TOURISM ECONOMICS
The Business and Finance of Tourism and Recreation

Article

Tourism Economics
1–24

DOI: 10.1177/1354816619840845
journals.sagepub.com/home/teu
SAGE

Modelling tourist expenditure at origin and destination

Emilio Gómez-Déniz
University of Las Palmas de Gran Canaria, Spain

Jorge V Pérez-Rodríguez
University of Las Palmas de Gran Canaria, Spain

José Boza-Chirino
University of Las Palmas de Gran Canaria, Spain

Abstract

This article proposes a model of foreign tourist expenditure, based on expenditure in the country of origin (i.e. reservation of accommodation and transport) and on goods and services at the destination. The study focuses on two measures reflecting the two types of expenditure: the tourist budget share and the difference in growth rates between expenditure at origin and at destination. The random nature of each of these variables is taken into account. The tourist budget share is determined using a fractional response model, based on the beta distribution. This approach allows us to accommodate certain aspects of the empirical budget share distribution, such as skewness, and to represent the results as bounded between 0 and 1, but also to include covariates. The empirical analysis was conducted using data obtained by the Canary Islands Tourist Expenditure Survey, focusing on German and British tourists in particular. The results obtained show that the fractional regression model proposed represents the behaviour of the relevant variables reasonably well and surpasses the performance of the linear regression model.

Keywords

beta regression, expenditure at origin and at destination, growth rates, tourist budget share

Corresponding author:
José Boza-Chirino, Department of Quantitative Methods and TIDES Institute, University of Las Palmas de Gran Canaria, 35017, Las Palmas de Gran Canaria, Spain.
Email: jchirino@ulpgc.es

Introduction

Tourist expenditure is a consumer choice, first among goods and services offered in different sectors, and then among options in the tourism industry (Tribe, 2005; Zheng and Zhang, 2013). According to Syriopoulos and Sinclair (1993), the assumption of weak separability, whereby there is independence only among groups of commodities, not among individual commodities, has various implications: first, that tourists allocate their budget between tourism activities and other goods and services; second, that the tourism budget is allocated to one or more destinations, including the home country; third, that tourists choose how to allocate their tourism budget among the different goods and services offered at the selected destination(s).

Most empirical studies of the determinants of expenditure at a single tourist destination focus on modelling total trip expenditure (aggregate or daily) from a microeconomic standpoint; see, for example, Aguiló and Juaneda (2000), Fredman (2008), Wang and Davidson (2010), Marcussen (2011), Thrane and Farstad (2011), García-Sánchez et al. (2013), Disegna and Osti (2016) and Aguiló et al. (2017). These researchers concluded that expenditure can be affected by many variables, including socio-economic status, nationality, age, occupation, income, length of stay, type of holiday accommodation, travel companion, destination loyalty and package holiday status.

However, to our knowledge, no previous studies have examined the distribution of different types of tourist expenditure.[1] This is an important question because recent empirical studies have shown, for example, that low-cost carriers (LCCs) have made air travel available to all budgets (Dobruszkes, 2013) and enable tourists to spend more at their destination, thus redistributing the trip budget. Therefore, policymakers and institutions in the tourism sector would benefit from knowing not only how tourists distribute their spending but also the effects of price changes and other exogenous variables on this distribution.

The present study analyses the above question in terms of tourists' allocation of expenditure at origin and at destination. The allocation of expenditure includes the different elements of the trip budget (transport, accommodation, external activities, food, shopping, etc.); see Ferrer-Rosell and Coenders (2017). We term this allocation the tourist budget share, that is, the budget allocated to goods and services consumed at the destination as a proportion of the total trip budget (i.e. expenditure at origin on accommodation and transport, and expenditure at destination on goods and services). Therefore, this variable is bounded between 0 and 1.[2] Moreover, the rate of change in expenditure can be evaluated in statistical terms.

This article contributes new insights to our understanding of tourism expenditure and to the analysis of this question in the following ways.

Firstly, we propose a model to represent tourists' allocation of expenditure between origin and destination, for a given destination. Specifically, we take into account the statistical nature of the tourist budget share and the quotient of types of expenditure. In other words, we consider dependence between expenditures (i.e. different categories of tourist expenditure are interdependent) and the non-negative right-skewed distribution of expenditures. Cárdenas et al. (2015) assumed a gamma distribution for the variable number of euros spent per person per day corresponding to the individual. Nevertheless, it is difficult to fit data using distributions such as the generalized gamma or the inverse Gaussian. Therefore, it might be better to consider the expenditure at destination as a proportion of total expenditure. This approach enables us to model the tourist budget share on a statistical basis, based on a variable bounded between 0 and 1.

Secondly, we model tourist budget share by allowing for covariates. However, we do not use a model related to the consumer demand theory, such as the almost ideal demand system proposed by Deaton and Muellbauer (1980), which starts from a specific cost function and gives the budget share equations in an n-good system depending on the prices for each good, and the total expenditure on goods deflated by a price index. Instead, we use a fractional response model, of the type which has been applied in many economic settings such as employee participation in pension plans (Papke and Wooldridge, 1996, 2008), financial issues (Pérez-Rodríguez and Gómez-Déniz, 2015) and insurance (Gómez-Déniz et al., 2013).[3] Therefore, the marginal effects of several factors can be analysed to account for the tourist budget share at destination. In this respect, for example, economic, socio-demographic and trip characteristic factors could be used (Disegna and Osti, 2016; Marcussen, 2011; Thrane, 2014; Wang and Davidson, 2010).

The empirical analysis presented in this article is based on a study of the Canary Islands, which is one of the major package holiday destinations in Spain. With respect to tourist expenditure and taking into account the information obtained in the Survey of Tourist Expenditure conducted by the Canary Islands Government, this article focuses on tourists who allocate part of their budget to a specific destination and who distribute their spending among the different goods and services offered at this destination. In this respect, we separate the expenditure incurred at origin (i.e. the accommodation and transport booked and paid for prior to arrival) from that at destination (i.e. the goods and services consumed after arrival, such as restaurant meals, transport and leisure activities).

The rest of this article is organized as follows. The second section reviews recent studies of tourist expenditure. The third section then presents the beta regression model used to estimate the determinants of this expenditure. This is empirically illustrated in the fourth section, with data for British and German tourists in the Canary Islands. These data are analysed to evaluate the factors determining the expenditure by these tourists. The results thus obtained are then compared with those provided by the ordinary least squares (OLS) procedure. Finally, the fifth section presents the main conclusions drawn.

Literature review

Many studies have been conducted to model individual travellers' expenditure and its determinants.

Empirical research in this respect usually distinguishes between macro- and micro-level expenditure by tourists. In fact, only a few have analysed expenditure at the macro level. One such was Pyo et al. (1991), who used a regression-based method in which the linear expenditure system specification model was applied to analyse US domestic tourism demand, with macroeconomic data. The model assumes consumer utility maximization under budget constraints and takes into account five commodity groups: transport, accommodation, food, entertainment/recreation and other goods and services. Statistical analysis was conducted by OLS and seemingly unrelated regression equations. The findings obtained suggest that among these tourist-oriented products, transport is the most price sensitive and food, the least. The widely held perception that tourist products are luxury items may result from the income effect of transport expenditure.

The majority of studies in this field focus on micro-level data. For example, Aguiló and Juaneda (2000), Fredman (2008), Thrane and Farstad (2011) and García-Sánchez et al. (2013) all analysed the daily expenditure of foreign tourists. See Table 1 for a brief description of some microdata studies.

For a review of this research, see Wang and Davidson (2010), Marcussen (2011), Brida and Scuderi (2012) and Disegna and Osti (2016).

Table 1. Brief description of some papers focusing on expenditure, using microdata.

