Socioeconomic and environmental factor analysis of the international tourism expenditure in Spain

Jesus Barreal*†, Adolfo Hernández*‡, Gil Jannes*§

draft

Abstract

The international tourist expenditure in Spain depends on several factors. This paper proposes a model in which socioeconomic features, environmental qualities, touristic destinations and arrival regions are taken into account. An OLS model is developed to measure the relation of these variables with the deflacted daily expenditures of international tourists. Correlation level and Variance Inflaction Factor are considered to select the independent variables. Also, the functional form is tested using the Box-Cox tools. With all these elements, a panel data analysis is constructed and tested to discriminate between fixed and random effects. Finally the Hausman test is used to select the correct Panel data model. The main conclusion is that the expenditure depends on the type of accomodation, the touristic season and environmental features.

Keywords: Environmental quality, Panel Data, Tourism Expenditure, Socieconomic factors.

*Department of Financial and Actuarial Economics & Statistics, Universidad Complutense de Madrid
†jbarreal@ucm.es
‡adolfher@ucm.es
§gil.jannes@ucm.es

1 Introduction

The tourism sector is one of the most important economic activities in Spain. For example, the World Travel and Tourism Council estimates that tourism and travel currently accounts for a total contribution of roughly 15% of the economic activity in Spain, and forecasts this number to increase to 16.4% of GDP and 16.9% of employment by 2028 [WTTC, 2018]. Spain is thus one of the main touristic destinations not just in Europe, but worldwide. Indeed, the World Travel Organization finds that Spain is the second country in the world in terms of tourist arrivals as well as in terms of total receipts from tourism in 2017 [UNWTO, 2018].

The growing economic importance of tourism worldwide has led to a dramatical increase in statistical research on tourism in the past few decades (see [Song and Li, 2008] and [Goh and Law, 2011] for reviews). The aim is manifold: one wishes to better predict tourism demand in order to anticipate it, but also to better understand the factors determining tourism demand with the hope of being able to influence these. This might stimulate the aim of developing specific regions for tourism, or of emphasizing certain specific tourist profiles one wishes to attract. The question of how to define "tourism demand" is therefore non-trivial: It is easier to collect relatively accurate data on the number of tourist arrivals, and a large majority of studies on tourism have therefore used this variable. However, in many cases, tourism expenditure might in fact be a more interesting variable, even though this is more difficult to measure.

In this study, we are interested in tourism expenditure and use 2008-2018 data on tourism in Spain, collected from Egatur (Encuesta de gasto turístico - Survey of touristic expenditure), to study the average daily expenditure of international tourists in Spain in that period. We are not merely interested in describing this expenditure and its possible trend, but in studying the explanatory factors that determine this expenditure. We therefore apply a panel data procedure using an OLS (Ordinary Least Squares) methodology. We focus on factors intrinsic to the Spanish region of destination, which can be directly related to the tourism industry (e.g., the type of accomodation on offer) or more general, either socioeconomic (population density) or environmental (percentage of forested land). We also study factors associated to the country of origin, and compare touristic expenditure profiles from the main countries of origin, and verify whether there is a difference in the spending profile depending on the time of year.

2 Literature review

As mentioned above, reviews about statistical researches on tourism demand can be found in [Song and Li, 2008] and [Goh and Law, 2011]. Data panel analyses of tourism are usually addressed in terms of the number of tourist arrivals or of overnight stays, not of expenditure. This is true in particular for Spain. For example, [Garín-Muñoz and Amaral, 2000] has studied the influence of real per capita income, exchange rates and real prices on the number of overnight stays of international tourists in Spain in the period 1985-1995. The number of tourist arrivals to the Canary and Balearic Islands were studied in [Garín-Muñoz, 2006] and [Garín-Muñoz and Montero-Martín, 2007], respectively, and of German tourists specifically to the whole of Spain in [Garín-Muñoz, 2007]. In all three cases, the authors emphasize the economic conditions in the country of origin and the cost of tourism in the destination, as well as a consumer loyalty effect, as determining factors. An even more specific study was carried out in [Ledesma-Rodríguez et al., 2001], focusing on the island of Tenerife, while a recent study [Rey et al., 2011] examines the impact of low-cost airlines on Spanish tourism.

