INFLUENCE OF TOURISM SECTOR IN ALBANIAN GDP: ESTIMATION USING MULTIPLE REGRESSION METHOD

Ph.D. Student **Eglantina HYSA**
Epoka University, Tirana, Albania

Abstract

During last years, tourism sector has significantly increased in Albania, since after year 1990 Albania has passed from a centralized economy to a liberal one. Tourism sector plays an important role in economic and social development. The contributions of this sector reflect directly into the generation of national income. The two main components matching the tourism movements are the number of tourists and the number of overnights in hotels. Investments done in this sector could be expected to have high positive influence in the country's GDP. This study seeks to identify the influence of tourists, their overnights in hotels and capital investment spending by all sectors directly involved in tourism sector on tourism total contribution to gross domestic product of Albania during 1996-2009. A regression analysis has been performed taking as dependent variable GDP generated by tourism sector and as independent variables, capital investment, tourist number and overnights in hotels. Even if all the variables have been found to be positivlye related, the variable 'overnights of foreigners and Albanians in hotels' have beenfound insignificant.

Key words: *Albania, Capital investment, Regression analysis, Tourism.*

JEL classification: *D12, L83*

1. INTRODUCTION

In general, tourism is considered one of the largest industries in the economy. It is seen as the key promoter of both economic growth and social development because of its potential to promote understanding and international relationships. Tourism is a mixture of natural resources such as, climate, natural beauty and environment, and a favorable climate put up by governmental structures, habitants' good willingness toward tourism development and appropriate infrastructure. The Albanian Riviera is considered very rich and special compared to the modest surface of the Albanian territories. The main difficulties that Albania is facing in the tourism sector are the issues that have to do with the infrastructure. Even if in these last years the government has invested in it, the infrastructure condition is pretty humble compared to the other countries.

Tourism development is closely connected to investment and employment. In this context during the years 2009-2010, the restaurant, bar-restaurants, shopping centers, transportation and construction of various infrastructure, have increased their activity (Albanian Institute of Statistics - INSTAT), having positive impact also in the increase in the employment level in tourist areas. The generated income from tourism sector is used in the development of important areas.

During 2010, the number of foreigners who have visited Albania has increased by 30.2 per cent compared to 2009, mostly from Kosovo, Macedonia and Montenegro and the number of clients accommodated in hotels increased by 13.8 percent in 2010 compared with the previous year. Average night staying in hotels for 2010 was 2.5 (INSTAT).

This study aims to find out the main components positively influencing in the contribution to the total gross domestic product of Albania. The econometric model used in this study uses three variables, which are supposed to explain the tourism contribution to GDP. Capital investment and arrivals of foreigners by means of transport result as significant for the applied model, whereas overnights of foreigners and Albanians in hotels are insignificant. This is probably because there exist a considerable number of private hosting homes in Albania that "act" as hotel without being officially registered as hotels, which are much cheaper and as a consequence, much more preferable.

2. DATA AND METHODOLOGY

The first part of this section will deal with the data analysis taken into consideration during this study. It is of great importance to analyze the variables path through the time. The period used for the variables trend is 1995-2010 since there are no available data before year 1995. The GDP generated by direct travel and tourism industries and capital investment spending by all sectors directly involved in the travel and tourism industries are taken from the World Travel & Tourism Council (WTTC). The

definitions of these two factors are important to be given since they are based on these definitions while taking their overall value. GDP generated by direct travel and tourism industries takes into consideration also the indirect and induced contributions, including the contribution of capital investment spending. Whereas, the capital investment spending by all sectors directly involved in the travel and tourism industries takes into consideration the investment spending by other industries on specific tourism assets such as new visitor accommodation, passenger transportation equipment, as well as restaurants and leisure facilities for specific tourism use. Whereas, the data for two other variables, precisely arrivals of foreigners by means of transport and overnights of foreigners and Albanians in hotels are taken from official website of Albanian Institute of Statistics (INSTAT). All the data are included in Table 1.

