Congestion affecting the dynamic of tourism demand. Evidence from the most popular destinations in Spain

Isabel P. Albaladejo[1]
Departamento de Métodos Cuantitativos para la Economía. Universidad de Murcia. Murcia. Spain.

Maribel González-Martínez
Departamento de Métodos Cuantitativos para la Economía. Universidad de Murcia. Murcia. Spain.

Corresponding author: Isabel P. Albaladejo. Campus de Espinardo, 30100 Murcia, Spain. isalba@um.es.

Abstract. In this study, we propose a dynamic econometric model for tourism demand which takes into account the implications of the Tourism Area Life Cycle (TALC) theory on tourism demand. Unlike other dynamic models, in our specification the effect of the lagged demand on the current tourism demand is not constant, but dependent on congestion. We estimate the model using disaggregated data from the most visited Spanish municipalities for the period 2006-2015. Two panel data estimations are carried out: one with the coastal tourist resorts and the other one with the inland municipalities. The results show that tourism congestion reduces the positive previous tourist effect on current arrivals, suggesting that increasing congestion could worsen the attraction of a tourist destination. Congestion is more negatively perceived in inland destinations than coastal ones. Finally, a strong persistence in tourism demand for coastal destinations is shown.

Keywords: Tourism Area Life Cycle Model (TALC), tourism demand, tourist congestion, dynamic panel data model.

JEL Classification: C23, L83, Z32.

1 Introduction

Spain has a large number of highly appealing sun and sea tourism resorts like Mallorca, Benidorm, Marbella, Las Palmas or Tenerife and also important cultural cities like Madrid, Barcelona, Seville or Granada. These destinations receive large volumes of tourist flows each year. For example, in Barcelona the number of tourists accommodated in hotels grew by 5.4% in 2015, exceeding 7 million tourists. These flows of tourists foster the growth of the Spanish economy. According to the National Statistics Institute (INE), in 2015 the tourism

[1]Isabel Albaladejo has been partially supported by MICINN under the project ECO2016-76178P.

sector in Spain contributed 11.1% to the gross domestic product (GDP) and generated 2.49 million jobs (13% of the total). But the tourist flows are also likely to threaten the welfare of the tourist destinations. Congestion or overcrowding of a destination increases when the number of visitors is excessive in relation to the space or capacity of the destination to accommodate those tourists, especially during peak periods. Nowadays, many destinations such as Barcelona, Mallorca or Benidorm, among others, are experiencing congestion. This affects the quality of life of local residents but also the tourists visiting the destination, which could damage their opinion.

Tourists' perceptions are important for the future development of a destination. It is well-known and accepted in demand analysis that previous tourists can affect the current tourism demand either because they repeat destinations (habit persistence), or because they influence other potential visitors (word-of-mouth recommendations) (Morley 2009, Garín-Muñoz 2006, 2007 and 2009, among others). This positive influence of previous tourists may be larger or smaller depending on the attractiveness of the tourism destination. Therefore a destination's feature, like tourism congestion, could affect the impact of previous visitors on current tourist flows.

The Tourism Area Life Cycle (TALC) theory (Butler, 1980), the most popular theory on tourism evolution, suggests that the rhythm of tourists arrivals at a destination is related to its level of congestion, assuming that the congestion depends on the number of tourists and carrying capacity (the social, physical or economics capacity) of the destination. The result is an S shape or logistic curve representing the evolution of a tourist area from its discovery to its final stage. The limits of growth and the shape of the curve represent the existence of congestion problems and upper carrying capacity limits. During the first stages, the number of visitors grows at an increasing rate. However, as it approaches its carrying capacity, the process slows down, suggesting that the subsequent congestion has a negative effect on arrivals (Albaladejo, González-Martínez and Martínez-García., 2016).

This paper aims to make an empirical contribution to the research into tourism demand of the most visited destinations where congestion can be a problem. We focus on the effect of the previous tourists on the tourism demand and we investigate how this effect can be influenced by the congestion at the destination. To model tourism demand, we consider a dynamic econometric model where current and lagged demand maintain a quadratic relationship. According to the TALC theory, this model considers a non-constant growth rate for tourism demand. Furthermore, it allows the effect of previous tourists not to be constant but to vary with the ratio between tourists and carrying capacity. Since this ratio is related to congestion of the destination, our econometric model allows the effect of previous tourists to depend on congestion. This non-constant effect is more suitable than a constant one, which is assumed in most empirical studies on tourism demand (Garín-Muñoz, 2006, 2007 and 2009; Garín-Muñoz and Montero-Martín, 2007; Massidda and Etzo, 2012; Capacci, Scorcu and Vici, 2015, among others).

Using this model, we analyze the determinants of tourism demand at the most popular tourist destinations in Spain. Our empirical analysis uses disaggregated data from the 93 Spanish municipalities designated by the INE as tourist destinations during the period 2006-2015. We perform two estimates: one with panel data from the 53 coastal municipalities and the other one with panel data from the remaining 40 municipalities. The advantage of considering a disaggregated data set is that it allows us to study the effect of tourism congestion at the most-visited tourist destinations in Spain. Our sub-sample analysis also shows the differences between the behavior of tourists arriving at sun and beach resorts, many of which are considered as mass tourism destinations, with those who visit the inland municipalities. A system GMM dynamic panel data analysis (Blundell and Bond, 1998) is run to estimate both models. The estimates show that tourism congestion negatively affects the impact of previous tourist over current demand. This influence would imply an evolution of the destination according with the TALC. Our results indicate that congestion is more negatively perceived at the inland destinations than at coastal ones. In addition, a strong persistency in tourism demand of coastal destinations is found.