There are several interesting examples of panel data analyses of worldwide tourism in the literature. For example, [Eilat and Einav, 2004] is perhaps the first comprehensive attempt to understand the determinants of international tourism through a panel data method. General panel data studies of the influence of tourism development on economic growth worldwide (or comparing different countries) can be found, for example, in [Chou, 2013] and [Sequeira and Nunes, 2008], and specifically for the Mediterrean region in [Dritsakis, 2012].

Finally, recent panel data studies for specific countries other than Spain can be found, for example, in [Narayan et al., 2010] for the Pacific Island Countries; [Surugiu et al., 2011] for Romania, and [Ibrahim, 2011] for Egypt.

3 Methodology & data

3.1 The Econometric model

The Ordinary Least Squares (OLS) model could be design to measure the relation between expenditures of international tourists (Y_t) and other external variables [Greene, 2012]. Following this econometric model is defined the next equation:

$$Y_t = \mu + \beta_j X_{jt} + \epsilon$$

in which μ represents the intercept of the equation, X_{jit} is the vector of exogenous variables, β is the coefficient that relate the independent and dependent variable and ϵ_t is the error term. The β coefficient represents the effects of changes in the exogenous variable j in the dependent variable. therefore, if presents positive results, then an increment in the independent variable has a direct effect in the dependent variable. On the other side, the subscripts (t, j) represents the time and the division of exogenous variables, respectively. In this case there are divided in 3 groups; socioeconomic features, environmental factors and microdata provided by EGATUR Survey. To select this variables is going to be considered the correlation between them. In this sense, the variables that present more than 80% of correlation values is going to be dropped from the studio. In order to test the functional form is going to use the Box-Cox transformation. This tool changes the the original variable y_t into a new variable $y_t^{(\lambda)}$, in which λ is the factor that transforms the functional form. If $\lambda = 1$, then the functional form is equal to linear OLS. However if $\lambda \neq 1$, then the functional form is not linear.

$$y_t^{(\lambda)} = \begin{cases} \dfrac{y_t^{\lambda} - 1}{\lambda} & (\lambda \neq 0) \\ ln(y_t) & (\lambda = 0) \end{cases}$$

This transformation improves the treatability of OLS regression, by stabilizing variance and improving the overall smoothness and linearity of the distribution. The λ values are summarize in the Results section.The Variance Inflaction Factor (VIF) is developed to analyse the existence of multicolineality. These was calculated following the next equation, in which R_k^2 is obtained by regressing the each variables on the remaining variables.

$$VIF_k = \frac{1}{1 - R_k^2}$$

Therefore, if the VIF presents values upper than 10, then there is multicolineality. On the other hand, if the value are around one, there is no correlation. However, if the value are upper than 5, further investigation should be developed.

As the OLS model does not consider heterogeneity across regions or time, then a Panel data should be developed. Then, the next equation include the Data Panel model in which i represents each region or entity where data are recorded:

$$Y_t = \mu + \beta_j X_{jit} + \epsilon$$

Previous model could include Fixed Effects (FE) or Random Effects (RE). The first model considers that the error term are decompose in two terms: ν_i and π_{it}. The first parameters is a fixed term according with i values that represents the Spanish regions, meanwhile, the later is considered a error term. The RE model has the same equation, but the ν is random, not depend on Spanish regions. This term has a $E[\nu] = \bar{\nu}$ and $Var(\nu 0) \neq 0$.

$$Y_{it} = \mu 0 + \beta_j X_{jit} + \beta_k X_{kit} + \beta_h X_{hit} + \nu_i + \pi_{it}$$

The Hausman test are used as a selection criteria to choose between FE and RE. In this test thee null hypothesis propose a random effects as a preferred model, otherwise, the alternative hypothesis propose a model with fixed effects (Greene, 2012).

$$H = (\beta_{RE} - \beta_{FE})'[Var\beta_{RE} - Var\beta_{FE}](\beta_{RE} - \beta_{FE})$$

$$H \sim \chi_n 2$$

3.2 Data & variables

The expenditures of international tourists in Spain is recorded from a Spanish survey of touristic expenditure, which is actually designed and collected by the National Statistical Institute (INE). The questionnaire is focused on study the tourist behaviour to non-residents visitors in Spain. However, the current research considers the period 2008-2018, so the data until 2015 is obtained from Turespaña data base. This department depended of Spain's State Secretary of Tourism and it was in charge of this questionnaire during the time preceding of being transfered to INE. Note that daily data could be obtained until 2015, whereas form this period only monthly data are provided by INE. For this reason, the current paper are developed by monthly data. On the other size, the data was added by Spanish regional entities because is the most disaggregated data that both surveys provide. Therefore, 17 Spanish autonomous communities and 2 cities autonomous are involved in current research.

