

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/235285894>

# Heckit modelling of tourist expenditure: Evidence from Spain

Article in *International Journal of Service Industry Management* · July 2005

DOI: 10.1108/09564230510601404

CITATIONS

86

READS

249

2 authors, including:



Juan Luis Nicolau

Virginia Tech

128 PUBLICATIONS 1,914 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Service Failure and Recovery [View project](#)



Innovation in Tourism and Hospitality [View project](#)

**HECKIT MODELLING OF A TWO-STAGE TOURIST CHOICE PROCESS:  
EVIDENCE FROM SPAIN**

Juan L. Nicolau\*

Francisco J. Más

Dpt. of Economía Financiera, Contabilidad y Marketing  
Faculty of Economics  
University of Alicante  
Campus de San Vicente  
Ap. Correos 99  
E-03080 Alicante  
SPAIN  
Phone and Fax: +34 965.90.36.21  
e-mail: [JL.Nicolau@ua.es](mailto:JL.Nicolau@ua.es)

Nicolau, J.L. and Más, F. (2005), "Heckit Modelling of Tourists Expenditures: Evidence from Spain", *International Journal of Service Industry Management*, vol. 16, 3, 271-293.

<http://www.emeraldinsight.com/doi/pdfplus/10.1108/09564230510601404>

\* This study has benefited from a "Spanish Tourism" grant from the Secretary of State for Commerce and Tourism of the Ministry of Economics to carry out the Doctoral Thesis of the first author.

# **HECKIT MODELLING OF A TWO-STAGE TOURIST CHOICE PROCESS: EVIDENCE FROM SPAIN**

## **Abstract**

The objective of this paper is to decompose the process of tourist choice into two stages: taking a holiday and holiday expenditures, using the Heckit model, which avoids the problems of the methodologies applied to date. We propose expectations on the effect on the above decisions of tourist characteristics relating to the destination, personal restrictions and socio-demographic and psychographic characteristics. The empirical application, which is carried out in Spain on a sample of 3,781 individuals, finds a two-stage tourist choice process as the expenditure decision is correlated with that of taking a holiday. In addition, these decisions are also explained by individual tourist characteristics.

**Key words:** Two-stage tourist choice process, going on holiday, tourists expenditures, Heckit model.

## INTRODUCTION

Tourism has become one of the world's largest and fastest growing industry. In fact, among the services that have absorbed more consumer money, expenditures on tourism has increased steadily. This growth trend in tourists expenditures is a stimulus for the business in the competitive markets of services (Cai et al., 1995).

On this account, the study of tourists expenditures (such as “travel costs”, “accommodation”, “food” and “tickets”) is a key element in the analysis of tourism returns in a destination and in the formulation of marketing strategies and policies. For example, a destination may decide to design marketing strategies which will attract “quality tourists”, who will spend more during their stay, as opposed to attracting numerous tourists with less intention to spend.

However, the determinants of tourists expenditures at an individual level have been little studied (Cai et al., 1995; Cannon & Ford, 2002), despite the fact that this expenditure represents an increasing proportion of the family budget (Melenberg & Van Soest, 1996). Essentially, literature is based on the examination of holiday expenses as an independent decision and has various approaches. The majority of studies are based on the classical regression methodology (Dardis et al., 1981; Davies & Mangan, 1992; Taylor et al., 1993; Fish & Waggle, 1996; Agarwal & Yochum, 1999; Aguiló & Juaneda, 2000; Cannon & Ford, 2002; Jang et al., 2004), which is not without limitations. The large number of null observations (nil tourists expenditures) assigned to individuals who do not take a holiday in the period analysed leads to a situation in which the application of the classical regression models does not guarantee consistent unbiased estimations of the parameters (Maddala, 1987).

Even the use of truncated regression models on the sub-sample resulting from the elimination of nil values only allows inferences to be made on the sub-sample itself and, moreover, has three problems (Cai, 1998): i) the loss of efficiency due to the reduction in sample size; ii) the sample is no longer probabilistic, even though the sampling procedure is; and iii) the exclusion of households which do not have any tourists expenditures in a specific period is an assumption that they have never taken nor will ever take a holiday.

A second research thread uses the censored regression or Tobit model (Dardis et al., 1994; Cai et al., 1995; Leones et al., 1998; Cai, 1998; 1999) in the analysis of

tourists expenditures, as it allows for inclusion of all expenditure observations –nil and positive–, thus minimising the problems of bias and inconsistency.

However, the Tobit model also presents problems. Firstly, this model is based on the assumption of censored data; which means that it is assumed that only realisations above a certain value are observed, which would be seen as a data defect (Greene, 1999 p. 817). This treatment of data as censored in the context of tourism implies assigning a nil value to households which do not provide their tourists expenditures in a questionnaire. Obviously, this approach is not correct. The existence of numerous households with nil tourists expenditures is not due to a censorial problem (unobservable values), but to the very nature of the data, given that the value zero *is* observable and has the qualitative meaning that an individual decides not to go on holiday. Secondly, Sigelman & Zeng (1999) show –in the context of policy decisions– that an application of the Tobit model on data with no censorial problems, gives a poor fit and produces significant bias in the estimations. Because of the above, the application of the Tobit model would be conceptually inappropriate (Sigelman & Zeng, 1999), and the correct method would be to model the decisions which cause zeros along with the expenditure decision (Maddala, 1992, p. 341). This implies a decomposition of the tourist choice process into two stages, “taking a holiday” and “holiday expenses”, due to the fact that both decisions are non independent and nested.

In order to solve the problems of the Tobit model, a third perspective, defended by Melenberg & Van Soest (1996) proposes the application of a system of equations [1]. Within this perspective, this study intends to examine this tourist choice sequential through the estimation of the Heckit model proposed by Heckman (1979). We propose various expectations on the dimensions which explain the above decisions in terms of tourist characteristics relating to the destination, personal restrictions and socio-demographic and psychographic characteristics. The empirical application is carried out in Spain on a sample of 3,781 individuals.

The rest of the paper is organised as follows: The second section reviews the relevant literature of tourists expenditures and the decision to take a holiday and proposes various expected effects. Section three describes the methodology and the sample used. Finally, sections four and five present the results and conclusions, respectively.

## EXPECTATIONS ON HOLIDAY DECISION AND EXPENDITURES.

The study of the determinants of expenditure on tourism products has largely taken taking an exploratory approach without proposing any direction on the impact of the dimensions studied. This section reviews the relevant literature in order to propose research expectations on the decision to go on holiday with regard to personal restrictions (income and family size), socio-demographic characteristics (age, education and size of the city of residence) and psychographic characteristics (opinion on taking holidays). We also propose expected effects on the determinants of tourists expenditures, which are placed into three groups: i) individual characteristics related to the destination (distance to the destination and type of accommodation); ii) personal restrictions (income, size of family and number of children); and iii) socio-demographic characteristics (age, marital status and education).

