

We believe that this problem may be caused by multiple collinearity and autocorrelation between various explanatory variables. It may be possible to use the variance expansion factor method for multicollinear remedies[10], use the white-test for heteroscedastic remedy, and use the DW test and the BG test for autocorrelation remedies. These methods should be able to fix these problems.

Policy recommendations. According to the analysis of the above model, I believe that the Chinese government should establish a strategic concept of tourism development, and improve tourism infrastructure construction, and strengthen the development and integration of tourism resources. Then, we should expand publicity, develop distinctive tourism brands, and introduce advanced management technologies to integrate with China's tourism industry.

The Reference

[1] Y.J. Shen. Thoughts on Tourism-led Economic Development. Tourism Journal, 2003(6): 21-24.(In Chinese)

[2] M. Yan. The relationship between tourism and economic development level. Tourism Journal, 1999 (5): 9-15.(In Chinese)

[3] H.L. Qi, H. Xiong. Cointegration Analysis of the Relationship between Tourism Income and GDP in China. Journal of Huazhong Normal University(Natural Science), 2014,01:136-141.(In Chinese)

[4] Y.D. Ye. Econometric analysis of domestic tourism income. Corporate Herald, 2012, 04: 98-99.(In Chinese)

[5] C.C. Li. *Research on regional differences in tourism concentration in Xinjiang and its interaction with the economy*. (MS. Xinjiang University, China, 2016).p.22.(In Chinese)

[6] Jacint Balaguer, Manuel Cantavella-Jorda. Tourism as a long-run economic growth factor:The Spanish case.Applied Economics, 2002, 34(7).

[7] Adams P. D., Parmenter B. R. An applied general equilibrium analysis of the economic effects Of tourism in a quite small, quite open economy. Applied Economics, 1995, 27(10): 985-994.

[8] Hacklin. The Tourism-Economy Causality in the United States: A Sub-Industry Level Examination. Tourism Management, 2010, 30(4): 553-558.

[9] Malory. Reflections on the Development of Mining Heritage Tourism in African Tourism Management, 2013, 28(3): 677-687.

[10] J.X. Wang. Research on fiscal policy of tourism development. Dongbei University of Finance and Economics, 2011.01:36-42.(In Chinese)