The paper is organized as follows. The following section provides the theoretical foundations of our model. Section 3 presents the data and variables considered in the study. Section 4 provides the empirical model and describes the econometric method used for the estimation. Section 5 offers the results and their interpretation. Finally, Section 6 draws some conclusions.

2 The theoretical framework

The TALC theory (Butler, 1980), according to the evolutionary theory of tourism areas, argues for the existence of an S-shaped lifecycle in the growth of tourism destinations with six key stages: exploration, involvement, development, consolidation and stagnation, arriving at a final post-stagnation stage where decline, rejuvenation or other intermediate solutions are possible (see Figure 1). Each stage is characterized by a different growth rhythm and the upper limit of the curve is determined by the carrying capacity of the destination.

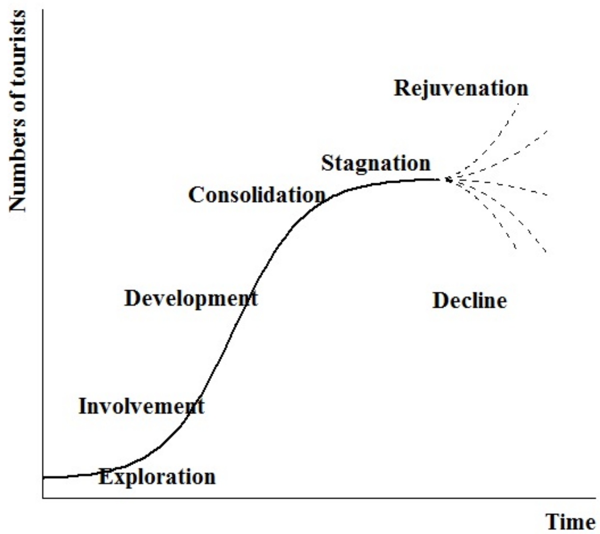

Figure 1: Evolution of a tourist area according to the TALC.
Source: Butler (1980).

Since 1980 many authors have applied, applauded, criticised and modified the TALC as a tool for understanding the evolution of tourism areas. Butler (2006a and 2006b) provide a collection of works giving an overview of the TALC, of its applications, weakness and some new contents and methods which have been proposed to enlarge or supplement the theory. Some mathematical models have also been presented to hypothesize and test the TALC model (Cole 2009, 2012; Lundtorp and Wanhill, 2001; Moore and Whitehall, 2005). Finally, the TALC model has been combined with several theories (teleological theory, regulation theory and chaos theory) to create valid frameworks to understand the evolution, growth and development of the destination (Oreja-Rodriguez, Parra-Lopez and Yanes-Estevez, 2008; Garay and Canoves, 2011; Russell and Faulker, 2004).

In this paper, we are interested in the implications of the TALC theory for the tourism demand. Examining the lifecycle model from tourists perspective, its S-shape evolution implies the number of tourists can indicate the stage of the tourist´s destination development (Lundtorp and Wanhill, 2001). In this sense, TALC theory relates the degree of congestion of a destination with the rate of growth in tourism demand. Popular or mature destinations suffering congestion or overcrowding problems should be in the final stages of the Butler lifecycle model, where the volume of tourists is still growing but at a declining rate until reaching the highest number of tourists, in the stagnation phase. The stagnation period implies the existence of a growth ceiling, which is interpreted as being the carrying capacity of the destination (Haywood, 1991).

The decreasing growth rate predicted by the TALC theory for mature destinations is not in accordance with standard dynamic econometric models for tourism demand. These models usually include the lagged tourism demand in a linear fashion, assuming an exponential growth for tourism demand (Albaladejo *et al.,* 2016). Consequently these models predict a constant growth rate for the number of tourists.

Albaladejo et al. (2016) proposed a dynamic econometric model for tourism demand where the growth rate of tourism demand is not constant. This model included the lagged demand using a quadratic form. Assuming that this quadratic function is influenced by a constant carrying capacity, the model can predict a positive tourism growth rate, but decreasing for congestioned destinations, in agreement with the TALC model. Then, congestion plays an important role in this tourism demand model.

Taking into account the above, the tourism demand of a destination could be defined by

$$T_t = \beta_1 T_{t-1} + \beta_2 \frac{T_{t-1}^2}{CC} + \gamma^{'} \cdot X_t + \varepsilon_t \tag{1}$$

where subscripts t denote the time period, the variable T_t is the current number of tourists, T_{t-1} is the previous number of tourist, CC is the carrying capacity of the destination and $X_t' = (x_t^1, x_t^2, ...x_t^k)$ is the vector of the other explanatory variables (price, income, exchange rate, etc.) that define the demand according to classical economic theory. The regression error term is ε_t.

Note that in Equation (1), we cannot interpret β_1 as measuring the effect of previous tourism on current tourism; we need to take into account β_2, T_{t-1} and CC as well. Therefore, our model allows a non constant effect of previous tourists. The marginal effect of T_{t-1} on T_t is measured by

$$\frac{\partial T_t}{\partial T_{t-1}} = \beta_1 + 2\beta_2 \frac{T_{t-1}}{CC} \tag{2}$$

which is interpreted as persistency and word-of-mouth effects on tourism demand (Albaladejo et al.*,* 2016; Massidda and Etzo, 2012; Capacci et al., 2015).