### Expectations relative to the decision to go on holiday

#### A) Personal restrictions

*Level of income.* Income is a personal budget restriction which determines the spending capacity of individuals and is taken into account in order to maximize utility (Crawford & Godbey, 1987). In fact, income has been proved to be highly explicative of holiday taking behaviour (Mergoupis & Steuer, 2003). Essentially, empirical literature shows that medium-high and high income groups are more likely to take vacations (Hay & McConnell, 1979; S.G.T., 1989a; 1992; 1993; Bardón, 1991; Walsh et al., 1992; I.E.T., 2000). This result corroborates the idea that tourism generally behaves as a *normal good* with positive demand-income elasticity, increasing its consumption as income increases (Davis & Mangan, 1992; Middleton, 1994). Along this line, it is expected a positive impact of income on the probability of going on holiday:

**E.1:** Greater levels of income are associated with greater probabilities of going on holiday.

*Household size.* Essentially, household size is a representative aspect of the so called *interpersonal barriers* (Crawford & Godbey, 1987). Therefore, Caswell & McConnell (1980), Eymann & Ronning (1992, 1995) and Walsh et al. (1992) consider that family size (a commonly used indicator of household size) plays an important and deterrent role in recreational decisions, both in the realisation of holidays and in the

determination of the destination, as large family size restricts holiday spending. Therefore, insofar as a reduced household size, characterised by a lack of children, implies more possibilities to travel and cover holiday costs (Collins and Tisdell, 2002), we expect that:

**E.2:** *Larger household size reduces the propensity to go on holiday.*

#### B) Socio-demographic characteristics

*Age.* One of the most important demographic dimensions which influence holiday demand is the age of the tourist (Mieczkowski, 1990). Authors generally agree that the assumption of a linear relationship between age and holiday travel seems excessively simplistic and unrepresentative of the real behaviour of individuals. Obviously, a linear impact implies that the marginal effect of a change in age on participation in a certain recreational activity is constant and independent of age, when in reality, the effect of an increase of a decade (on the predisposition to take part in an activity holiday, for example) varies according to whether the individual is twenty or fifty years old.

Authors such as Hay & McConnell (1979), Miller & Hay (1981) and Walsh et al. (1992) propose a non-linear relationship between age and propensity to take holidays, in such a way as to show a positive (negative) marginal effect up to a certain point, and a negative (positive) marginal effect after that point. Eymann & Ronning (1992) and Eymann (1995) suggest further stretching of the age-propensity to take holidays relationship, allowing non-linear impacts by defining age group variables. This allows them to represent any behaviour pattern in function of age; such as the *bimodal relationship* proposed by Becker (1992), Lawson (1991) and Oppermann (1995) of a greater propensity to travel among both younger and older people. This is basically due to a lack of children and the support given by public institutions to these two age groups (Núñez de Cella, 1998). We, therefore, propose the following statement:

**E.3:** *Age exerts a non-linear U shaped effect on the probability of going on holiday.*

*Education.* According to Parker (1976), there is a positive link between the realisation of tourist activities and an individual's educational level. Higher levels of education foment interest in tourism. Firstly, this allows better access to information and knowledge (Cai et al., 1995) and, secondly, higher educational levels may provide

training and preparation for some types of recreation activities (Dardis et al., 1981). Consequently, we expect that:

**E.4:** *Greater educational levels are related to greater propensity to go on holiday.*

*Size of the city of residence.* The size of the city of residence could also justify the decision to go on holiday. At an empirical level, the work of the S.G.T. (1989a, 1992) finds that the proportion of the population which takes holidays reaches the lowest levels in towns with lower populations. This is due to the fact that inhabitants of high population density cities have a greater need to escape in search of relaxation (Eymann & Ronning, 1997). Along this line, it is expected that:

**E.5:** *A larger city of origin brings about greater propensity for travel during holiday periods.*

#### C) Psychographic factors.

*Favourable opinion of going on holiday:* Although the previous characteristics are of great use in explaining tourist behaviour, Plog (1994) suggests incorporating dimensions which allow representation of other internal aspects of the individual. In fact, Ashok et al (2002) and Seddighi & Theocharous (2002) show that the choice can be influenced by non-product related aspects. Along this line, González & Díaz (1996) suggest that values and life styles (psychographic variables) provide a global description of the cognitive structure of the individual, therefore the examination of this variable represents a fundamental complement of socio-demographic characteristics in order to properly configure holiday products. However, these psychographic factors are not widely used in the literature of choice as they are not directly observable by the analyst, who would have to make additional effort in the collection of information (Plog, 1994) through databases and VALS (*Value and Life Styles*), LOV (*List of Values*) or AIO (*Activities, Interests and Opinions*) studies.

In any case, certain one-dimensional indicators -also known as primary dimensions or life style parameters (Lehmann, 1993; Bigné et al., 2000.)- allow the capture, as proxies, of the psychographic aspects of the individual. Chief among them being the favourable/unfavourable opinion of the product, as a person with a favourable opinion of going on holiday presents greater probability of tourist travel (Plog, 1994; Ryan, 1995). Therefore, we propose the following statement:



**E.6:** *Favourable opinions of going on holiday positively affect the option of leaving the habitual place of residence.*

### **Expectations relative to tourists expenditures**

A) Individual characteristics related to the destination:

*Distance between the place of origin and the destination.* Distance is a fundamental component of the consumption of tourism products, due to the marked spatial character of tourism. However, to the best of the authors' knowledge, this dimension has only been considered by Leones et al. (1998), who found that distances travelled over 150 miles had a positive impact on expenditures. However, in agreement with Cannon & Ford (2002), the inclusion of day-trip visitors in the Leones et al. data set accounts for much of the difference. In this study, we propose a positive link between distance and expenditure, based on the following argumentation: Firstly, that greater service provided by a travel company or, where appropriate, extended use of a private vehicle, brings about the payment of higher prices [2], which increase the total price of the holiday; and, secondly, the traditional research perspective holds that distance -or the tourist's geographical position relative to destinations- is a restriction or dissuasive variable of destination choice, as the displacement of an individual entails physical, temporal and financial effort (Taylor & Knudson, 1976). Following this approach, which considers distance to be a factor which reduces utility, Silberman (1985) suggests that as distance increases length of stay will increase. This is due to the fact that travel costs are fixed and independent of the number of days spent at the destination, meaning that longer stays allow individuals to spread these fixed costs over a longer period. In other words, a tourist will be prepared to make a long journey if s/he stays at the destination for at least the minimum number of days which will compensate for the effort made in the journey. Consequently, given that long distances lead to longer holidays and, in turn, to higher expenditure –as argued in the previous expected effect –, we can expect that distant destinations should be associated with higher costs. We, therefore, propose the following expectation:

**H.8:** *Long distances are associated with higher tourists expenditures.*

*Accommodation.* The analysis of accommodation type and its effect on tourists expenditures is necessary in certain countries, such as Spain, where the number of organised holidays is low due to the high percentage of private holiday apartments

