FEW DETERMINATS OF THE AVERAGE AMOUNTS OF MONEY TOURISTS SPEND WHILE VISITING ALBA COUNTY. A REGRESSION MODEL CONSIDERING WHITE'S HETEROSKEDASTICITY-CONSISTENT STANDARD ERRORS & THE COVARIANCE AND DIFFERENT WEIGHTS TECHNIQUES.

Andreea Cipriana Muntean[1],
Iulian Bogdan Dobra [2]

ABSTRACT: *The aim of this research is to analyse the average expenditure of the tourists visiting Alba County during 2013-2015, and the way this is influenced by different independent variables. Into Equation Estimation, Least Squares Options, we outlined the next specific estimation settings: for Coefficient covariance matrix we selected "White" and for weights options we pointed out Inverse standard deviation for Type, we entered NO_DAYS in the Weight series field, and for Scaling we choose Average mode. In the log-log regression model, half of the predictors will determine an increase and the other half will determine a decrease in the tourists' average expenditure. According to the model, the results show that if the number of days at the destination increases by 1 %, the expenditures will increase by 0.58%, and if the number of visits increases by 1 % than the expenditures will decrease by 0.64%. Of all the exogenous dichotomous variables, the one related to 56_65 age (i.e. tourist between 56 and 65 years old) contributes most to the decrease of the estimated average expenditure, by about 1.16%, and the one that contributes most to their increase is "staff amiability", by about 0.70%. In conclusion, we accomplished to estimate the equation by using White's heteroskedasticity-consistent standard errors & covariance and different Weights options. The results show that in 49% of the cases, tourists' average expenditure increases and it decreases in 51%.*

Keywords: *expenditure, tourist, coefficient covariance matrix, WLS, log-log model, ANCOVA.*

JEL Codes: *Z31, Z32*

Introduction

Motto: "In almost every other race of animals, each individual, when it is grown up to maturity, is entirely independent, and in its natural state has occasion for the assistance of no other living creature. But man has almost constant occasion for the help of his brethren, and it is in vain for him to expect it from their benevolence only. He will be more likely to prevail if he can interest their self-love in his favour, and show them that it is for their own advantage to do for him what he requires of them" (Adam Smith, 1902:56).

This study is a review of expenditure analysed from the tourism perspective presenting a variety of factors (i.e. our model started with 23 predictors, and ended with 12, except the intercept) that could affect tourism expenditure.

According to (Chapter 4, entitled *The demand perspective: tourism expenditure*, paragraph number 4.2 of) United Nation World Tourism Organization (i.e. UNWTO) Recommendations on Tourism Statistics "tourism expenditure" refers to "the amount paid for the acquisition of

[1] "1 Decembrie 1918" University of Alba Iulia, Faculty of Economic Sciences, Strada Unirii Street, no.15-17, Alba Iulia, Email: andreea.muntean@uab.ro;
[2] "1 Decembrie 1918" University of Alba Iulia, Faculty of Economic Sciences, Strada Unirii Street, no.15-17, Alba Iulia, Email: dobra.iulian@uab.ro.

consumption goods and services, as well as valuables, for own use or to give away, for and during tourism trips" (UNWTO, 2010:31).

Sampol and Perez (2000:3) pointed out that in a tourism market, the fundamental variable for undertaking a profitability analysis is the tourist expenditure of its visitors.

It is well known that the reporting of statistics related to tourists must be in a "common approach", which in fact, should be applied to any research area. From this point of view, specialists Peter Laimer and Jürgen Weiß emphasize that (2006:22): "a common language for tourism statistics in general, and tourism expenditure in particular, is indispensable to the work of policy makers and for the private sector".

Specialized journals enfold a large number of empirical studies intended at measuring the demand for tourism, both in terms of number of tourists and their expenditure. A reference work is that of Lim (1997), who highlighted 100 models of international tourism demand according to the type of data used for estimation.

Consequently, the main objective of this study is to analyse the average expenditure of tourists visiting Alba County. Our review is partially comparable to the study of Taylor, Fletcher and Clabaugh (1993), who examined the variety features of tourists, reliant on whether they visited the heritage sites and on their different amount of expenditure.

Literature reviews

It is well known that in a tourism market, the central variable to undertake a profitability analysis is the amount of expenditure registered by its tourists. This paper aims to analyse this expenditure in an unbalanced, undated structure in order to identify those characteristics that describe different tourist profiles according to expenditure levels.

In tourism literature, there are several works dealing with tourists' expenditure based on several predictors, variables that we've had in mind when we developed the regression model.

Asgary et al. (1997) established that, by adding social and demographic variables to the model, along with economic variables, the explanatory power of the model improved noticeably.

Poon (1993) has outlined that the tourism market has become gradually segmented along demographic, socio-economic and psycho-graphic lines, and along tourist's motivations, hobbies, opinions, etc.

In terms of the effect demographics and travel variables have on overspending, there are authors (Hong, Fan and Palmer 2005, Peerapatdit, 2004) who state that a significant positive relationship exists between income, assets and leisure travel spending.

In their paper entitled *A Linear Expenditure Model for Tourism Demand*, Pyo, Uysal and McLellan (1991) highlighted that among the tourism-oriented products, transportation is the most price sensitive product.

In a research regarding visitor expenditures in mountain tourism, author Peter Fredman (2008), highlighted that male visitors spend more money outside the mountain region compared to female, and trips by train or plane are associated to higher expenditures outside the final destination in the mountain region compared to travels by car/bus.

Sampol and Perez (2000) emphasize that factors such as nationality, age, profession, type of accommodation, type of booking, payment for items, opinion on prices and the trip itself, etc., influence the Average Tourist Expenditure.

Literature shows that there are authors (Nguyen, 2016) who use age, gender and education among the determinants of tourism overspending.

In terms of the length of stay (i.e. number of days in our case), Agarwal and Yochum (1999) conducted the survey data on overnight visitors and they found that this was a significant determinant of visitors' expenditures.

Kevin Barry and John O'Hagan, in their article entitled *An Econometric Study of British Tourist Expenditure*, state (1971:147): ”There are a great number of non-economic factors, some of which are measurable and others not… These include the following: population increase and changes in the age structure of population; the increase in the degree of urbanisation and the concomitant increase in the desire of people to get into open spaces; the increase in the length of paid holidays; the increase in the level of education giving people a greater interest in travelling abroad and learning about other people first-hand”.

So, it can be concluded that independent variables like number of visits, number of days, socio-demographic variables, respectively, can influence tourist expenditure, and also, these covariates in all the regressions considerably better fit the data.

Research Design and Methodology

All the information on indicators (i.e. Expenditure, Number of Visits, Number of Days, Means of Transport, variables that highlighted the point of view of tourists on the statement: ”*Staff amiability can make this hotel / hostel become one of the places preferred by tourists*” - inquiry number 25 from the research contract no. 4579/162/19.03.2014, and socio-demographic variables, respectively), was collected from a market research contract in the tourism sector in Alba County, (i.e. Contract no. outlined above).

The total number of tourists who are subject to our research and who responded to questionnaires amounts to 349. It should be noted that respondents are tourists from Romania (i.e. Alba County and other Counties) and abroad.

The period submitted for analysis is 2013 – 2015, when there have been identified 365 tourists. The difference of 16 out of 365 are the respondents who registered expenditures over 3,001 Lei.

As far as the dependent variable EXPENDITURE is concerned, we should mention that, there were tourists reporting expenditure between 10 and 9,000 lei, while the average of the entire sample was about 768 lei. After the tabulation of this variable (see Appendix A), out of 365 observations analysed, over 95% were included in categories up to 3,000 Lei, and almost 5% between 3,000 Lei and 9,000 Lei. Therefore, we adjusted the sample up to 3.001 Lei tourist expenditures, which determined a total number of observations up to 349. Finally, following the data processing, it has been discovered that there are some high-values observations, which could influence both the variables' statistical significance in the regression model we wanted to elaborate, and the coefficient of multiple determinations for multiple regressions. To conclude, we used logarithmic transformation of regressand (i.e. LOG (EXPENDITURE)), and it has been found that the regression model has improved.