The average daily expenditure per international tourist (DDEIT) is the variable selected. This is obtained from the Egatur database as

$$\text{DEIT} = \frac{\text{total expenditure during visit}}{\text{number of overnight stays}}$$

	mean	sd	min	max	range
España	182.57	19.77	138.80	235.20	96.41
Andalucia	168.74	40.90	100.78	374.20	273.42
Aragón	174.04	49.93	97.43	374.17	276.74
Asturias	168.89	60.27	78.29	491.20	412.91
Baleares	175.58	65.31	68.36	474.90	406.54
Canarias	172.79	55.66	60.35	412.96	352.61
Cantabria	155.85	37.83	89.22	350.73	261.52
Cataluña	206.75	46.82	96.57	328.39	231.83
Castilla y León	167.29	38.41	100.49	288.51	188.02
Castilla la Mancha	160.13	41.50	92.47	346.62	254.15
Extremadura	144.45	34.54	65.86	237.44	171.58
Galicia	160.41	49.44	85.46	489.00	403.54
Madrid	241.07	52.58	110.64	354.97	244.33
Murcia	135.41	42.99	61.94	399.62	337.69
Navarra	185.30	64.25	63.78	445.56	381.77
Pais Vasco	209.18	54.99	120.89	491.59	370.70
Rioja	173.52	72.96	58.77	413.96	355.19
Valencia	153.75	30.20	103.34	268.75	165.41
Ceuta	203.26	161.62	57.46	530.14	472.68
Melilla	130.00	96.68	37.43	333.94	296.50

Table 1: Summary of Data by total and Spanish regions: mean, standard deviation, minimum, maximum and range of the daily expenditure per international tourist in Spain (in €) from 2008 to 2018.

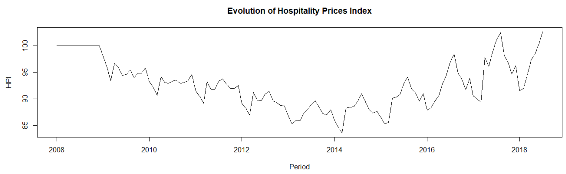

Figure 1: HPI evolution from 2008 to 2018 (Data source:INE)

where both quantities are measured per tourist. A summary data of DDEIT by Spanish regional entities is given in Table 1.

As well as, one euro not have the same value than tomorrow, then the expenditures of international tourists should consider this economic criterion. Therefore, the expenditures should be corrected the according to Consumer Prices Index (CPI).In this way, there is provided several index to indicate the evolution of Spanish prices. In this way, the Hospitality Prices Index (HPI) is considered to consider the evolutions of prices related with tourists expenditures. The base year to calculate this index is 2008 and it considers the accommodation cost per costumer depending on hosting class, Spanish region and month. The Figure 1 shows the evolution of HPI during the last ten years.

In previous graph could be observed the influences of national and international crise in hospitality prices. This index records his lower level during 2014 and the prices is not similar to 2008 until now. Then, this effect should be considered in the current research because it is important to compare the different prices level to a specific point of time.

So, the next equation is used to calculate the deflacted daily expenditure for international tourists (DDEIT) in Spain.

$$\text{DDEIT} = \frac{\text{daily expenditure for intenational tourists}}{\text{HPI}}$$

A histogram of previous variable, namely the deflected daily expenditure by international tourists in Spain, can be found in Figure 2. The size of the bars in this histogram are therefore proportional to the number of months in which the daily average expenditure lied within the corresponding range of values. The Figure 3a) and Figure 3b) shows the