(Bote et al., 1991). To date, this dimension has hardly been studied. We have only found the work of Agarwal & Yochum (1999) which shows that hotel accommodation is associated with higher expenditures, while apartments/villas and staying with friends and family are linked to lower expenditures. This result seems to be reasonable since accommodation with lower costs per person per night, both commercial (rented apartments and villas) and private (own or friend/family's apartments and villas), implies a lower total expenditure, while more expensive accommodation such as hotels are linked to higher costs. Consequently, we propose the following statement:

**E.8:** *Hotel accommodation is associated with higher tourists expenditures.*

B) Personal restrictions

*Level of income.* In line with the argument of expectation E.1, income influences tourism consumption patterns. Empirical literature shows a positive relationship between income and expenditure on tourism products (Dardis et al., 1981, 1994; Cai et al., 1995; Fish & Waggle, 1996; Mudambi & Baum, 1997; Cai, 1998, 1999; Agarwal & Yochum, 1999; Cannon & Ford, 2002). Along this line, expectation E.9 is as follows:

**E.9:** *Greater levels of income are associated with higher tourists expenditures.*

*Household size.* With regard to the effect of household size on tourists expenditures, the effect is uncertain. While large families might be expected to spend more on recreation, expenditures on necessities would also increase, thus reducing the amount available for discretionary items such as recreation (Dardis et al., 1981). However, this reasoning appears to be more closely linked to the initial decision to go on holiday taken by a family in order to control total spending and the family budget. With regard to the family size/spending relationship, it is logical to expect that, once the initial decision to go on holiday has been taken; larger families will spend more, given that the services required are greater. Therefore, we expect that:

**E.10:** *Larger household size increases tourists expenditures.*

Alternatively, when the literature of tourists expenditures analyses this interpersonal barrier, it specifically concentrates on the presence of children on the holiday. Thus, given that children are normally not income earners, an increase in the number of children in a given party is anticipated to lead to lower total party expenditures (Agarwal & Yochum, 1999). These authors, along with Cannon & Ford

(2002), obtain a negative effect of the number of children on tourists expenditures. Consequently, we make the following statement:

**E.11:** *The presence of a large number of children on the holiday reduces expenditure per travel group.*

#### C) Socio-demographic characteristics

*Age.* With regard to the relationship between age and tourists expenditures, Opaschowski (1990) shows that age generates different points of view towards leisure and tourism consumption. Young and middle age people consider holidays to be a part of their lifestyle and that extra holiday spending has to be saved for during the rest of the year. In fact, the study of the B.A.T. (1989; 1990) detects a tendency among young people to prioritise holiday expenditure, although once at the chosen destination half of these holidaymakers restrict their spending. Moreover, Dardis et al. (1981) show that tourists expenditures decrease with age, which can be explained by the fact that these costs include durable and non durable goods. In other words, in the early years, families tend to store durable goods associated with holidays (e.g. an apartment or caravan); therefore, expenditure on durable goods is reduced with age.

With reference to the above points, we can expect a non-linear effect, which means that with younger people there is an increase in holiday spending until they arrive at a threshold age and begin to reduce their spending. Consequently, we propose the following statement:

**E.12:** *Age exerts a non-linear inverted U shaped effect on tourists expenditures.*

*Marital status.* Marital status is considered to be a determinant factor in tourists expenditure behaviour (Cai et al., 1995). In particular, the tourism activities of both partners are complementary and non-substitutional. The spending pattern differences between married and single people may be attributable to the incremental expenses of the spouse on vacations taken as joint activities by husbands and wives (Cai, 1998). Along this line, Dardis et al. (1981), Cai et al. (1995) and Cai (1998, 1999) find a positive relationship between tourists expenditures and marriage. In accordance with this argument, it is stated that:

**E.13:** *Marriage increases holiday expenditures.*

*Education.* Apart from the effect it has on the decision to go on holiday, Dardis et al. (1981), Cai et al. (1995) and Cai (1998, 1999) find a positive relationship between higher educational levels and greater tourists expenditures. This result can be explained by, firstly, the fact that people with higher educational qualifications usually find higher paid occupations, which allows them higher holiday budgets, and secondly, because people with higher educational levels take a greater number of foreign holidays (S.G.T., 1989a; 1992; 1993; Bardón, 1991; I.E.T., 2000), which usually cost more than national holidays. Therefore, we expect that:

**E.14:** *Higher levels of education lead to greater tourists expenditures.*

## RESEARCH DESIGN

### Methodology

The methodology proposed to decompose the tourist choice process into two stages (taking a holiday and tourists expenditures) is based on the estimation of a Heckit model.

Let  $z_{tk}$  be is a group of variables  $k$  which represent the characteristics of individual  $t$  which determine the decision to take a holiday –measured by a latent variable  $d_t^*$ –, and  $\gamma_k$  are the coefficients which reflect the effect of these variables on this decision; and  $x_{ts}$  is a group of variables  $s$  which represent the characteristics of individual  $t$  which determine the level of tourists expenditures  $E_t$ , and  $\beta_s$  are the coefficients which reflect the effect of these variables on this spending level. Thus, the Heckman model (1979) takes the following form:

$$d_t^* = \sum_{k=1}^K \gamma_k z_{tk} + u_t \quad (1)$$

$$E_t = \sum_{s=1}^S \beta_s x_{ts} + \varepsilon_t \quad \text{observed only if } d_t^* > 0 \quad (2)$$

where the disturbances  $u_t$  and  $\varepsilon_t$  follow a bivariate normal distribution with a zero mean, variances  $\sigma_u$ ,  $\sigma_\varepsilon$  respectively, and covariance  $\sigma_{u\varepsilon}$ . Therefore, we define a dichotomic variable  $d_t$ , which takes a value of one when the latent variable  $d_t^* > 0$ , and zero if not. In this way,  $d_t=1$  indicates the decision to take a holiday and  $d_t=0$  that of staying at home

during the holiday period. This interpretation causes equation (1) to represent a Binomial model.

Considering the joint density function of  $u_t$  and  $\varepsilon_t$ , the estimation of the parameters of the previous system of equations by the method of maximum likelihood is immediate. However, the maximisation of the likelihood function is complex due to problems of convergence (Greene, 1999; Sigelman & Zeng, 1999), which leads us to use the two-stage estimation method proposed by Heckman (1979), which allows us to obtain consistent estimations for the parameters. This method starts off with the expression of conditional expectation of  $y$ , so that:

$$E(E_t/d^*>0)=x\beta+\sigma_{\varepsilon u}\sigma_{\varepsilon}\lambda(-\gamma z) \quad (3)$$

where  $\lambda$  is the inverse ratio of Mills, defined as  $\lambda(-\gamma z)=\phi(-\gamma z)/(1-\Phi(-\gamma z))$ ;  $\beta$  and  $\gamma$  are the vectors of parameters which measure the effect of the previously mentioned variables  $x$  and  $z$ ; and  $\phi$  and  $\Phi$  are the functions of density and distribution of a Normal, respectively. The expression of conditional expectation shows that both decisions –to take a holiday and level of spending– are related; in other words, the expectation of  $E_t$  equals  $x\beta$  only when the errors  $u_t$  and  $\varepsilon_t$  are non correlated,  $\sigma_{\varepsilon u}=0$ ; otherwise, the expectation of  $E_t$  is affected by the variables of equation (1). The significance of parameter  $\sigma_{\varepsilon u}$  ratifies the superiority of this model over others.