Equation (2) shows that the slope of the relationship between current and previous tourists is determined by the coefficient β_2. This coefficient captures the impact of congestion on the marginal effect, because congestion is related with the ratio between tourists and carrying capacity of the destination ($\frac{T_{t-1}}{CC}$). Three scenarios can be considered related to the value of β_2. These allow different types of preferences with respect to congestion by part of the tourists (Marsiglio, 2015).

In the first scenario, congestion does not matter. In this case, β_2 is equal to zero and the marginal effect of previous tourism is constant and equal to β_1. Hence, the resulting tourism demand model corresponds to the most common dynamic model, which includes previous demand in a linear fashion (Garín-Muñoz, 2006, 2007 and 2009; Garín-Muñoz and Montero-Martín, 2007; Massidda and Etzo, 2012; Capacci et al., 2015, among others). Since empirical evidence shows a positive β_1, this specification assumes a benefical effect of the previous visitors on current demand that does not vary with the number of tourists at the destination nor with the tourism congestion. Consequently, in this scenario, the development of a tourism destination is not limited by its carrying capacity, as predicted by TALC.

In the second scenario, $\beta_2 > 0$. In this case, congestion is a tourist attractor and therefore, the effect of previous tourists is positively influenced by the ratio $\frac{T_{t-1}}{CC}$. The quadratic function has a parabolic shape opening upwards, implying an increasing unlimited growth rate. This scenario disagrees with the S-shaped curve. Perhaps, a destination with this type of growth is not usual, but it could define the evolution of tourism demand for certain type of events, like musical concerts (Arenal Sound in Castellón or Benidorm Sound).

If $\beta_2 < 0$ the congestion is negatively perceived by tourists, as is usual, then we are in the third scenario. In this case, the effect of previous tourists is negatively influenced by the ratio $\frac{T_{t-1}}{CC}$. If β_1 is positive, as expected, the negative β_2 implies that the positive effect from previous tourists decreases as tourists increase because of the congestion. In this case, the quadratic function has a parabolic shape opening downwards, implying a diminishing marginal effect of T_{t-1} on T_t. Figure 2 plots a possible quadratic relationship between T_t and T_{t-1} and Figure 3 plots the marginal effect of T_{t-1} on T_t.

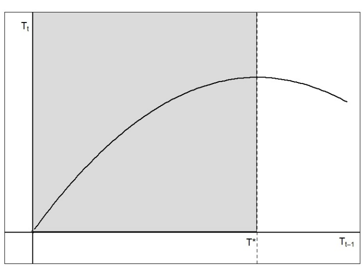

Figure 2: Quadratic relationship between T_t and T_{t-1} if $\beta_2 < 0$

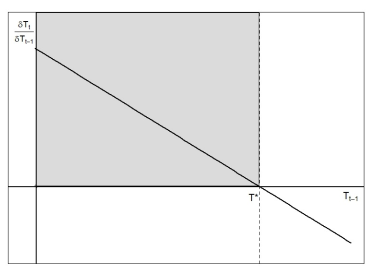

Figure 3: Marginal effect of T_t and T_{t-1} if $\beta_2 < 0$

The maximum of the parabola, the number of previous tourists for which the marginal effect is equal to zero, $T^* = \frac{-\beta_1}{2\beta_2}CC$, is positive since $\beta_1 > 0$ and $\beta_2 < 0$. Figure 3 shows a negative marginal effect when the tourism flows are higher than T^*. This situation is not in agreement with the empirical evidence, which shows a positive effect of previous tourists. Therefore, in practice, we can consider the shaded area in Figures 2 and 3 as the region of interest, and T^* as an upper limit of tourism. T^* depends positively on the carrying capacity of the destination and the ratio $\frac{\beta_1}{\beta_2}$. The higher the coefficient β_2, the more negatively the congestion is perceived, and the more rapidly the marginal effect decreases as the number of tourists increases.

The quadratic function of our model (Equation (1)) with $\beta_2 < 0$, implies a logistic growth for the tourism demand modified by the evolution of explanatories variables. We can easily rewrite Equation (1) as

$$T_t = aT_{t-1}(1 - \frac{b}{a}T_{t-1}) + c_t \tag{3}$$

where $a = \beta_1,\ b = -\frac{\beta_2}{CC}$ and $c_t = \gamma^{'} \cdot X_t + \varepsilon_t$.

From this expression, the following difference equation is obtained

$$\Delta T_t = (a - 1)T_{t-1}(1 - \frac{b}{a-1}T_{t-1}) + c_t. \tag{4}$$

which, when formulated in continuous time agrees with the logistc differential equation.[2] Therefore, the growth rate of the number of tourists can be positive but decreasing for congested tourism areas, as predicted by the TALC theory.

3 Data and variables

We consider tourism demand in the most visited Spanish destinations. Our sample includes 93 Spanish municipalities chosen as tourist destinations by the INE for the period 2006-2015. A tourist destination is a municipality where the concentration of tourism supply is significant. All these municipalities have some important tourism attraction (beaches, artistic and historical monuments, museums, palaces, parks, etc) or are near to an attraction. They are important destinations for domestic and international tourism in Spain. In 2015, about 60 million tourists stayed at hotels in these municipalities, accounting for almost two thirds (63.6 %) of the total number of tourists arriving in Spain and staying in hotels (INE).