As far as the independent variables of the regression model are concerned, we shall mention the following aspects. We selected two quantitative variables (i.e. stage six of our research, in which we established the final selection for our predictors): Number of Visits (i.e. NO_VISITS) and Number of Days (i.e. NO_DAYS) and 10 qualitative variables: Mean of Transport – Train (i.e. MT_T), Mean of Transport – Other (i.e. MT_O), Total Disagree (i.e. TD), Disagree (i.e. D_01), Total Agree (i.e. TA), WOMAN, tourist aged between 26 and 35 (i.e. _26_35), tourist between 56 and 65 years old (i.e. _56_65), UNMARRIED, Master's Degree (i.e. MD).

Concerning the tabulation of the NO_VISITS and NO_DAYS control variables, it has been observed that most of the tourists had chosen to visit Alba County 4 times (i.e. 166 tourists, 45.48%) and 5 times (i.e. 71 tourists, 19.12%); to stay 3 days (i.e. 107 tourists, 29.32 %), 1 day (i.e. 81 tourists, 22.19 %) and 2 days (i.e. 70 tourists, 19.18%), respectively.

In term of the qualitative variables, we mentioned that data were classified into three categories, as follows: 1) variables that highlight the means of transport used for travel (i.e. Alba County) are: train, car , bus or other means of transport; - benchmark variable car; 2) variables that highlight tourists' point of view on the statement "*Staff amiability can make this hotel / hostel to*

become one of the preferred places for tourists" (i.e. Inquiry number 25 from the research contract no. 4579/162/19.03.2014) are : Total Agree, Agree, Neither Agree Nor Disagree, Disagree and Total Disagree) - benchmark variable Agree; 3) demographic variables, respectively: sex, men and women - benchmark variable men; age - this category has the following 6 subcategories: up to 25 years old, between 26 and 35 years old, between 36 and 45 years old, between 46 and 55 years old, between 56 and 65 and over 65 years old - benchmark variable 36_45; marital status: Married, Unmarried, Divorced/ DIV, Widowed/ WID - benchmark variable Married; level of education: Middle School/ MID_S, Vocational School/ VOC_S, High School/ H_S, Bachelor's Degree, Master's Degree/ MD, Doctorate Degree/ PhD - benchmark variable Bachelor's Degree.

In our scientific approach we want to determine the average expenditure of the tourists (i.e. 349) in Alba County for the period of time 2013-2015, and the way they react to the independent variables highlighted above. The options we have chosen in the equation estimation (i.e. Coefficient covariance matrix and Weights) headed us, in the end, to introduce only part of these variables in the regression model. Therefore, the regression model contains three quantitative variables (i.e. Expenditure, No_Visits, No_Days) and ten qualitative variables (i.e. Mean of Transport – Train, Mean of Transport – Other, Total Disagree, Disagree, Total Agree, Woman, Tourists aged between 26 and 35, tourists between 56 and 65 years old, Unmarried, Master's Degree). Thus, the specific function is:

EXPENDITURE = F (MT_T, MT_O, TD, D_01, TA, WOMAN, 26_35 AGE, 56_65 AGE, UNMARRIED, NO_VISITS, NO_DAYS) (1.0)

In order to compare the average values of the expenditure, a framework of the regression analysis has been used. We have also tried to use the ANCOVA model which provides a method of statistically controlling the effect of the quantitative regressor (i.e. covariate). For the completion of the analysis, the following *first model* was considered:

(1.1)

Where:

Log (Z) - (average) tourist expenditure;
u - error term.

We wish to specify that the above equation is the "basis" of our research and from this relation we started our analyses, but changes in the methodology are highlighted in the research stages below. Consequently, in order to complete the analysis, the following *final model* was considered:

(1.2)

Data were inserted in an unbalanced undated worksheet and subsequently processed by means of the Eviews 7.1. Therefore, according to the application software, into Equation Estimation, Least Squares Options, we had the possibility to specify two additional settings for the estimation panel:

a) *Coefficient covariance matrix* (i.e. Estimation default, White and Heteroskedasticity and Autocorrelation Consistent-HAC Newey-West) - for this option we selected "White" (i.e. d.f. adjustment) in every specification;

b) *Weights* – before highlighting the options for this method, we considered useful Richard Startz' (2015:342) specifications, the renowned professor of economics:"Ordinary least squares attach equal weight to each observation. Sometimes you want certain observations to count more than others. One reason for weighting is to make sub-population proportions in your sample mimic sub-population proportions in the overall population. Another reason for weighting is to down-weight the high error variance observations". There are three basic weight options in our software package that we may specify: *Type*, *Weight series* and *Scaling*. For *Type* we selected *Inverse standard deviation*, for *Weight* series we entered NO_DAYS in the Weight series field, and for *Scaling* we chose *Average* mode.

Long and Ervin (1998) highlighted that tests based on a Heteroscedasticity Consistent Covariance Matrix (i.e. HCCM) are consistent, and the specific literature that treats this estimator considers that there are three additional small sample versions of the HCCM as follow: a) HC1 (Hinkley, 1977) resulted from a calculus of the degree of HC0 freedom correction (White, 1980), b) HC2 (MacKinnon and White, 1985) elaborated taking into account that the covariance matrix will be a less biased estimator, and c) HC3 particularized by MacKinnon and White (1985).

In this paper, we used the HC1 estimator and the standard errors for the WLS estimator, and we noticed its advantages and disadvantages (e.g. WLS estimators may have worse finite-sample properties than unweighted estimators). This way, Joseph P. Romano and Michael Wolf (2014), both Professors of Economics, outlined that sensibly weighting the data can lead to noticeable efficiency gains over OLS, and combining Weighted Least Squares (i.e. WLS) with HC standard errors allows a valid inference, even if the conditional variance model is misspecified.

It is well known that the EViews software package offers built-in tools for estimating the coefficient covariance under the assumption that the residuals are conditionally heteroskedastic. In this case, the coefficient covariance estimator is named a Heteroskedasticity Consistent Covariance (White).

Regarding HC1, we considered the formula of Long and Ervin (1998), based on Lemma 2 - *Consistency of variance estimate* by Hinkley[3] (1977), and the degree-of-freedom White heteroskedasticity consistent covariance matrix estimator. Finally we outlined the following estimator:

(1.3)

Where:

u_t^2 - the estimated residuals,

n - the number of observations (i.e. in our case 365 and in the last stage 349),

k - the number of regressors (i.e 23 at the start and in the final stage 12), and

$\frac{n}{n-k}$ -is degree-of-freedom correction

In our WLS performing, the estimator (1.4) and the default estimated coefficient covariance matrix (1.5) may be written as follows (Eviews, 2010):

[3]David V. Hinkley (1977) *Jackknifing in Unbalanced Situations*, Technometrics, Vol. 19, No. 3 (Aug.,), pp. 285-292.

(1.4)


(1.5)

Where:

D - a diagonal matrix containing the scaled *w* along the diagonal
z and *X* - matrices associated with z_t and x_t

In the *first stage* of our research, 365 observations had been introduced in an unbalanced undated worksheet type. The selected method in the equation estimation settings was the Least Squares. Our model started with EXPENDITURE as regressand and with 23 predictors, except the intercept, as follows: NO_VISITS and NO_DAYS – quantitative variables; MT_T, MT_B, MT_A, TD, D_O1, NA_ND, TA, WOMAN, tourist up to 25 years old, aged between 26 and 35, between 46 and 55 years old, between 56 and 65 of age, over 65 years old, UNMARRIED, DIV, WID, MID_S, VOC_S, H_S, MD, PhD – dichotomous variables, respectively.

Following the boxplot graphical representation (i.e. Fig. no. 1 tourist expenditures) and performing tabulation of expenditures series, we noted: a) near outliers (i.e. green circles in Fig no. 1) stands at around 3,000 lei and far outliers over 3,000 lei ((i.e. purple stars in Fig. no. 1); b) over 95% of the categories/tourists are spending up to 3,001 lei estimated expenses (i.e. Appendix A). Consequently it made an adjustment of the sample by reducing the number of categories (i.e. lower estimated expenses by 3,001 Lei) and far outliers dropped below 3,000 lei (i.e. purple stars in Fig. no 1 second part).