From expression (3), we find that

$$E_t/d^*>0=E(E_t/d^*>0)+v_t=x\beta+\sigma_{\varepsilon u}\sigma_{\varepsilon}\lambda(-\gamma z)+v_t \quad (4)$$

where  $v_t$  is the distributed error term  $N(0, \sigma_{\varepsilon}(1-\sigma_{\varepsilon u}(\lambda(\lambda-\gamma z))))$  [3].

Starting off with equation (4), Heckman proposes the following two-stage procedure: i) estimate by maximum likelihood the coefficients  $\gamma$  of the Probit model represented by equation (1), and calculate the Mills' inverse ratio for each observation of the sample  $\hat{\lambda}$ ; and ii) estimate  $\beta$  and  $\beta_{\lambda}=\sigma_{\varepsilon u}\sigma_{\varepsilon}$ , with an OLS regression of  $E_t$  over  $x$  and the estimation of  $\hat{\lambda}$ .

Therefore, in our two-stage choice context we propose the application of the Heckit model estimated by the above method in order to simultaneously model the decisions to take a holiday and the level of tourists expenditures.

## Sample and Variables

To reach the proposed research objectives, we use information on tourist choice behaviour obtained from a national survey called “Holiday behaviour of the Spanish (III)” carried out by the Spanish Sociological Research Centre. This is due to the following reasons: i) The availability of information on tourist behaviour; and ii) The survey is home based and directed at a sample of individuals (over 18 years old) which avoids the selection bias characteristic of samples obtained in destinations and allows the incorporation of the decision processes of individuals who do not take holidays; leading to a more precise analysis of tourist demand.

The sample is taken with multistage sampling, stratified by conglomerations, with proportional selection of primary units -cities- and of secondary units -censorial sections-. The collection of the information was made through personal, at home, interviews with a structured questionnaire. The final sample is of 3,781 individuals, with a sample error of  $\pm 1.24\%$  for a confidence level of 95.5%.

In order to make the proposed choice model operative, we define the variables used, identifying the dependent and independent variables.

1) *Dependent variables*: The discrete nature of the decision “going on holiday” leads to use a dichotomous variable, in such a way that it takes a value of 1 if the individual goes on holiday and 0 otherwise. The survey focuses on the main holidays only, but it does not consider other types of vacation (e.g. honeymoon). The variable relative to tourists expenditures is found by a quantitative variable which represents costs incurred during the holiday. This variable includes the components “travel costs”, “accommodation”, “food” and “tickets”.

2) *Independent variables*: a) Personal characteristics relating to the destination:  
i) *Distance between origin and destination*. In accordance with the literature of choice in tourism, we use the physical separation in kilometres between the place of origin and the chosen destination (Wennergren & Nielsen, 1968; Stopher & Ergün, 1979; Moutinho & Trimble, 1981; Louviere & Hensher, 1983; Peterson et al., 1983; Silberman, 1985; Perdue, 1986; Borgers et al., 1988; Fesenmaier, 1988; Adamowicz et al., 1994; Dellaert et al., 1997; Schroeder & Louviere, 1999; Kemperman et al., 2000). The information on distances between origins and destinations is found in the Interactive Campsa Guide, for national destinations and through the Euclidean distance

in the case of international destinations. ii) *Accommodation type*. The type of accommodation selected by the tourist is classified by literature through different categorical variables (Alegre & Pou, 2003). In particular, our study considers the following five dummy variables: “hotel”, “campsite” “own apartment or villa”, “rented apartment or villa” and “family or friends’ house”. The final one is used as a reference category. In all of the above, a value of 1 indicates the presence of each alternative, and 0 otherwise.

b) Personal restrictions: i) *Income*. This dimension considers different income levels in order to observe the possible lack of linearity to their effect (Eymann & Ronning, 1997). Monthly income levels are placed into the following categories: *Income 1*, up to 600€ per month; *Income 2*, between 600 and 1200€; *Income 3*, between 1200 and 2400€; *Income 4*, between 2400 and 4500€; and *Income 5*, more than 4500€. *Income 1* is taken as the base reference. ii) *Household size*. This is measured by the number of people living in the house (Caswell & McConnell, 1980; Eymann & Ronning, 1992; 1997; Walsh et al., 1992). iii) *Children*. The number of children under sixteen who go on holiday (Moutinho, 1987).

c) Socio-demographic characteristics: i) *Age*. With the object of testing for possible non-linear effects, and in order to give more flexibility to the effect of age, we follow Cai’s (1998) approach by constructing an age group variable in which we define four categorical variables thus: *Age 1*, under 25 years old; *Age 2*, between 26 and 45; *Age 3*, between 46 and 65; and *Age 4*, over 65 years old. As a reference category we take *Age 4*. This piecewise definition allows us to represent any pattern in function of age. (Eymann & Ronning, 1992; 1997; Cai, 1998). The grouping is based on the World Tourism Organisation’s recommendations (Smith, 1995, p. 28).

ii) Marital status. Dummy variable where married=1 and single=0 (Hay & McConnell, 1979; Eymann & Ronning, 1997); iii) *Education*. We establish three educational levels through three categorical variables: *Education 1*, Basic Education; *Education 2*, Secondary education; and *Education 3*, University Education. Category *Education 1* is taken as a base reference. (Caswell & McConnell, 1980; Eymann & Ronning, 1997; Riera, 2000);

iv) *Size of City*. The size of the place of residence is defined by the following categorical variables: *Size of City 1*, up to 10.000 inhabitants; *Size of City 2*, between 10,000 and 100,000 inhabitants; *Size of City 3*, between 100,000 and 1,000,000

inhabitants; *Size of City 4*, over 1,000,000 inhabitants. The category *Size of City 1* is taken as a base reference (Eymann & Ronning, 1997; Smith & Munley, 1978).

d) Psychographic factors. As *one-dimensional indicators of the internal aspects of an individual* we include the following dimension: *An individual's favourable/unfavourable opinion of going on holiday at least once a year*. This is measured with a dichotomous variable and takes a value of one if an individual has a favourable opinion of going on holiday at least once a year, and zero if the person has the opposite view (Plog, 1994).

e) Control variable: *Duration of stay*. Length of stay is represented by a quantitative variable of the number of days that a tourist spends outside the usual place of residence, in line with Mak & Moncur (1979) and Silberman (1985). If we consider that the number of days that a tourist spends away from the usual place of residence is "holiday quantity" (Silberman, 1985), we can expect a relationship between the duration of stay and both, total expenditure and per-day expenditure incurred during the holiday. At an empirical level, the importance of length of stay to tourists expenditures has been shown in various studies (Spotts & Mahoney, 1991; Taylor et al., 1993; Nogawa et al., 1996; Saeton & Palmer, 1997; Van Limburg, 1997; Leones et al., 1998; Mules, 1998; Agarwal & Yochum, 1999; Aguiló & Juaneda, 2000; Cannon & Ford, 2002). On this account, we control for this dimension.

