

Nonconstant reputation effect in a dynamic tourism demand model for Spain*

Abstract

Following the ideas of the Tourism Area Life Cycle (TALC) theory, we propose a dynamic econometric model for tourism demand where the reputation effect (the effect of the lagged demand on the current tourism demand) is not constant, but dependent on congestion. We test the model using panel data from Spanish regions during the period 2000-2013. Two estimations are performed depending on whether the tourists' origin is domestic or international. The results show that the reputation effect is not constant in both estimates, supporting the idea that tourism congestion influences tourist arrivals in Spain.

JEL Classification: O41, C61, F43, C32.

Keywords: Tourism Area Life Cycle Model (TALC), reputation effect, congestion, dynamic panel data.

1 Introduction

Research in tourism economics has been dominated by demand analysis (Sinclair, Blake and Sugiyarto, 2003). For a review of the methods used to analyze tourism demand see, for example, Li, Song, and Witt (2005), Song and Li (2008) and Song, Dwyer, Li and Cao (2012). Since the 1990s, demand modelling studies have shifted from static regression models to more sophisticated dynamic specifications. Dynamic models aim to avoid potential problems such as spurious regression, poor predictions and structural instability (Witt and Song, 2000, and Song and Turner, 2006), and take into account important factors like repeat visits, habit persistence, and word-of-mouth recommendations or reputation (Morley, 2009). These models support the idea that previous visitors have an impact on the current tourism demand. In this paper we call this intertemporal link the Persistence or Reputation effect, and it can be caused by a wide branch of different factors.

*The authors have been partially supported by MICINN under projects ECO2008-01551/ECON, ECO2011-24352 (María Pilar Martínez-García) and ECO2013-45698-P (Isabel Albaladejo). The third author also acknowledges the support by COST Action IS1104 "The EU in the new economic complex geography: models, tools and policy evaluation".

The most common way to incorporate dynamics into demand models is to include the lagged demand in a linear fashion as an explanatory variable. However, this may not be sufficient to account for the dynamics of tourism demand (Morley, 1998, 2009). This simple inclusion of the lagged demand assumes a constant persistence or reputation effect.

Nevertheless, the Tourism Area Life Cycle (TALC) theory (Butler, 1980), the most popular one on tourism evolution, suggests that the effect of lagged demand on current demand is not constant, but changes as the level of occupation approaches the destination carrying capacity. According to the TALC theory, during the first stages, the number of visitors increases at an increasing speed. However, as it approaches the carrying capacity the process slows down. Lundtorp and Wanhill (2001, 2006) consider the ratio of visitors over the carrying capacity as the source of this slowing down. This ratio is called tourism congestion along this paper. Following this idea, we propose an econometric demand model where the reputation effect is not constant but varies with congestion.

Our research improves the literature in two ways. First, by specifying a non linear relationship between current and lagged demand we allow for a nonconstant reputation effect. Second, the tourism congestion is considered as the key to this nonlinearity. There may be other factors involved in this non linearity, however. We focus on congestion inspired by the TALC theory. We test the model with panel data from Spanish regions during the period 2000-2013. We perform two different estimates depending on whether the origin of the tourists is domestic or international. The results show a satisfactory performance. The reputation effect is not constant in both estimates, supporting the idea that congestion influences tourist arrivals in Spain.

Panel data have been used in recent studies on tourism demand. However, they simply include the lagged demand in a linear fashion in order to take into account habit persistence or reputation. As examples, we have the work by Maloney and Montes Rojas (2005) for tourist demand at Caribbean destinations, Naudé and Saayman (2005) for tourist demand in 43 African states, Garín-Muñoz (2006, 2007 and 2009) for tourism demand at different Spanish destinations, Garín-Muñoz and Montero-Martín (2007) for tourism demand in the Balearic Islands (Spain), Massidda and Etzo (2012) for domestic tourism in Italy, and Rodríguez et. al. (2012) for academic tourism demand in Galicia (Spain). All these studies assume a constant persistence or reputation effect. Our econometric specification is more flexible as it allows the reputation effect to vary with congestion. Furthermore, it allows us to analyze the effect of tourism congestion on a destination's appeal. To the best of the authors' knowledge, there are few empirical studies analyzing this supply-side factor.

The paper is organized as follows. Section 2 provides the theoretical foundations of the model. Section 3 assesses the congestion in Spain, as a tourist destination. Section 4 presents the data and variables used in our estimates. Section 5 provides the empirical model and describes the econometric methods used for estimation. Section 6 contains the results of our estimations and their interpretations. Finally, Section 7 offers some conclusions.