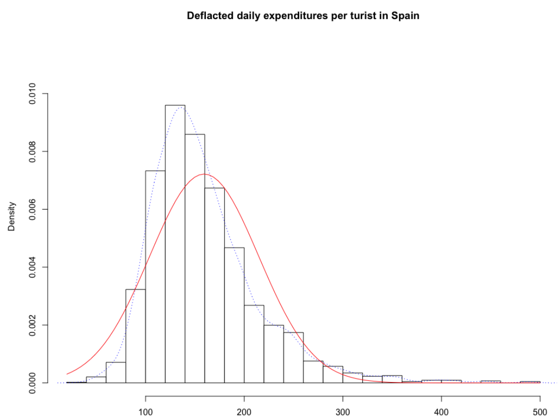

Figure 2: Histogram for the DDEIT in Spain by month. The blue line shows the density curve, the red line is the normal distribution approximation.(Data source:INE and Turespaña)

different patterns of DDEIT depending of temporal or spatial features. The Figure3a) remarks the volatility of expenditures in the Spanish autonomous cities. This happens because in this area is recorded few surveys. However, in the remaining regions present variability in the DEEIT data. The monthly heterogeneity is also recorded. Following the Figure 3b), there are months with more spread on recorded DDEIT and other with lower values. Then, there is evidences that the temporal and spatial heterogeneity influences the expenditures of foreign Spanish tourists.

The next list resume the selected variables to analyze the daily expenditures by international tourist:

- *Airport:* represents the percentage of EGATUR responses that the tourist arrive by plane.
- *EEUU:* is the percentage of tourists that response EGATUR and they are from United States of America.
- *United Kindom:* is the percentage of tourists that response EGATUR and they are from United Kingdom.
- *Catalonia:* represents the percentage of responses in which Catalonia is the main destination.

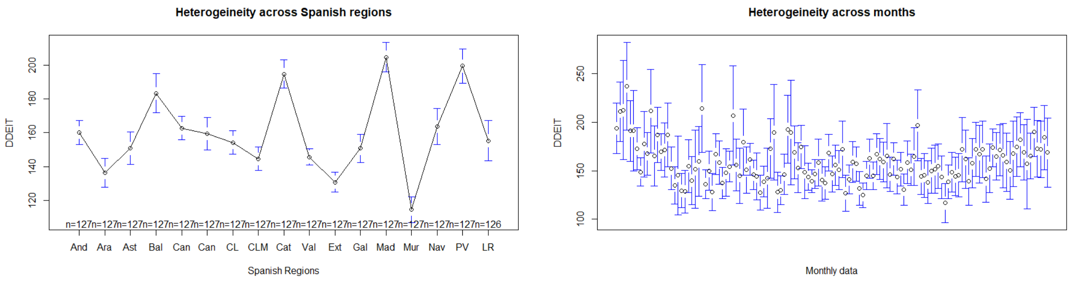

Figure 3: a) Heterogeneity across Spanish regions; b) Heterogeneity across monthly data

- *Madrid:* represents the percentage of responses in which Madrid is the main destination.
- *Balearic Islands:* represents the percentage of responses in which Balearic Islands is the main destination.
- *Summer:* is dichotomous variable that represents the summer months - July, August and September.
- *November:* is dichotomous variable that represents November.
- *Apartments:* represents the number of apartments available by month and Spanish region.
- *Campings:* represents the camping sites of each Spanish region by month.
- *Hotels:* represents the number of beds provided by hotels in each month and Spanish region.
- *Protected areas:* is the percentage that represents the protected areas over total surface in each Spanish region.
- *Population density:* is the number of inhabitants per square kilometre by region.

These variables are selected considering their correlation factor. This could be seen in the Figure 4. In this matrix could be observed the correlation value and the plot of both variables. Hotels and Apartments should be mentioned as well as the correlation index is high, however it not overcome the previous threshold and this variables are considered important to the analytical model. The same conclusion is applied to Catalonia and Campings, which presents a moderate correlation level and both are included in the model

The summary of all variables are in Table 2. This table contain the minimum, maximum, quartiles, mean and median values of all variables.

	Airport	USA	United Kingdom	Catalonia	Madrid	Balearic Islands	Summer	November	Apartments	Campings	Hotels	Protected areas	Population density	GDDeflactado
Min.	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	760	3.307	13.56	37.23
1st Qu.	0.07329	0.00000	0.06551	0.00000	0.00000	0.00000	0.0000	0.00000	565	1744	18987	23.579	75.56	122.27
Median	0.21831	0.02099	0.12907	0.00000	0.00000	0.00000	0.0000	0.00000	1182	3517	35492	27.440	97.60	148.72
Mean	0.32579	0.03784	0.15200	0.05818	0.05818	0.05818	0.2442	0.07879	7235	8772	82810	28.450	229.39	159.40
3rd Qu.	0.56131	0.05392	0.21718	0.00000	0.00000	0.00000	0.0000	0.00000	5668	8730	112512	30.514	223.56	183.56
Max.	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	66248	98855	360987	46.742	6624.61	490.52