## RESULTS AND DISCUSSION

The identification of the determinants of the decision to go on holiday in terms of the variables corresponding to expectations H.1-H.6 (income, household size, age, education, size of city and opinion of going on holiday) and the level of tourists expenditures during the holiday in terms of the variables corresponding to expectations H.7-H.14 (distance between origin and destination, type of accommodation, income, household size, number of children, age, marital status and education), implies the estimation by Heckman's two-stage estimator of the Heckit model, which is shown in Table I.

Take in Table I

Before applying the model, we carry out a detailed study of the correlations between the explanatory variables in order to avoid possible collinearity. Also, we test for collinearity among variables by calculating the variance inflation factor (VIF) for each



of the regression coefficients. The equations presented for each model constitute different combinations of them, in such a way that they all are well below the cut off figure of 10 recommended by Neter et al. (1985).

The results obtained show that the coefficient associated with the inverse Mill's ratio is significant in all the models, which indicates that the correlation between the error terms of the decision to go on holiday ( $u_t$ ) and that of tourists expenditures ( $\varepsilon_t$ ) is different than zero,  $\sigma_{\varepsilon u} \neq 0$ . This shows that the Heckit model is appropriate for this type of analysis since it allows for covariance in the errors and, therefore, that tourist choice can be decomposed into a two-stage process (going on holiday and tourists expenditures).

With regard to the individual and joint significance of the explanatory variables of the estimation, the following aspects stand out. Firstly, the likelihood ratio test for the Binomial model (equations GH) and the statistic  $F$  for the regression (equations E) are significant in all the equations at a level below 0.1%, which means that individual characteristics related to the destination, personal restrictions, socio-demographic characteristics and the psychographic variables provide relevant information on the decisions considered. Likewise, in order to determine the specification which best represents the model we calculate the Schwarz Bayesian Information Criterion (Schwarz BIC) for each equation. In virtue of this, the Heckit model 2 presents the optimum specification, which is corroborated by reaching the largest McFadden  $\rho$  coefficient (20.64%) in the Binomial model (which is considered more than acceptable for this type of model (Hensher & Johnson, (1981)), and the largest  $R^2$  (16.5) in the regression equation (which, although it is a moderated fit, represents an percentage of explained variation which is higher than that reached in literature (Fish & Waggle, 1996; Cannon & Ford, 2002)).

Secondly, the significance tests of the individual parameters show that, the determinants of the decision to go on holiday are income, household size, size of city and opinion of going on holiday. The determinant factors of tourists expenditures seem to be the following: distance between origin and destination, type of accommodation, income, household size, age, marital status and length of stay. We should stress that the estimations of these coefficients in both decisions show robust results in all equations given that the variables are significant in all of them [4].

The analysis of the significance of the parameters allows us to test the proposed expectations for each group of variables:

A) Going on holiday

i) Personal restrictions. Firstly, we examine income. For the decision to go on holiday, all the categorical variables relative to income levels show a positive sign. Moreover, all the parameters are significantly greater than that of the reference category of low income (*Income 1*), showing that the category *Income 4* has the greatest impact on the probability of going on holiday, which points to the existence of a saturation point [5]. This confirms expectation E.1 that the probability of going on holiday has a positive relationship with income, in line with Hay & McConnell (1979), S.G.T. (1989a; 1992; 1993), Bardón (1991) and Walsh et al. (1992), showing that this type of product is a *normal product* with a saturation point.

The effect of household size is significant and negative for the decision to go on holiday, which allows us to accept expectation E.2 that *larger household size reduces the propensity to go on holiday*. This result suggests that the number of members of a household is an interpersonal barrier at the moment of leaving the usual place of residence during the holiday period, as large family size restricts holiday spending, in line with Caswell & McConnell (1980), Eymann & Ronning (1992, 1997) and Walsh et al. (1992).

ii) Socio-demographic characteristics. As regards age, we find no influence of this variable on the decision to go on holiday as the parameters estimated which relate it to this decision are not significant, which leads us to reject expectation E.3. This lack of significance of age is also found by the studies of Collins & Tisdell (2000, 2002), the S.G.T. (1989a, 1992), Bardón (1991) and the I.E.T. (2000), and is explained by the fact that motivation can exert a greater influence than age when going on holiday (Collins & Tisdell, 2000, 2002). For example, an individual makes a journey to visit family regardless of age. Therefore, this non-significance of age suggests that there are other personal factors which push an individual to go on holiday regardless of age.

With regard to education, the positive and significant sign of the parameters of *Education 2* and *Education 3* in the equation of going on holiday shows that higher educational levels are associated with a greater propensity to go on holiday, which supports expectation E.4. This result proves the argumentation of Parker (1976), Dardis

et al. (1981) and Cai et al. (1995) that education foments an interest in tourism, by facilitating access to information and providing training and preparation for some types of recreational activities.

The size of the city of residence shows a positive sign for larger cities (categories 3 and 4), whose coefficients are significantly greater than those of the small size categories (1 and 2); which is indicative of the existence of a need to escape from large urban centres (Eymann & Ronning, 1992), and corroborates expectation E.5. Finally, the positive sign of the variable relating to the favourable/unfavourable opinion of going on holiday supports expectation E.6 that a favourable opinion foments holidays. Therefore, this psychographic dimension of individuals determines holiday decisions, in line with Plog (1994), González & Díaz (1996), Ashok et al., (2002) and Seddighi & Theocharous (2002).

#### B) Level of tourists expenditures

i) Characteristics of individuals related to the destination. The distance between the origin and the destination is positive and significant, which lends support to expectation E.7 that longer distances bring about higher costs due to higher transport costs (either public or private), and that the dissuasive effect of distance leads tourists to spend more time at a destination in order to spread the fixed costs of longer distances over a longer period and hence, to higher costs.

With regard to accommodation type, the results obtained indicate that staying in hotels, own apartments/chalets and rented apartments/chalets leads to higher costs than staying with family and friends. This is as expected as the coefficients are positive and significant. However, the important result is the order of magnitudes of the parameters of the three accommodation types: rented apartment/chalets (69.627), hotel (62.087) and own apartment/chalets (26.635 [6]). This result leads us to reject expectation E.8 that staying at hotels leads to higher spending. One possible explanation for this result could be length of stay. Staying at a hotel is expensive per day so stays at hotels tend to be shorter, whereas, stays at rented apartments/chalets are longer, which means that the total price paid will be higher. For this reason, an exhaustive analysis of the effect of accommodation type on spending should be controlled by the variable of “number of days”. In fact, the only study found which examines accommodation type is that of Agarwal & Yochum (1999), whose dependent variable is daily spending.