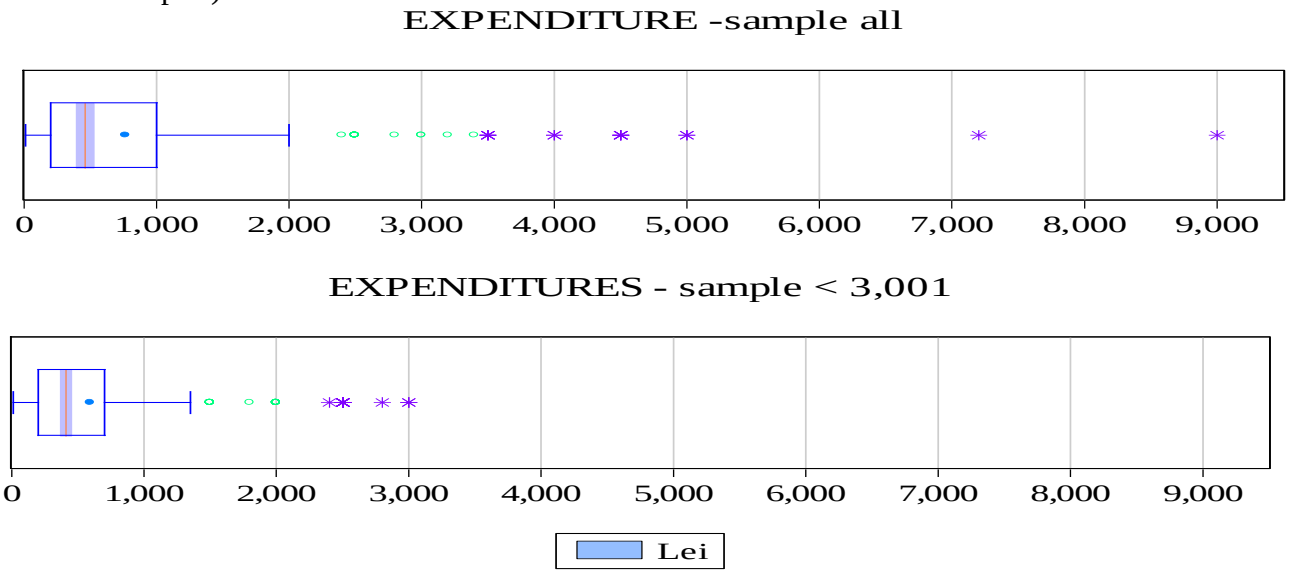

Figure 1. Tourist expenditures boxplot

Source: authors' own processing data in Eviews 7.2

In stage three, the insignificant covariates were eliminated (i.e. MT_B, tourist between _46_55 age, over _65 age, DIV, WID, MID_S, VOC_S, H_S, PhD). One can notice that the model kept those predictors that have improved the probability after the second (i.e. WOMAN, _26_35 age) or this stage (i.e. MT_T, MT_O, TD, _56_65 age, MD).

In the next stage, it was conducted the logarithm of the controlled variable and independent variables respectively (i.e. NO_VISITS and NO_DAYS,).

In order to ”improve” the covariates probability, and considering the relation (1.3) in the *fifth stage* of our research in equation estimation (1.2), coefficient covariance matrix, we have chosen the White cross-section standard errors & covariance option (d.f. corrected). Here the total number of stacked observations are 349 for each of the variables in the model, and the total number of estimated parameters are 14 (i.e. C, Log(NO_VISITS), Log(NO_DAYS), MT_T, MT_O, TD, D01, NA_ND, TA, WOMAN, _26_35, _56_65, UNMERR, MD).

In stage six, the NA_ND predictor was dropped, and in the end we wanted to see the effect of Weights options in equation estimation. Consequently, in those ”circumstances”, in Least Square for equation (1.2), we specified settings for Weights (i.e. considering relation (1.4) and (1.5)) as follows: type – *Inverse standard deviation* – relation; Weight series *NO_DAYS* and Scaling – *Average* has computing coefficient covariance method.

Comparison between real data and the forecast from the model

At the end of our research we highlighted a comparison between the actual data on the average/estimated expenditure for tourists. In other words, it was analysed the pair Log(Expenditure)&Log(Expenditure)F, but with information data from the forecasting model using (1.2) from stage seven. The forecast sample is 1 to 365 tourists, the adjusted sample is expenditure below 3,001 Lei and the total number of included observations is 349.

Results and discussion

1. Using the data from the unbalanced undated worksheet and the regression (1.1), we acquired the following results:

Table 1

Explanatory variables	Coefficient	Standard Error	t-Statistic	Prob.
Intercept	1264.958	217.1972	5.82400	0.0000
NO_VISITS	-155.9150	40.34350	-3.8646	0.0001
NO_DAYS	65.2231	10.32224	6.3186	0.0000
MT_T	-383.3590	293.8473	-1.3046	0.1929
MT_B	-80.6402	157.6928	-0.5113	0.6094
MT_O	217.7114	256.0755	0.8501	0.3958
TD	51.8408	363.7085	0.1425	0.8867
D_01	256.5465	395.8056	0.6481	0.5173
NA_ND	-329.7710	181.9225	-1.8127	0.0708
TA	375.6136	112.8223	3.3292	0.0010
WOMAN	-33.3678	101.2636	-0.3295	0.7420
_<25 age	-154.9971	173.9722	-0.8909	0.3736
_26-35 age	13.8910	154.0751	0.0901	0.9282
_46-55 age	71.8951	173.8959	0.4134	0.6795
_56-65 age	-448.603	198.5756	-2.2591	0.0245
_>65 age	-414.583	304.9085	-1.3596	0.1748
UNMARRIED	-431.235	135.7386	-3.1769	0.0016
DIV	-344.736	212.3491	-1.6234	0.1054
WID	-90.1044	304.6223	-0.2957	0.7676
MID_S	-369.643	387.9091	-0.9529	0.3413
VOC_S	-477.628	221.8781	-2.1526	0.0320
H_S	-9.70780	132.3496	-0.0733	0.9416
MD	-47.0617	132.0875	-0.3562	0.7218
PhD	12.0862	238.8683	0.0505	0.9597
		$R^2 = 0.2780$		

Source: authors' own processing data in Eviews 7.2

As these regression results show, the estimated coefficients in (1.1) are highly statistically significant for NO_VISITS, NO_DAYS, as the *p* value is very low. The "slope" for TA, 56-65 age UNMARRIED, VOC_S is statistically significant at about 3 percent, and the "slope" for NA_ND is significant at the .10 level. The intercept coefficient is also statistically significant as its *p* value is 0.0000. The "slope" in (1.1) is not statistically significant for MT_T, MT_B, MT_O, TD, D_01, WOMAN, _<25 age, _26-35 age, _46-55 age, _>65 age, DIV, WID, MID_S, H_S, PhD, respectively, as the *p* value exceeds 0.1054 (n.a. starting with exogenous variable DIV (*p* =0.1054), and ending with the control variable PhD (*p* =0.9597).

We are aware of that tourists' average expenditure is distorted by the near outliers and far outliers (see Fig. no. 1). Please note that in stage II these values were "dropped" from the work-file because we adjusted the sample.

The coefficient of determination R^2 shows that the sample regression line does not fit the data, as its value is 0.2780.

2. Under these conditions and results, in stage two we carried out the adjustment of the sample from the previous stage, and consequently the regression model is as follows:

Table 2

Explanatory variables	Coefficient	Standard Error	t-Statistic	Prob.
Intercept	1011.199	127.1060	7.9555	0.0000
NO_VISITS	-127.7979	23.6121	-5.4123	0.0000
NO_DAYS	42.8586	6.1078	7.0169	0.0000
MT_T	-315.0978	167.4657	-1.8815	0.0608
MT_B	-34.0299	91.4942	-0.3719	0.7102
MT_O	-83.2580	155.6437	-0.5349	0.5931
TD	204.0701	207.2496	0.9846	0.3255
D_01	160.3495	225.6566	0.7105	0.4778
NA_ND	-239.0033	103.8226	-2.3020	0.0220
TA	184.9131	65.4070	2.8271	0.0050
WOMAN	-53.2192	58.9631	-0.9025	0.3674
_<25 age	-57.0694	100.0421	-0.5704	0.5688
_26-35 age	57.3003	89.8276	0.6378	0.5240
_46-55 age	-96.8687	105.8786	-0.9149	0.3609
_56-65 age	-137.9349	114.3654	-1.2060	0.2287
_>65 age	-264.0316	175.7472	-1.5023	0.1340
UNMARRIED	-242.8830	78.0283	-3.1127	0.0020
DIV	-50.2533	121.7028	-0.4129	0.6799
WID	-198.2546	176.1658	-1.1253	0.2613
MID_S	-64.7756	221.9046	-0.2919	0.7705
VOC_S	-173.2806	127.5111	-1.3589	0.1751
H_S	-85.7060	77.8516	-1.1008	0.2718
MD	127.6981	76.7989	1.6627	0.0973
PhD	48.5600	139.8585	0.3472	0.7287
		R^2 = 0.3144		

Note: relation (1.1) with adjusted sample –expenditure <3,001 lei

Source: authors' own processing data in Eviews 7.2

It can be noticed that the "slope" is highly statistically significant for NO_VISITS, NO_DAYS, and for intercept, as their *p* value is 0.0000. Also, the "slope" for NA_ND, TA,

UNMARRIED, is statistically significant at about 2 percent, respectively. The "slope" for MT_T and MD is significant at the .10 level.

