Analysis of Influencing Factors of Tourism Revenue Based on Multiple Linear Regression Model

Na Li and Gang Fang*
Business School, Beijing Institute of Fashion Technology, Beijing, China

Keywords: Tourism Development, Economic Model, DPI.

Abstract: The report of the 20th National Congress of the Communist Party of China points out that "accelerating the high-quality development of culture and tourism, and constantly meet the people's new expectations for a better life". In the new era, people are more eager to live a better life, and tourism is gradually becoming an important channel to enhance people's sense of well-being. Tourism, as an important part of the tertiary industry, plays a major role in enhancing the overall economy of China. In order to explore the economic factors affecting the development of tourism in China, this paper uses computer technology to conduct numerical simulations to analyse the relationship between the total income of the tourism industry and the per capita disposable income of residents, the number of travel agencies, the number of tourists and other factors in China from 1994 to 2019. Multiple regression analysis, tests and corrections are performed by Econometrics Views. The final model regression results are used to make informative recommendations for the recovery of the tourism industry.

1 INTRODUCTION

In the report of the 20th Party Congress, "promoting cultural self-confidence and self-improvement, forging new glories of socialist culture", it is clearly stated that "insisting on shaping tourism with culture, highlighting culture with tourism, and promoting the deep integration and development of culture and tourism" is an important requirement for the prosperous development of cultural undertakings and cultural industries. The cultural construction in a prominent position, the work of culture and tourism to make important arrangements, fully reflects the Party Central Committee with Comrade Xi Jinping as the core to the cultural construction and tourism development of high importance. Although the epidemic in China has gradually stabilized, there is still a huge gap between the overall development of the industry before the outbreak. Then, the development of the tourism industry in the context of the normalization of the epidemic is also a major issue that the country urgently needs to address at present. By analyzing the factors affecting the development of China's tourism industry, we put forward targeted suggestions for the development of China's tourism industry, thus promoting the healthy development of China's tourism industry and at the same time promoting the development of China's overall economy.

This paper uses the computer software Econometric Views (EViews), which refers to the observation of quantitative patterns of socio-economic relations and economic activities, using econometric methods and techniques. The core of EViews is through designing models, collecting information, estimating models, testing models, and applying models (structural analysis, economic forecasting, and policy evaluation). In this paper EViews processes time series data and performs multiple linear regression using least squares to assign real economic meaning to the regression results. Based on the model regression results, recommendations can be made in the context of the current market environment on the one hand, and forecasting the future development of the industry on the other.

* Corresponding author

Li, N. and Fang, G.
Analysis of Influencing Factors of Tourism Revenue Based on Multiple Linear Regression Model.
DOI: 10.5220/0012029200003620
In *Proceedings of the 4th International Conference on Economic Management and Model Engineering (ICEMME 2022)*, pages 270-275
ISBN: 978-989-758-636-1

2 REVIEW OF THE LITERATURE

Many domestic and foreign scholars have conducted many analyses of the impact of tourism development in recent years. SHI, P. H. et al (2020) selected panel data from 60 tourism cities. Using the double difference method to analyze the impact of tourism model on the level of economic development of tourism economy and constructing a region-wide tourism model can increase tourism investment and promote tourism development in the region. Zhou Li (2019) constructs tourism consumer price index for urban and rural residents in the process of tourism consumption forecasting research found that tourism consumer price index fluctuations in the short term for urban and rural residents have an impact, while long-term fluctuations only have an impact on rural residents. DING Y.F. (2021) selected the tourism data of Anhui Province from 2000-2018 for the development pattern of tourism in Anhui Province, the relationship between tourism revenue and per capita GDP, road mileage, analyzed empirically using EViews, and it was finally obtained that per capita GDP and railroad mileage have a significant impact on domestic tourism in Anhui Province. CAI L.J. (2015) analyzed the data of domestic tourism industry from 1994 to 2013 and modeled that the economic influences that significantly affect the development status of domestic tourism industry are the per capita income of urban and rural residents and the number of domestic population.