Table 1 - Main economic indicators in tourism sector for the period 1995-2010

	Travel & Tourism Total Contribution to GDP *(in billion ALL)*	Capital Investment *(in billion ALL)*	Arrivals of Foreigners by Means of Transport *(in thousands)*	Overnights of Foreigners and Albanians in Hotels *(in thousands)*
1995	24.222	1.418	304	211
1996	29.774	2.686	287	394
1997	33.461	3.098	119	108
1998	36.337	2.651	184	154
1999	59.624	2.781	371	215
2000	82.264	3.966	317	326
2001	99.114	9.843	353	395
2002	120.16	13.61	470	532
2003	126.671	20.473	558	277
2004	131.681	19.734	645	243
2005	157.981	20.973	748	344
2006	181.277	20.699	937	459
2007	218.373	21.703	1,126	547
2008	238.895	21.895	1,419	490
2009	285.757	19.471	1,856	539
2010	319.524	20.697	2,418	610

Based on the Figure 1, the travel and tourism total contribution to GDP is found to have a positive and almost constant trend from year to year. Capital investment has experienced stagnation during 1995-2000 and 2003-2010, showing up the only immediate increase during 2000-2003. Based on this configuration, it can strongly support the fact that the Albanian economy did not generate the same incomes through the years in the tourism sector. The third variable, arrivals of foreigner by means of transport, shows a fall down in year 1997, corresponding with the political problems in Albania during this year. Beginning from 1998 and on, the arrivals of foreigners has increased slowly and just after 2005 this variable has significantly increase. The last variable, the overnights of foreigners and Albanians in hotel, show up a more complicated trend. In year 1997, 2004 and 2008 the overnights of tourist in hotels have experienced significant decline. However, year 2010 shows the largest number of tourists in the hotels inside Albania.

The second part of the section tries to explain the methodology used in this study. Multiple regression is a technique that allows more than one factor to enter the analysis separately so that the effect of each can be estimated. This method gives the opportunity to identify the impact of various simultaneous influences upon a single dependent variable. There are five assumptions that must be met in multiple regression and correlation: linearity, homoscedasticity, normality, multicollinearity and autocorrelation (Chow, 1989).
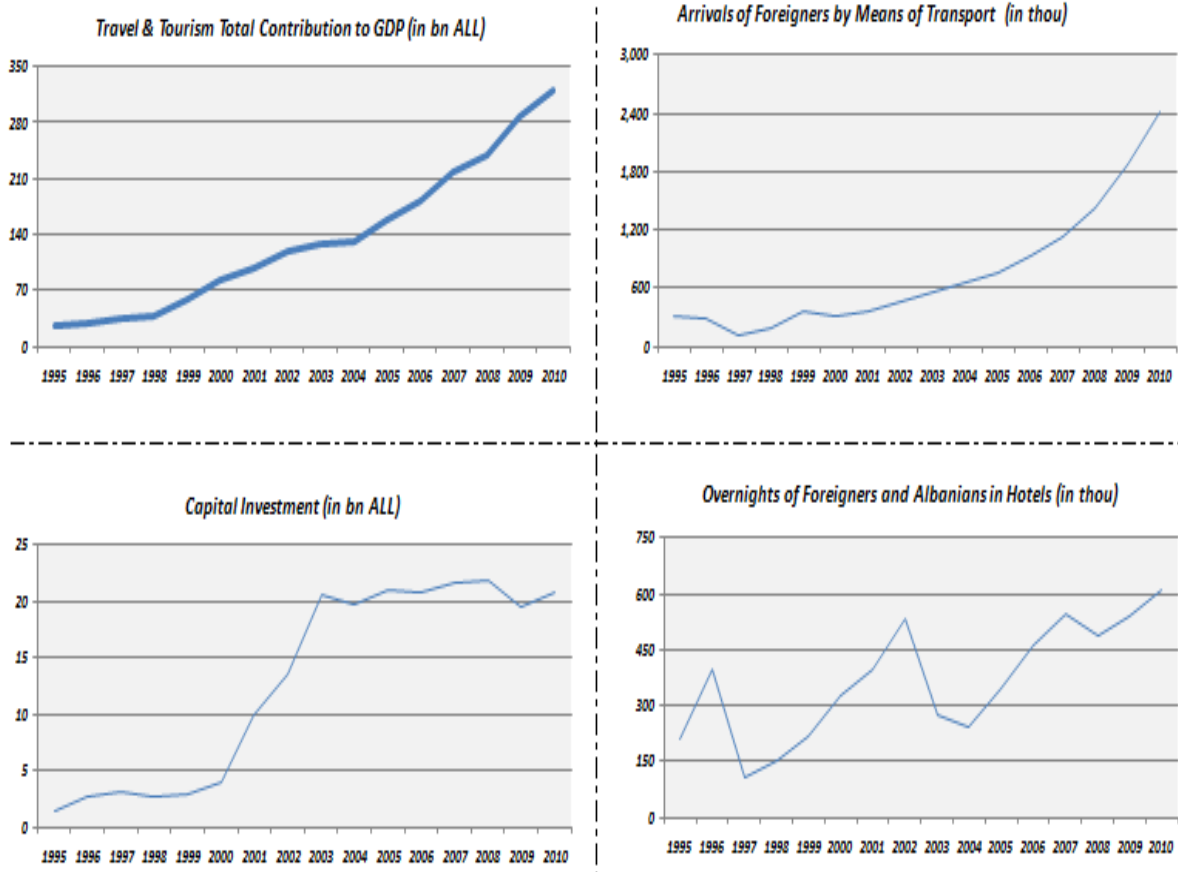
Figure 1 – Variables trend for Albania (1995-2010)