In Spain the weight of inland and cultural tourism has increased in recent years, but the main reason tourists come to Spain is still the so-called sun and sand tourism, and beaches are the preferred destinations. Accordingly, more than half (53) of the 93 Spanish tourist destinations considered in our study are coastal municipalities and many of them are on the Mediterranean, which is the main destination for international tourism in Spain. The remaining 40 municipalities are spread around inland Spain and many of them are urban areas.

Figure 4 shows the evolution of the number of tourists who chose hotels and similar establishments as accommodation from 2006 to 2015 in the 53 coastal municipalities and in the 40 inland ones. The number of tourists is greater in coastal localities, but its evolution is similar in both types of destination. Tourists increased from 2006 to 2015, but the growth was not continuous

[2]Lundtorp and Wanhill (2001) showed that a logistic curve can be quite a good theoretical representation of Butler´s lifecycle path.

throughout the period. A decline is observed in 2009 and 2012, as a consequence of the global financial crisis and the economic recession in Spain. The consequences of the crisis seem to have been more negative in the coastal destinations. Since 2013, the number of tourists seems to be experiencing a new growth phase.

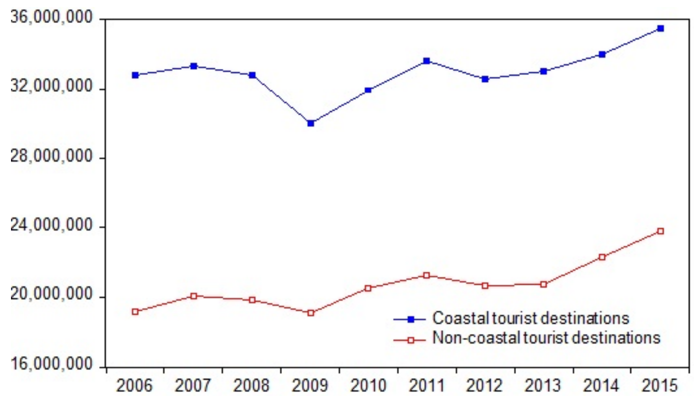

Figure 4: Tourists lodged at hotels

Many Spanish coastal municipalities are mass tourism destinations. Due to the limited extensions of coastal environments, they are often amongst the first places to experience spatial tourism congestion. In 2015, the 53 coastal municipalities identified as tourist destinations represented 60% of the number of tourists staying at hotels in the 93 Spanish tourist destinations. However, they represent only 30% of the total geographical space of the tourist destinations of Spain (INE). The Spanish inland municipalities identified as tourist destinations have lower tourism congestion. However the space requirements of tourists arriving at these municipalities, where multiple functions are performed, can interfere with those of local residents and produce overcrowding and congestion (La Rocca, 2005).

To analyze the main determinants of tourism demand at the most popular tourist locations in Spain, the model proposed in Section 2 (Equation (1)) is estimated using annual data disaggregated by municipality of destination.[3] We estimate the tourism demand at sun and beach resorts and inland destinations separately. Accordingly, two balanced panel data sets are used. One consists of the 53 coastal municipalities for the period 2006-2015, and the other of the 40 non-coastal municipalities for the same period. Panel data have some advantages over cross sectional or time series data. One is that it enables us to control for unobservable cross-sectional heterogeneity. Time series and cross-section studies not controlling for this heterogeneity run the risk of obtaining

[3]Seasonality is an interesting topic related to congestion. However, since we use annual data, it is not possible to take into account the seasonality effect. This is a limitation of our paper that we consider as a possibility for a future research.

biased results. Moreover, panel data usually give a large number of data points, so increasing the degrees of freedom, reducing the collinearity among explanatory variables and improving the efficiency of econometric estimates (Hsiao, 2003 and Baltagi, 2008).

Our model includes economic demand variables, such as income and prices, variables accounting for destination attributes (hotel beds and area), and a quadratic form to capture the effect of the past tourists. Additionally, we include two dummy variables to control for the effects of the economic crisis.

The dependent variable is the number of annual tourists who choose hotels and similar establishments as accommodation (T).[4] This variable is employed as an indicator of tourism demand at each destination. Data are taken from the INE.

Among the explanatory variables, we consider two traditional economic factors: origin income and price. To measure origin income (GDP), we use the real per capita GDP of EU-28, since Europe is by far the main origin of most of the international tourism flows to Spain. This variable was taken from the OCDE. The price variable (IP) included in our model reflects the cost of living of tourists at the different destinations relative to the cost of living in the country of origin:

$$IP_{destination} = \frac{CPI_{destination}}{CPI\_EU28 \cdot EX\_EU28}$$

where $CPI_{destination}$ is the consumer price index (CPI) for each of the destinations considered. For each municipality we consider the CPI corresponding to the province where it is located. CPI_EU28 is the CPI for EU-28, and EX_EU28 is the nominal effective exchange rate of Spain vs EU-28. Data on exchange rates and CPI for EU-28 were collected from Eurostat. Data on CPI for the provinces in Spain were collected from the INE.

Two variables reporting on the characteristics of the destination are also included among the explanatory variables. The first is the total area in km^2 of the municipality ($AREA$). This variable accounts for the different geographic dimensions. The second is the ratio between the number of hotel beds for tourist accommodation at the destination and the square kilometers of the destination municipality ($BEDKM$):[5]

$$BEDKM_{destination} = \frac{hotel\ beds_{destination}}{km^2_{destination}}$$

This variable measures the intensity of tourism supply at each destination. It allows us to analyze the role of the degree of tourism vocation of the destination,

[4] We are aware that using only the data of tourists staying at hotels is a limitation. However, data for demand at other types of accommodations are not available for a disaggregated study like ours. In addition, the total number of tourists who choose hotels and similar establishments as accommodation in Spain represents 82% of total arrivals in 2015, according to the INE.