In order to refine the test of this expected effect, we estimate the effect of accommodation type on spending per day, in line with Agarwal & Yochum (1999). In this case, the only significant and positive coefficients are those of hotels and rented apartments/chalets, with the former having a higher magnitude (7.872; prob<0.001) than the latter (4.851; prob<0.001) [7]; which is in line with expectation E.8.

In any case, this result shows that tourists with higher total spending during their holidays are those which stay in rented apartments. Therefore, longer stays lead to higher spending, which compensates for the lower price per day of rented apartments/chalets compared to hotels.

ii) Personal restrictions. Income: the results obtained for holiday expenditure show that only the coefficient of high income (*Income 5*) is significant and positive, which backs up expectation E.9 that greater levels of income are associated with higher tourists expenditures, in line with the evidence found in literature (Dardis et al., 1981, 1994; Cai et al., 1995; Fish & Waggle, 1996; Mudambi & Baum, 1997; Cai, 1998, 1999; Agarwal & Yochum, 1999; Cannon & Ford, 2002). Comparing this result with that obtained for the decision to go on holiday, we find a saturation point for the decision to go on holiday, but not for spending levels; in other words, *Income 5* individuals present a slight reduction in their number of holidays with respect to *Income 4*, but when they go on holiday, they spend much more than other groups.

The effect of household size is positive for the volume of expenditure, which leads us to accept expectation E.10 that *larger household size increases tourists expenditures*. This means that, family size is a barrier at the moment of going on holiday (see expectation E.2), but families which decide to go on holiday spend more as family size increases.

The number of children on the holiday does not have a significant coefficient (except in model 4, with a level of 10%), which leads us to reject expectation E.11. This suggests that the presence of a large number of children under 16 years of age does not influence tourists expenditures. This result, along with the previous one, implies that holiday expenditures increase with household size and, in particular, with an increase in the number of people over 16 years of age. Interestingly, a remaining issue would be to determine the age at which children no longer form a part of the holidaying group.

iii) Socio-demographic characteristics.

As regards age, spending patterns appear to differ with the age of the individual, as evidenced by the significance of the parameters of age groups 2 (between 26 and 45) and 3 (between 46 and 65). The positive signs of both variables suggest that higher spending tourists are found among these age groups; moreover, the magnitude of the parameter of age 3 is lower than that of age 2. In summary, from 26 years old onwards, tourists increase their holiday spending until they reach 46, when they begin to reduce spending until they reach the lowest spending levels of the over 65s (*Age 4*); which reflects an inverted U shaped effect. Underlying this is the higher spending on durable tourism goods in the early years of family life (Dardis et al., 1981). Hence, this result leads us to accept expectation E.12 that age has a non-linear influence on holiday expenditures, in line with Dardis et al. (1981) and Opaschowski (1990).

With regard to marital status, the positive and significant sign of the coefficient supports expectation E.13 that marriage increases tourists expenditures, which suggests that the tourist activities of both partners are complementary and we observe incremental expenses of the spouse, on vacations taken as joint activities by husbands and wives, in line with Cai (1998).

With regard to educational level, the non-significant parameters of the spending equation do not allow us to accept expectation E.14 that higher educational levels lead to higher tourists expenditures. This result coincides with that of Cannon & Ford (2002). One possible explanation could be that there are other factors influencing the effect of education on holiday spending. For example, a large family will have higher holiday spending even though their educational level may be low.

Length of stay is a significant and positive dimension, in such a way that a longer stay leads to higher spending levels as, by increasing the “holiday quantity”, the total price will be higher. Additionally, in order to refine the analysis of this control variable, we break it down into three groups: *stay1*, up-to-3-day stays, *stay2*, between 4 and 7 days and *stay3*, more than 7 days. Regarding the spending per day, shortest stays (up to 3 days) lead to greater tourists expenditures (319.14€ per day), compared to the other two groups: 78.03€ for stays between 4 and 7 days and 41.36€ for more than 7 days (The Anova test corroborates the significant differences ( $F\text{-statistic}=7.80;p<0.000$ ) among the three groups). As regards total expenditures, the differences are less important: 593.80€ for *stay1*, 484.27€ for *stay2* and 721.51€ for *stay3* (The Anova test finds significant differences as well ( $F\text{-statistic}=82.82;p<0.000$ )). Therefore, in general,

longer stays lead to greater total expenditures and shorter stays brings about greater per-day expenditures.

## CONCLUSIONS

The idea that the choice of tourists expenditures should be seen as a two-stage process, through which the tourist first decides whether or not to go on holiday and then decides on the level of holiday spending, has allowed us to analyse this aspect in the context of a sample of 3,781 Spanish individuals obtained in origin. We propose the use of a Heckit model which allows for the simultaneous modelling of both decisions and the testing of various expectations on the decision to go on holiday (personal restrictions and socio-demographic and psychographic characteristics) and the decision on the level of expenditure (individual characteristics related to the destination, personal restrictions and socio-demographic characteristics).

The empirical application carried out on the sample allows us to reach the following conclusions: a) *Joint Modelization*. The tourist choice process can be decomposed into two stages: going on holiday and tourists expenditures. Therefore, the spending decision should be modelled jointly with the decision to go on holiday due to the dependency between them, as evidenced by the significant correlation shown by Mill's inverse ratio parameter.

b) *Decision to go on holiday*. The dimensions which appear to have an effect on this decision are income, household size, education, size of the city of origin and opinion of going on holiday. We can conclude that a greater propensity to go on holiday is associated with high income (meaning that holidays are normal goods, though with a saturation point), with smaller household size (due to the monetary restrictions of households with many members), with higher educational levels (as they foment interest in tourism), with residence in large cities (because of the need to escape), and with a favourable opinion of going on holiday (psychographic dimension).

c) *Expenditure level*. The determinant factors which have an effect on the level of expenditure are distance between origin and destination, type of accommodation, income, household size, age, marital status and length of stay. This results allows us to conclude that higher spending levels are associated with further distance (due to greater transport costs - public or private- and the association between longer stays and distant

destinations), with accommodation in rented apartments/chalets (due to the fact that these tourists spend more time at the destination and, therefore, have a greater total expenditure), with higher income groups (showing, once again, that tourism is a normal good), with household size (due to the higher costs implied by larger families), with ages between 26 and 65 (especially between 26 and 45, due to the large amount spent on durable tourism products in the early years of the family life cycle; suggesting that, in any case, age has a non-linear inverted U shaped effect on holiday spending), with marriage (as the tourism activities of both partners are complementary) and with longer stays (due to the higher price of an increased holiday quantity).