Based on Figure 2, the first two assumptions can be confirmed as fulfilled conditions. The normality and the autocorrelation assumptions also held for our regression equation (Dougherty, 2007).
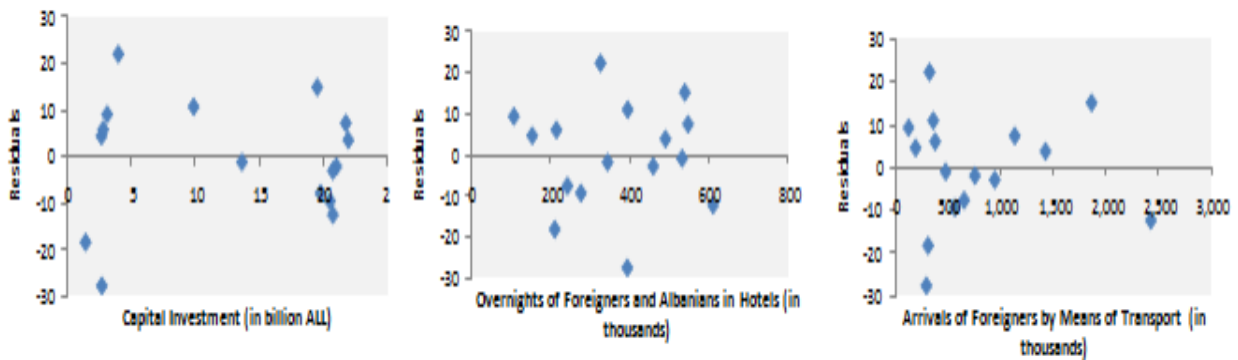

Figure 2 – Residual Plot of there independent variables

The only problematic issue can be evident in the multicollinearity assumption, which exists when the independent variables are correlated. For this reason the correlation matrix between variables has been constructed (Table 2). According to the results, the most powerful correlation exists between arrivals of foreigners by mean of transports and overnights of foreigners and Albanians in hotels since it has the value of 0.762. The second strong correlation exists between arrivals of foreigners by mean of transports and capital investment that is 0.703 and finally the correlation between capital investment and overnights of foreigners and Albanians in hotels which is 0.642. According to different authors, there is not a special problem in case the correlation coefficient between the independent variables vary from +0.70 to -0.070. The first correlation can be assumed to be pretty near to this interval. Hence no special problem is related to our model referring the multicollinearity assumption.

Table 2 - Correlation matrix between variables

	Travel & Tourism Total Contribution to GDP *(in billion ALL)*	Capital Investment *(in billion ALL)*	Arrivals of Foreigners by Means of Transport *(in thousands)*	Overnights of Foreigners and Albanians in Hotels *(in thousands)*
Travel & Tourism Total Contribution to GDP *(in billion ALL)*	*1.000*			
Capital Investment *(in billion ALL)*	*0.846*	*1.000*		
Arrivals of Foreigners by Means of Transport *(in thousands)*	*0.958*	*0.703*	*1.000*	
Overnights of Foreigners and Albanians in Hotels *(in thousands)*	*0.810*	*0.642*	*0.762*	*1.000*

3. BUILDING THE REGRESSION MODEL

In this study, the multiple regression equation includes 3 independent variables, such as X_1, X_2, and X_3 (Reinsel, 1997).

$$\ddot{Y} = a + b_1 X_1 + b_2 X_2 + b_3 X_3 \quad (1)$$

Where:
- Y - is travel and tourism total contribution to GDP
- X_1 - is capital investment
- X_2 - is arrivals of foreigners by means of transport
- X_3 - is overnights of foreigners and Albanians in hotels
- a - is the Y-intercept, the value of Y when all the X's are zero

b_j - is the net change in $\hat{Y}$ for each unit change in X_j, holding all other X's constant.