[5] We are aware that using only hotel beds data is a limitation. Unfortunately, data from other types of lodgement supply, as second homes or rented apartments are not available for a disaggregated study like ours.

derived from the intentions of residents of a municipality to develop or promote tourism. Data on area and on hotel beds are taken from the INE.

Additionally, a measure of the carrying capacity (CC) of the destinations has to be defined to build the quadratic form of our model. In the literature, there are many definitions of carrying capacity in tourism. In its most traditional sense, it is understood as the maximum number of tourists or the tourist use that can be accommodated within a specific geographic destination (O'Reilly, 1986). This capacity has been identified in terms of limits of environmental, social, economical or physical factors (Butler, 1980; Saveriades, 2000; Cole, 2009; Diedrich and García-Buades, 2009).[6] In this paper, the carrying capacity has been identified as a physical limit. The square kilometers of each destination are used as a proxy of its carrying capacity. The bigger a destination, the more possibilities of its offering an extensive and diversified tourist supply, thus the higher the chance of accommodating visitors suitably. The advantage of using this measure is that its homogeneous character allows comparison among several municipalities. In addition, the relationship between the tourist flows of a destination and its geographic area is related to the congestion or overcrowding suffered by the destination. Thus, using square kilometers as a measure of carrying capacity allow us to test the influence of tourism congestion on tourism demand.

Finally, based on Figure 4, our model also includes two dummy variables, $Y2009$ and $Y2012$, to capture the influence on tourism of the financial and economic crisis in Spain. $Y2009$, takes the value 1 in 2009 and 0 in other years, and $Y2012$ takes the value 1 in 2012 and 0 in other years. 2009 and 2012 were the years of the Spanish recession when GDP growth rate decreased most.

4 Methodology and model specification

Following the model proposed in Section 2 and considering the variables defined above, the econometric model to be estimated with panel data is represented as

$$\begin{aligned} T_{it} &= \eta_i + \beta_1 T_{it-1} + \beta_2 \frac{T_{it-1}^2}{CC_i} + \beta_3 GDP_t + \beta_4 IP_{it} + \\ &\quad + \beta_5 AREA_i + \beta_6 BEDKM_{it} + \beta_7 Y2009_t + \beta_8 Y2012_t + \varepsilon_{it} \end{aligned} \tag{5}$$

where the subscript i denotes the destination municipality, t indicates the time period ($t = 2006 - 2015$), η_i is the unobserved destination-specific variable (or fixed effects) that varies across destinations but is invariable within a destination over time, and ε_{it} is a disturbance term. A key assumption throughout this paper is that the disturbance ε_{it} is uncorrelated across destinations, but heteroskedasticity across time and destinations is allowed for. The number of

[6]Note that the carrying capacity refers to the maximum number of tourists the hosting destination can accommodate. However, for long-term sustainable development the optimal number of tourists is lower than this maximum according to Marsiglio (2017).

tourists, the real per capita GDP, the relative price and the beds-area ratio are in logs, and therefore their coefficients may be interpreted as elasticities.

As discussed in Section 2, Equation (5) allows the effect of the previous tourist on current tourism demand to be nonconstant. Since panel data are used, the effect of previous tourists varies not only over time but also across destinations

$$\frac{\partial T_{it}}{\partial T_{it-1}} = \beta_1 + 2\beta_2 \frac{T_{it-1}}{CC_i} \tag{6}$$

This effect represents the elasticity of current tourism demand with respect to previous demand. Both the previous number of tourists and the carrying capacity can modify this elasticity. A positive sign is expected for β_1, so a negative β_2 would imply that this elasticity decreases with the ratio between previous tourists and the geographic dimension of the destination. That is, if congestion is negatively perceived, as tourism congestion increases, the elasticity of previous tourists decreases. If β_2 is zero, congestion does not matter and the elasticity is constant. With regard to the remaining coefficients, we expect a positive sign for β_3, β_5 and β_6 and a negative sign for β_4, β_7 and β_8.

A generalized method of moments (GMM) panel data estimation (Arellano and Bover, 1995; Blundell and Bond, 1998) is used to conduct our empirical analysis. Ordinary Least Squares (OLS) is not appropriate to estimate dynamic panel models with the lagged dependent variable among the regressors. The lagged dependent variable is correlated with the unobserved effect (η_i) which gives rise to "dynamic panel bias" (Nickell, 1981). The within groups and random effects estimators do not eliminate the "dynamic panel bias" and are also biased and inconsistent. To solve this problem, Arellano and Bond (1991) suggest first-differencing the model to remove the unobserved fixed effects (η_i). As the differenced lagged dependent variable is still potentially endogenous, it is instrumented with lagged levels of the endogenous variable to solve the problem of autocorrelation. If the ε_{it} are not serially correlated, we can use lags 2 and upwards of the endogenous variable as instruments. Blundell and Bond (1998) extended this estimator by building a system of equations formed by the equation in first differences and the equation in levels. The extended GMM estimator, called system GMM, uses lagged first-differences as instruments for the equation in levels in addition to the usual lagged levels as instruments for the equation in first-differences.