2 A nonlinear dynamic demand model

Econometric models studying tourism demand are based on the classical economic theory which postulates that income and price-type factors are likely to play a central role in determining the demand. Moreover, theoretical and empirical studies suggest that the behavior of tourism demand may also be affected by dynamic elements (Morley 2009). Accordingly, most tourism demand modelers have included the lagged demand as an explanatory variable (Salman, 2003, Song and Witt, 2003, Croes and Vanegas, 2005, Garín-Muñoz 2006, Garín-Muñoz and Montero-Martín 2007, among others). These models assume a habit persistence or reputation effect that boosts current demand. This intertemporal effect between current demand, T_t , and lagged demand, T_{t-1} , is mathematically measured by the partial derivative $\partial T_t / \partial T_{t-1}$.

The standard dynamic econometric model formally obeys the specification

$$T_t = \beta_0 + \beta_1 T_{t-1} + \gamma' \cdot X_t + \varepsilon_t \quad (1)$$

where the lagged dependent variable T_{t-1} is an explanatory variable and $X_t' = (x_t^1, x_t^2, \dots, x_t^k)$ is the vector of the remaining k explanatory variables (price, income, etc.), which can also include lagged explanatory variables and dummy variables. β_0, β_1 and $\gamma' = (\gamma_1, \gamma_2, \dots, \gamma_k)$ are parameters. The regression error term is ε_t . The proper procedure of estimation and properties of the resulting estimators will depend mainly on the statistical properties of ε_t and of the variables. The demand for tourism, T_t , is measured as the number of nights, number of visitors or tourists' expenditures. See Song et. al. (2010) for a recent review of tourism demand measures. The dependent and explanatory variables can be either in levels or log transformed.

Equation (1) assumes an exponential trend for tourism demand, modified by the evolution of the explanatory variables X_t . The reason for this is that

$$\frac{\partial T_t}{\partial T_{t-1}} = \beta_1 \text{ is constant.} \quad (2)$$

That is, this model assumes that the persistence or reputation effect is constant (lagged demand has a constant effect on the current demand). If variable T_t measures the logarithm of tourism demand, equation (2) means that the elasticity of current tourism demand with respect the lagged demand is constant.

However, the theoretical literature argues that this effect may not be constant (Butler, 1980, 2009, 2011; Morley, 1998, 2000, 2009). Morley suggest a diffusion model, which shares some properties with Butler's (1980) tourism area life cycle (TALC) model. The TALC theory is one of the most widely accepted descriptions of the temporal evolution of tourism areas. The theory argues that resorts evolve over an S-shape curve. Lundtorp and Wanhill (2001) show that this evolution might be satisfactorily approximated by the logistic growth model¹

¹Although Lundtorp and Wanhill (2001) formulated the model in continuous time, here we present its discrete version to fit the econometric analysis better.

$$T_t - T_{t-1} = \sigma T_{t-1} \left(1 - \frac{T_{t-1}}{CC} \right) \quad (3)$$

where parameter σ is the intrinsic rate of tourism growth, assumed as positive, and CC refers to the carrying capacity.

The S-shape pattern is due to the interaction of two opposite effects. First, the less uncertainty associated with holidaying at a known destination and the spreading of the knowledge about destinations as people talk about their holidays lead to a positive autocorrelation of past visitors and current tourists. Secondly, the subsequent congestion has a negative effect on arrivals.

Rearranging the terms in equation (3) gives

$$T_t = \beta_1 T_{t-1} + \beta_2 \frac{T_{t-1}^2}{CC}, \quad (4)$$

with $\beta_1 > 0$, $\beta_2 < 0$. Mathematically, equation (4) is a Riccati equation with constant coefficients which has been used to describe diffusion processes. Note that from equation (4)

$$\frac{\partial T_t}{\partial T_{t-1}} = \beta_1 + 2\beta_2 \frac{T_{t-1}}{CC}. \quad (5)$$

That is, contrary to (2), the persistence or reputation effect is not constant. There exists a positive but diminishing marginal effect of past visitors, T_{t-1} , on current tourism, T_t . This reputation effect decreases with the amount of past tourism.

The non-constant effect (5) is essential for TALC theorists. If the carrying capacity is constant, as the number of visitors grows, the speed of growth decreases. That is, tourism areas "... carry with them the potential seeds of their own destruction, as they allow themselves to become more commercialized and lose their qualities which originally attracted tourists" (Plog 1974:58).

Traditionally, tourism carrying capacity has been considered as a given and static value. However, several authors argue that it can change. (Saveriades, 2000; Cole, 2012, Albaladejo and Martínez-García, 2014). Carrying capacity, CC_t , could evolve along time due to changes in tourists' preferences, tourism supply, or the evolution of environmental or social restrictions. Moreover, destinations can expand their capacity simply by rejuvenating the products and services, by investing in developing new ones, opening up to new markets or improving their communication infrastructures.