The values of b_1, b_2 and b_3 are called the ***regression coefficients***. They indicate the change in the estimated value of the dependent variable for a unit change in one of the independent variables, when the other independent variables are held constant. By using the coefficients, the regression equation is as below:

$$Y = -4.8112 + 3.4444 \, Cap.Inv. + 0.0929 \, Arriv.Foreign. + 0.0663 \, Overnigh.Hot. \quad (2)$$

Table 3 gives a clear picture of the regression of GDP generated by tourism sector and capital investment, arrivals of foreigners and tourists' overnights in hotels.

Table 3 - OLS Regression of GDP generated by tourism sector and 3 independent variables

	Capital Investment *(in billion ALL)*	Arrivals of Foreigners by Means of Transport *(in thousands)*	Overnights of Foreigners and Albanians in Hotels *(in thousands)*
Coefficient:	3.4444*	0.0929*	0.0663
P-value:	0.0001	0.0000	0.1058
	(5.5801)	(9.5993)	(1.7491)
R^2	0.9812		
F	208.7603		
N	16		

*Notes: 1 - *Significant at the 5% level*
2- The values in brackets are t-values

The ANOVA Table

A convenient means of showing the regression output is to use an ANOVA table. The variation in the dependent variable is separated into two components: the explained variation and unexplained variation (Wonnacott, 1990).

The column headed "*DF*" refers to the degrees of freedom associated with each category. The total degrees of freedom is $(n − 1) = 15$. The degrees of freedom for regression is k, the number of independent variables. The degrees of freedom associated with the error term is $n − (k + 1) = 12$. The SS in the middle of the top row of the ANOVA table refers to the sum of squares, or the variation.

$$\text{Total variation} = SS\,total = \Sigma\left(Y - \overline{Y}\right)^2 \quad (3)$$

$$\text{Error variation} = SSE = \Sigma\left(Y - \hat{Y}\right)^2 \quad (4)$$

$$\text{Regression variation} = SSR = \Sigma\left(\hat{Y} - \overline{Y}\right)^2 = \left(SS\,Total - SSE\right) \quad (5)$$

The column headed MS refers to the mean square and is obtained by dividing the SS term by the *df* term. Thus, MSR, the mean square regression, is equal to SSR/*k*, and MSE equals SSE / [*n* − (*k* + 1)]. The results of ANOVA table are as below:

ANOVA

	df	SS	MS	F	Significance F
Regression	k=3	SSR=128162.0473	MRS=SSR/k=42720.68244	MSR/MSE=208.76035	1.28E-10
Residual	n-(k+1)=12	SSE=2455.677955	MSE=SEE/(n-(k+1))=204.6398296		
Total	n-1=15	SS Total=130617.7253			

Analysis of Variance

It is likely that the estimation can include some error. The error in the predicted value of the dependent variable is measured by the **multiple standard error of estimate**.

$$s_{y.12\ldots k} = \sqrt{\frac{\Sigma\left(Y - \hat{Y}\right)^2}{n - (k + 1)}} = 14.305 \quad (6)$$

Another measure of the effectiveness of the regression equation is the **coefficient of multiple determination**, which is the proportion of the variation in the dependent variable, *Y*, that is explained by the set of independent variables $x_1, x_2, x_3, \ldots x_k$

The coefficient of multiple determination, R^2, takes the values from 0 to 1, which is the percent of the variation explained by the regression. The closer R^2 is to 1, the stronger the association between Y and the set of independent variables, $x_1, x_2, x_3, \ldots x_k$.

The ANOVA table is used to calculate the coefficient of multiple determination. It is the sum of squares due to the regression divided by the sum of squares total.

$$R^2 = \frac{SSR}{SSTotal} = 0.981 \quad (7)$$

In our regression model, the coefficient of multiple determination is 0.98, indicating that the three independent variables, considered jointly, explain 98 percent of the variation in billions of tourism sector contribution to GDP.

Adjusted Coefficient of Determination

As the number of independent variables in the regression model increases, the coefficient of multiple determination increases. Even if the additional independent variable is not a good predictor, its inclusion in the model decreases SSE which in turn increases SSR and R^2. In this case R^2_{adj} is used to measure the effectiveness of multiple regression model.

$$R^2_{adj} = 1 - \frac{\dfrac{SSE}{n - (k + 1)}}{\dfrac{SSTotal}{n - 1}} = 0.976 \quad (8)$$

Global Test: Testing Whether the Multiple Regression Model is Valid

The overall ability of the independent variables $X_1, X_2, \ldots X_k$, to explain the behavior of the dependent variable *Y* can be tested. Two tests of hypotheses are considered. The first one is called the **global test**, which investigates the possibility that all the regression coefficients are equal to zero.