It can be seen that many scholars have done a lot of research on the development of the tourism industry, most of them have done research on the development of tourism in a certain provincial area, and there is a great limitation on the number of samples selected. Therefore, this paper selects data from 1994-2019 based on the research of many scholars for the factors influencing the overall level of development of the domestic tourism industry in China, and proposes reference suggestions for tourism development through the results of empirical analysis combined with China's national conditions.

3 CURRENT STATUS OF DOMESTIC TOURISM DEVELOPMENT

In recent years, with the development of China's social economy and the continuous improvement of people's living standards, people's demand for tourism is also increasing, the number of domestic tourism shows a steady increment, and the total income level of the tourism industry is also increasing.

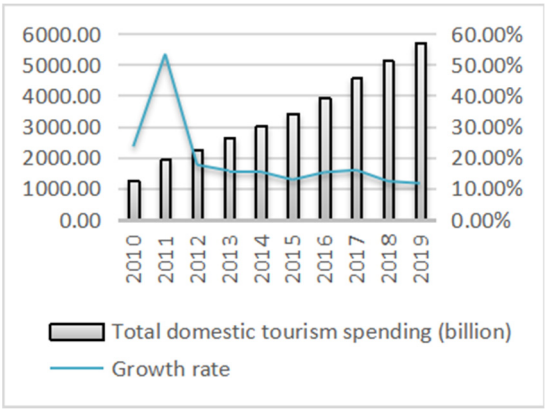

Figure 1: Gross domestic tourism receipts and tourism growth rate in the last decade

According to the analysis of China's tourism industry revenue from 2010-2019, it can be seen that China's tourism revenue has shown a rapid growth trend in the last decade, from 1257.977 billion yuan in 2010 to 5,7250.92 billion yuan in 2019, an increase of about 4.56 times. In 2009, the "Opinions of the State Council on Accelerating the Development of Tourism" for the first time clearly positioned tourism as "a strategic pillar industry of the national economy and a modern service industry that is more satisfying to the people", resulting in a growth rate of 23.53% in tourism revenue in 2010 and up to 53.46% The tourism industry was hit hard by the sudden worldwide epidemic in 2020, with a total domestic tourism revenue of 222.8630 billion yuan, and since then China's tourism industry has entered a moment of silence.

4 ECONOMETRIC ANALYSIS OF DOMESTIC TOURISM REVENUE

The empirical part of this paper is divided into identifying variables, modeling, regression parameter estimation, economic significance and statistical inference

4.1 Model Variables and Modeling

For the analysis of the factors influencing domestic tourism income studied in this paper, the national disposable income, the number of domestic tourists,

and the number of travel agencies are used as influencing variables, so as to explore the main factors of domestic tourism income.

4.1.1 Analysis of Model Variables

The variables are set as in the table 1. Domestic tourism industry revenue is also affected by other unpredictable factors, let's say unquantifiable variables such as consumption perceptions, so they are included in the random disturbance term, denoted by μ.

Table 1: Variable settings.

Y	X1	X2	X3	X4	X5	X6	X7
Time	Domestic tourism revenue (billion)	DPI (RMB)	Number of travel agencies (pcs)	Number of domestic tourists (million)	Total number of star-rated hotels (pcs)	Inbound Visitors (million)	Private car ownership (million)

4.1.2 Model Assumptions

To better explain the economic variables, the model (1) is therefore treated as logarithmic.

$$Y=\beta_0+\beta_1 \ln X_1+\beta_2 \ln X_2+\beta_3 \ln X_3+\beta_4 \ln X_4+\beta_5 \ln X_5+\beta_6 \ln X_6+\beta_7 X_7+\mu \quad (1)$$

4.1.3 Data Sources

Data in this paper are from the National Statistics Office and are recorded from 1994 to 2019 for completeness. Inflation using CPI to eliminate economic variables, i.e. Y/CPI (Xi/CPI), and the data obtained after processing are shown in Table

4.2 Regression Parameter Estimation

The OLS regression estimation of the data in Table 2 using EViews yielded the results as in Table.