The coefficient is not statistically significant for MT_B, MT_O, TD, D_01, WOMAN, _<25 age, _26-35 age, _46-55 age, _56-65 age, _>65 age, DIV, WID, MID_S, H_S, PhD, respectively, as the *p* value exceeds 0.1340 (n.a. starting with exogenous variable _>65 age (*p* =0.1340), and ending with the control variable MID_S (*p* =0.7705).

The R^2 coefficient of determination registered a slight improvement, from 0.2780 up to 0.3144, however the sample regression line does not fit the data, as its value is around 0.28.

3. In stage three the effect of eliminating the insignificant variables in the equation estimation (1.1) is highlighted in Table number 3:

Table 3

Explanatory variables	Coefficient	Standard Error	t-Statistic	Prob.
Intercept	846.5995	107.0388	7.9092	0.0000
NO_VISIT	-119.6812	23.3877	-5.1172	0.0000
NO_DAYS	42.9790	5.9858	7.1801	0.0000
continuous				
Explanatory variables	Coefficient	Standard Error	t-Statistic	Prob.
MT_T	-301.2493	167.0923	-1.8028	0.0723
MT_O	-79.9790	153.4423	-0.5212	0.6025
TD	211.6403	206.2465	1.0261	0.3056
D_01	95.8144	222.4746	0.4306	0.6670
NA_ND	-199.4095	102.7772	-1.9402	0.0532
TA	200.3466	63.9269	3.1339	0.0019
WOMAN	-66.1670	57.7573	-1.1456	0.2528
_26_35 age	133.6662	71.0794	1.8805	0.0609
_56_65 age	-87.8999	105.1028	-0.8363	0.4036
UNMARRIED	-209.9773	62.5222	-3.3584	0.0009
MD	189.4970	69.1914	2.7387	0.0065
R^2 = 0.2903; S.E. of regression 529.8171; Sum squared residuals 94036567;				
F-statistic 13.3537; Prob. (F-statistic) 0.0000				

Note: adjusted sample –expenditure <3,001 lei

Source: authors' own processing data in Eviews 7.2

The table shows a change of the *se* and *t* values for independent variables. Most of the explanatory variables registered a decrease by about half of the value, each compared with stage one. However, we are aware of the *se*'s and *t*'s high values.

Also, one can notice that it improved the statistical significance of the dichotomous variables MT_T, MT_O, D_01, WOMAN, _26-35 age, and MD. Yet, out of 13 independent variables, excluding the intercept, only 6 are statistically significant at this stage.

Regarding the summary statistics of the regression we notice that the Standard Error of regression and the Sum squared residuals showed higher values. The *p* value of F-statistic is less than the significance level of 5%, so we reject the null hypothesis that all the slope coefficients are equal to zero.

4. The effects of conducting the logarithm of controlled variable and independent variables are as it follows:

Table 4

Explanatory variables	Coefficient	Standard Error	t-Statistic	Prob.
Intercept	5.7307	0.1684	34.028	0.0000
Log(NO_VISIT)	-0.3831	0.0980	-3.9090	0.0001

Log(NO_DAYS)	0.6476	0.0696	9.3027	0.0000
MT_T	-0.5447	0.2710	-2.0099	0.0452
MT_O	-0.5482	0.2482	-2.2087	0.0279
TD	0.1286	0.3338	0.3852	0.7003
D_01	0.5859	0.3605	1.6254	0.1050
NA_ND	-0.2089	0.1658	-1.2599	0.2086
TA	0.2970	0.1040	2.8549	0.0046
WOMAN	-0.0329	0.0934	-0.3528	0.7244
_26_35 age	0.1256	0.1150	1.0928	0.2752
_56_65 age	-0.0357	0.1706	-0.2096	0.8341
UNMARRIED	-0.4050	0.1010	-4.0078	0.0001
MD	0.2462	0.1120	2.1982	0.0286
R^2 = 0.3413; S.E. of regression 0.8573; Sum squared residuals 246.2637;				
F-statistic 13.3537; Prob. (F-statistic) 0.0000				

Note: adjusted sample –expenditure <3,001 lei

Source: authors' own processing data in Eviews 7.2

It can be seen that out of 13 independent variables, excluding the intercept, only 8 are statistically significant at this stage.

The Standard Error of regression has improved substantially, from 529.8171 in stage three to 0.8573 in this stage, but the Sum squared residuals had still registered a higher value (i.e. 246.2637). The probability of F-statistic is less than the significance level 0.0000, so we reject the null hypothesis that all the slope coefficients are equal to zero. Under these circumstances, R^2 has improved from 0.2903 in stage three, to 0.3413 in the current stage, but all variables, regressand and the regressors, are not highly positively correlated.

5. White cross-section standard errors & covariance option (d.f. corrected), took effect on Standard Error, t-statistic, probability respectively, as follows:

Table 5

Explanatory variables	Coefficient	Standard Error	t-Statistic	Prob.
Intercept	5.7307	0.1742	32.8855	0.0000
Log(NO_VISIT)	-0.3831	0.1032	-3.7122	0.0002
Log(NO_DAYS)	0.6476	0.0784	8.2597	0.0000
MT_T	-0.5447	0.2134	-2.5516	0.0112
MT_O	-0.5482	0.3147	-1.7420	0.0824
TD	0.1286	0.2774	0.4636	0.6432
D_01	0.5859	0.2183	2.6839	0.0076
NA_ND	-0.2089	0.1472	-1.4188	0.1569
TA	0.2970	0.1094	2.7138	0.0070
WOMAN	-0.0329	0.0936	-0.3522	0.7249
_26_35 age	0.1256	0.1188	1.0578	0.2909
_56_65 age	-0.0357	0.1823	-0.1961	0.8446
UNMARRIED	-0.4050	0.1066	-3.7991	0.0002
MD	0.2462	0.1010	2.4371	0.0153
R^2 = 0.3413; S.E. of regression 0.8573; Sum squared residuals 246.2637;				
F-statistic 13.3537; Prob. (F-statistic) 0.0000				

Note: adjusted sample –expenditure <3,001 lei

Source: authors' own processing data in Eviews 7.2

It can be seen that the standard error values for 9 variables had increased insignificantly, and it had insignificantly decreased for 5 variables, while t-statistic is also approximately equal to the

previous stage. Probability is significant for 5 dichotomous variables, for the quantitative variable, and for the intercept, respectively. However 5 qualitative variables aren't statistically significant (i.e. TD, NA_ND, WOMAN, _26_35 age, _56_65 age) and one is significant at 0.10 level (i.e. MT_O)

6. In the sixth stage, the effect of eliminating the insignificant variable NA_ND in the equation estimation (1.1) is highlighted in Table number 6:

Table 6

Explanatory variables	Coefficient	Standard Error	t-Statistic	Prob.
Intercept	5.6772	0.1669	34.0072	0.0000
Log(NO_VISIT)	-0.3734	0.1034	-3.6089	0.0004
Log(NO_DAYS)	0.6470	0.0780	8.2849	0.0000
MT_T	-0.5524	0.2239	-2.4667	0.0141
continuous				
Explanatory variables	Coefficient	Standard Error	t-Statistic	Prob.
MT_O	-0.5509	0.3166	-1.7401	0.0828
TD	0.1660	0.2735	0.6068	0.5444
D_01	0.6263	0.2130	2.9399	0.0035
TA	0.3407	0.0988	3.4460	0.0006
WOMAN	-0.0358	0.0933	-0.3840	0.7012
_26_35 age	0.1407	0.1179	1.1933	0.2336
_56_65 age	-0.0341	0.1823	-0.1870	0.8517
UNMARRIED	-0.4079	0.1063	-3.8355	0.0001
MD	0.2439	0.1029	2.3696	0.0184
R^2 = 0.3382; S.E. of regression 0.8581; Sum squared residuals 247.4307; F-statistic 14.3092; Prob. (F-statistic) 0.0000				

Note: adjusted sample –expenditure <3,001 lei

Source: authors' own processing data in Eviews 7.2

The elimination of predictor variable NA_ND has led to a slight improvement of probabilities of other insignificant stimulus variables (i.e. TD, WOMAN, 26-35 age and _56-65 age). Yet, out of 12 independent variables, excluding the intercept, only 8 are statistically significant at this stage.