Taking into account both the tourism area life cycle theory and a dynamic carrying capacity, the econometric model we propose to analyze the tourism demand is

$$T_t = \beta_0 + \beta_1 T_{t-1} + \beta_2 \frac{T_{t-1}^2}{CC_{t-1}} + \gamma' \cdot X_t + \varepsilon_t. \quad (6)$$

Likewise in Morley (1998, 2000, 2009), model (6) is a quadratic form where the square of lagged demand is divided by a time-dependent variable, in our case,

the carrying capacity. Moreover, according to (6),²

$$\frac{\partial T_t}{\partial T_{t-1}} = \beta_1 + 2\beta_2 \frac{T_{t-1}}{CC_{t-1}}, \quad (7)$$

which means that the persistence or reputation effect is not constant, but is affected by the ratio T_{t-1}/CC_{t-1} . In this paper we call this ratio tourism congestion.

Note that in this model, not only lagged demand, like in (5), but also past carrying capacity, which is not constant in this specification, can modify the reputation effect. If tourism demand is measured with logarithms, equation (7) means that the elasticity of tourism demand with respect to lagged demand is not constant but dependent on lagged congestion. According to (7), if past visitors perceived congestion, the destination reputation worsens and the current tourism demand is negatively affected.

Alleviating tourism congestion requires purposive efforts from entrepreneurs and governments (Albaladejo and Martínez-García, 2014). As is illustrated in the following section, Spain, as a tourist destination, has increased the number of tourism spots throughout its territory, which can be interpreted as an enlargement of the country's tourism carrying capacity. We shall study to what extent this policy has had positive effects on tourism demand.

3 Tourism congestion in Spain

In order to define tourism congestion of a destination a measure of its carrying capacity must be given. 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). However, a measure of the carrying capacity of a destination is difficult to define; there are many factors involved and they are not all quantifiable.

Traditionally, the main reason tourists come to Spain is its sunny climate close to the coast, the so-called "sun and beach" tourism. Over the last 10 or 15 years, the tourists' preferences have changed and they are showing a desire for more activities and alternative forms of tourism (Aguiló et al, 2005). This heterogeneity of the demand has given rise to an increase and diffusion of the supply, with new alternatives being developed in the coastal areas and also in other regions and in many cities of Spain (Ivars, 2004). The spread of the supply through Spain from 2001 to 2013 can be seen in Figures 1 and 2. In 2001, the number of tourism spots was considerably lower than in 2013 and they were

²Note that parameters β_1 and β_2 in equation (6) do not have the same meaning as β_1 and β_2 in equation (4). The two models are different because (4) assumes a constant carrying capacity while the carrying capacity varies in (6) as time passes.

mainly situated on the coasts of the Mediterranean and Atlantic and its two archipelagoes. In 2013, all regions have some tourism spot.



Figure 1: Tourism spots in Spain in 2001



Figure 2: Tourism spots in Spain in 2013

The number of tourism spots is a quantitative measure of the tourism supply of a destination but also of the space distribution of the services offered. In Spain, the National Statistics Institute of Spain (INE) identifies as "tourism spot" a municipality where the concentration of tourism supply -not only lodgings- is significant. All these spots count on some important tourism attraction (beaches, monuments, etc.) or are near to an attraction; the greater the number of tourism spots, the larger the spatial dispersion of supply and therefore the higher the chance to accommodate visitors, that is, lower congestion. The number of tourism spots in Spain can therefore be used as a proxy of its carrying capacity, assuming the limitation of each tourism spot can serve equal numbers of tourists. The advantage of using this measure is that its homogeneous character allows comparison among several destinations.

In this paper, the tourism congestion of a destination at time t is taken as

$$\frac{T_t}{CC_t} = \frac{\text{number of tourists in the destination at time } t}{\text{number of tourism spots in the destination at time } t}$$

The congestion of a destination fluctuates due to changes in the number of tourists and/or in the number of tourism spots. In addition, it increases when the number of tourists grows or when the number of tourism spots decreases.

4 Data and variables

In order to analyze the main determinants of Spanish tourism demand, we estimate the model proposed in Section 2 (equation 6) using data from inter-

national and domestic visitors arriving in Spanish Autonomous Communities during 2000-2013. Spain is an important destination for foreign tourists but also for domestic tourists. In fact, in 2013 just over half of the tourists in Spain were domestic (51% of tourists who chose hotels as accommodation). However, their evolution has varied greatly over the period studied (Figures 3 and 4). In both cases, tourism rose sharply from 2002 to 2007. After that, a decline is observed in both types of tourists in 2008 and 2009, as a result of the global financial crisis and economic recession. Since 2010, the demand for foreign and domestic tourists has had different behaviors. The number of foreign tourists seems to be experiencing a new growth phase, while the domestic tourists continues to decrease, probably because the 2008-2014 crisis hit harder in Spain.