d) *Differentiated effect of the dimensions.* An important conclusion of this analysis is the differentiated effect of a given dimension on each decision. In particular, this is observed in the following dimensions: i) *Income.* Income has a positive relationship with the decision to go on holiday and with expenditure; however, with the first decision, we find a saturation point which does not exist with the second decision. Individuals from the highest income group (*Income 5*) present a slight reduction in the number of holidays with respect to those of the group below them (*Income 4*), but when they go on holiday, they spend much more than the other groups. ii) *Household size.* The variable of family size is significant for both decisions, but with opposite signs. This result suggests that the number of people in a household is a restriction (negative sign for decision to take a holiday) at the moment of deciding to go on holiday, due to the higher costs (positive sign of the spending equation) when large families go on holiday. iii) *Age.* Age does not seem to have any influence on the decision to go on holiday, whereas it is a determinant of spending and shows a non-linear effect. And iv) *Education.* Higher levels of education are associated with a greater propensity to go on holiday, whereas they are not determinants of the level of spending.

With regard to implications to management, we can point out that, generally, the knowledge of this two-stage choice process gives an alternative perspective to the segmentation of the tourism market in order to characterise the profile of tourists with the greatest propensity to go on holiday and to find their spending patterns (e.g. most profitable tourist type); which, in turn, is fundamental for the formulation of marketing strategies by tourism organisations.

Furthermore, the results obtained suggest the following specific implications: i) the promotion of destinations by Spanish local tourism agencies should be developed

with special attention paid to some faraway markets of origin, due to the expected propensity for these tourists to spend longer periods at the destination. ii) The specialisation of destinations in terms of accommodation type and length of stay, through agreements between local tourism agencies and hotel operators acting in Spain. Tourists who stay in hotels have higher daily costs than those in rented apartments/chalets, whereas in the latter there is greater total spending (due to the higher number of days at the destination). Evidently, the optimal situation would be to have the maximum number of tourists staying in hotels, which would be of interest to those destinations aiming to direct their promotions at high income tourists. However, if the objective of a destination is to attract a wider range of people –high and medium income groups–, it could combine the promotion of both accommodation types and thus adapt itself to the spending capacity of each group. Moreover, this combined use of accommodation types could compensate for the opportunity costs of the fact that hotels are not fully occupied. And iii) the design of some holiday packages by tour operators should be adapted to the needs of tourists between the ages of 26 and 65, married tourists and large families with a large number of over 16s, as they could represent the most profitable tourist profiles.

Among the limitations of the study is the fact that we do not consider the impact of important explanatory dimensions such as price variables (fares, accommodation costs, etc.) and personal motivations, due to the lack of information on them. Among future lines of research, it can be said that the results presented here should be supported by other studies on other geographical areas. Likewise, it would be interesting to test the proposed expectations from a longitudinal perspective, which would allow an observation of the temporal evolution of the two-stage choice process.



## REFERENCES

- Adamowicz, W., Bocal, P., Williams, M. and Louviere, J. (1998), "Stated Preference Approaches for Measuring Passive Use Values: Choice Experiments and Contingent Valuation", *American Journal of Agricultural Economics*, 80, february, 64-75.
- Agarwal, V.B. and Yochum, G.R. (1999) "Tourist Spending and Race of Visitors", *Journal of Travel Research*, 38, 173-176.
- Aguiló, E. and Juaneda, C. (2000) "Tourism Expenditure for Mass Tourism Markets", *Annals of Tourism Research*, 27, 3, 624-637.
- Alegre, J. and Pou, LL. (2003) "La Estancia Media de los Turistas en Islas Baleares: Determinantes Microeconómicos e Implicaciones sobre la Evolución del Gasto Agregado", XII Simposio Internacional de Turismo y Ocio, Esade, Barcelona.
- Ashok, K., Dillon, W.R. and Yuan, S. (2002), "Extending Discrete Choice Models to Incorporate Attitudinal and Other Latent Variables", *Journal of Marketing Research*, 34, february, 31-46.
- B.A.T. Freizeit-Forschungsinstitut (1989), *Wie Arbeiten wir Nach dem Jar 2000? Freizeit-Impulse für die Arbeitswelt von Morgen*, Hamburg.
- B.A.T. Freizeit-Forschungsinstitut (1990), *Urlaub 89/90. Die Grenzen Grenzenlosen Reisens: Trendwende im urlaubsverhalten?*, Hamburg.
- Bardón, E. (1991) "Resumen del Estudio sobre el Grado de Satisfacción de la Demanda Turística Nacional and Extranjera en relación con el Producto Turístico Español", *Estudios Turísticos*, 110, 65-123.
- Becker, C. (1992) "Lebenslanges Urlaubsreiseverhalten: Erste Ergebnisse einer Pilotstudie", *Materilisen zur Fremdenverkehrsgeographie*, 25, 70-82.
- Bigné, E., Font, X. and Andreu, L. (2000) *Marketing de Destinos Turísticos: Análisis y Estrategias de Desarrollo*, Esic, Madrid.
- Borgers, A.W. J., Van Der Heijden, R.E.C.M. and Timmermans, H.J.P. (1989) "A Variety Seeking Model of Spatial Choice-behaviour", *Environment and Planning A*, 21, 1037-1048.
- Bote, V., Huescar, A. and Vogeler, C. (1991) "Concentración e Integración de las Agencias de Viajes Españolas ante el Acta Única Europea", *Papers de Turisme*, 5, 5-43.
- Cai, L.A (1998) "Analyzing Houshold Food Expenditure Patterns on Trips and Vacations: A Tobit Model", *Journal of Hospitality & Tourism Research*, 22, 4, 338-358.
- Cai, L.A. (1999) "Relationship of Houshold Characteristics and Lodging Expenditure on Leiusure Trips", *Journal of Hospitality & Leisure Marketing*, 6, 2, 5-18.
- Cai, L.A., Hong, G. and Morrison, A.M. (1995) "Household Expenditure Patterns for Tourism Products and Services", *Journal of Travel & Tourism Marketing*, 4, 4, 15-40.
- Cannon, T.F. and Ford, J. (2002) "Relationship of Demographic and Trip Characteristics to Visitor Spending: An Analysis of Sports Travel Visitors across Time", *Tourism Economics*, 8, 3, 263-271.
- Caswell, M.F. and McConell, K.E. (1980). "Simultaneous Estimation of Jointly Dependent Recreation Participation Function", *Journal of Environmental Economics and Management*, 7, 65-76.