Authors	Data	Method	Dependent variable	Independent variables	Conclusions
Aguiló and Juaneda (2000)	A sample of 5500 tourists were interviewed in 1996 at the end of their holiday at Mallorca, Menorca and Ibiza airports	Two regression models with dummy variables	1. Expenditure in the country of origin 2. Tourism expenditure in the Balearic Islands	Dummy variables and two quantitative variables (number of days, number of people included in the declared expenditure)	The tourism product may be understood not only as potential demand, but also in terms of profitability considering the expenditure. The idea that at business level it is necessary to develop a product strategy that takes into account the productivity in terms of expenditure
Laesser and Crouch (2006)	The data used in this analysis stem from the 1998 International Visitor Survey in Australia conducted by AC Nielsen. The data were collected by means of face-to-face interviews at airports	Log-linear regression	Total expenses for a trip of one person-day	Visitors were further segmented according to country of residence and when the overall reason for travel was holiday. Group travel, accommodation, duration of trip	Hotel accommodation is associated with higher expenditure, whereas staying at a friend's or relative's house is associated with comparably less spending. Group travel is associated with 10% lower expenditure than average. Holiday trip duration and number of travel companions are associated with increased overall spending
Craggs and Schpfield (2009)	A total of 392 useable questionnaires were obtained from a convenience sample	χ^2, *t* tests and ANOVA to test for differences between subjects' expenditure	Expenditure by local category	Age, gender, level of education, visit purpose, frequency of visit, group size	Young visitors (18–34) tend to be light or medium spenders compared with members of the older age groups (45–64), who tend to be medium to heavy spenders

(continued)

Table 1. (continued)

Authors	Data	Method	Dependent variable	Independent variables	Conclusions
Marcussen (2011)	11,000 leisure and business visitors staying in commercial accommodation in Denmark during 2004	OLS regression analysis	Seven different dependent: Transportation spending per person per night, accommodation spending per person per night, other spending per person per night, total spending per person per night, spending per night per travel party, spending per person per stay and spending per travel party per stay (all in EUR)	Accommodation, nights, party size, distance, transportation mode, age, income, repeat visitor, shopping, package tour, online info, gender	When total spending is the sum of transport + accommodation + other. The sum of the regression coefficients for the three spending components is exactly equal to the regression coefficient for total spending (per person per night)
Thrane and Farstad (2011)	3913 observations from Norwegian families who had been on holiday in Norway during the summer of 2008 and who had spent at least one night away from home	OLS regression analysis	Personal tourism expenditure	Length of stay, travel party size, purpose of trip, transportation mode, destination type, household income, age	The relationship between length of stay and personal tourism expenditure is inelastic. Specifically, a 10% increase in length of stay produces a 6.4% increase in personal tourism expenditure. The effect of travel party size on personal tourism is clearly significant and suggests that a 10% increase in travel party size produces a 1.6% decrease in personal tourism expenditure

(continued)

Table 1. (continued)

Authors	Data	Method	Dependent variable	Independent variables	Conclusions
Brida et al. (2013)	A total of 1193 visitors to northern Italy were interviewed in three different cities: Trento (Trentino), Bolzano and Brunico (South Tyrol). The sample consists of 63.68% tourists, that is, people who declared a positive expenditure on accommodation and 36.32% day visitors	Double-hurdle model estimated by MLE	Expenditure at Christmas Market by local categories	Socio-demographic and economic characteristics of the respondents: age, gender, day visitors	Visitors who are 45 or older spend more than the younger visitors. Positive impact of length of stay on expenditure
Chang et al. (2013)	137 first-time and 133 repeat inbound visitors from the 2009 Annual Survey Report on Visitor Expenditures and Trends, published by the Taiwan Tourism Bureau	Almost Ideal Demand System to analyse visitors' expenditure choices	Price and expenditure elasticity: first-time visitors and repeat visitors	Accommodation, food, transport, entertainment, shopping	Visitors' total expenditure increases when their budgets increase. Previous travel experience does not contribute to significant differences in visitors' preference and expenditure pattern
García-Sánchez et al. (2013)	Data set of nearly 184,000 observations covering 2004–2009. Database 'Survey of Tourist Expense' (EGATUR), Spanish Institute for Tourism Studies	OLS, and robust-to-heteroscedasticity standard errors calculated using pooled cross-sectional analysis	Log daily expenditure	Travel-related variables and socio-economic characteristics: income, age, length of stay, accommodation, repetition, group	Length of stay influences daily expenditure. There is an inverted U-shaped relationship between tourist age and his or her daily expenditure. The trip duration has a negative effect on daily expenditure, but this effect is smaller the longer the trip. There are scale economies in the group size

(continued)

Table 1. (continued)

Authors	Data	Method	Dependent variable	Independent variables	Conclusions
Zheng and Zhang (2013)	7474 households in the 1996 data and 8124 households in the 2006 data from the Consumer Unit Characteristics and Income files (FMLY) and the expenditure sections of the Interview survey files (EXPN)	Tobit model estimation	The total expenditure is divided into food/ beverage, lodging, transportation and sightseeing/ entertainment to examine more specific changes	Annual income after tax, age, family size, children, gender, married, education, occupation, race	There is no apparent change in the consumption pattern in regard to the age of household head, number of children and marital status. The difference in the total expenditure between Black people and others was more substantial than that between White people and others
Disegna and Osti (2016)	Foreign visitors who visited the provinces of Bolzano, Trento and Belluno in 2011 (1030 observations). Annual Survey Banca d'Italia, (2011) the Bank of Italy (Banca d'Italia) entitled International Tourism in Italy	Tobit model	Total visitor expenditure in the different product categories	Length of stay, repetition, accommodation, transportation, food and beverages, shopping and other services	For a 1-day increase, they expect to see a 3.2% decrease in the total visitor spending. Higher the number of repeat visits, the lower the total expenditure: first-time visitors and visitors who have already visited Italy up to 5 times, respectively, spend 29.6% and 14.4% less than visitors who have already visited Italy more than 5 times in the past
Eugenio-Martin and Inchausti-Sintes (2016)	Canary Islands Tourist Expenditure Survey (53,608 observations)	Simultaneous equations estimated by 3SLS	Expenditure at origin and at destination	Income, price, length of stay, package, island visited, group, motivation, previous visits	The hypothesis proposed 'Low-cost travel savings from tourists' place of origin are transferred, at least partially, to higher tourism expenditures at the destination' holds for most tourist profiles, with savings-transfer ratios that range between 10.3% and 46.1%

(continued)

Table 1. (continued)

Authors	Data	Method	Dependent variable	Independent variables	Conclusions
Aguiló et al. (2017)	Sample of 1217 observations. Study data were obtained at the Palma de Mallorca Airport during summer 2009 when the visitors were leaving the island	Robust Least Squares	Total expenditure is separated into the impact on daily expenditure and the impact on length of stay	Income, employee, student, accommodation, group, satisfaction, package, all inclusive	The positive relationship comes from the results that high-income tourists present a high daily expenditure. High-income tourists present a larger length of stay
Gómez-Déniz and Pérez-Rodríguez (2018)	Tourists visiting the Canary Islands during the period from 2009 to the first quarter of 2012	Subordinated process for the aggregate expenditure (continuous variable) conditional on the length of stay (discrete variable) of tourists at a given location, in the spirit of statistical mixtures	Expenditure at origin and at destination	Socio-economic status, nationality, age, occupation, income, length of stay, type of holiday accommodation, travel companion, destination loyalty and package type	They propose two distribution models, for which closed-form expressions are obtained. Covariates are introduced in order to study the factors that affect the aggregate expenditure. The results obtained indicate that the models achieve a reasonably good fit, with and without covariates

OLS: ordinary least squares; 3SLS: Three-stage least squares; ANOVA: Analysis of variance; MLE: Maximum likelihood estimator.

According to Disegna and Osti (2016), most microdata studies of tourism expenditure (aggregate or daily) have used the context of general linear specification models to examine which factors best account for this expenditure. The estimation method most commonly used is OLS (see, for example, Wang and Davidson, 2010; Brida and Scuderi, 2012 and Wu et al., 2013). However, other specifications and estimators have also been used. For example, Leones et al. (1992), Lee (2001), Zheng and Zhang (2013), Barquet et al. (2011) and Kim et al. (2011), among others, estimated tourism expenditure by means of the Tobit model. More recent applications include the Heckman model, by Alegre et al. (2013), the double-hurdle model, by Brida et al. (2013), and a subordinated process for the aggregate expenditure (continuous variable) conditional on the length of stay (discrete variable) of tourists at a given location, in the spirit of statistical mixtures (Gómez-Déniz and Pérez-Rodríguez, 2018). According to these studies, the main factors underlying expenditure are socio-economic status, nationality, age, occupation, income, length of stay, type of holiday accommodation, travel companion, destination loyalty and package type. In an alternative approach, other studies, such as Lew and Ng (2012) and Thrane (2014), have analysed the conditional quantile (median and other percentiles), rather than the conditional mean, using conditional quantile regression to measure tourist spending and providing additional insights into the question of micro-level tourism.[4] However, to our knowledge, no previous studies have been undertaken to analyse the relationship between expenditure at destination and total trip expenditure.

Models of tourist expenditure at a single destination

For simplicity, we distinguish between two types of expenditure, by a tourist visiting a single destination: expenditure at origin (pre-booked transport and accommodation, usually paid to travel agencies and tour operators) and expenditure at the destination, mainly for good and services, including restaurant meals, transport and recreation activities. From these differentiated expenditures, we propose a model to represent the tourist budget share.