Table 2: EViews regression results.

Dependent Variable: LOG(Y); Method: Least Squares;Sample: 1994 2019;Included observations: 26				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5.15429	5.144394	-1.001924	0.3297
LOG(X1)	-3.513656	1.838875	-1.910764	0.0721
LOG(X2)	0.066679	0.349255	0.190919	0.8507
LOG(X3)	1.363423	0.531704	2.564252	0.0195
LOG(X4)	-0.052115	0.227301	-0.229277	0.8212
LOG(X5)	0.067663	0.560575	0.120703	0.9053
LOG(X6)	1.521488	0.771876	1.971155	0.0643
X7	0.03299	0.134708	0.2449	0.8093
R-squared	0.993102	Mean dependent var		0.480515
Adjusted R-squared	0.99042	S.D. dependent var		1.05365
S.E. of regression	0.10313	Akaike info criterion		-1.457995
Sum squared resid	0.191444	Schwarz criterion		-1.070888
Log likelihood	26.95394	Hannan-Quinn criter.		-1.346522
F-statistic	370.2196	Durbin-Watson stat		0.580444
Prob(F-statistic)	0			

4.3 Economic Significance and Statistical Inference

4.3.1 Economic Significance Test

According to the regression results, the sign of the coefficients of X1 and X4 is negative, which is contrary to the economic significance and fails the economic significance test. All other variables have positive signs, which are in line with economic theory, and therefore pass the economic significance test.

4.3.2 Statistical Significance Tests

(1) Goodness-of-fit test
From Table 2, we can obtain R2 =0.993129, Adjusted R-squared=0.990457. This shows that the estimated

regression equation is a good fit for the sample observations.

Significance test of the regression equation. In order to test the significance of the linear relationship between the explanatory variables and the explained variables in the model from the overall, the original hypothesis (2) tested H_0:

$$H_0: \beta_1=\beta_2=\beta_3=\beta_4=\beta_5=\beta_6=\beta_7 \quad (2)$$

Given significance level, α=0.1, n=26, k=7, distribution side quantile$F_{0.1}$ (6, 18)=2.13, $F_{0.1}$(8, 18) =2.04, Take its average value, $F_{0.1}$(7, 18)=2.085. F =371.6717 > $F_{0.1}$ (7, 18)=2.085. Therefore, H_0 is negated and there is a significant linear relationship in the overall regression equation.

(2) Significance test of explanatory variables

The coefficients of the model explanatory variables were tested for significance at the level of significance α=0.1, and the hypothesis (3) was formulated.

$$H_0: \beta_i=0 \ (i=1,2,3,4,5) \quad (3)$$

Checking the t-distribution table, when the critical value $t_{0.1/2}(24)$ =1.711 for degree of freedom of 24. $|t_1|$==1.8073> $t_{0.1/2}(24)$ = 1.711, Therefore, the original hypothesis is rejected and β1 is significantly not 0. After analysis, it can be seen that β3 and β6 are significantly not 0, while β2, β4 and β7 do not pass the significance test, and the original hypothesis $\beta_2=0$ 、$\beta_4=0$ and$\beta_7=0$ is accepted.

4.3.3 Econometric Tests

(1) Multicollinearity test

The explanatory variables were tested for covariance using EViews and the test results are shown in Table 3. From Table 3, it can be seen that there is a serious multicollinearity between the explanatory variables, so the stepwise regression method was used to correct for it. The results obtained by stepwise regression using EViews are shown in Table 4.

Table 3: Multicollinearity test

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
X_1	1.000	0.964	0.946	0.231	0.513	0.979	0.881
X_2	0.964	1.000	0.932	0.425	0.407	0.907	0.844
X_3	0.946	0.932	1.000	0.299	0.227	0.926	0.838
X_4	0.231	0.425	0.299	1.000	-0.338	0.041	0.328
X_5	0.513	0.407	0.227	-0.338	1.000	0.563	0.347
X_6	0.979	0.907	0.926	0.041	0.563	1.000	0.824
X_7	0.881	0.844	0.838	0.328	0.347	0.824	1.000

Table 4: Regression between lnY and each explanatory variable.