Regarding the summary statistics of regression, we have noticed that the Standard Error of regression and the Sum squared residuals registered almost the same values compared to the prior stage. The *p* value of F-statistic is less than the significance level of 5%, so we reject the null hypothesis that all the slope coefficients are equal to zero. The R^2 coefficient of determination slightly decreased, from 0.3413 to 0.3382. However, the sample regression line does not fit the data, as its value is around 0.34.

7. In the last stage, the effect of the specified settings for Weights in equation estimation (1.2) is underlined as follows:

Table 7.1

Explanatory variables	Coefficient	Standard Error	t-Statistic	Prob.
Intercept	6.1776	0.2237	27.6081	0.0000
Log(NO_VISIT)	-0.6475	0.1542	-4.1992	0.0000
Log(NO_DAYS)	0.5805	0.0475	12.2145	0.0000
MT_T	-0.5897	0.2484	-2.3737	0.0182
MT_O	-0.4389	0.1170	-3.7501	0.0002
TD	0.6953	0.2033	3.4195	0.0007

D_01	0.6230	0.2376	2.6216	0.0091
TA	0.5313	0.1772	2.9973	0.0029
WOMAN	-0.3479	0.1722	-2.0205	0.0441
_26_35 age	0.3262	0.1465	2.2270	0.0266
_56_65 age	-1.1602	0.4721	-2.4572	0.0145
UNMARRIED	-0.4933	0.1609	-3.0659	0.0023
MD	0.2739	0.1284	2.1324	0.0337

R^2 = 0.7177; S.E. of regression 0.5428; Sum squared residuals 99.015;
F-statistic: 71.1953; Prob. (F-statistic) 0.0000

Note: relation (1.2) with adjusted sample –expenditure <3,001 lei

Source: authors' own processing data in Eviews 7.2

As these regression results indicate, the estimated coefficients in (1.2) are highly statistically significant for log(NO_VISITS), log(NO_DAYS), MT_O, TD, D_01, TA, UNMARRIED, and intercept, as their *p* value is very low. At the same time the "slope" for MT_T, WOMAN, _26-35 age, 56-65 age and MD, is statistically significant.

Also, the interpretation of 1.2 is that the elasticity of EXPENDURE with respect to NO_DAYS is about 0.53, suggesting that if the total number of days goes up by 1 percent, on average, the expenditure of tourists goes up by about 0.53 percent. Thus, tourists' expenditure is quite responsive to changes in number of days. Similarly, the interpretation of 1.2 is that the elasticity of EXPENDURE with respect to NO_VISITS is about -0.64, suggesting that if the total number of visits goes up by 1 percent, on average, the tourists' expenditure goes down by about 0.64 percent. Therefore, tourists' expenditure is reasonably responsive to changes in number of visits but this is a "negative reaction" from the tourism service and local economy point of view. In our opinion, this should determine the local policy makers (i.e. Town Council of Alba Iulia, Alba County Council, etc.) to identify new tourist offers or to improve the existing ones.

In terms of the dichotomous variables related to transport, one can notice that MT_T and MT_O cause a decrease in expenditure. The elasticity of EXPENDURE with respect to MT_T is about -0.59, suggesting that if the total number of tourists who travel by *train* goes up by 1 percent, on average, the tourists' expenditure goes down by about 0.59 percent, for an actual average expenditure of 275.89 lei (i.e. $e^{5.62}$, 5.62 =6.18-0.59). The elasticity of EXPENDURE with respect to MT_O is about -0.44, suggesting that if the total number of tourists who travel by other mean of transport goes up by 1 percent, on average, the tourists' expenditure goes down by about 0.44 percent, for an actual average expenditure of 311.06 lei (i.e. $e^{5.74}$, 5.74 =6.18-0.44,). In this case, we want to remind that the benchmark category is the Mean of Transport by Car (i.e. MT_C). This can be justified, perhaps, by lower fares practised in our country, issues that entail some savings or small expenditure. The European Commission Detailed Average Prices Report[4] shows that Romania recorded the lowest price level of a train ticket (i.e. 1.69 euros, average in 2015), ranking our country on the last place out of a total of 10 reporting countries. In the case of urban bus transport, single ticket, Romania is on the last place out of 13 reporting countries (i.e. 0.43 euro, average in 2015).

As regards the dichotomous variables related to staff amiability of hotel/ hostel staff (i.e. Inquiry no. 25), it can be noticed that TD, D_01 and TA determine an increase in expenditure. The elasticity of EXPENDURE with respect to TD is about 0.70, suggesting that if the total number of tourists who are in Total Disagree with Inquiry no. 25 goes up by 1 percent, on average, the

[4] The *Detailed Average Prices Report* (august, 2016), of European Commission, Eurostat, is based on the price data collected in 2015 for a set of consumer products. The products were specified using a reduced version of the PPP standard product definitions. *General Rail travel description*: Type: single ticket, domestic trip; Bought: same day at the station; Class: 2nd ("regular" if not rated); Fare for: adult; Time: week day; Trip length: approx. 50 km; Exclude: seat reservations, reduced tickets, high speed trains, tickets bought with rail card; Average 2015. General *Urban bus transport, single ticket description*: Bought: in advance; Fare for: adult; Trip distance: 5 km or two zone; reference quantity: 1 ticket; Average 2015.

tourists' expenditure goes up by about 0.70 percent, for an actual average expenditure of 239.85 lei (i.e. $e^{5.48}$, 5.48 =6.18-0.70). The elasticity of EXPENDURE with respect to D_01 is about 0.62, suggesting that if the total number of tourists who are in Disagree with Inquiry no. 25 goes up by 1 percent, on average, the tourists' expenditure goes up by about 0.62 percent, for an actual average expenditure of 257.24 lei (i.e. $e^{5.55}$). The elasticity of EXPENDURE with respect to TA is about 0.53, suggesting that if the total number of tourists who are in Total Agree with Inquiry no. 25 goes up by 1 percent, on average, the tourists' expenditure goes up by about 0.53 percent, for an actual average expenditure of 284.29 lei (i.e. $e^{5.65}$) In this case, we wish to remind that the benchmark category is *Agree*.

What is extremely interesting and "unexpected" for us, is the sign of these predictor variables. All variables have positive coefficients, which implies that expenses increase both for tourists who do not agree with question no. 25, and for those who agree. The prior category of tourists probably tries to identify hotels/ hostels that meet their needs, and the last category is probably willing to spend more because their needs are satisfied.

It can be observed that the average expenditure of man tourist is 481.83 lei (i.e. $e^{6.1776}$), and the "slope" coefficient β_9 indicates that the average expenditure of woman tourist is lower by about 141 lei, for an actual average expenditure of 340.46 lei (i.e. $e^{5.83}$)

Considering the tourists age, one can observe that the average expenditure of tourists aged between 26 and 35 is 665.14 lei (i.e. $e^{6.50}$), and the "slope" coefficient β_{10} indicates that the average expenditure of tourists of 26-35 years old is higher by about 183 lei. The mean expenditure of tourist of 56-65 years old is 151.41 lei (i.e. $e^{5.02}$), and the "slope" coefficient β_{11} tells us that the average expenditure of tourists aged between 56 and 65 is lower by about 330 lei. From our point of view, these values can be explained in two ways: 1) most tourists are from Romania and have a relatively low income level; 2) generally younger tourists are willing to consume more, especially if they have children, compared to older tourists who usually have a lower level of income.

As for the tourists' marital status and their level of education, one can notice that the mean expenditure of an unmarried tourist is 292.95 lei (i.e. $e^{5.68}$), and the "slope" coefficient β_{12} tells us that the average expenditure of unmarried tourists is lower by about 189 lei. The mean expenditure of the tourist with a Master's Degree is 632.70 lei (i.e. $e^{6.45}$), and the "slope" coefficient β_{13} indicates that the mean expenditure of the tourist with a Master's Degree is higher by about 151 lei.