In this paper, we apply the system GMM (Blundell and Bond, 1998) procedure to estimate the model (Equation (5)). A crucial assumption for the validity of GMM is that the instruments are exogenous. We conduct two diagnostic tests: Hansen (1982) J tests of the over identifying restrictions for the GMM estimators[7], and the Arellano and Bond (1991) test for autocorrelation in the disturbance term, ε_{it}.

[7]The Hansen statistics is a chi-squared test to determine if the residuals are correlated with the instrument variables. If nonsphericity is suspected in the errors, the Hansen overidentification test is theoretically superior to the Sargan (1958) test.

5 Results

Given the scope of the present study, the model in Equation (5) is estimated twice, once for coastal municipalities and once for non-coastal ones. In both cases, we calculate two-step system GMM estimators. The lag of the dependent variable, the quadratic term and the hotel beds-area ratio are treated as endogenous in all estimates. Since the usual formulas for coefficient standard errors in two-step GMM tend to be downward biased when the instrument count is high, we use the Windmeijer (2005) standard errors correction.

The empirical results from the estimation of the model for coastal and non-coastal destinations are shown in Table 1.[8] Two main findings are worth noting in Table 1. First, the analysis confirms that tourism congestion matters. In both estimates, the estimated coefficient of the lagged dependent variable is significant and positive, and the estimated coefficient of the quadratic term is significant and negative, revealing a non-constant effect of the previous tourists on current tourists. This effect is negatively affected by the congestion. Second, the disaggregated analysis highlights interesting differences in the behavior of tourists arriving at coastal and non-coastal municipalities. Additionally, the results reveal a generally satisfactory performance of the econometric models. The autocorrelation tests (Arellano and Bond, 1991) do not detect any serial correlation problem in the residuals. As expected, the residuals in differences are autocorrelated of order 1, while there is no autocorrelation of second order. In addition the Hansen (1982) J-test does not reject the null hypothesis for joint validity of the instruments.

Focusing on the effect of the previous tourists, the significant estimated β_2 indicates that the effect of previous tourists is not constant and proves the need for a quadratic specification in the model. Since estimated β_2 is negative the quadratic funtion is a parabola opening downwards, providing evidence of a growth of the tourism destinations in line with the TALC theory. As in absolute value the estimated value of β_1 is well above the estimated value for β_2, the estimated elasticity of tourism demand with respect to the previous demand is positive and decreases slowly with the ratio between previous tourists and geographic area of the destination. Since a higher ratio implies more tourism congestion, this result means that, in the most visited Spanish tourists destinations, tourism congestion reduces the beneficial influence of previous tourists on the current demand. However, the effect of tourism congestion is different at sun and beach destinations and inland destinations. Estimated β_2 is rather more negative in the tourism demand for inland municipalities (-0.2224) than in the coastal municipalities (-0.0371). Consequently, our results show that tourists looking for sun and beach are less concerned about congestion. In inland municipalities there are a higher adversion to overcrowding. Thus, sustainaible tourism is more likely to occur in these municipalities as Cerina (2007) and Marsiglio (2015) argue.

[8] We have also calculated the one-step robust to heteroskedasticity system GMM estimator for comparison. Results of the one-step version are very similar to those shown in Table 1, and are available to interested readers.

Table 1: *Estimation results for tourism demand to Spanish municipalities, 2006-2015*

Dependent variable: T_{it}	System GMM	
Explanatory variables	coastal municipalities	non-coastal municipalities
T_{it-1}	0.7144***	0.3804***
$\frac{T_{it-1}^2}{CC_i}$	-0.0371***	-0.2224***
GDP_t	0.1467***	0.4111***
IP_{it}	-1.0949***	-1.8351***
$AREA_i$	0.0015***	0.0011***
$BEDS\_AREA_{it}$	0.3056***	0.6670***
$Y2009_t$	-0.0881***	-0.0174
$Y2012_t$	-0.0545***	-0.0678***
Hansen test (p-value)	0.269	0.922
AR(1) (p-value)	0.000	0.023
AR(2) (p-value)	0.339	0.299
Number of observations	477	360
Number of groups	53	40

Note: *,**,*** denote significant at the 10%, 5%and 1% level respectively. All estimations are made using the xtabond2 command in STATA10 (Roodman, 2009).

Additionally, our empirical analysis shows that one of the most important determinants of the sun and beach tourism demand seems to be the lagged dependent variable. Estimated β_1 is far greater in the equation of coastal municipalities (0.7144) than in the equation for non-coastal municipalities (0.3804). This result, together with a lesser impact of congestion at the coastal destinations, provides evidence of stronger persistence in sun and beach tourism demand than in inland tourism demand. A large portion of sun and beach tourists are not willing to give up their annual visit to the Spanish beaches, and decide to repeat destinations. This persistence of sun and sand tourists is in line with previous studies on Spanish resorts. Aguiló, Alegre and Sard (2005) and Ivars, Rodríguez and Vera (2013), analysing the Balearic Islands and Benidorm, two of the leading Spanish sun and sand tourism markets, also find high repeat visit rates at both destinations. According to Aguiló et al. (2005) in the year 2000, only 33.5% of all tourists to the Balearics were first-time visitors, showing a considerable degree of loyalty to the Balearics.