It tests the overall ability of the set of independent variables to explain differences in the dependent variable. The null and the alternative hypothesis are as below:

H_0: $b_1 = b_2 = b_3 = 0$
H_1: Not all the bs = 0

The test statistic used is the *F* distribution is calculated to be the following which corresponds also with Table 3.

$$F = \frac{MRS}{MSE} = \frac{SSR/k}{SSE/[n - (k + 1)]} = 208.76 \quad (9)$$

The critical value of F is found to be 3.12 (the significance level equal to 0.05, degree of freedom at nominator 3 and degree of freedom at denominator 12). By comparing F-stat to critical value of F, the null hypothesis is rejected, meaning that not all the regression coefficient is equal to zero. Hence, at least one or more of independent variables is useful in explaining differences in the dependent variable, without identifying which regression coefficients are not zero and how many.

Evaluating Individual Regression Coefficients

The second test of hypothesis identifies which of the set of independent variables are significant predictors of the dependent variable. That is, it tests the independent variables individually rather than as a unit. This test is useful because unimportant variables

can be eliminated from the regression model. The null and the alternative hypothesis are as below:

H_0: $b_2 = 0$
H_1: $b_2 \neq 0$

The test statistic is the Student t distribution with $n - (k + 1)$ degrees of freedom. In this study the degree of freedom is equal to 12. The critical value for alpha equal to 0.05, two-tailed test, is then equal to +/- 2.179. in order to find out if the null hypothesis should be rejected or not, the critical value is compared to student test, which is computed by the software and found in Table 3. The results of student tests are as below:

$$t_1 = \frac{b_1 - 0}{s_{b_1}} = 5.58 \tag{10}$$

$$t_2 = \frac{b_2 - 0}{s_{b_2}} = 9.60 \tag{11}$$

$$t_3 = \frac{b_3 - 0}{s_{b_3}} = 1.75 \tag{12}$$

In the case of the first and the second variable, the null hypothesis is rejected, whereas in the third variable, the hypothesis test finds that the null hypothesis cannot be rejected. Hence, the variable should be dropped from the model. Based also on Table 3, if the first two variables are foundsignificant at 0.05 significance level, the third variable is not significant even for 0.1 significance level. After the variable "overnights of foreigners and Albanians in hotels" is removed from the model, a new regression model is constructed using the other two remaining variables. The remaining variables are both significant and the new model is explained 97.64%.

Table 4 - **OLS Regression of GDP generated by tourism sector and 2 independent variables**

	Capital Investment (in billion ALL)	Arrivals of Foreigners by Means of Transport (in thousands)
Coefficient:	3.6953*	0.1026*
P-value:	0.0001	0.0000
	(5.7186)	(11.9730)
R^2	0.9764	
F	269.00	
N	16	

*Notes: 1- *Significant at the 5% level*
2- The values in brackets are t-values

4. CONCLUSION

After running the regression equation the following results showed up:
a. The most powerful correlation of independent variables has been found between arrivals of foreigners by mean of transports and overnights of foreigners and Albanians in hotels of 0.762;
b. During 1995-2010, the independent variables, "capital investment" and 'arrivals of foreigners by mean of transports' are significant predictors for the dependent variable "travel and tourism total contribution to GDP";
c. The independent variable, overnights of foreigners and Albanians in hotels, is insignificant to explain the contribution to GDP coming from tourism sector.

As a consequence, the model helps us to identify the impact factors in the contribution of GDP in tourism sector. An intense emphasizes should be provided to the capital investment in tourism sector since high level of investment in this sector and smartly defined the pivotal parts of the Albanian tourism would lead to higher incomes from tourism sector. The second issue to underline is that both, the foreigner and the Albanians are potential tourists. It is crucial the establishment of different strategies, new techniques and innovative ideas for tourism marketing inside and outside Albania.

BIBLIOGRAPHY

1. Chow, G. (1989) *Econometrics*, McGraw Hill, New York.
2. Dougherty, C. (2007) *Introduction to Econometrics*, Oxford University Press.
3. Hannan, E.J., Deistler, M. (1988) *The Statistical Theory of Linear Systems*, New York Wiley.
4. Reinsel, G.C. (1997) *Elements of Multivariate Time Series Analysis*, Springer Varleg, New York.
5. Wonnacott, T.H., Wonnacott, R.J. (1990) *Introductory Statistics for Business and Economics*, New York Wiley.
6. *** Albanian Institute of Statistics (INSTAT), www.instat.gov.al.
   *** World Travel & Tourism Council (WTTC), www.wttc.org.