Tourist budget share: Expenditure at destination in relation to total trip expenditure

Let E_1 and E_2 be the random variables associated with tourist expenditure at origin and at destination, regarding vacation plans in, for example, the Canary Islands.

Let $W_2 = E_2/(E_1 + E_2)$ be the tourist budget share at destination and let $W_1 = E_1/(E_1 + E_2)$ be the trip budget share at origin.

If we assume the two expenditure variables are independent[5] and follow a gamma distribution with parameters (a, c) and (b, c), $a > 0$, $b > 0$ and $c > 0$, respectively, then the variable $W_2 = E_2/(E_1 + E_2)$, with a support in the interval $[0, 1]$, follows a beta distribution with parameters a and b. When the variable under study has bounded support, the classic beta distribution is the best distribution with which to model this situation. Observe that although E_1 and E_2 are assumed to be independent, both types of expenditure, at origin and at destination, share the same parameter $c > 0$ since $E(E_1) = a/c$ and $E(E_2) = b/c$, where $E(\cdot)$ represents the expectation of the corresponding random variable. Thus, $E(E_2) = \frac{b}{a}E(E_1)$ and the expectation of expenditure at destination will be higher (lower) than that of expenditure at origin provided that $a < b$ ($a > b$). The latter assumption is corroborated in the empirical analysis section of this article.

Due to the reflection symmetry of the beta distribution, we have $f(w_2; a, b) = f(1 - w_2; b, a)$ and therefore information can be derived about tourist expenditure not only at the destination but also at the origin.

In order to study the behaviour of this variable, or to determine its mean value, the best approach is that proposed by Ferrari and Cribari-Neto (2004), who made use of a fractional response model to consider the following parameterization of the probability density function (pdf) of the classical beta continuous distribution

$$f(w_2; \mu, \phi) = \frac{{w_2}^{\mu\phi-1}(1-w_2)^{(1-\mu)\phi-1}}{B(\mu\phi, (1-\mu)\phi)}, \quad 0 < w_2 < 1 \tag{1}$$

where $0 < \mu < 1$ and $\phi > 0$, which provides a simple model into which covariates can be introduced, making use of the fact that $E(W_2) = \mu$. This parameterization of the classical beta distribution retains the flexibility of the pdf, which continues to adopt multiple forms according to the values of the two parameters on which it depends. Here, $B(m, n)$ is the Euler beta function given by $B(m, n) = \Gamma(m)\Gamma(n)/\Gamma(m+n)$, and $\Gamma(m)$ is the Euler gamma function given by $\Gamma(m) = \int_0^\infty t^{m-1}e^{-t}dt$.

Let us now assume that μ depends on k covariates x_i and a vector k of unknown regression coefficients $\beta = (\beta_1, \ldots, \beta_k)^T$ and then consider the logit link function

$$\mu(\boldsymbol{x}_i, \beta) = \frac{\exp(\boldsymbol{x}'_i\beta)}{1 + \exp(\boldsymbol{x}'_i\beta)} \in (0, 1) \tag{2}$$

Then, pdf (1) can be rewritten as

$$f(w_{2i}; \mu(\boldsymbol{x}_i, \beta), \phi) = \frac{\Gamma(\phi){w_{2i}}^{\mu(x_i,\beta)\phi-1}(1-w_{2i})^{(1-\mu(x_i,\beta))\phi-1}}{\Gamma(\mu(\boldsymbol{x}_i, \beta)\phi)\Gamma((1-\mu(\boldsymbol{x}_i, \beta))\phi)} \tag{3}$$

for $i = 1, \ldots, n$, where $0 < w_{2i} < 1$, $0 < \mu(\boldsymbol{x}_i, \beta) < 1$, $\phi > 0$. Using this parameterization, the conditional mean is $E(w_{2i}|\boldsymbol{x}_i) = \mu(\boldsymbol{x}_i, \beta)$ and the variance is given by

$$\mathrm{var}(w_{2i}|\boldsymbol{x}_i) = \frac{\mu(\boldsymbol{x}_i, \beta)(1-\mu(x_i, \beta))}{1+\phi} \tag{4}$$

and the cumulative distribution function is

$$F(w_{2i}; \mu(\boldsymbol{x}_i, \beta), \phi) = \frac{B(w_{2i}, \mu(\boldsymbol{x}_i, \beta)\phi, (1-\mu(\boldsymbol{x}_i, \beta))\phi)}{B(\mu(\boldsymbol{x}_i, \beta)\phi, (1-\mu(\boldsymbol{x}_i, \beta))\phi)} \tag{5}$$

where $B(z, m, n)$ is the incomplete beta function given by $B(z, m, n) = \int_0^z t^{m-1}(1-t)^{n-1}dt$.

ϕ is termed the precision parameter because, from equation (4) we have for fixed μ_i, the larger the value of ϕ the smaller the variance of w_{2i}. In this case, ϕ^{-1} is a dispersion parameter.

From equation (3), we obtain the log-likelihood of the model with covariates, which is given by

$$\begin{aligned}\ell_i(\mu(\boldsymbol{x}_i, \beta), \phi) = {} & \log(\Gamma(\phi)) - \log(\Gamma(\mu(\boldsymbol{x}_i, \beta)\phi)) - \log(\Gamma((1-\mu(\boldsymbol{x}_i, \beta))\phi)) \\ & + (\mu(\boldsymbol{x}_i, \beta)\phi - 1)\log w_{2i} \\ & + ((1-\mu(\boldsymbol{x}_i, \beta))\phi - 1)\log(1-w_{2i})\end{aligned}$$

The normal equations and Fisher's information matrix for this model are discussed in detail in Ferrari and Cribari-Neto (2004).

The beta regression model could be further developed to explicitly allow for covariates to model the scale parameter ϕ. That is likely to further improve the fit in real data applications with respect to standard linear models, allowing us to incorporate features such as heteroscedasticity or skewness. Formally introduced by Simas et al. (2010), this is termed the variable dispersion beta regression model.

In this model, the scale parameter is not constant for all observations but is modelled in a similar way to the mean parameter, depending on the covariates. This additional model can be specified by using equation (2) together with the usual log link for the scale parameter ϕ. That is, we consider

$$\phi(\boldsymbol{x}_i, \zeta) = \exp(x'_{\mathbf{i}}\zeta) \in (0, \infty)$$

for $i = 1, \ldots, n$, where $\zeta = (\zeta_1, \ldots, \zeta_r)'$ is a vector r of unknown regression coefficients, different from those considered for μ_i. Clearly, both μ and ϕ may be influenced by different characteristics and variables. For this reason, the explanatory variables used to model them may not be the same.

Marginal effects and interpretation of the coefficients. The marginal effect reflects the variation of the conditional mean of w_2 due to a one-unit change in the kth covariate.

It can be calculated as $\delta_{ik} = \frac{\partial \mu_i}{\partial x_{ik}} = \beta_k \mu_i (1 - \mu_i), i = 1, 2, \ldots, n$. Thus, the marginal effect indicates that a one-unit change in the kth regressor increases or decreases the expectation of the proportion of expenditure at destination for the ith tourist by δ_{ik} units, $i = 1, 2, \ldots, n$. This expression is similar to that for the logit marginal effect, but it is not constant as in the linear regression model, which can be directly represented by β_k.

To evaluate marginal effects for continuous exogenous variables, we can calculate expressions at the sample means of the data or evaluate the marginal effects at every observation and then calculate the sample average of the individual marginal effect (average partial effects).

On the one hand, for example, the response for an individual with average characteristics can be calculated as follows

$$\bar{\delta}_k \equiv \frac{\partial \mu_i}{\partial x_{ik}}\bigg|\bar{x} = \hat{\beta}_k \frac{\exp(\bar{\boldsymbol{x}}'\hat{\beta})}{(1 + \exp(\bar{\boldsymbol{x}}'\hat{\beta}))^2} = \hat{\beta}_k \bar{\mu}(1 - \bar{\mu})$$

where $\bar{\mu} = \exp(\bar{\boldsymbol{x}}'\hat{\beta})/(1 + \exp(\bar{\boldsymbol{x}}'\hat{\beta}))$ and $\hat{\beta}$ is the estimated vector of parameters.

On the other hand, average partial effects can be calculated by the expression, $\hat{\bar{\delta}} = \frac{1}{n}\sum_{i=1}^{n} \hat{\delta}_{ik}$, which aggregates all the individuals.