In connection with the lifecycle theory, our estimates of β_1 and β_2 suggest an upper limit of the tourism growth that is greater at the sun and beach destinations than the inland ones. The high degree of repetition is one of the factors behind the success of the Spanish sun and beach resorts. But there are other positive factors related to the efforts by the tourist industry and local government to restructure the markets. In this sense, Aguiló et al. (2005) highlights the role of the quality of the hotel infrastructure. They argue that the Balearic Islands are still competitive due to a tourism supply that has undergone a restructuring process that uses quality-based criteria. As a result, the evidence suggests that Spanish coastal destinations have been able to address issues such as the negative image of a mass tourism destination, and to postpone any stagnation phase.

Regarding economic variables, sun and sand tourists seem comparatively less price and income sensitive than inland tourists. The estimated income elasticity is positive and significant in both equations, showing that the arrival of tourists depends positively on the economic situation of the European Union, which is the main market of origin. According to the estimated elasticities, the effect of per capita GDP of EU-28 is larger at non-coastal destinations. A rise of 1% in per capita GDP raises the tourism demand in non-coastal municipalities almost three times more (0.4111%) than in the coastal municipalities (0.1467%). Tourist arrivals are also responsive to price changes. As expected, the estimated elasticity for relative prices is negative and significant with values of -1.0949 for coastal municipalities and -1.8351 for non-coastal municipalities.

The geographic area (in km^2) of each destination has been included in the model to account for the size effect. The estimated β_5 is positive, significant and very similar at both destinations: 0.0015 for coastal municipalities and 0.0011 for non-coastal ones. A positive and significant coefficient is also estimated for the ratio capturing the capacity of different tourist municipalities to respond to tourism demand, $BEDKM$. According to our results, investment in hotel accommodation has an important impact on tourism demand. The estimated β_6 is 0.3056 for coastal municipalities and more than twice as high, 0.6670, for non-coastal locations. The higher elasticity estimated for non-coastal municipalities implies that the positive consequences of tourism investment are greater at these destinations.

The dummy variables representing the impact of the global crisis, $Y2009$ and $Y2012$, have the expected negative sign. At coastal destinations, their estimated coefficients are significant and indicate a drop in tourist arrivals of around 9% in 2009 and 5.5% in 2012, as a consequence of the crisis. However, at non-coastal destinations the only significant coeficient is that for $Y2012$, suggesting a drop in tourist arrivals of around 7% in 2012.

Finally, since the effect of previous tourists depends negatively on tourist congestion at the destination, this effect varies across the municipalities and over time. At a particular destination, the positive influence of the past visitors decreases as tourist flows increases, and the level of occupation approaches its maximum capacity. In any given year, the effect of the previous tourists varies between the different destinations depending on the relationship between previous tourism demand and the geographic area of the destination, showing that tourism congestion is negatively perceived by tourists. Therefore, to enhance the Spanish destinations' competitiveness, one strategy would be to gear toward reducing the damaging effects of congestion.

6 Conclusions

This paper seeks to investigate the determinants of the tourism demand at the most popular Spanish destinations, with special emphasis on the role of congestion. In line with the TALC theory, our study takes into account that a mature destination could grow at a lower rate as congestion increases. These destinations receive a high number of tourists, which contributes to the growth of the economy, but the space requirements of tourists can lead to problems of congestion, and detract from their attractiveness as tourism destinations. Therefore, the positive influence of previous tourists on tourism demand may be negatively affected by the congestion of the tourism destination.

We consider a dynamic model for tourism demand where current and lagged demand maintain a quadratic relationship. We show that this nonlinear model allows for a positive but decreasing growth rate for tourism demand, as predicted by the TALC theory. In this scenario, the effect of previous tourists on current demand is not constant. It decreases with congestion, measured as the intensity of tourist demand per square kilometre. Another advantage of the quadratic specification is that it nests the most common linear dynamic specification for tourism demand model which implies both a constant growth rate for the number of tourists and a constant marginal effect of the previous tourists.

The determinants of the Spanish tourism demand are analyzed using data from the 93 municipalities chosen by the INE as tourist destinations, for the period 2006-2015. Since the tourism congestion is mainly relevant at the most visited destinations, employing geographically disaggregated arrival data is crucial to ascertain its impact. The system GMM procedure is carried out to estimate the proposed model with two panel data sets, one for the 53 coastal municipalities and the other for the 40 inland municipalities. This sub-sample analysis highlights the differences between the behavior of tourists arriving at both types of destinations.

Our empirical analysis confirms that tourism congestion damages destinations reputation, proving the need for a quadratic specification in the model. We find that the positive effect of the previous visitors is non-constant and decreases as tourist flows increase at each destination. This result supports the TALC theory, since it implies that demand for the most visited destinations could grow at a decreasing rate because of congestion. As a consequence, the stagnation stage of a destination is closer or further, depending on its tourism congestion. Once this stage is reached, the destination has various possible ways forward, from its decline to its rejuvenation. Preventing the decline or stagnation will depend on the efforts and actions of local governments and agents. This will imply changes at the destination that can come about in many different ways, but always as a reaction to the congestion and the needs and wishes of the demand.

In addition, our sub-sample analysis reveals interesting differences in estimated tourism demand for coastal and non-coastal municipalities. The positive effect of the past visitors is higher for sun and beach resorts, suggesting that these more traditional destinations in Spain have a well-established reputation as tourist destinations. Accordingly, their tourism demand is less elastic. Price and income elasticities are significantly lower than elasticities for non-coastal destinations. The negative impact of tourism congestion on tourist demand is also lower at coastal destinations. These results confirm that the sun and beach tourism is what represents the tourism industry in Spain and that, for the moment, it has little real competition. Other types of tourism developed at inland destinations, such as cultural, rural or health, do not seem to making any inroads.