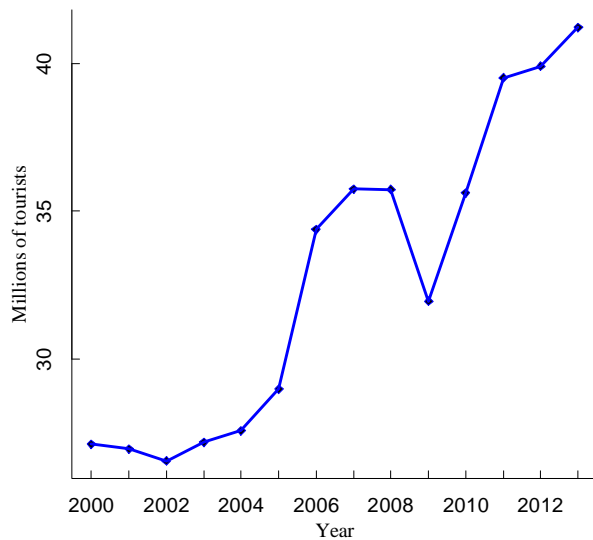


Figure 3: Evolution of number of international visitors staying one or more nights in a hotel in Spain 2000-2013. Source: INE

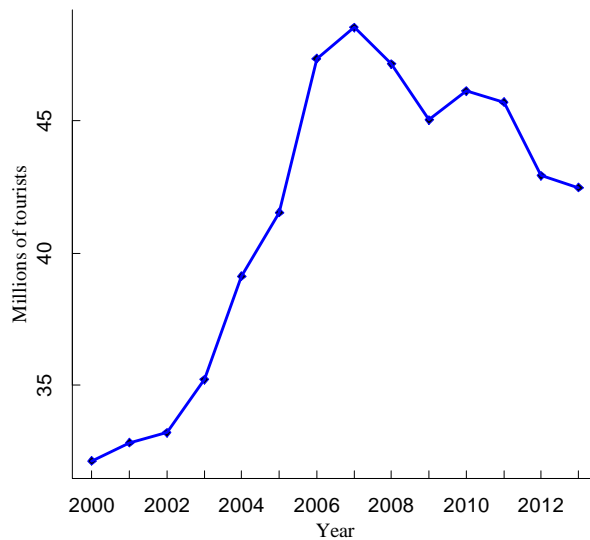


Figure 4: Evolution of number of domestic visitors staying one or more nights in a hotel in Spain 2000-2013. Source: INE

Tourism arrivals in Spain are not homogeneously distributed among regions (Autonomous Communities) either in terms of volume and composition: domestic versus international. In 2013 the regions of the Mediterranean (Catalonia, Valencia, Murcia and Andalusia), the two archipelagos (the Balearic Islands and the Canary Islands) and Madrid accounted for almost 80% of tourism (Figure 5). Moreover, in most Spanish regions, domestic tourism is higher than international tourism. However, the Balearic Islands, Canary Islands and Catalonia, receive higher percentages of foreign tourists than domestic, 86.8%, 76.2% and 63.3% respectively.



Figure 5: Percentage of visitors in 2013 staying one or more nights in a hotel in Spain by Autonomous Communities.

Two models, one for domestic tourism and another for international tourism, are estimated using data disaggregated by region of destination. We use a balanced panel data set consisting of the 17 Autonomous Communities of Spain for the period 2000-2013. The panel data has some advantages over cross sectional or time series data. One is that it enables us to control for unobservable cross sectional heterogeneity, which is common in regional data. Time series and cross section studies not controlling for this heterogeneity run the risk of obtaining biased results. Moreover, panel data usually give a large number of data points, increasing the degrees of freedom, reducing the collinearity among explanatory variables and improving the efficiency of econometric estimates (Hsiao, 2003 and Baltagi, 2008).

The models include economic demand variables, such as income and prices, dummy variables for controlling the effects of the economic crisis of 2008 onwards, and a quadratic form to capture the effect of the lagged demand. Our quadratic relationship allows the positive reputation effect to be non constant, but dependent on the previous congestion (equation 7).

According to the model, the dependent variable is the number of domestic or international tourists (DT and IT , respectively) who choose hotels and similar establishments as accommodation. Data are taken from the Encuesta de Ocupación Hotelera (EOH) of the INE. Two traditional economic factors are included among the explanatory variables: origin income and price. To measure origin income, we use the per capita real GDP of Spain ($GDPSP$) in the domestic tourism model and the per capita real GDP of EU28 ($GDPEU$) in the international tourism model. Both variables were taken from the OCDE. The price variable 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.

We construct two relative price variables, one for the domestic demand model (DP) and one for the international demand model (IP) :

$$\begin{aligned} DP_{it} &= \frac{CPI_{it}}{CPI_{SPt}} & i = 1, \dots, 17 \\ IP_{it} &= \frac{CPI_{it}}{CPI_{EUt} \cdot EX_t} & i = 1, \dots, 17 \end{aligned}$$

where CPI_{it} , CPI_{SPt} and CPI_{EUt} are the consumer price indices (CPIs) for each of the 17 destinations considered, Spain and EU28, respectively; EX_t is the nominal effective exchange rate Spain vs EU28. Data on exchange rates and CPIs for Spain and EU28 were collected from Eurostat. Data on CPI for the 17 Autonomous Communities in Spain were collected from the INE.