Table 2: Summary variables

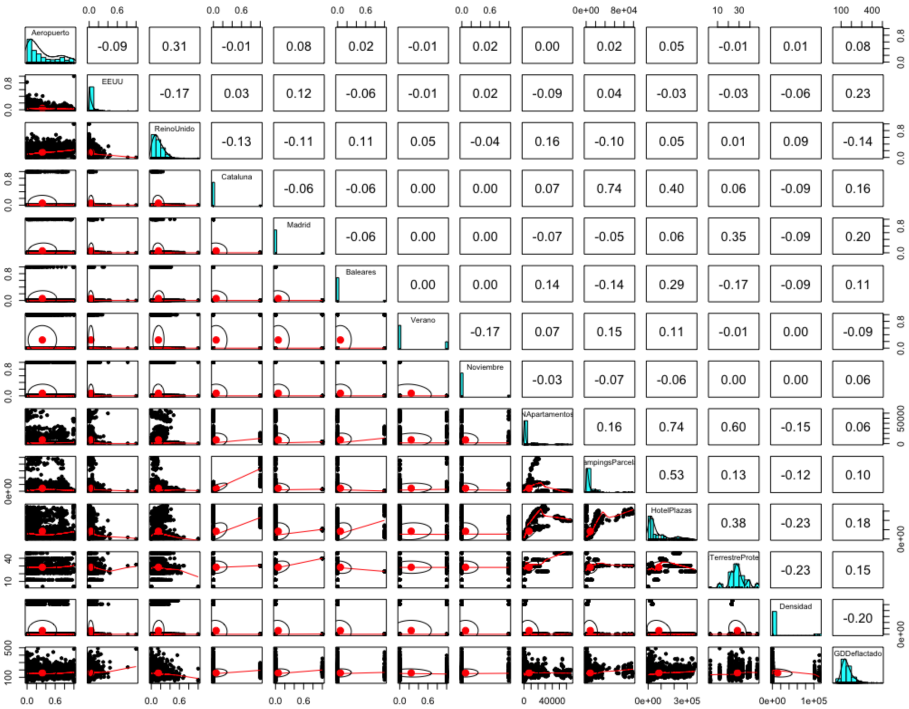

Figure 4: Correlation matrix

4 Results & analysis

The Table 3 presents the OLS results in which is obtained a $R - Adjusted^2$ of 0.1838 and the F-statistic shows that the regression model is better than a model with no independent variables $(p < 2.2e - 16)$. All of model coefficients are relevant $(p < 0,05)$, so all its estimations should be considered. In this way, the tourist expends per day more in Catalonia, Madrid or Balearic Islands in comparison with the mean of foreign tourists. On the other hand, the touristic sector has significants effect into the international expenditures, so the hotels has a positive value and apartments or campings has a negative effects. Finally, the summer tourist has a negative effect into the expenditures.

The boxcox test is developed and a global result is shown in the Figure 5. As well as, the result is next to zero, so a logarithmic OLS model is suggested.

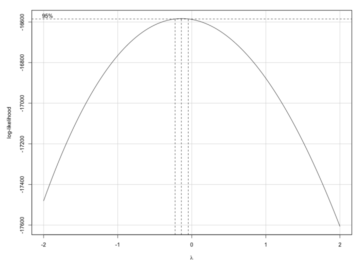

Figure 5: BoxCox

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.242e+02	6.418e+00	19.356	< 2e-16 ***
Airport	2.073e+01	3.772e+00	5.497	4.32e-08 ***
USA	2.012e+02	1.936e+01	10.395	< 2e-16 ***
United Kingdom	-5.108e+01	9.951e+00	-5.133	3.10e-07 ***
Catalonia	3.432e+01	7.019e+00	4.890	1.08e-06 ***
Madrid	2.261e+01	5.873e+00	3.849	0.000122 ***
Balearic Islands	2.758e+01	5.628e+00	4.901	1.02e-06 ***
Summer	-8.829e+00	2.597e+00	-3.400	0.000687 ***
November	6.968e+00	4.038e+00	1.725	0.084591 .
Apartments	-4.923e-04	1.855e-04	-2.653	0.008028 **
Campings	-3.255e-04	1.323e-04	-2.460	0.013957 *
Hotels	9.502e-05	2.790e-05	3.406	0.000671 ***
Protected areas	9.139e-01	2.187e-01	4.178	3.06e-05 ***
Population density	-1.745e-04	3.292e-05	-5.302	1.26e-07 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 3: OLS Results