In connection with the TALC theory, our results indicate a different lifecycle model in both types of destinations. The higher persistence of tourism demand found at the sun and beach destinations suggests a higher growth limit than at the inland destinations. In consequence, in equal geographical spaces, the coastal destinations can cater for a greater number of tourists than inland destinations. And, this is the reality in Spain, where mass tourism destinations are located mainly on the coast. However, in a sustainable tourism context these mass tourism destinations are questionable because tourism also has a negative impact on residents and areas receiving tourists. From this perspective, in Spain a more equitable distribution of the demand for tourism between coastal and inland areas would be desirable. Thus, the political implications of our results point to a need for efforts and actions to improve the image and attractiveness of the types of tourism developed at inland destinations.

7 References

Aguiló, E., Alegre, J. and Sard, M. (2005). The persistence of the *sun and sand* tourism model. *Tourism Management* 26(2), 219-231.

Albaladejo, I. P., González Martínez, M. I. and Martínez-García, M. P. (2016). Nonconstant reputation effect in a dynamic tourism demand model for Spain. *Tourism Management,* 53, 132-139.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58, 277–297.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68, 29–51.

Baltagi, B. (2008). *Econometric analysis of panel data* (4th ed.). Chichester: Wiley.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87, 115–143.

Butler, R.W. (1980). The concept of a tourist area cycle of evolution: implications for management of resources, *Canadian Geographer*, 24(1), 5–12.

Butler, R. W. (Ed.). (2006a). *The tourism area life cycle. Applications and modifications*. Clevedon; Buffalo: Channel View Publications.

Butler, R. W. (Ed.). (2006b). *The tourism area life cycle. Conceptual and Theoretical Issues*. Clevedon; Buffalo: Channel View Publications.

Capacci, S., Scorcu, A. E. and Vici, L. (2015). Seaside tourism and eco-labels: The economic impact of Blue Flags. *Tourism Management*, 47, 88–96.

Cerina, F. (2007). Tourism specialization and environmental sustainability in a dynamic economy, *Tourism Economics* 13, 553-582.

Cole, S. (2009). A logistic tourism model – Resort Cycles, Globalization and Chaos. *Annals of Tourism Research*, 36 (4) 689-714.

Cole, S. (2012). Synergy and congestion in the tourist destination life cycle. *Tourism Management*, 33 (5) 1128–1140.

Diedrich, A. and García-Buades, E. (2009). Local perceptions of tourism as indicators of destination decline. *Tourism Management* 30, 512–521.

Garay, L., and Canoves, G. (2011). Life cycles, stages and tourism history: The Catalonia (Spain) experience. *Annals of Tourism Research*, 38, 651–671.

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

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

Garín-Muñoz, T. (2009). Tourism in galicia: domestic and foreign demand, *Tourism Economics*, 15(4), 753-769.

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.

Hansen, L.P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50, 1029-1054.

Haywood, K. (1991) Revisiting Resort Cycle. *Annals of Tourism Research* 19:351–354.

Hsiao, C. (2003). *Analysis of panel data* (2nd ed.). Cambridge: Cambridge University Press.

Ivars, J. A., Rodríguez, I. and Vera, J. F., (2013). The evolution of mass tourism destinations: New approaches beyond deterministic models in Benidorm (Spain), *Tourism Management*, 34, 184-195.

La Rocca, R.A. (2005). Mass Tourism and Urban System: Some Suggestions to Manage the Impacts on the City, *e-Review of Tourism Research* (eRTR), 3(1), 8-17.

Lundtorp, S., and Wanhill, S. (2001). The resort life cycle theory. Generating processes and estimation, *Annals of Tourism Research*, 28(4), 947-964.

Marsiglio, S. (2015). Economic growth and environment: tourism as a trigger for green growth, *Tourism Economics* 21, 183-204.

Marsiglio, S. (2017). On the carrying capacity and the optimal number of visitors in tourism destinations, *Tourism Economics* 23, 632-646.

Massidda, C. and Etzo, I. (2012). The determinants of Italian domestic tourism: A panel data analysis. *Tourism Management*, 33(3), 603-610.

Moore, W., and Whitehall, P. (2005). The tourism area life cycle and regime switching models. *Annals of Tourism Research*, 32, 112–126.

Morley, C. L. (2009). Dynamics in the specification of tourism demand models, *Tourism Economics*, 15(1), 23-39.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 49, 1417-1426.

Oreja Rodriguez, J. R., Parra-Lopez, E., and Yanes-Estevez, V. (2008). The sustainability of island destinations: Tourism area life cycle and teleological perspective. The case of Tenerife. *Tourism Management*, 29, 53–65.

Roodman, D.M. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, 9(1), 86-136.

Russell, R. and Faulkner, B. (2004). Entrepreneurship, Chaos and the tourism area lifecycle. *Annals of Tourism Research*, 31, 556–579.

Sargan, J.D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica*, 26, 393-415.

Saveriades, A. (2000). Establishing the social tourism carrying capacity for the tourist resorts of the east coast of the Republic of Cyprus. *Tourism Management*, 21, 147-156.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators, *Journal of Econometrics*, 126(1), 25-51.