Table 7.2

Estimation output for (1.2)

	Weighted	Unweighted
Statistic	Statistics	Statistics
R-squared	*0.7177*	*0.1840*
Adjusted R-squared	0.7076	0.1548
S.E. of regression	0.5428	0.9528
Sum squared residuals	99.0152	305.0704
F-statistic	71.1952	
Prob(F-statistic)	0.0000	

Note: relation (1.2) with adjusted sample –expenditure <3,001 lei

Source: authors' own processing data in Eviews 7.2

The R^2 coefficient of determination in *weighted statistics* shows that the sample regression line does fit the data, as its value is 0.72, but in the case of *unweighted statistics*, it is only 0.18, which indicates that the predictand and the predictors are not highly positively correlated. In the case of weighted statistics, the Standard Error of regression registered a low value and the Sum squared residuals decreased to 99.01. In the case of unweighted statistics, the Standard Error of regression registered a low value and the Sum squared residuals reported a higher value. The *p*

value of F-statistic was less than the significance level of 5%, so we reject the null hypothesis that all the slope coefficients are equal to zero.

Comparing the real data and the forecast

The forecast evaluation details are presented in Table no. 8 *Forecast evaluation of (1.2).*

The forecast sample is 1 to 365, the adjusted sample is expenditure <3001, and the number of observations included is 349.

The reported forecast statistics indicate that our forecasting model does perform well out-of-sample.

The Root Mean Squared Error is small when compared to the standard deviation of Log(expenditure) series (i.e. 1.04).

Table 8

Forecast evaluation of (1.2)

Indicators	Value
Root Mean Squared Error	0.934948
Mean Absolute Error	0.720289
Mean Absolute Percentage Error	13.44154
Theil Inequality Coefficient	0.078415
Bias Proportion	0.000712
Variance Proportion	0.072562
Covariance Proportion	0.926727

Note: Forecast sample: 1 to 365 if expenditure is <3,001, included observations: 349.

Source: authors' own processing data with EViews 7.2

The Theil Inequality Coefficient (i.e. TIC) shows an average error of about 0.078 which is relative minor, but the value is smaller than 1. Also the Mean Absolute Percentage Error (i.e. MAPE) is higher, but it is well known that MAPE is scale sensitive and should not be used when working with low-volume data (i.e. 349 observations). The Bias and Variance Proportions are small, which implies that the error of prediction is concentrated in covariance proportion (i.e. 0.93) and shows that the forecast is quite "good". In conclusion, TIC gives a more acceptable indicator than MAPE for measuring the "fit" of the model.

Comparing the real data and the forecast

According to real data, there is an increase of tourists' expenditure in 49% of the analysed observations and a decrease in 51% of the analysed observations.

Table 9

Real data and forecast for dependent variable

Series	Number of values>0	Number of values<0
Log(expenditure)	349	349
Log(expenditure)F	171	178
%	49%	51%

Note: Forecast sample: 1 to 365 if expenditure<3001;

Source: authors' own processing data in EViews7.2

Due to the forecast results based on (1.2) for adjusted sample by expenditure <3,001 lei, it is established that Log(expenditure)F increased in only 49% of the cases and the expenditure has fallen in 51% of observations. Diagrammatically, we have the situation shown in Figure number 2.

Therefore, using the elaborated model (1.2), the actual data and the predicted results of this simulation, it can be noticed that the predictor variables are not a "strong response" for tourists

expenditure in the town of Alba Iulia and in Alba County. Actually, half of the explanatory variables (i.e. Log(No_Days), TD, D_01, TA, 26-35 age, MD) contribute to the increase in spending, and the other half (i.e. Log(No_Visits), MT_T, MT_O, Woman, 56-65 age, Unmarried) negatively affect the estimated expenditure of tourists.

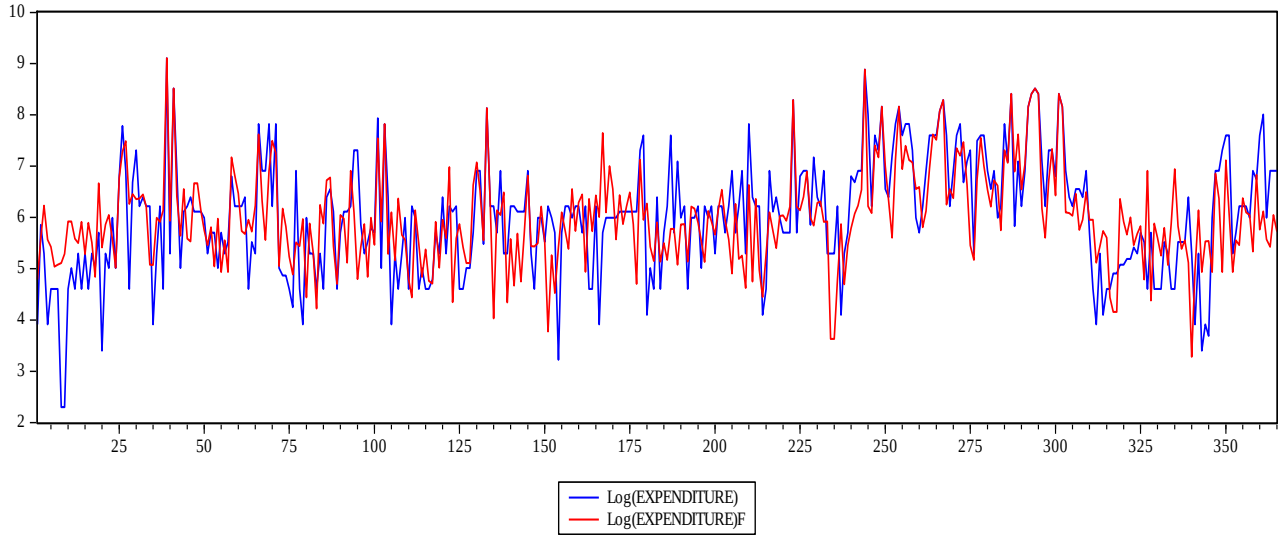

Figure 2. The actual and the predicted data of tourist's expenditure plotting

Source: authors' own processing data in EViews 7.2

Study Limitations

In this article we reviewed the tourist expenditure literature and we established a regression model with specific predictors. Our study here, however, has its limitations, which determined the approach for further research in this field. For example, the investigation on the influence of the perceived value (Bradley & Sparks, 2012) and the service quality (Um et al., 2006) upon tourist satisfaction and implicitly upon tourist spending. Also, we intend to use this potential regressors which may provide some forecastability.

The article admits the next limitations:

- there are other relevant concepts as follows: decision-making (Quintal, et. al., 2010), values (Li & Cai, 2012), motivations (Kang, et. al., 2012), self-concept and personality (technology (Cohen, et al., 2014), that we do not address within the scope of our research, and in addition, to better understand the impacts of socio-demographic and travel-related variables, other consumer behaviour specific variables may deserve further research efforts;
- the study had an unbalanced, undated work-file which determined a "limited view" of the regressand's dynamic and an estimation for the next periods. Indeed, it is well known that a longitudinal research in tourism expenditure would provide better image, "thus offering a unique perspective on how the behaviour and its influences evolve over time" (Cohen, et al., 2014:898);
- the article had a relatively small sample which, probably, influenced the Unweighted statistics (i.e. included observations: 349).

Conclusions

The results of this research provided some explanation for different predictors of the average expenditure of the tourists (i.e. 349) visiting Alba Iulia and Alba County during 2013-2015. Twelve major determinants in tourism expenditures were identified: Number of Visits, Number of Days, Means of Transport - Train, Means of Transport - Other, Total Disagree, Disagree, Total Agree, Woman, and tourists aged between 26 and 35, tourists aged between 56 and 65, Unmarried, and Master's Degree.

Regarding the methodological approach, it can be noticed that the following steps were recorded: first regression with all variables in stage one; adjustment of the sample in stage two (i.e. expenditure<3,001 Lei); elimination of insignificant predictors in stage three and six; conducting the logarithm of controlled variable and independent variables in stage four; White cross-section standard error and covariance option –d.f. corrected, in stage five, and in the last stage setting for Weights in the equation estimation.

As these regression results indicate, the estimated coefficients in (1.2) are highly statistically significant for log(NO_VISITS), log(NO_DAYS), MT_O, TD, D_01, TA, UNMARRIED, and intercept, as their *p* value is close to zero. At the same time, the "slope" for MT_T, WOMAN, _26-35 age, 56-65 age and MD, is less than the significance level of 5%.