For indicator variables such as x_{ik} which takes only the value 0 or 1, the marginal effect is

$$\frac{E(w_{2i}|x_{ik} = 1, x_{i1}, \ldots, x_{i,k-1})}{E(w_{2i}|x_{ik} = 0, x_{i1}, \ldots, x_{i,k-1})} = \exp(\beta_k)$$

Therefore, the conditional mean is $\exp(\beta_k)$ times larger if the indicator variable is 1 rather than 0.

To compute the standard errors, we can use the delta method (or other robust methods). The estimator of the asymptotic covariance for the average partial effects can be defined by the expression

$$\mathrm{asy}(\mathrm{var}(\hat{\bar{\delta}})) = G(\hat{\beta})\hat{V}G'(\hat{\beta})$$

where $\hat{V}$ is the estimator of the asymptotic covariance matrix of $\hat{\beta}$ and $G(\hat{\beta})$ is the matrix of partial derivatives of marginal effects with respect to $\hat{\beta}$. These derivative matrices are averaged over the observations rather than being computed at the mean of the data.

Modelling the relation between the two types of expenditure

In this section, we develop a distribution function for the relation between expenditure at destination and at origin for n tourists. This distribution function is related to the growth rates of each type of expenditure.

Let $G_2 = (E_2 - E_1)/E_1$ be the growth rate of expenditure at destination in relation to expenditure at origin. G_2 can also be formulated as $G_2 = E_2/E_1 - 1$. By taking $Z_2 = E_2/E_1$ as the rate between expenditure at destination and at origin, we can statistically model the quotient of the two types of expenditure, in other words, the quotient of tourist budget share, using distribution functions. Therefore, $E_2/E_1 = W_2/W_1$, where $(W_1 = 1 - W_2)$. Note, too, that $Z_2 = 1 + G_2$, that is, the quotient is equal to 1 plus the growth rate between the two expenditures.

It is straightforward to confirm that if W_2 follows the pdf given in equation (1) then the random variable $Z_2 = W_2/(W_1)$ has a pdf that is given by

$$g(z_2; \mu, \phi) = \frac{z_2^{\mu\phi-1}}{B(\mu\phi, (1-\mu)\phi)(1+z_2)^{\phi}}, \quad z_2 > 0 \tag{6}$$

that is, the beta prime distribution, which is also termed a beta distribution of the second kind. The relation between expenditure at destination and at origin, $Z_2 = W_2/(W_1)$, then follows the pdf given in equation (6), with a mean that is given by

$$E\left(\frac{W_2}{W_1}\right) = \frac{\mu\phi}{\phi(1-\mu) - 1}$$

It can readily be shown that

$$\frac{dE(W_2)}{dE(W_1)} = \frac{dE(W_2)}{d\mu}\frac{d\mu}{dE(W_1)} = -1 \tag{7}$$

where d represents the derivative. This is consistent with the fact that the greater the expectation of expenditure at destination, the lower the expenditure at origin. On the other hand, we have

$$\frac{\partial}{\partial\mu}E\left(\frac{W_2}{W_1}\right) = \frac{\phi(\phi-1)}{[\phi(1-\mu)-1]^2} = \frac{\phi-1}{\phi\mu^2}E^2\left(\frac{W_2}{W_1}\right)$$

which is positive for $\phi > 1$, negative for $\phi < 1$ and constant for $\phi = 1$. That is, for $\phi > 1$, if the average expenditure at destination increases then the rate of expenditure between origin and destination also increases, which seems evident. However, for values of $\phi < 1$, this rate decreases as the average expenditure at destination increases.

The pdf given in equation (6) reaches the maximum (the mode) at the point $z_{20} = (\mu\phi - 1)/(1 + (1-\mu)\phi)$. Then, the probability of exchanging expenditure at destination for expenditure at origin increases for values less than z_0, while it decreases for values greater than z_0.

As equation (7) is constant, it is preferable to use the geometric mean since it has been shown (see for instance Johnson et al. (1995)) that for this statistic the following relation is sustained

$$G\left(\frac{W_2}{W_1}\right) = \frac{G(W_2)}{G(W_1)} \tag{8}$$

where G represents the geometric mean. If $Y \sim B(\mu\phi, (1-\mu)\phi)$, that is, $W_1 \sim B((1-\mu)\phi, \mu\phi)$, then from equation (8) we obtain

$$G\left(\frac{W_2}{W_1}\right) = \frac{\exp[\psi(\mu\phi) - \psi(\phi)]}{\exp[\psi((1-\mu)\phi) - \psi(\phi)]} = \exp[\psi(\mu\phi) - \psi((1-\mu)\phi)]$$

where $\psi(\cdot)$ represents the digamma function (the derivative of the logarithm of the gamma function). It follows straightforwardly that

$$\frac{dG(W_2)}{dG(W_1)} = -\frac{\psi_1(\mu\phi)}{\psi_1((1-\mu)\phi)} \exp[\psi(\mu\phi) - \psi((1-\mu)\phi)]$$

where $\psi_1(\cdot)$ is the derivative of the digamma function.

Empirical analysis

The data for this analysis were obtained from the Tourist Expenditure Survey conducted by the Canary Islands regional government during the period January 2009 to March 2012. This survey is based on personal interviews of tourists who visited any of the Canary Islands (Gran Canaria, Fuerteventura, Lanzarote, Tenerife, La Palma, La Gomera and El Hierro) at the time of departure. The data collection was carried out by the Canary Islands Institute of Statistics (ISTAC, Spanish initials).

The survey obtains quarterly information about the total expenditure by tourists who arrive by air, either on a package holiday or travelling independently.[6]

Specifically, the present analysis includes tourists from Germany and the United Kingdom, staying at least one night and no more than 30 consecutive nights.[7] This population represents over 47% of total tourist visits to the Canary Islands during the study period. After data cleansing, 10,425 and 9805 pooled observations remained, for German and British tourists, respectively.

The study variables from the ISTAC Survey that were included in our analysis were length of stay, reason for coming to the Canary Islands, number of previous visits, vacation characteristics and personal characteristics.

The following variables finally included in our analysis are in line with those used in many previous studies in this field, cited in the second section.

1. Tourist budget data, obtained by ISTAC: (a) expenditure at origin (€), that is, flights and accommodation. No distinction is made between types of accommodation (i.e. full board, half board or only accommodation); (b) expenditure at destination (€), for example, restaurant meals, transport and recreation.
2. Length of stay (days).
3. Repetition. The total number of times the tourist has visited the Canary Islands before the present trip. A value of 0 is possible, indicating no previous visits.

Table 2. Descriptive statistics for all tourists, filtered database.

	German		British	
Variables	Mean	Standard deviation	Mean	Standard deviation
Expenditure share (%)	26.5	15.32	34.43	17.68
Length of stay	10.25	4.60	8.50	3.38
Household income (categorical variable)	3.58	1.95	3.41	1.89
Holiday (%)	96.15	–	96.46	–
Sun-and-beach (%)	84.69	–	87.45	–
Booking in advance (%)	62.02	–	65.69	–
Low-cost carrier (%)	17.72	–	48.16	–
Travel party size	2.07	0.87	2.60	1.19
Repetition	2.69	2.80	3.34	2.90
Accommodation in 4–5 star hotels (%)	50.00	–	36.96	–
Accommodation in less than 3-star hotels (%)	20.40	–	12.90	–
Age of the respondent	43.28	13.71	42.92	13.45
Number of tourists after data cleansing	10,425		9805	

4. Reason for visiting the Canary Islands. Two variables are considered: (a) sun-and-beach. This is a dummy variable which takes the value 1 if the main reason for visiting is sun-and-beach, and 0 otherwise; (b) the main reason for coming. This is a dummy variable which takes the value 1 when the tourist arrives on holiday, and 0 otherwise.
5. Personal characteristics. Two variables are considered: (a) household income, an ordered categorical variable which takes the following values: (1) income €12,000–24,000, (2) income €24,001–36,000, (3) income €36,001–48,000, (4) income €48,001–60,000, (5) €60,001–72,000, (6) income €72,001–84,000, (7) income greater than €84,000; (b) the respondents' age (years).
6. Vacation characteristics. (a) Type of accommodation, expressed via a dummy variable which takes the value 1 if the accommodation is a 4–5 star hotel/aparthotel, and 0 otherwise, and by another dummy variable which takes the value 1 if the accommodation is a hotel/aparthotel with fewer than 3 stars, and 0 otherwise. In both cases, the reference category is accommodation consisting of extra-hotels, own home, friends or family, or other. (b) Size of travel party, that is, the number of persons in the holiday package paid for at origin. (c) Booking in advance. This is a dummy variable which takes the value 1 if the tourist made the holiday booking, independently of tour operators, in advance.
7. Finally, we controlled for quarterly seasonality, including dummy variables for summer (June–September), autumn (October–December) and winter (January–March).