	$\ln X_1$	$\ln X_2$	$\ln X_3$	$\ln X_4$	$\ln X_5$	$\ln X_6$	X_7
R^2	0.9878	0.8984	0.9854	0.4764	0.7664	0.9849	0.7718

Since the economic significance of the explanatory variables in the regression process as well as the p-value test are significant, Therefore, the explanatory variables with larger R2 values were selected as the main regressors of the model, and then stepwise regression was performed to find the best regression equation, as shown in Table 5

Table 5: Stepwise regression table.

	$\ln X_1$	$\ln X_1$, $\ln X_2$	$\ln X_1$,$\ln X_3$	$\ln X_1$,$\ln X_3$,$\ln X_4$	$\ln X_1$,$\ln X_3$,$\ln X_5$	$\ln X_1$,$\ln X_3$,$\ln X_6$	$\ln X_1$,$\ln X_3$, X_7
t	√	×	√	×	×	√	×
Economic significance	√	×	√	√	√	×	√
Adjusted R-squared	0.9873	0.9880	0.9900	0.9894	0.9897	0.9921	0.9893

The results of the stepwise regression are shown in Table 5, passing the t-test as well as the economic significance test for the variables X_1 and X_3.

The adjusted model R2 =0.990778, Adjusted R-squared=0.98998, F=1235.452, indicating a good fit of the model. Regression model (4):

$$LnY=-3.84306+0.87498*\ln X_1 +0.56804*\ln X_3+\mu \quad (4)$$

(2) Heteroskedasticity test

A further test for the presence of heteroskedasticity in the model using the White's test, the constructed auxiliary function (5)

$$\lambda^2=\alpha_0+\alpha_1 LnX_1+\alpha_2 LnX_3+\alpha_4 LnX_1 LnX_3+\alpha_5(LnX_1)^2+\alpha_6(LnX_3)^2 \quad (5)$$

Formulate a hypothesis (6)

$$H_0: \alpha_i = 0,\ i= 1,\ldots, 6$$
$$H_1: \text{At least one } \alpha_i \text{ is not } 0 \quad (6)$$

The results of the White test are shown in Table 6.

Table 6: White test results.

Heteroskedasticity Test: White; Method: Least Squares; Sample: 1994 2019; Included observations:26				
F-statistic	2.050364	Prob. F(5,20)		0.1148
Obs*R-squared	8.81095	Prob. Chi-Square(5)		0.1168
Scaled explained SS	4.424215	Prob. Chi-Square(5)		0.4901
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.323536	3.424603	-0.094474	0.9257
LOG(X1)	-0.383148	1.855374	-0.206507	0.8385
(LOG(X1))^2	-0.057874	0.242156	-0.238996	0.8135
(LOG(X1))*(LOG(X3))	0.100552	0.434789	0.231267	0.8195
LOG(X3)	0.261064	1.633977	0.159772	0.8747
(LOG(X3))^2	-0.039413	0.193569	-0.203611	0.8407

From Table 6, we get $nR^2=8.810950$, Tested by White, $\alpha=0.1$, $\chi^2_{0.1}(5)=9.24$, $nR^2=8.810950<\chi^2_{0.1}(5)=9.24$,Therefore, accepting the original hypothesis H_0, it can be assumed that there is no heteroskedasticity in the model.

(3) Autocorrelation test

The DW test is used to test whether there is autocorrelation in the error term μ. The hypothesis is proposed

$$H_0: \mu \text{ is not autocorrelated}$$
$$H_1: \mu \text{ is autocorrelated} \quad (7)$$

Table 7: Regression results of lnY with lnx1 and lnX3.