As some variables contains negative or zero values, then a logarithm could not be applied to this variables. Therefore, logarithmic is applied to dependent variable but individual boxcox are calculated to selected the functional form of independent variables. In this way, the Table 4 summarize the obtained values. The results shows that Boxcox transformation should be made for all strictly positive independent variables $(X_{jit} > 0)$. The hotels, protected areas and population density should be analyzed by including a exponential of 4.32,4.4 and 13, respectively. Considering previous Box-Cox analysis, the results suggest that the functional form should be:

$$\begin{aligned} log(Y_{it}) = &\beta_0 + \beta_1 Airport_{it} + \beta_2 EEUU_{it} + \beta_3 UnitedKingdom_{it} + \beta_4 Catalonia_{it} + \\ &+ \beta_5 Madrid_{it} + \beta_6 Baleares_{it} + \beta_7 Summer_{it} + \beta_8 November_{it} + \\ &+ \beta_9 Apartments_{it} + \beta_{10} Campings_{it} + \beta_{11} Hotels_{it}^{4.4} + \\ &+ \beta_{12} ProtectedArea_{it}^{4.32} + \beta_{13} PopulationDensity_{it}^{13} + \epsilon it \end{aligned}$$

The lineal OLS model and Log OLS model is compared by using Akaike Information Criteria (AIC) to select the correct functional form. Therefore, the AIC before transformation is 23285.87 and the AIC after transformation is 782.94. As the model with lower AIC should be selected, the AIC confirms that Log OLS model is better than lineal OLS model.

	MLE of lambda	Score Statistic (z)	Pr(>|z|)
Hotels	4.3233	5.1397	2.752e-07 ***
Protected areas	4.3961	2.2800	0.02261 *
Population density	13.0011	-4.4845	7.309e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 4: Boxcox Results

With previous functional specification is calculated the a new OLS model. The main results are shown in the Table 5. In these case the $R - Adjusted^2$ raise to 0.2064 and the obtained F-statistic conclude the same than previous OLS model. The results show relevant coefficients, however, in this case, November is not significant. So, the socioeconomic features and environmental qualities are important to explain the expenditures of international tourists in Spain.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	4.762e+00	3.707e-02	128.457	< 2e-16 ***
Airports	1.082e-01	2.179e-02	4.967	7.33e-07 ***
USA	1.098e+00	1.118e-01	9.821	< 2e-16 ***
United Kingdom	-3.001e-01	5.747e-02	-5.222	1.94e-07 ***
Catalonia	1.824e-01	4.054e-02	4.499	7.19e-06 ***
Madrid	1.094e-01	3.392e-02	3.225	0.001277 **
Balearic Islands	1.457e-01	3.250e-02	4.482	7.78e-06 ***
Summer	-6.036e-02	1.500e-02	-4.025	5.90e-05 ***
November	3.534e-02	2.332e-02	1.515	0.129896
Apartments	-3.984e-06	1.072e-06	-3.718	0.000206 ***
Campings	-2.096e-06	7.641e-07	-2.743	0.006138 **
Hotels	8.129e-07	1.611e-07	5.045	4.91e-07 ***
Protected areas	7.234e-03	1.263e-03	5.727	1.17e-08 ***
Population density	-1.162e-06	1.901e-07	-6.112	1.16e-09 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 5: OLS Results following the new functional form

The VIF results are shown in the Table 6. This values are lower for the main independent variables, then most of then present lower correlation. However, the Apartments and Hotels variables present high values and multicollinearity could affect this model. As well as, this values are closely to VIF threshold, the variables are considered into the final model.

Airport	1.1512
EEUU	1.0563
United Kingdom	1.2348
Catalonia	2.3626
Madrid	1.6542
Balearic Islands	1.5188
Summer	1.0893
November	1.0359
Apartments	5.3172
Campings	3.5044
Hotels	5.6636
Protected areas	2.6562
Population density	1.1241

Table 6: VIF Results following the new functional form

The Table 7 shows the Hausman test result. It records a p-value upper than 0.05, therefore, the Random Effects is the best model to analyze the DDEIT instead of a model with Fixed Effects. The functional form used in this estimation is the same as previous models.