Additionally, as seen in Figures 3 and 4, we also consider dummy variables to capture the influence on tourism of the financial and economic crisis of 2008.³ In the domestic tourism model we include a dummy variable ($D2008$) that takes the value of 1 from 2008 onward and zero otherwise. As Spain has undergone a deeper and prolonged economic crisis than its main outbound tourism countries (United Kingdom, France and Germany), international tourist arrivals have been less affected by the crisis of 2008 than domestic arrivals. Figure 3 shows that the break in the upward trend of international arrivals lasts just two years: 2008 and 2009, so in the international model we include two dummy variables: $Y2008$ and $Y2009$. Each takes the value 1 in the year mentioned and 0 in other years.

Our dynamic econometric model also includes the lagged carrying capacity, CC_{t-1} , which, as defined in section 3, is the number of tourism spots in a region. These were collected from the Encuesta de Ocupación Hotelera (EOH) of the INE.

5 Methodology and model specification

Following the model proposed in Section 2 and considering the variables defined in Sections 3 and 4, the econometric models as represented as:

$$DT_{it} = \eta_i + \beta_1 DT_{i,t-1} + \beta_2 \frac{DT_{i,t-1}^2}{CC_{i,t-1}} + \beta_3 GDPSP_t + \beta_4 DP_{it} + \beta_5 D2008_t + \varepsilon_{it} \quad (8)$$

$$\begin{aligned} IT_{it} &= \eta_i + \beta_1 IT_{i,t-1} + \beta_2 \frac{IT_{i,t-1}^2}{CC_{i,t-1}} + \beta_3 GDPEU_t + \beta_4 IP_{it} \\ &+ \beta_5 Y2008_t + \beta_6 Y2009_t + \varepsilon_{it} \end{aligned} \quad (9)$$

where the subscripts i ($i = 1, \dots, 17$) and t ($t = 2000 - 2013$) denote the destination region and time period, respectively. η_i is the unobserved regional-specific

³We appreciate this recommendation from an anonymous reviewer.

variable (or fixed effects) that varies across regions but is invariable within a region over time, and ε_{it} is a disturbance term. A key assumption throughout this paper is that the disturbance ε_{it} is uncorrelated across regions, but regional heteroskedasticity and serial correlation is allowed for. Given a specific time period t , the dummies variables and the origin incomes ($GDPSP$ and $GDPEU$) are common to all destinations. Therefore, these variables only vary throughout time, while the others vary both throughout time and across regions. The number of domestic and international tourists, per capita real GDPs and prices are in logs, and therefore coefficients may be interpreted as elasticities.

As discussed in Section 2, the relation between current and past tourism depends on β_1 , β_2 and the previous level of congestion⁴. Since a positive sign is expected for β_1 , a negative β_2 would imply that the elasticity between current and past tourism is positive but decreasing with the previous congestion. If β_2 is zero, the elasticity is constant. As usual in demand models, we expect a positive sign for β_3 and a negative sign for β_4 , β_5 and β_6 .

A generalized method of moments (GMM) panel data estimation (Arellano and Bond, 1991, Arellano and Bover, 1995, Blundell and Bond, 1998) was applied 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 regional 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. This estimator is called the difference GMM. Blundell and Bond (1998) extended the difference GMM 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 equation in levels in addition to the usual lagged levels as instruments for equation in first-differences. Blundell and Bond (1998) showed that the system GMM estimator has better finite sample properties with highly persistent series. They demonstrated that if the dependent variable is close to a random walk the difference GMM performs poorly because lagged levels are weak instruments for first differences. Their Monte Carlo analysis finds both a large finite sample bias and poor precision for the difference GMM procedure in this case. System GMM not only improves the precision but also reduces the finite sample bias.

In this paper, we apply the difference GMM (Arellano and Bond, 1991)

⁴When we refer to congestion in the domestic model it means the ratio between the number of domestic tourists and the number of tourism spots. Likewise, congestion in the international model means the ratio between the number of international tourists and the number of tourism spots.

and the system GMM (Blundell and Bond, 1998) procedures to estimate the models (8) and (9).⁵ In both procedures, we use the one-step robust to heteroskedasticity estimator and the two-step estimator for comparison.⁶ Although the two-step estimator is theoretically preferred, it is appropriate to consider the one-step results when making inferences since the asymptotic standard errors of one-step GMM estimators are virtually unbiased (Arellano and Bond, 1991).

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 estimators⁷, and the Arellano and Bond (1991) test for autocorrelation in the disturbance term, ε_{it} .

6 Results

As a preliminary analysis of the data, we checked for the integration properties of the variables involved. The Levin, Lin and Chu (2002) panel unit root tests (LLC test) was carried out.⁸ This procedure tests the null of a unit root against the alternative of a stationary process for all cross sections. After a graphical diagnosis, the equations for the unit root test have been specified with fixed effects and individual time trends in the data generating process. Table 1 shows the results of the LLC test. The hypothesis that each variable has a unit root is rejected, although some series do show high levels of persistence.