The profile of the tourist who visited the town of Alba Iulia and Alba County during 2013-2015, and the register of the increase of tourist spending may be sketched as follows: man aged between 26 and 35 years, between 56 and 65 years old, married, master's degree and using the car.

In terms of explanatory variables, from the interpretation of 1.2, one can understand that the elasticity of EXPENDURE with respect to NO_DAYS is about 0.53, suggesting that, on average, the expenditure of tourists goes up by about 0.53 percent. Therefore, the tourist expenditure is quite responsive to the changes in number of days. Likewise, the interpretation of 1.2 is that the elasticity of EXPENDURE with respect to NO_VISITS is about -0.64, signifying that, on average, the expenditure of tourists goes down by about 0.64 percent. Hence, the tourist expenditure is sensibly responsive to the changes in number of visits, but this is a "negative reaction" from the point of view of touristic service and local economy. In our opinion, this, should determine the local policy makers (i.e. Town Council of Alba Iulia, Alba County Council, etc.) to establish new touristic offers or to develop existing ones.

With regard to the dichotomous variables related to transport, it was highlighted that MT_T and MT_O cause a decrease in expenditure. On average, the expenditure of tourists goes down by about 0.59 percent, for an actual average expenditure of 275.89 lei (i.e. almost 61 euros), in the case of tourists travelling by train; and it goes down by about 0.44 percent, for an actual average expenditure of 311.06 lei (i.e. approx. 69 euros) for those who travel by other mean of transport (i.e. benchmark category is the Mean of Transport by Car -MT_C). Perhaps, this can be justified by the lower fares practised in our country, issues that entail some savings or small expenditure. According to the Report of European Commission (2015), Romania recorded the lowest level of a train ticket. Also, in the case of a single ticket for urban bus transport, Romania is on the last place out of 13 reporting countries.

The dichotomous variables related to staff amiability of hotel/ hostel staff (i.e. Inquiry no. 25), determine an increase in expenditure. The elasticity of EXPENDURE with respect to TD is about 0.70 (i.e. benchmark category is *Agree),* for an actual average expenditure of 239.85 lei (i.e. 53 euros); the one of EXPENDURE with respect to D_01 is about 0.62, for an actual average expenditure of 257.24 lei (i.e. 57 euros); and the elasticity of EXPENDURE with respect to TA is about 0.53, for an actual average expenditure of 284.29 lei (i.e. almost 63 euros), respectively.

We may notice the sign of these predictor variables. All variables have positive coefficients, which implies that expenses increase both for tourists who do not agree with question no. 25, and for those who agree. This results may be due to the limits of our research.

Furthermore, the average expenditure of man tourist is 481.83 lei (i.e. 107 euros), and the mean expenditure of woman tourist is lower by about 141 lei (i.e. 31 euros).

Considering the tourists' age, one can observe that the mean expenditure of the tourist aged between 26 and 35 is higher than of the 36_45 category, by about 183 lei (i.e. almost 41 euros), and the mean expenditure of the tourist aged between 56 and 65 is lower by about 330 lei (i.e. approx. 73 euros). We think that these values can be explained in two ways: 1) most tourists are from Romania and they have a relatively low income level; 2) generally younger tourists are willing to

consume more, especially if they are tourists who have children, compared to older Romanian tourists who, usually, have a lower level of income.

Taking into account the tourist's marital status and level of education, one can notice that the mean expenditure of unmarried tourists is lower than of married ones by about 189 lei (i.e. 42euros). Also, from the point of view of the education level, the mean expenditure of tourists with Master's Degree is higher than the benchmark variable Bachelor's Degree by about 151 lei (i.e. 34 euros).

In the case of *weighted statistics*, the R^2 coefficient of determination shows that the sample regression line does fit the data, but in the case of *unweighted statistics*, the predictand and the predictors are not highly positively correlated. Therefore, in our future research, it is necessary to identify one or more independent variables that can improve the coefficient of determination.

Looking at the weighted statistics, the Standard Error of regression registered a low value and the Sum squared residuals decreased to 99.01. In terms of the unweighted statistics, the Standard Error of regression registered a low value and the Sum squared residuals reported a higher value. The *p* value of F-statistic was lower than the significance level of 5%, so we rejected the null hypothesis that all the slope coefficients are equal to zero.

The forecast evaluation details for the adjusted sample (i.e. expenditure <3,001 Lei) indicated that the Bias and Variance Proportions are small, which implies that the error of prediction is concentrated in covariance proportion and shows that the forecast is quite "acceptable". In conclusion, the Theil Inequality Coefficient gives a more acceptable indicator for measuring the "fit" of the model than the Mean Absolute Percentage Error (i.e. almost 13%).

Finally, the comparison between the real data and the forecast, using the elaborated model (1.2), outlined that half of the explanatory variables (i.e. Log(No_Days), TD, D_01, TA, 26-35 age, MD) contributed to the increase in spending and the other half (i.e. Log(No_Visits), MT_T, MT_O, Woman, 56-65 age, Unmarried) negatively affected tourists' estimated expenditure.

In conclusion, even if we managed to point out a profile of tourists visiting Alba County, we need to identify other methods, regression models to highlight a better measure of the average expenditure of tourists.

Acknowledgements

The research assistance of our renowned colleague Professor Ph.D. Nicoleta Breaz is greatly appreciated by authors. In writing this paper, we have greatly benefited from discussions with our friend Alina Popa. We are grateful to anonymous referees for helpful debates and suggestions.

Bibliography

1. Agarwal, V.B., & Yochum, G.R., 1999. Tourist Spending and Race of Visitors. Journal of Travel Research, (38), 173-176
2. Asgary, N., De Los Santos, G., Vincent, V., & Davila, V., 1997. The determinants of expenditures by Mexican visitors to the border cities of Texas. Tourism Economics, 3(4), 319-328.
3. Barry K., O' Hagan, J., 1972. Econometric study of British tourist expenditure in Ireland, Economic and Social Research Institute, Economic and Social Review, 3(2), 143-161.
4. Boksberger, P., Dolnicar, S., Laesser, C., & Randle, M., 2011. Self-congruity theory: To what extent does it hold in tourism? Journal of Travel Research, 50(4), 454–464.
5. Bradley, G., & Sparks, B., 2012. Antecedents and consequences of consumer value: A longitudinal study of timeshare owners. Journal of Travel Research, 51(2), 191–204.
6. Cohen, S. A., Prayag, G. and Moital, M., 2014. Consumer behaviour in tourism: Concepts, influences and opportunities, Current Issues in Tourism, 17(10), 872-909.
7. Efron, B., 1982. The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia, PA.