Table 2 shows descriptive statistics for the dependent and explanatory variables associated with the filtered database. In brief, expenditure share at destination was slightly higher for British than for German tourists. The main reason for the trip was to take a holiday. These tourists preferred a sun-and-beach holiday and the family size was two persons. On average, the tourists had visited on three previous occasions, and their household income was €36,000–48,000. Half of the German tourists stayed in 4- or 5-star hotels. The tourists' average age was 43 years and the majority booked via a tour operator. Almost half of the British tourists travelled by a LCC.

Estimation results

Model estimates for budget share and quotient between expenditures without covariates. Table 2 shows the initial estimation results (not including covariates), obtained using Mathematica statistical software (see Wolfram (2003) and Ruskeepaa (2009); among others). Taking into account that $a = \theta\phi$ and $b = (1-\theta)\phi$, for the sample of German tourists these values are given by $a = 2.02747$, $b = 5.5442$, so that $a/b = 0.365689$, close to the empirical value 0.360654 given by the quotient between average expenditure at destination and average expenditure at origin. For the British tourists, the resulting values of a and b are $a = 2.02554$ and $b = 3.89146$, with $a/b = 0.52051$, which are also close to the empirical value given by 0.5251. Accordingly, the hypothesis that E_1 and E_2 follow initially independent gamma distributions does not seem inconsistent with the results obtained.

Figure 1 shows graphs (histogram and theoretical densities, QQ plot and PP plot) which reflect the suitability of the beta distribution for modelling this type of data.

Figure 2 shows the empirical distribution of the quotient between expenditure at destination and at origin (Z_2), and the pdf given in equation (6) for estimated values shown in Table 3.

The slope of the distribution of the quotient between the two expenditures is positive until values close to 0.20 are reached for Germans and 0.25 for Britons (the modes of the two pdfs). After these points, the slope of the distribution of the quotient becomes negative, meaning that these tourists spent more at destination than those with a quotient of less than 0.2 and 0.25, respectively. The growth rate is -0.8 and -0.75. In other words, German tourists spent 80% of their expenditure at origin, while the figure for their British counterparts was 75%. The quotient was above 0.75 for 25% of each population of tourists.

Budget share estimates with covariates. The model with covariates was analysed using the R package. Table 3 shows the results obtained for OLS coefficients for the linear regression model and maximum likelihood (ML) estimates for the non-linear beta model with covariates. In addition, two samples are distinguished: in- and out-of-sample, respectively. The first arbitrarily considers 90% of the total observations for each nationality. In this sample, we compare the models estimated by OLS and ML using the non-nested Vuong's test. The second comprises the remaining 10% of total observations and allows us to evaluate model forecasts using both the EMA and the MSE statistics.

It is important to note that the coefficients of linear regression (OLS) and the beta model (ML) cannot be compared directly. Therefore, we include the marginal effects and their standard errors obtained by the delta method.[8] In each case, the results for German and British tourists are shown separately. Before disaggregating the data, we also performed a poolability test to evaluate the subsample stability of the coefficients estimated. If two subsamples are poolable, then the pooled analysis could also be reported. Accordingly, we performed this poolability test, in which the null hypothesis of poolability assumes homogeneous slope coefficients. An F test was then applied to test for poolability across our two cross sections. The results obtained lead us to reject the null hypothesis of poolability, because $F = 595.35$ (p value $= 0.00$). Therefore, the two subsamples are not combined in Table 4.[9] All models include quarterly dummies.

In general, the marginal effects obtained for the in-sample results are statistically significant. The ML result is higher for the beta estimates than for the OLS estimates, both for the German and for the British tourists. Vuong's test to compare non-nested models shows that the beta estimates are preferred to the OLS estimates. Accordingly, the beta model was mainly used to interpret these

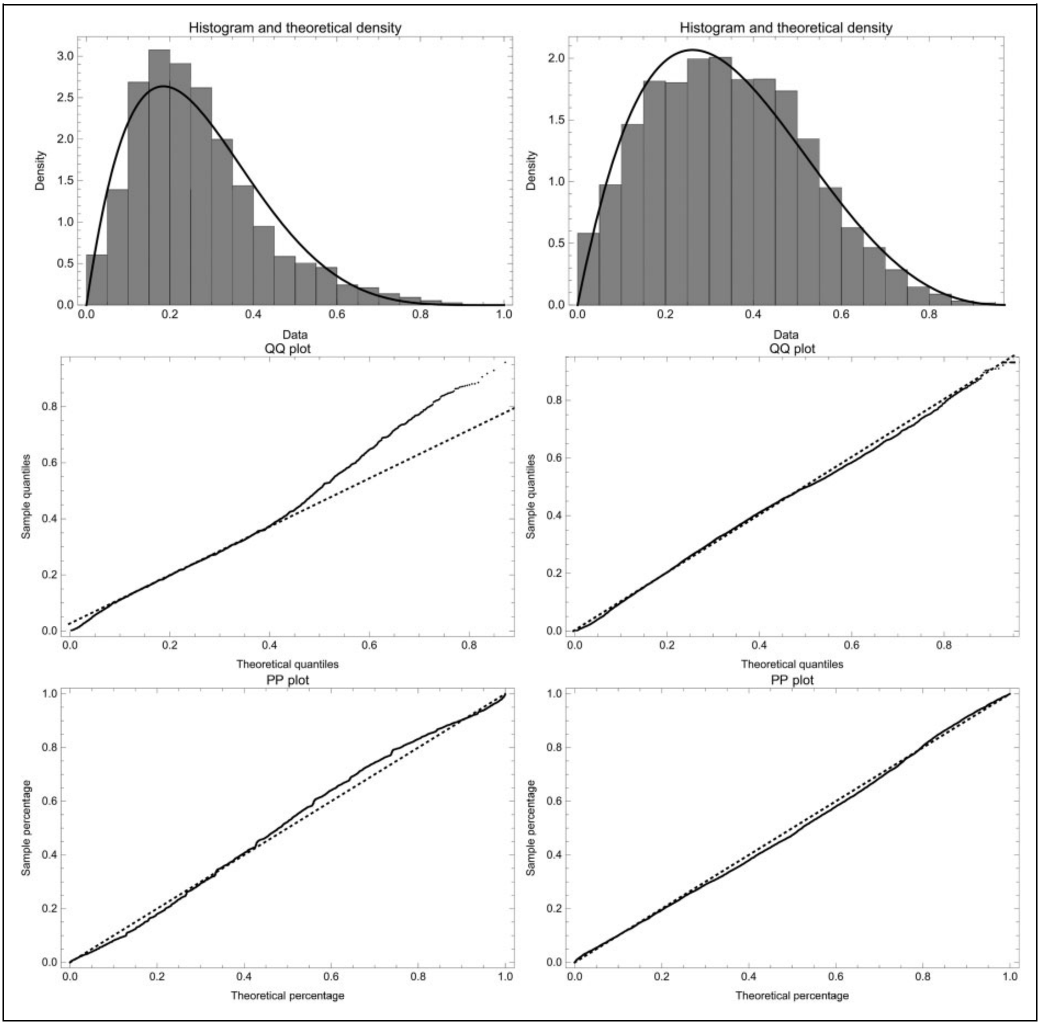

Figure 1. Diagnostic plots for tourist budget share at destination (W_2) data without including covariates. Data for Germans (left) and Britons (right).

coefficients, although it should be noted that the EMA and MSE statistics for the beta models were larger than the OLS values in the out-of-sample forecasts, reflecting the poorer forecasting capability of the beta model.

To our knowledge, no previous research findings in this respect have been reported, and so we were unable to compare our findings with those of other researchers. Nevertheless, we include below a discussion of papers in which the total aggregate or daily expenditure has been analysed.