Table 1. *Levin, Lin and Chu (2002) panel unit root tests*

Domestic variables	DT	$\frac{DT^2}{CC}$	$GDPSP$	DP
$t - star$	-1.93**	-3.37***	-1.89**	-9.94***
International variables	IT	$\frac{IT^2}{CC}$	$GDPEU$	IP
$t - star$	-6.86***	-5.04***	-2.85**	-2.49***

Note: t-star is the LLC statistic in the most general model with fixed effects and individual time trends. It is asymptotically distributed as a standard normal under the null hypothesis of nonstationarity. The optimal lag length is selected using the Akaike Information Criterion (AIC). *, **, *** denote significant at the 10%, 5% and 1% level respectively.

We show four different GMM estimates for each model: one-step and two-step versions of the difference GMM (GMM-DIF) and system GMM (GMM-SYS). In all estimates the lag of the dependent variable and the quadratic term

⁵We appreciate the recommendation about using the system GMM from an anonymous reviewer.

⁶One-step GMM estimator is based on the assumption that the ε_{it} are i.i.d. In this paper, we use one-step robust estimators, where the resulting standard errors are consistent with panel-specific autocorrelation and heteroskedasticity.

⁷The 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 over-identification test is theoretically superior to the Sargan (1958) test.

⁸We have applied a panel unit root test following the recommendation of an anonymous reviewer.

are treated as endogenous. Due to the small number of regions, all the estimates are obtained using only the second lag of the variables as instrument.⁹ Because 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. Since the panel unit root test suggests tourism series are moderately persistent, we think that system GMM estimator is more suitable for our study.

The empirical results from the estimation of the models (8) and (9) are shown in Tables 2 and 3, respectively. According to the system GMM estimates the coefficient of the lagged dependent variable is significant and positive and the coefficient of the quadratic term is significant and negative, revealing a nonconstant reputation effect which is negatively affected by previous level of tourism congestion. Additionally, the results reveal a general 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 for joint validity of the instruments.

Table 2: *Estimation results for domestic tourism model, 2000-2013*

Dependent variable: DT_{it} Explanatory variables	GMM-DIF		GMM-SYS	
	one-step	two-step	one-step	two-step
$DT_{i,t-1}$	0.842***	0.832***	0.969***	0.953***
$\frac{DT_{i,t-1}^2}{CC_{i,t-1}}$	-0.0008**	-0.0009*	-0.0004**	-0.0005**
$GDPSP_t$	0.799***	0.787***	0.688***	0.652***
DP_{it}	-	-	-	-
$D2008_t$	-0.072***	-0.070***	-0.089***	-0.087***
Hansen test (p-value)	0.794	0.794	1.000	1.000
AR(1) (p-value)	0.003	0.004	0.002	0.003
AR(2) (p-value)	0.110	0.163	0.113	0.120
Number of observations	204	204	221	221
Number of groups	17	17	17	17

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

⁹Roodman (2009b) argues that finite sample problems caused by a large number of instruments are of two sorts. First, numerous instruments can overfit instrumented variables, biasing coefficient estimates towards those from non-instrumenting estimators. Second, instrument proliferation can take two-step GMM far from the theoretically efficient ideal and can weaken the Hansen test.

Table 3: *Estimation results for international tourism model, 2000-2013*

Dependent variable: IT_{it}	GMM-DIF		GMM-SYS	
	one-step	two-step	one-step	two-step
Explanatory variables				
$IT_{i,t-1}$	0.532***	0.528***	0.989***	0.982***
$\frac{IT_{i,t-1}^2}{CC_{i,t-1}}$	0.0003	0.0003	-0.0003**	-0.0004*
$GDPEU_t$	2.374***	2.312***	0.629***	0.535**
IP_{it}	-1.332***	-1.285**	-	-
$Y2008_t$	-0.076***	-0.074***	-0.083***	-0.082***
$Y2009_t$	-0.113***	-0.114***	-0.158***	-0.169***
Hansen test (p-value)	0.794	0.794	1.000	1.000
AR(1) (p-value)	0.006	0.005	0.002	0.003
AR(2) (p-value)	0.304	0.313	0.362	0.428
Number of observations	204	204	221	221
Number of groups	17	17	17	17

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

Table 2 shows estimation results for the domestic tourism model (equation 8). We find that all variables are significant except relative price. So, we estimate the model without price. All estimates (GMM-DIF and GMM-SYS) yield similar results. Focusing on the system GMM procedure, since estimated β_1 (0.969 and 0.953) is positive and β_2 (-0.0004 and -0.0005) negative, it is clear that, for current levels of tourists and tourism spots, the reputation effect is positive and decreases slowly with the previous congestion. The estimated income elasticity (0.688 and 0.652) is positive and significant, showing that the national demand for tourism in Spanish regions depends positively on the wealth of Spain. Finally, estimated β_5 is negative and significant (-0.089 and -0.087). So, as expected, domestic tourism has been negatively affected by the economic crisis. These results show that domestic tourists arrivals depend heavily on the Spanish economic situation.