8. Fredman, P., 2008. Determinants of visitor expenditures in mountain tourism. Tourism Economics, 14 (2), 297–311.
9. Hinkley, D., V., 1977. Jackknifing in Unbalanced Situations. Technometrics, 19(3), 285-292.
10. Horn, S.D., R.A. Horn, and D.B. Duncan, 1975. Estimating heteroscedastic variances in linear model, Journal of the American Statistical Association, 70, 380-385.
11. Kang, E. J., Scott, N., Lee, T. J., & Ballantyne, R., 2012. Benefits of visiting a dark tourism site: The case of the Jeju April 3rd peace park, Korea. Tourism Management, 33, 257–265.
12. Quintal, V., Lee, J., & Soutar, G., 2010. Risk, uncertainty and the theory of planned behavior: A tourism example. Tourism Management, 31, 797–805.
13. Lim, C., 1997. Review of International Tourism Demand Models, Annals of Tourism Research, 24 (4), pp. 835-849.
14. Laimer, P., and Weiß, J., 2006. Data Sources On Tourism Expenditure. The Austrian Experiences Taking Into Account The Tbop Requirements. International Workshop On Tourism Statistics, Jointly organized by UN Statistics Division and UNWTO Department of Statistics and Economic Measurement of Tourism UNWTO Headquarters, Madrid/Spain 17-20 July 2006.
15. Li, M., & Cai, L. A., 2012. The effects of personal values on travel motivation and behavioral intention. Journal of Travel Research, 51(4), 473–487.
16. MacKinnon and White, 1985. Some heteroskedasticity consistent covariance matrix estimators with improved finite sample properties, Journal of Econometrics, 29, 305-325.
17. Nguyen, Q., 2016. Linking loss aversion and present bias with overspending behaviour of tourists: Insights from a lab-in-the-field experiment. Tourism Management, (54), 152-159.
18. Peerapatdit, N., 2004. The impact of selected socio-demographic, travel-related and psychographic variables on travel expenditure (Orlando, Florida).
19. Romano J. P., Wolf M., 2014. Resurrecting Weighted Least Squares. Technical Report No. 2014-1, Department of Statistics Stanford University, https://statistics.stanford.edu/sites/default/files/2014-11.pdf (May, 2016).
20. Sampol, C. J, and Perez E. A., 2000. Tourist expenditure for mass tourism markets. Annals of Tourism Research, 27(3), 624–637.
21. Scott Long, J., S. and Ervin L., H., 1998. Correcting for Heteroscedasticity with Heteroscedasticity Consistent Standard Errors in the Linear Regression Model: Small Sample Considerations, Working Paper, Indiana University, Bloomington, IN 47405, September 23, 5-6.
22. Starzs, R., 2015. Chapter 14, A Taste of Advanced Estimation, EViews Illustrated for Version 9, IHS Global Inc., 341-344.
23. White H., 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. Econometrica, Vol. 48(4), 817-838.
24. Um, S., Chon, K., & Ro, Y. H., 2006. Antecedents of revisit intention. Annals of Tourism Research, 33(4), 1141–1158.
25. ***EViews 7 User's Guide II, 2010. Chapter 35, Pooled Time Series, Cross-Section Data; Generalized Least Squares; Robust Coefficient Covariances Quantitative Micro Software, LLC, USA.
26. ***EViews 7 User's Guide II, 2010. Chapter 37, Panel Estimation, Least Squares Panel Options, Quantitative Micro Software, LLC, USA.
27. ***The Department of Economic and Social Affairs of the United Nations Secretariat, Statistics Division, 2008. Chapter 4 The demand perspective: tourism expenditure, International Recommendations for Tourism Statistics, Studies in Methods Series M No. 83/Rev.1, ISBN 978-92-1-161521-0.
28. ***EUROPEAN COMMISSION, EUROSTAT web site, Railway transport measurement, Railway transport - Passenger transport by type of transport/ detailed reporting only; 1, 000

passengers: http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=rail_pa_typepas&lang=en June, 2016.
29. ***EUROPEAN COMMISSION, EUROSTAT, Directorate C: National Accounts, Prices and Key Indicators, Unit C-4: Price statistics. Purchasing Power Parities. Housing statistics, DETAILED AVERAGE PRICES REPORT, August 2016 http://ec.europa.eu/eurostat/documents/272892/272992/Consumer-price-research-2015/ June, 2016.

Appendix A

Tabulation of EXPENDITURE

Sample: 1 to 365
Included observations: 365, Number of categories: 49

Value	Count	Percent	Cumulative Count	Cumulative Percent
10	2	0.55	2	0.55
25	1	0.27	3	0.82
30	2	0.55	5	1.37
40	1	0.27	6	1.64
50	9	2.47	15	4.11
60	4	1.10	19	5.21
70	1	0.27	20	5.48
100	38	10.41	58	15.89
120	1	0.27	59	16.16
130	2	0.55	61	16.71
135	2	0.55	63	17.26
150	13	3.56	76	20.82
160	2	0.55	78	21.37
180	2	0.55	80	21.92
200	34	9.32	114	31.23
225	1	0.27	115	31.51
240	1	0.27	116	31.78
250	9	2.47	125	34.25
300	27	7.40	152	41.64
330	1	0.27	153	41.92
340	1	0.27	154	42.19
350	3	0.82	157	43.01
400	23	6.30	180	49.32
450	18	4.93	198	54.25
500	47	12.88	245	67.12
600	14	3.84	259	70.96
700	5	1.37	264	72.33
800	3	0.82	267	73.15
850	1	0.27	268	73.42
900	5	1.37	273	74.79
1000	30	8.22	303	83.01
1200	5	1.37	308	84.38
1300	1	0.27	309	84.66
1350	1	0.27	310	84.93
1500	10	2.74	320	87.67
1800	1	0.27	321	87.95
2000	14	3.84	335	91.78
2400	1	0.27	336	92.05
2500	10	2.74	346	94.79
2800	1	0.27	347	95.07
3000	2	0.55	349	95.62
3200	1	0.27	350	95.89
3400	1	0.27	351	96.16
3500	4	1.10	355	97.26
4000	2	0.55	357	97.81
4500	4	1.10	361	98.90
5000	2	0.55	363	99.45
7200	1	0.27	364	99.73
9000	1	0.27	365	100.00
Total	365	100.00	365	100.00

Source: authors' own processing data with EViews7.2

Appendix B

Tabulation of Number of visits and days

Tabulation of NO_VISIT Sample: 1 to 365 Included observations: 365 Number of categories: 5					Tabulation of NO_DAYS Sample: 1 to 365 Included observations: 365 Number of categories: 18				
Valu	Cou	Percen	Cumu-lative Count	Cumu-lative Percent	Valu	Coun	Percent	Cumu-lative Count	Cumu-lative Percent
1	40	10.96	40	10.96	1	81	*22.19*	81	22.19
2	52	14.25	92	25.21	2	70	19.18	151	41.37
3	36	9.86	128	35.07	3	107	*29.32*	258	70.68
4	166	*45.48*	294	80.55	4	32	8.77	290	79.45
5	71	19.45	365	100	5	24	6.58	314	86.03
Total	365	100	365	100	6	6	1.64	320	87.67
					7	29	7.95	349	95.62
					8	2	0.55	351	96.16
					9	2	0.55	353	96.71
					10	3	0.82	356	97.53
					11	1	0.27	357	97.81
					12	1	0.27	358	98.08
					13	1	0.27	359	98.36
					14	1	0.27	360	98.63
					15	2	0.55	362	99.18
					21	1	0.27	363	99.45
					30	1	0.27	364	99.73
					80	1	0.27	365	100
					Total	365	100	365	100

Source: authors' own processing data with EViews7.2

Appendix C

Tabulation of LOG(EXPENDITURE)F and LOG(EXPENDITURE)

Tabulation of LOG(EXPENDITURE)F

Sample: FROM 1 TO 365 IF EXPENDITURE<3001 Lei
Included observations: 349
Number of categories: 5

Value	Count	Percent	Cumu-lative Count	Cumu-lative Percent
[3, 4)	4	1.15	4	1.15
[4, 5)	40	11.46	44	12.61
[5, 6)	159	45.56	203	58.17
[6, 7)	117	33.52	320	91.69
[7, 8)	29	8.31	349	100
Total	349	100	349	100

Tabulation of LOG(EXPENDITURE)

Sample: FROM 1 TO 365 IF EXPENDITURE<3001 Lei
Included observations: 349
Number of categories: 41

Value	Count	Percent	Cumu-lative Count	Cumu-lative Percent
2.302585	2	0.57	2	0.57
3.218876	1	0.29	3	0.86
3.401197	2	0.57	5	1.43
3.688879	1	0.29	6	1.72
3.912023	9	2.58	15	4.3
4.094345	4	1.15	19	5.44
4.248495	1	0.29	20	5.73
4.60517	38	*10.89*	58	16.62
4.787492	1	0.29	59	16.91
4.867534	2	0.57	61	17.48
4.905275	2	0.57	63	18.05
5.010635	13	3.72	76	21.78
5.075174	2	0.57	78	22.35
5.192957	2	0.57	80	22.92
5.298317	34	9.74	114	32.66
5.4161	1	0.29	115	32.95
5.480639	1	0.29	116	33.24
5.521461	9	2.58	125	35.82
5.703782	27	7.74	152	43.55
5.799093	1	0.29	153	43.84
5.828946	1	0.29	154	44.13
5.857933	3	0.86	157	44.99
5.991465	23	6.59	180	51.58
6.109248	18	5.16	198	56.73
6.214608	47	*13.47*	245	70.2
6.39693	14	4.01	259	74.21
6.55108	5	1.43	264	75.64
6.684612	3	0.86	267	76.5
6.745236	1	0.29	268	76.79
6.802395	5	1.43	273	78.22
6.907755	30	8.6	303	86.82
7.090077	5	1.43	308	88.25
7.17012	1	0.29	309	88.54
7.20786	1	0.29	310	88.83
7.31322	10	2.87	320	91.69
7.495542	1	0.29	321	91.98
7.600902	14	4.01	335	95.99
7.783224	1	0.29	336	96.28
7.824046	10	2.87	346	99.14
7.937375	1	0.29	347	99.43
8.006368	2	0.57	349	100
Total	349	100	349	100

Source: authors' own processing data with EViews7.2