In our study, the income effect is always positive and statistically significant for German and British tourists. Therefore, the higher the tourist income, the higher the proportion of expenditure at destination. In general, this finding is consistent with previous results, according to which higher household income is positively associated with greater spending (Aguiló et al., 2017; Brida et al.,

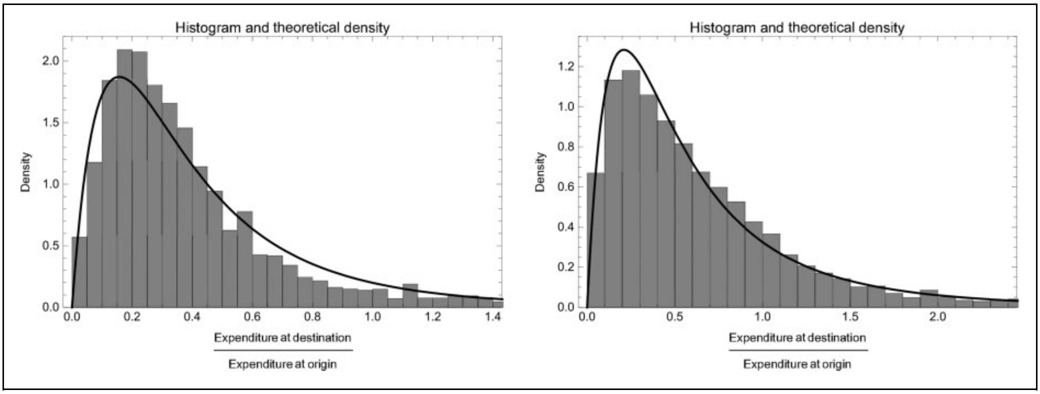

Figure 2. Fitted and empirical quotient between expenditure at destination and at origin for German (left) and British tourists (right).

Table 3. Estimated parameter, standard error (in parentheses) and value of the log-likelihood function in the maximum for the two populations of tourists without covariates using the beta distribution.

	Germans	Britons
$\hat{\theta}$	0.268	0.342
	(0.001)	(0.002)
$\hat{\phi}$	7.571	5.917
	(0.101)	(0.079)
$\ell_{\max}$	5663.62	3400.61

2013; Marrocu et al., 2015). For example, García-Sánchez et al. (2013) reported that high-income tourists spend 50% more than those with low incomes.

With respect to personal characteristics, the following results were obtained. The coefficient of age (in logs) is negative and statistically significant for German and British tourists, suggesting that older tourists tend to spend less than younger ones. In another study, García-Sánchez et al. (2013) suggested there might be an inverted-U relationship between tourists' age and their daily spending, that is, that young and old tourists spend less than middle-aged ones (Craggs and Schpfield, 2009). However, Brida and Scuderi (2012), in a review of studies of the determinants of tourist expenditure, examining 86 papers and 354 estimates from econometric models based on individualized data, found no conclusive effects of age on expenditure (thus, of the 87 significant metric indicators, 61 were directly related whereas 26 were inversely related, and 103 were not significant).

With respect to the vacation characteristics, the following results were obtained. The coefficient of the variable for previous visits to the Canary Islands (repetition) is positive and statistically significant only for the British tourists. Therefore, with repeated visits British tourists increase their share of expenditure at destination. Opinions are divided in prior studies as to the possible effects of repeated visits and tourist spending. García-Sánchez et al. (2013) obtained negative values, suggesting that repeated visits were associated with less expenditure (daily expenditure decreased by 2.5%). However, Chang et al. (2013) examined the spending of first-time and repeat visitors and found that previous travel experience did not generate significant differences in consumption patterns. This behaviour might be explained by the degree of satisfaction experienced, that is, tourists will repeat their holiday destination if their expectations are met and they are satisfied (Disegna and Osti, 2016).

The coefficient for length of stay (in logs) is positive and statistically significant. Thus, the more nights spent at the destination, the greater the tourist budget share. Aguiló et al. (2017) derived a model of tourists' daily expenditure that included the length of stay as an independent variable. These authors recorded a significant positive relationship between tourist spending and length of stay, which they explained as being due to the fact that high-income tourists not only present high levels of daily expenditure but also stay for longer periods. In a recent paper, Gómez-Déniz and Pérez-Rodríguez (2018) propose a model based on a positive correlation between aggregate expenditure at origin plus destination and length of stay. However, Brida and Scuderi (2012) reviewed 188 estimations and found that 98 reported a significant positive association between length of stay and expenditure. On the other hand, 45 estimations reported a negative association in this respect. That is, expenditure decreases when length of stay increases (see, for example, Thrane and Farstad (2011)).

The coefficient for the reason for visiting the Canary Islands was found to be statistically significant and positive, but for the sun-and-beach motivation it was negative. These results suggest that when the reason for the trip is other than a sun-and-beach holiday, there is a greater predisposition to spend at the destination. In this respect, García-Sánchez et al. (2013) indicate that daily tourist expenditure at the destination is greater when tourists perform non-beach-related activities.

The coefficient for accommodation, usually the most important expense at origin, was negative in all cases. In other words, the greater the cost of accommodation (i.e. the higher the rating (number of stars) of the hotel), the lower the share of tourist expenditure at destination with respect to the use of extra-category hotels, own home or staying with friends or relatives (the reference category). Previous studies of daily tourist expenditure have reported finding a positive relationship between accommodation and daily expenditure (Laesser and Crouch, 2006; García-Sánchez et al., 2013).

For German and British tourists, group size is inversely related to tourist budget share at destination. Therefore, travelling as a group reduces expenses at destination and, therefore, the tourist budget share. In this respect, García-Sánchez et al. (2013) estimated that the per capita expense of travelling alone is 45% greater than when travelling in a group of four people. Similarly, Wang and Davidson (2010) found that group size was inversely associated with total expenditure per person.

For German tourists, a significant negative coefficient was obtained for the variable booking in advance, but this association was positive for the British visitors. In this respect, it has been reported that customers who book their accommodation closer to the date of departure are usually willing to pay more (Schwartz, 2010). Moreover, when customers book in advance, they can often obtain a slight reduction in cost, whilst maintaining their preferences. Strikingly, the price of accommodation decreased if the booking was made 10 days prior to departure (Maseiro et al., 2015). In this respect, our results show that the tourist budget share only increased for the British tourists. This result shows that these tourists have more funds available for spending at their destination.

Finally, the coefficient for LCC is positive and significant for both German and British tourists. Thus, when tourists obtain cheaper flights, the proportion of their spending at the destination increases, compared to spending at origin. As explained by Ferrer-Rosell et al. (2015), the reduced price of transport provided by the emergence of low-cost airlines has redistributed tourists' budgets

Table 4. Marginal effects and p values for OLS and using the beta regression model for German and British tourists.

	German				British			
	OLS		Beta		OLS		Beta	
	Marginal effect	p Value	Marginal effect	p Value	Marginal effect	p Value	Marginal effect	p Value
In-sample								
Income	0.0018	0.0134	0.0010	0.1082	0.0023	0.0035	0.0019	0.0139
Log (age)	−0.0217	0.0000	−0.0190	0.0000	−0.0372	0.0000	−0.0360	0.0000
Repetition	−4.7E–5	0.9317	−0.0004	0.3159	0.0031	0.0000	0.0027	0.0000
Log (length of stay)	0.0457	0.0000	0.0484	0.0000	0.0261	0.0000	0.0278	0.0000
Holiday	0.0081	0.0413	0.0076	0.0259	0.0252	0.0000	0.0245	0.0000
Sun-and-beach	−0.0133	0.0607	−0.0116	0.0653	−0.0181	0.0268	−0.0178	0.0177
Accommodation, 4- or 5-star hotel	−0.1147	0.0000	−0.1052	0.0000	−0.1752	0.0000	−0.1726	0.0000
Accommodation, 3-star hotel	−0.0972	0.0000	−0.0863	0.0000	−0.0991	0.0000	−0.0953	0.0000
Booking in advance	−0.0146	0.0000	−0.0148	0.0000	0.0039	0.2215	0.0038	0.2205
Low-cost carrier	0.0421	0.0000	0.0379	0.0000	0.0460	0.0000	0.0452	0.0000
Travel party size	−0.0318	0.0000	−0.0331	0.0000	−0.0233	0.0000	−0.0249	0.0000
Winter	−0.0068	0.0848	−0.0088	0.0112	−0.0162	0.0000	−0.0172	0.0000
Summer	−0.0181	1.8E–5	−0.0162	0.0000	−0.0250	0.0000	−0.0244	0.0000
Autumn	−0.0162	0.0000	−0.0148	0.0000	−0.0036	0.4382	−0.0048	0.2804
Constant	0.3979	0.0000			0.5306	0.4382		
$\hat{\phi}$			9.2424	0.0000			8.3631	0.0000
$\ell_{\max}$	5795.4817		6669.8059		4755.3302		5066.4773	
Vuong's test			−30.8928				−12.9201	
p Value			[0.00]				[0.00]	
Observations	9382		9382		7494		7494	
Out-of-sample								
MAE	0.1053		0.7415		0.1329		0.6456	
MSE	0.0179		0.5710		0.0283		0.4489	
Observations	1043		1043		981		981	

OLS: ordinary least squares; MAE: Mean absolute error; MSE: Mean squared error.

between spending at origin and at destination. Indeed, if these tourists fly low-cost and do not book an all-inclusive holiday, then their spending at destination will be even higher (see Eugenio-Martin and Inchausti-Sintes (2016)).