Table 3 shows the results for the international tourism model (equation 9). System GMM results are very similar to those of the domestic model. As relative price is not significant, we conduct system GMM estimation without price. This estimation show a positive sign for β_1 (0.989 and 0.982) and negative for β_2 (-0.0003 and -0.0004). Therefore, the elasticity of international tourism demand with respect to the lagged demand is positive and decreasing with the previous congestion, as in the estimation carried out for domestic tourism. As regards the estimated income elasticity, the results are consistent with economic theory. As expected, a positive elasticity is estimated for per capita real GDP with values of 0.629 and 0.535, suggesting that international tourist arrivals in the Spanish regions depend on the economic situation of the European Union, which is the main market of origin. Finally, the dummy variables representing the impact of the global crisis, $Y2008$ and $Y2009$, have the expected negative signs and both are significant.

To summarize, the main determinants of the domestic and international tourism in Spain seem to be per capita tourist income and the lagged dependent variable, which controls the role of the persistence or reputation. The estimated coefficients β_1 and β_2 indicate that reputation has played an active role in regional Spanish tourism. This reputation effect is not constant but varies depending on the level of congestion. Higher levels of previous congestion worsen the reputation of the destination and lead to a lower reputation effect. In line with the TALC theory, our result implies a positive but decreasing reputation effect. This decreasing reputation effect is revealed most clearly in the domestic demand model, where both GMM procedures lead to the same results. In the international demand model, only the system GMM estimator shows this result.¹⁰

This less conclusive result for international tourists may be due to the characteristics of international tourism in Spain. On one side, international tourist arrivals are not homogeneously distributed between Spanish regions in terms of volume, so the effect of congestion may be different for regions with high levels of international tourism (Balearic Islands, Canary Islands and Catalonia) than for the rest. Moreover, tourist perception of congestion may depend on the country of origin. In a recent paper, Santana-Jiménez and Hernández (2011) analyze the influence of the tourist perception of overcrowding on the tourist affluence. Using the population density to measure overcrowding in a tourist area, they estimate a tourism demand model with data of tourists coming from UK and Germany to the Canary Islands. Their results show opposite signs of the density effect on demand across the different islands and countries of origin, revealing that tourists' perception of overcrowding depends on consumer characteristics and destination.

7 Concluding remarks

There is general agreement on the desirability of taking into account the reputation of a destination among the factors explaining tourism demand. To date, most empirical studies on tourism demand include the lagged demand to measure this persistence effect in a linear fashion. These specifications assume a reputation effect that is constant over time, i.e. independent of variables like the level of tourism congestion that could affect tourist arrivals. This assumption contrasts with accepted theories on tourism, like the TALC theory.

In this paper, we follow the TALC theory to propose a new dynamic specification to estimate demand elasticities. Our tourism demand model includes a quadratic form of the lagged demand and allows a nonconstant reputation effect which depends on congestion.

¹⁰In the GMM-DIF estimation the quadratic term is not significant and estimated β_1 (0.532 and 0.528) is considerably less than that obtained with the GMM-SYS estimator. This low value is likely to be the result of the expected downward bias when the available instruments are weak (Blundell and Bond, 1998).

We use a panel data of tourists arrivals in the 17 Spanish Autonomous Communities during the period 2000-2013 to test the proposed model. The analysis was performed separately for domestic tourist arrivals and international arrivals, using two different GMM estimators: difference GMM and system GMM. In both cases the econometric model includes traditional economic factors, such as income and relative prices, and the quadratic function of the lagged demand. Our dynamic specification is more flexible than that used elsewhere. The reputation effect is not fixed but depends on congestion, defined as the ratio between the number of tourists and the number of tourism spots at the destination. Nonetheless, the number of tourism spots quantifies not only the supply but also its spatial dispersion at the destination. To the best of our knowledge, there are few contributions that include tourism congestion in their model specification, and none in the way we do.

In both models, the system GMM estimation reveals a positive reputation effect which decreases with the previous level of tourism congestion, as theoretically suggested by the TALC model. For the domestic tourism model, the different GMM estimates lead to the same results, reinforcing the conclusion. However, the results are not so clear for the case of international tourism. This opens up new lines of research. Is it because for most Spanish regions domestic tourism is considerably larger than the international? Is it because international tourism in Spain is concentrated in a few regions? Do international tourists look for a different tourism product than domestic tourists? In the future it would be interesting to conduct a similar study for those Spanish regions where international tourism has greater weight: the Balearic Islands, Canary Islands and Catalonia. Would we obtain a different conclusion? Moreover, allowing the possibility of spatial autocorrelation between regions could enrich tourism demand analysis and it is another area to work on in the future.

8 References

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

Albaladejo, I. P., Martínez-García, M. P. 2014. An R&D-based endogenous growth model of international tourism, forthcoming in *Tourism Economics*.