Hausman Test	
chisq	5.8592
df	13
p-value	0.9511

Table 7: Hausman Test

The Table 8 shows the results of RE model. The $R - Adjusted^2$ is equal to 0.45937, this value similar to previous research obtain (REFERENCES). The F-statistic is equal to

143.614 and its p-value is lower than 0.05 ($p_{F-statistic} < 2.22e-16$). The coefficients in the model are highly significant ($p < 0.001$) and its signs are the same as previous models. In this way, the Airport, USA, Catalonia, Madrid, Balearic Islands, November, Hotels and Protected areas coefficient is positive and the remaining negatives.

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	4.7587e+00	3.9358e-02	120.9069	< 2.2e-16 ***
Airport	1.0021e-01	2.8523e-02	3.5135	0.0004514 ***
USA	1.0703e+00	1.1106e-01	9.6374	< 2.2e-16 ***
United Kingdom	-2.7407e-01	5.7168e-02	-4.7941	1.745e-06 ***
Catalonia	1.8191e-01	4.0192e-02	4.5260	6.337e-06 ***
Madrid	1.1053e-01	3.3897e-02	3.2607	0.0011286 **
Balearic Islands	1.3982e-01	3.2209e-02	4.3411	1.483e-05 ***
Summer	-5.8402e-02	1.4792e-02	-3.9482	8.126e-05 ***
November	3.8847e-02	2.3035e-02	1.6864	0.0918574 .
Apartments	-4.2767e-06	1.0676e-06	-4.0059	6.388e-05 ***
Campings	-2.2337e-06	7.5484e-07	-2.9591	0.0031184 **
Hotels	8.5826e-07	1.6046e-07	5.3488	9.785e-08 ***
Protected areas	7.3324e-03	1.2527e-03	5.8533	5.552e-09 ***
Population density	-1.1721e-06	1.8855e-07	-6.2164	6.083e-10 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 8: Estimation of Random effects model

The Breusch-Pagan test presents a value of152.45 with 13 degree freedom, which is associated to a $p-value < 2.2e-16$. Then, there is heteroskedasticity in the model and a robust covariance matrix should be used. This implies that a robust estimation is calculated from previous random effects model. The results of Table 9 controls the heteroskedasticity to calculate consistent coefficients. The Arellano method is using to calculate the robust standard errors for estimators [Arellano, 1987]. The robustness results shows that the main variables are significant, however, November now is not significant. Therefore, now the coefficient of November could not be used to achieve some conclusions.

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	4.762e+00	3.707e-02	128.457	< 2e-16 ***
Airport	1.082e-01	2.179e-02	4.967	7.33e-07 ***
USA	1.098e+00	1.118e-01	9.821	< 2e-16 ***
United Kingdom	-3.001e-01	5.747e-02	-5.222	1.94e-07 ***
Catalonia	1.824e-01	4.054e-02	4.499	7.19e-06 ***
Madrid	1.094e-01	3.392e-02	3.225	0.001277 **
Balearic Islands	1.457e-01	3.250e-02	4.482	7.78e-06 ***
Summer	-6.036e-02	1.500e-02	-4.025	5.90e-05 ***
November	3.534e-02	2.332e-02	1.515	0.129896
Apartments	-3.984e-06	1.072e-06	-3.718	0.000206 ***
Campings	-2.096e-06	7.641e-07	-2.743	0.006138 **
Hotels	8.129e-07	1.611e-07	5.045	4.91e-07 ***
Protected areas	7.234e-03	1.263e-03	5.727	1.17e-08 ***
Population density	-1.162e-06	1.901e-07	-6.112	1.16e-09 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 9: Random effects model with robustness coefficients

5 Conclusions

The daily expenditure of foreign tourists in Spain depends on several factors. The previous results highlight the importance of socioeconomic characteristics to explain the expenditures of international tourists. Hence, the population density is a variable that influences the expenditures: where bigger population densities are recorded, the international tourist expenses are lower. Also, the environmental quality is an important factor. In this way, regions where a major percentage of lands are protected record larger expenditures. These factors suggest that one could design a public policy which promotes tourist destinations with lower population density as well as to promote more forest protected areas.