Following our analysis of the beta model, see Table 4, we now evaluate one assuming that the precision parameter depends on certain regressors (in this case, for simplicity, the same as the mean parameter). This model also incorporates a link function.[10] Table 5 shows the estimates obtained for the total sample. In this case, also for simplicity, the results shown are based on coefficients, rather than marginal effects. The precision parameter, which could be related to features such as heteroscedasticity or the skewness of the beta model, can also be explained by reference to a set of regressors. Many of the coefficients are statistically significant at 5% and,

Table 5. Parameters and p values for maximum likelihood fitted to the data using the beta regression model for German and British tourists.

	Germans				Britons			
	Mean, μ		Precision, ϕ		Mean, μ		Precision, ϕ	
	Coefficient	p Value	Coefficient	p Value	Coefficient	p Value	Coefficient	p Value
Income	0.0122	0.0000	−0.0385	0.0000	0.0097	0.0000	−0.0336	0.0000
Log (age)	−0.1025	0.0000	0.0702	0.0000	−0.1667	0.0000	−0.0266	0.0000
Repetition	−0.0018	0.2963	−0.0134	0.0000	0.0128	0.0000	−0.0033	0.2008
Log (length of stay)	0.2384	0.0000	−0.0281	0.0000	0.1341	0.0000	0.0487	0.0000
Holiday	0.0368	0.0000	−0.0135	0.2419	0.1189	0.0000	−0.0071	0.5584
Sun-and-beach	−0.0591	0.0000	−0.0092	0.3976	−0.0798	0.0000	−0.0429	0.0002
Accommodation, 4- or 5-star hotel	−0.5665	0.0000	0.2115	0.0000	−0.8024	0.0000	−0.1404	0.0000
Accommodation, 3-star hotel	−0.4753	0.0000	0.0553	0.0179	−0.4335	0.0000	−0.3219	0.0000
Booking in advance	−0.0803	0.0000	0.0345	0.0116	0.0167	0.0418	0.0304	0.0292
Low-cost carrier	0.2132	0.0000	−0.2317	0.0000	0.1991	0.0000	−0.0967	0.0000
Travel party size	−0.1733	0.0000	0.0200	0.0000	−0.1109	0.0000	−0.0287	0.0000
Winter	−0.0340	0.0069	−0.1082	0.0000	−0.0741	0.0000	−0.0696	0.0011
Summer	−0.0996	0.0000	0.0993	0.0000	−0.1196	0.0000	0.1059	0.0000
Autumn	−0.0873	0.0000	0.0310	0.1568	−0.0195	0.2327	−0.0710	0.0073
Constant	−0.3947	0.0000	2.1040	0.0000	0.1523	0.0000	2.5060	0.0000
ℓ_{max}	6768.7783				5135.9578			
Observations	10,425				9805			

although the sample size is slightly different, the signs of the coefficients in Table 5 are similar to those in Table 4 without including covariates in the precision parameter.

The precision parameter should be interpreted with caution. The larger the value of this precision, the smaller the variance of w_{2i}, provided that μ_i is taken as fixed. This is a question that merits further investigation in future research.

Discussion and conclusions

This article describes a study conducted to model tourist budget share for expenditure at destination and at origin, by tourists visiting the Canary Islands. Analysis of the results obtained reveals the determinants of the allocation of expenditure.

This model was constructed taking into account the statistical nature of the two types of expenditure, considering dependence, that is, the fact that the expenditure categories are interdependent, and also their non-negative right-skewed distribution. This approach is important, because it enabled us to model the expenditure share on a statistical basis based on a variable bounded between 0 and 1, from a microeconomic perspective.

Another important feature of this study is that expenditure share was modelled allowing for covariates, both in the mean and in the precision or scale parameter. In consequence, we were able to analyse the marginal effects of diverse economic and socio-demographic factors, and thus account for expenditure share at destination.

In our analysis of the study data (obtained by the Canary Islands Tourism Expenditure Survey), the expenditure incurred in the country of origin (accommodation and flights) was separated from that at the destination (restaurant meals, transport, recreation, etc.).

The main conclusions drawn from this study are that the beta model provides a closer fit to the data than the OLS estimates, and that the determinants of tourist budget share at the destination are in line with those reported in previous research into aggregate or daily tourist expenditure and the causality between the variables.

Given the multiplier effects of tourist spending on the local economy, it is important to know how this expenditure is shared between destination and origin in order to develop policies encouraging tourists to maximize the share of their budget allocated to local spending.

Policymakers in the Canary Islands seeking to increase foreign tourists' expenditure at destination should take into account that accommodation (the main payment made in the country of origin), the size of the travel party and the age of the individuals are all inversely associated with the tourist budget share at destination. Furthermore, the Canary Islands is characterized as a 'sun-and-beach' destination, and our results show that this form of tourism has a negative relationship with the distribution of the tourist budget share. Accordingly, in order to foster the arrival of tourists with a higher proportion of expenditure at destination than at origin, policies should be adopted to promote activities representing alternatives to the traditional sun-and-beach formula.

On the other hand, tourists' income and the use of LCCs are both positively associated with the tourist budget share at destination. The impact of low-cost airlines is still recent and in future studies it would be interesting to examine the substitution effect between expenditure at destination and at origin as a result of access to lower priced flights. According to Eugenio-Martin and Inchausti-Sintes (2016), 'Low-cost travel savings from tourists' place of origin are transferred, at least partially, to higher tourism expenditures at their destination'. In other words, in a micro-economic demand model, a decrease in the cost of air travel could be interpreted as a real increase in the budget and hence greater licence for spending at the destination.

Finally, in the distribution of the quotient between expenditure at destination and at origin, it is noteworthy that although the most frequent distributions are 20% for German tourists and 25% for British ones, as this quotient increases, so does the preference for spending at the destination. Future research should be undertaken to investigate this question in greater detail.

Acknowledgements

The authors would like to express their gratitude to the Referees for their relevant and useful comments.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Emilio Gómez-Déniz was partially funded by grant ECO2013–47092 (Ministerio de Economía y Competitividad, Spain) and ECO2017–85577–P (Ministerio de Economía, Industria y Competitividad. Agencia Estatal de Investigación).

ORCID iD

Jorge V Pérez-Rodríguez https://orcid.org/0000-0002-6738-9191
José Boza-Chirino https://orcid.org/0000-0002-8494-8864

Notes

1. Trip expenditure can be disaggregated into various categories (such as restaurant meals, local transport and leisure) to examine the determinants of disaggregated tourist expenditure (see Lee (2001) and Wang et al. (2006)).
2. The tourist budget share is not the budget share of an individual or household on all goods and services formed by the quotient between real expenditure and real income. However, both types of budget share, whether for all goods or exclusively for tourist goods, are non-negative real numbers that are bounded between 0 and 1.
3. Fractional response based on the beta regression model proposed by Ferrari and Cribari-Neto (2004) is currently used in many fields, perhaps because these authors have supplied a package based on R project which facilities the use of this interesting model.
4. Thrane (2014) gave practical advice on how to improve the modelling of micro-level tourism expenditure, recommending quantile regression for this purpose.
5. Although this assumption may appear unrealistic, empirical studies have shown that the correlation between these two random variables is below 0.25.
6. It should be noted that a problem of sample selection might arise, as reliance on data for individuals travelling by air is likely to cause sample selection bias, endogeneity and possibly inconsistent estimates.
7. This truncation was performed because according to the empirical distribution of the length of stay, the 99th percentile is equal to 29 nights. Therefore, the remaining 1% of data were truncated to avoid problems of interpretation, as a major characteristic of the Canary Islands as a tourist destination is that blocks of holidays are commonly for 7 and 14 days, and therefore 1% of the values will be for stays of 30–180 nights.
8. The marginal effects were calculated using the R package *mfx*, which allows us, via the *betamfx* function, to use the delta method to obtain the marginal effects associated with the maximum likelihood estimators, as well as the standard errors. The argument $Robust = TRUE$ reports white/robust standard errors. Marginal effects are also accessible in STATA.
9. This poolability analysis is not detailed here, but is available on request from the authors.
10. This new regression model can be computed using the R package provided in Grün et al. (2012).

References

Aguiló E and Juaneda C (2000) Tourist expenditure for mass tourism markets. *Annals of Tourism Research* 27(3): 624–637.