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. 2009. Tourism in the future: Cycles, Waves or Wheels? *Futures*, 41, 352
- Butler, R.W. 2011. *Tourism area life cycle*. Goodfellow Publishers Limited, Woodeaton, Oxford.
- 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.
- Croes, R.R. and Vanegas, M. 2005. An econometric study of tourist arrivals in Aruba and its implications. *Tourism Management*, 26(6), 879-890.
- Diedrich, A., García-Buades, E.. 2009. Local perceptions of tourism as indicators of destination decline. *Tourism Management* 30, 512–521.
- 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.
- Hsiao, C. 2003. *Analysis of panel data* (2nd ed.). Cambridge: Cambridge University Press.

- Ivars, J.A. 2004. Tourism planning in Spain. Evolution and Perspectives. *Annals of Tourism Research*, 31(2), 313–333.
- Levin, A., Lin, C.-F. and Chu, C.-S. J. 2002. Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, 108, 1–24.
- Li, G., Song, H., and Witt, S.F. 2005. Recent developments in econometric modeling and forecasting. *Journal of Travel Research*, 44, 82-99.
- Lundtorp, S., and Wanhill, S. 2001, The resort life cycle theory. Generating processes and estimation, *Annals of Tourism Research*, 28(4), 947–964.
- Lundtorp, S., and Wanhill, S. 2006. Time path analysis and TALC stage demarcation, in Butler, R (Ed.), *The Tourist Area Life Cycle: Conceptual and Theoretical Issues*, Channel View Publications, Clevedon.
- Maloney, W.F., and Montes Rojas, G.V. 2005. How elastic are sea, sand and sun? Dynamic panel estimates of the demand for tourism. *Applied Economics Letters*, 12, 277-280.
- Massidda, C. and Etzo, I. 2012. The determinants of Italian domestic tourism: A panel data analysis. *Tourism Management*, 33(3), 603-610.
- Morley, C. L. 1998. A dynamic international model, *Annals of Tourism Research*, 25(1), 70-84.
- Morley, C. L. 2000. Demand modelling methodologies: integration and other issues. *Tourism Economics*, 6, 5-20.
- Morley, C. L. 2009. Dynamics in the specification of tourism demand models, *Tourism Economics*, 15(1), 23-39.
- Naudé, W. and Saayman, A. 2005. Determinants of tourist arrivals in Africa: a panel data regression analysis. *Tourism Economics*, 11, 365-391.
- Nickell, S. 1981. Biases in dynamic models with fixed effects. *Econometrica*, 49, 1417-1426.
- O'Reilly, A. M. 1986. Tourism carrying capacity -concepts and issues. *Tourism Management*, 7(3), 154-167.
- Plog, S. C. 1974. Why destination areas rise and fall in popularity? *The Cornell Hotel and Restaurant Administration Quarterly*, 14, 55-58.
- Rodríguez, X.A., Martínez-Roget, F. and Pawlowska, E. 2012. Academic tourism demand in Galicia, Spain. *Tourism Management*, 33(6), 1583-1590.

- Roodman, D.M. 2009a. How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, 9(1), 86-136.
- Roodman, D.M. 2009b. A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1), 135-158.
- Salman, A.K. 2003. Estimating tourist demand through cointegration analysis: Swedish data. *Current Issues in Tourism*, 6(4), 323-339.
- Santana-Jiménez, Y., and Hernández, J.M. 2011. Estimating the effect of overcrowding on tourist attraction: The case of Canary Islands. *Tourism Management*, 32(2), 415-425.
- 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.
- Sinclair, M.T., Blake, A., and Sugiyarto, G. 2003. The economics of tourism, in C. Cooper eds, *Classic reviews in tourism* (pp. 22-54). Channel View Publications, Clevedon.
- Song, H., Dwyer, L., Li, G., and Cao, Z. 2012. Tourism economics research: a review and assessment. *Annals of Tourism Research*, 39(3), 1653-1682.
- Song, H., and Li, G. 2008. Tourism demand modelling and forecasting-A review of recent research. *Tourism Management*, 29(2), 203-220.
- Song, H., Li, G., Witt, S.F., and Fei, B. 2010. Tourism demand modelling and forecasting: how should demand be measured? *Tourism Economics*, 16(1), 63-81.
- Song, H. and Turner, L. 2006. Tourism demand forecasting. In Dwyer, L., and Forsyth, P., eds, *International Handbook on the Economics of Tourism* (pp. 89-114). Edward Elgar, Cheltenham.
- Song, H. and Witt, S. F. 2003. Tourism forecasting: The general-to-specific approach. *Journal of Travel Research*, 42(1), 65-74.
- Windmeijer, F. 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators, *Journal of Econometrics*, 126(1), 25-51.
- Witt, S.F. and Song, H. 2000. Forecasting future tourism flows, in S. Medlik and A. Lockwood (Eds.), *Tourism and hospitality in the 21st century* (pp. 106-118). Oxford: Butterworth-Heinemann.