Dependent Variable: LOG(Y); Sample: 1994 2019; Included observations: 26				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.843062	0.882877	-4.352883	0.0002
LOG(X1)	0.874978	0.239102	3.659432	0.0013
LOG(X3)	0.568045	0.208885	2.719411	0.0122
R-squared	0.990778	Mean dependent var		2.782924
Adjusted R-squared	0.989976	S.D. dependent var		1.053909
S.E. of regression	0.10552	Akaike info criterion		-1.551673
Sum squared resid	0.256091	Schwarz criterion		-1.406508
Log likelihood	23.17175	Hannan-Quinn criter.		-1.509871
F-statistic	1235.452	Durbin-Watson stat		0.485045
Prob(F-statistic)	0			

s.e.=0.105520, DW=0.485045, Check the DW statistics table, K=2, n=26, α=0.1, dl=1.224, Du =1.553, DW=0.485045<dl =1.224< du =1.553 in this model. Reject the original hypothesis, Therefore, μ is autocorrelated Standard error correction is performed using the newey-west method in EViews software to eliminate autocorrelation problems.

5 ANALYSIS OF RESULTS AND RECOMMENDATIONS

Through quantitative analysis, the factors that affect our tourism revenue are PDI and the number of domestic tourists. All other things being equal, i.e.,

for every 1 percentage point increase in per capita disposable income of residents, domestic tourism income increases by an average of 0.87 percentage points. For every 1 percentage point increase in the number of domestic tourists, domestic tourism revenue increases by an average of 0.57 percentage points, according to which the following countermeasures are proposed.

PDI significantly affects the total income of the tourism industry, and the steady increase of residents' income makes residents pay more attention to the pursuit of spiritual life and choose to travel for such enjoyment-oriented consumption, this will further promote the increase of tourism income. With the established level of per capita income, the government should strengthen as well as improve the social security system to increase the marginal propensity to consume and thus expand the consumption support (CAI, L. J. 2015). With the increase of residents' per capita income, people will pursue high-quality tourism products more, so the tourism industry can also be transformed to high-end tourism. For example, business tourism, health care tourism, cruise ships, Marine leisure tour. (CHEN, Q. J. & RAO, W. Y. 2019).

The results of the study show that the number of domestic tourism affects tourism income significantly and positively, while the main countermeasures that affect the number of travellers in China can be proposed in the following aspects. Firstly, guidance from the ideological level, while vigorously developing the economy, to improve the thinking of the residents of tourism consumption, to encourage and spread positive tourism ideas and concepts. Secondly broaden the marketing channels (YU, Tong & YE, Y. L. 2016). the background of the epidemic, soft marketing more into the hearts of people, users can read after the empathy and take the initiative to share a city's tourism information, not only has a lower cost of publicity, but also in the current context to play the most effective role.

ACKNOWLEDGEMENTS

Support by: The construction program of innovation team at Beijing Institute of Fashion Technology (BIFTTD201901); “The first batch of new liberal arts research and reform practice projects of the Ministry of Education” project (Project No.: 2021140009).

REFERENCES

CAI, L. J. (2015). An empirical analysis of the economic influences on the development status of domestic tourism. J. Commercial Accounting. 14, 110-112.

CHEN, Q. J. & RAO, W. Y. (2019). Empirical analysis of influencing factors of tourism in provinces and cities of China. J. New Finance Economics. 11,51-53.

DING, L. F. & Li, C. S. (2021). Construction and Empirical analysis of Tourism development Mode in Anhui Province. J. J. BaojiColl.Arts&Sci. (Nat.Sci.). 41(01), 52-58

SHI, P. H. et al (2020). Evaluation on the effect of the establishment of all-region tourism demonstration zone on the development of regional tourism economy-Empirical analysis based on difference and difference. J. Macroeconomics. 06,122-132+175.

YU, Tong & YE, Y. L. (2016). An empirical study on influencing factors of domestic tourism income in Guangdong Province. J. Journal of Wuhan Business University. 30(04),19-22.

ZHOU, Li (2019). An empirical analysis of the construction of tourism consumer price Index on the tourism consumption forecast of urban and rural residents -Test based on VAR model. J Prices Monthly. 06,1-5.