Aguiló E, Rosselló J and Vila M (2017) Length of stay and daily tourist expenditure: A joint analysis. *Tourism Management Perspectives* 21: 10–17.

Alegre J, Mateo S and Pou L (2013) Tourism participation and expenditure by Spanish households: The effects of the economic crisis and unemployment. *Tourism Management* 39: 37–49.

Barquet A, Brida J, Osti L, et al. (2011) An analysis of tourists' expenditure on winter sports events through the Tobit censored model. *Tourism Economics* 17(6): 1197–1217.

Brida J, Disegna M and Osti L (2013) Visitors' expenditure behaviour at cultural events: The case of Christmas markets. *Tourism Economics* 19(5): 1173–1196.

Brida J and Scuderi R (2012) Determinants of tourist expenditure: A review of microeconometric models. *Tourism Management Perspectives* 6: 28–40.

Cárdenas P, Pulido J and Pulido M (2015) The influence of tourist satisfaction on tourism expenditure in emerging urban cultural destinations. *Journal of Travel & Tourism Marketing* 33(497): 497–512.
Chang K, Chen C and Meyer T (2013) A comparison study of travel expenditure and consumption choices between first-time and repeat visitors. *Tourism Management* 35: 275–277.
Craggs R and Schpfield P (2009) Expenditure-based segmentation and visitor profiling at The Quays in Salford, UK. *Tourism Economics* 15(1): 243–260.
Deaton A and Muellbauer J (1980) An almost ideal demand system. *American Economic Review* 70: 312–326.
Disegna M and Osti L (2016) Tourists' expenditure behaviour: The influence of satisfaction and the dependence of spending categories. *Tourism Economics* 22(1): 5–30.
Dobruszkes F (2013) The geography of European low-cost airline networks: A contemporary analysis. *Journal of Transport Geography* 28: 75–88.
Eugenio-Martin J and Inchausti-Sintes F (2016) Low-cost travel and tourism expenditures. *Annals of Tourism Research* 57: 140–159.
Ferrari S and Cribari-Neto F (2004) Beta regression for modelling rates and proportions. *Journal of Applied Statistics* 31(7): 799–815.
Ferrer-Rosell B and Coenders G (2017) Airline type and tourist expenditure: Are full service and low cost carriers converging or diverging? *Journal of Air Transport Management* 63: 119–125.
Ferrer-Rosell B, Coenders G and Martńez E (2015) Determinants in tourist expenditure composition. The role of airline types. *Tourism Economics* 21(1): 9–32.
Fredman P (2008) Determinants of visitor expenditures in mountain tourism. *Tourism Economics* 14(2): 297–311.
García-Sánchez A, Fernández-Rubio E and Collado M (2013) Daily expenses of foreign tourists, length of stay and activities: Evidence from Spain. *Tourism Economics* 19(3): 613–630.
Gómez-Déniz E and Pérez-Rodríguez J (2018) Modelling distribution of aggregate expenditure on tourism. *Economic Modelling* (in press). Available at: https://doi.org/10.1016/j.econmod.2018.09.027
Gómez-Déniz E, Sordo M and Calderín-Ojeda E (2013) The Log–Lindley distribution as an alternative to the beta regression model with applications in insurance. *Insurance: Mathematics and Economics* 54: 49–57.
Grün B, Kosmidis I and Zeileis A (2012) Extended beta regression in R: Shaken, stirred, mixed, and partitioned. *Journal of Statistical Software* 48(11): 1–25.
Johnson N, Kotz S and Balakrishnan N (1995) *Continuous Univariate Distributions*. New York: Wiley.
Kim WG, Kim TT, Gazzoli G, et al. (2011) Factors affecting the travel expenditure of visitors to Macau, China. *Tourism Economics* 17(4): 857–883.
Laesser C and Crouch G (2006) Segmenting markets by travel expenditure patterns: The case of international visitors to Australia. *Journal of Travel Research* 44(4): 397–406.
Lee C (2001) Determinants of recreational boater expenditures on trips. *Tourism Management* 22(6): 659–667.
Leones J, Codby B and Crandall K (1992) Tracking expenditures of the elusive nature tourists of Southeastern Arizona. *Journal of Travel Research* 36(3): 56–64.
Lew A and Ng P (2012) Using quantile regression to understand visitor spending. *Journal of Travel Research* 51(3): 278–288.
Marcussen C (2011) Determinants of tourist spending in cross-sectional studies and at Danish destinations. *Tourism Economics* 17: 833–855.
Marrocu E, Paci R and Zara A (2015) Micro-economic determinants of tourist expenditure: A quantile regression approach. *Tourism Management* 50: 13–30.
Maseiro L, Nicolau J and Law R (2015) A demand-driven analysis of tourist accommodation price: A quantile regression of room bookings. *Journal of Travel Research* 50: 1–8.
Papke L and Wooldridge J (1996) Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of Applied Econometrics* 11(6): 619–632.

Papke L and Wooldridge J (2008) Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics* 145(1-2): 121–133.
Pérez-Rodríguez J and Gómez-Déniz E (2015) Spread component costs and stock trading characteristics in the Spanish stock exchange. Two flexible fractional response models. *Quantitative Finance* 15(12): 1943–1962.
Pyo S, Uysal M and McLellan R (1991) A linear expenditure model for tourism demand. *Annals of Tourism Research* 18: 443–454.
Ruskeepaa H (2009) *Mathematica Navigator. Mathematics, Statistics, and Graphics*, 3rd ed. Cambridge: Academic Press.
Schwartz Z (2010) Changes in hotel guests' willingness to pay as the date of stay draws closer. *Journal of Hospitality & Tourism Research* 24(2): 180–198.
Simas A, Barreto-Souza W and Rocha A (2010) Improved estimators for a general class of beta regression models. *Computational Statistics & Data Analysis* 54(2): 348–366.
Syriopoulos T and Sinclair M (1993) An econometric study of tourism demand: The AIDS model of US and European tourism in Mediterranean countries. *Applied Economics* 25: 1541–1552.
Thrane C (2014) Modelling micro-level tourism expenditure: Recommendations on the choice of independent variables, functional form and estimation technique. *Tourism Economics* 20(1): 51–60.
Thrane C and Farstad F (2011) Domestic tourism expenditures: The non-linear effects of length of stay and travel party size. *Tourism Management* 32(1): 46–52.
Tribe J (2005) *The Economics of Recreation, Leisure & Tourism*, 3rd ed. Amsterdam: Elsevier.
Wang D and Davidson M (2010) A review of micro-analyses of tourist expenditure. *Current Issues in Tourism* 13(5): 507–524.
Wang Y, Rompf P, Severt D, et al. (2006) Examining and identifying the determinants of travel expenditure patterns. *International Journal of Tourism Research* 8(5): 333–346.
Wolfram S (2003) *The Mathematica Book*. Champaign: Wolfram Media.
Wu L, Zhang J and Fujiwara A (2013) Tourism participation and expenditure behaviour: Analysis using a scobit based discrete-continuous choice model. *Annals of Tourism Research* 40: 1–7.
Zheng B and Zhang Y (2013) Household expenditures for leisure tourism in the USA, 1996 and 2006. *International Journal of Tourism Research* 15(2): 197–208.

Author biographies

Emilio Gómez-Déniz is Professor of Mathematics at University of Las Palmas de Gran Canaria (Spain). His research focuses on the distributions theory, Bayesian statistics, robustness, Bayesian applications in economics with emphasis in actuarial statistics (credibility, ruin theory, etc.). He has widely published in peer-reviewed national and international journals on these topics, such as *Insurance: Mathematics and Economics*, *Communications in Statistics*, *ASTIN Bulletin*, *Quantitative Finance*, *Journal of Productivity Analysis*, and *Physica A*, among others.

Jorge V Pérez-Rodríguez is Professor of Econometrics at University of Las Palmas de Gran Canaria (Spain). His research focuses on the financial statistics and econometrics and tourism economics. He has widely published in peer-reviewed national and international journals on these topics, such as *Empirical Finance*, *Quantitative Finance*, *Journal of Productivity Analysis*, *Tourism Management* and *Tourism Economics*, among others.

José Boza-Chirino is Professor of Statistics at the University of Las Palmas de Gran Canaria (Spain). His research focuses on tourism economics and sustainable tourism. He belongs to the Institute of Tourism and Sustainable Economic Development (Tides), collaborating as a researcher in several projects related to the development of sustainable tourism activities in Macaronesia. He is also a member of the ECOMAS research team, in which he participates in a project based on tourism and social representations, subalternativity and social empowerment of some villages near the protected areas of Senegal.