Another important conclusion is that tourists from the USA spend more than the average tourist, so it would make sense to organize promotional campaigns focusing specifically on US tourism. Meanwhile, tourists from the United Kingdom have a negative incidence of daily expenditures, i.e.: they spend less than the average tourist. Further research should be developed to analyze the factors that could influence UK tourists to increase their daily expenditures.

The tourist sector itself is also important to explain the expenditures. The number of hotels is directly related with the daily expenditures, whereas campings and apartments are negative correlated with the expenditure. This suggest that if the sector wants to increase the daily expenditures, then the number of hotels should be increased.

The temporal trends show that the expenditures in the month of November present a positive effects, whereas the summer daily expenditures have negative incidence. This could be explained by the fact that many Christmas presents are bought during and just before the Christmas holidays, whereas in Summer there is more competition to engage the tourists. Additionally, the main touristic destinations record positive effects on the daily expenditures, so these regions should maintain their touristic attractiveness.

Finally, the current research could be improved if more socioeconomic and environmental variables are considered. Also, more individual variables could be taken into account, since the daily expenditure profile could be influenced by personal decisions, such as a tourist's hobbies. Finally, the model could be developed by using a weight matrix to include the spatial effects in the new econometric model.

References

UNWTO Tourism Highlights, 2018 (Madrid).

World Travel and Tourism Council: Travel & Tourism Economic Impact, 2018.

Arellano, M. (1987). Computing robust standard errors for within-group estimators. *Oxford Bulletin of Economics and Statistics*, 49(4):431–434.

Chou, M. C. (2013). Does tourism development promote economic growth in transition countries? a panel data analysis. *Economic Modelling*, 33:226 – 232.

Dritsakis, N. (2012). Tourism development and economic growth in seven mediterranean countries: A panel data approach. *Tourism Economics*, 18(4):801–816.

Eilat, Y. and Einav, L. (2004). Determinants of international tourism: a three-dimensional panel data analysis. *Applied Economics*, 36(12):1315–1327.

Garin-Munoz, T. and Amaral, T. P. (2000). An econometric model for international tourism flows to spain. *Applied Economics Letters*, 7(8):525–529.

Garín-Mun¯oz, T. (2006). Inbound international tourism to canary islands: a dynamic panel data model. *Tourism Management*, 27(2):281 – 291.

Garín-Muñoz, T. and Montero-Martín, L. F. (2007). Tourism in the balearic islands: A dynamic model for international demand using panel data. *Tourism Management*, 28(5):1224 – 1235.

Garín-Muñoz, T. (2007). German demand for tourism in spain. *Tourism Management*, 28(1):12 – 22.

Goh, C. and Law, R. (2011). The methodological progress of tourism demand forecasting: A review of related literature. *Journal of Travel & Tourism Marketing*, 28(3):296–317.

Greene, W. H. (2012). *Econometric Analysis, 7th ed.* Prentice Hall/Pearson, Upper Saddle River, New Jersey.

Ibrahim, M. (2011). The determinants of international tourism demand for egypt: Panel data evidence. *European Journal of Economics, Finance and Administrative Sciences*, 30:50 – 58.

Ledesma-Rodríguez, F. J., Navarro-Ibáñez, M., and Pérez-Rodríguez, J. V. (2001). Panel data and tourism: A case study of tenerife. *Tourism Economics*, 7(1):75–88.

Narayan, P. K., Narayan, S., Prasad, A., and Prasad, B. C. (2010). Tourism and economic growth: A panel data analysis for pacific island countries. *Tourism Economics*, 16(1):169–183.

Rey, B., Myro, R. L., and Galera, A. (2011). Effect of low-cost airlines on tourism in spain. a dynamic panel data model. *Journal of Air Transport Management*, 17(3):163 – 167. Developments in Air Transport and Tourism.

Sequeira, T. N. and Nunes, P. M. (2008). Does tourism influence economic growth? a dynamic panel data approach. *Applied Economics*, 40(18):2431–2441.

Song, H. and Li, G. (2008). Tourism demand modelling and forecasting—a review of recent research. *Tourism Management*, 29(2):203 – 220.

Surugiu, C., Leitão, N. C., and Surugiu, M. R. (2011). A panel data modelling of international tourism demand: Evidences for romania. *Economic Research-Ekonomska Istraživanja*, 24(1):134–145.