- Collins, D. and Tisdell, C. (2000) "Changing Travel Patterns with Age: Australian Evidence and the Need to Modify Current Theories", *Australian Journal of Hospitality Management*, 7, 1, 15-25.
- Collins, D. and Tisdell, C. (2002) "Aged-Related Lifecycles: Purpose Variations", *Annals of Tourism Research*, 29, 3, 801-818.
- Crawford, D.W. and Godbey, G. (1987) "Reconceptualizing Barriers to Family Leisure", *Leisure Sciences*, 9, 119-128.
- Dardis, R., Derrick, F., Lehfeld, A. and Wolfe, K.E. (1981) "Cross-Section Studies of Recreation Expenditures in the United States". *Journal of Leisure Research*, 13, 3, 181-194.
- Dardis, R., Soberon, H. and Patro, D. (1994) "Analysis of Leisure Expenditures in the United States". *The Proceedings of the American Council on Consumer Interests*, 39, 194-200.
- Davis, B. and Mangan, J. (1992) "Family Expenditure on Hotels and Holiday", *Annals of Tourism Research*, 19, 691-669.
- Dellaert, B.G.C., Borghers, A.W.J. and Timmermans, H. J.P. (1997) "Conjoint Models of Tourist Portfolio Choice: Theory and Illustration", *Leisure Sciences*, 19, 31-58.
- Eymann, A. (1995) *Consumers' Spatial Choice Behavior*, Physica-Verlag, Heidelberg.
- Eymann, A. and Ronning, G. (1992) "Discrete Choice Analysis of Foreign Travel Demand" in Vosgerau, H.J., ed., *European Integration in the World Economy. Studies in International Economics and Institutions*, Springer, Berlin.
- Eymann, A. and Ronning, G. (1997) "Microeconometric Models of Tourists' Destination Choice", *Regional Science and Urban Economics*, 27, 735-761.
- Fesenmaier, D.R. (1988) "Integrating Activity Patterns into Destination Choice Models", *Journal of Leisure Research*, 20, 3, 175-191.
- Fish, M. and Waggle, D. (1996) "Current Income versus Total Expenditure Measure in Regression Models of Vacation and Pleasure Travel", *Journal of Travel Research*, 35, 2, 70-74.
- González, A.M. and Díaz, A.M. (1996) "Análisis del Comportamiento del Turista a partir de las Variables de Estilos de Vida", in Valdés, L. and Ruiz Vega, A., (Eds.), *Turismo y Promoción de Destinos Turísticos: Implicaciones Empresariales*, Universidad de Oviedo, Gijón, pp.127-150.
- Greene, W.H. (1999) *Análisis Económico*, Prentice Hall, Madrid.
- Hay, M.J. and McConnell, K.E. (1979) "An Analysis of Participation in Nonconsumptive Wildlife Recreation", *Land Economics*, 55, 4, 460-471.
- Heckman, J.J. (1979) "Sample Selection Bias as a Specification Error", *Econometrica*, 47, 1, 153-161.
- I.E.T. (2000), *Movimientos Turísticos de los Españoles en 1999*, Instituto de Estudios Turísticos, Madrid.
- Jang, S., Bai, B., Hong, G-S., and O'Learly, J.T. (2004) "Understanding Travel Expenditure Patterns: A Study of Japanese Pleasure Travelers to the United States by Income Levels", *Tourism Management*, forthcoming.
- Kemperman, A., Borghers, A.W.J., Oppewal, H. and Timmermans, H.J.P. (2000) "Consumer Choice of Theme Parks: A Conjoint Choice Model of Seasonality Effects and Variety Seeking Behavior", *Leisure Sciences*, 22, 1-18.
- Lawson, R. (1991) "Patterns of Tourist Expenditure and Types of Vacation Across the Family Life Cycle", *Journal of Travel Research*, 29, 4, 12-18.
- Lehmann, D.R. (1993) *Investigación y Análisis de Mercado*, Ceca, México.

- Leones, J., Colby, B. and Crandall, K. (1998) "Tracking Expenditures of the Elusive Nature Tourists of Southeastern Arizona", *Journal of Travel Research*, 29, 3,. 32-36.
- Louviere, J.J. and Hensher, D.A. (1983). "Using Discrete Choice Models with Experimental Design Data to Forecast Consumer Demand for a Unique Cultural Event", *Journal of Consumer Research*, 10, December, 348-361.
- Mak, J. and Moncur, J.E.T. (1979) "The choice of Journey Destinations and Length of Stay: A Micro Analysis", *The Review of Regional Studies*, 10, 2, 38-48.
- Melenberg, B. and Van Soest, A. (1996) "Parametric and Semi-parametric Modelling of Vacation Expenditures", *Journal of Applied Econometrics*, 11, 59-76.
- Mergoupis, T. and Steuer, M. (2003) "Holiday Taking and Income", *Applied Economics*, 35, 269-284.
- Middleton, V. (1995) "Managing the Marketing Mix", in Witt, S. F. and Moutinho, L., eds., *Tourism Marketing and Management Handbook*, Prentice Hall, Hertfordshire.
- Mieczkowski, Z. (1990) "World Trends in Tourism and Recreation", en American University Studies Series XXV Geography, vol. 3. Peter Lang, New York.
- Miller, J.R. and Hay, M.J. (1981). "Determinants of Hunter Participation: Duck Hunting in the Mississippi Flyway", *American Journal of Agricultural Economics*, 63, 677-684.
- Moutinho, L. (1987) "Consumer Behaviour in Tourism", *European Journal of Marketing*, 21, 10, 1-44.
- Moutinho, L. and Trimble, J. (1991) "A Probability of Revisitation Model: The Case of Winter Visits to the Grand Canyon", *The Service Industries Journal*, 11, 4, 439-457.
- Mudambi, R. and Baum, T. (1997) "Strategic Segmentation: An Empirical Analysis of Tourist Expenditure in Turkey", *Journal of Travel Research*, 36, 1, 29-34.
- Mules, T. (1998) "Descomposition of Australian Tourist Expenditure", *Tourism Management*, 19, 3, 267-271.
- Neter, J., W. Wasserman, and M. Kutner (1985): *Applied Linear Statistical Models: Regression Analysis of Variance and Experimental Design*, Irwin.
- Nogawa, H., Yamaguchi, Y. and Hagi, Y. (1996) "An Empirical Research Study on Japanese Sport Tourism in Sport-for-All Events: Case Studies of a Single-Night Event and a Multiple-Night Event", *Journal of Travel Research*, Fall, 46-54.
- Opaschowski, H.W. (1990) "¿Turismo de Masas o Turismo a Medida? Límites Económicos, Ecológicos and Psicológicos", *Papers de Turisme*, 4, 68-80.
- Oppermann, M. (1995) "Travel Life Cycle", *Annals of Tourism Research*, 22; 535-552.
- Parker, W.M. (1976) *The Sociology of Leisure*. George Allen & Unwin, London.
- Perdue, R.R. (1986) "Traders and Nontraders in Recreational Destination Choice", *Journal of Leisure Research*, 18, 1, 12-25.
- Peterson, G. L., Dwyer, J.F. and Darragh, A. J. (1983) "A Behavioral Urban Recreation Site Choice Model", *Leisure Sciences*, 6, 1, 61-81.
- Plog, S.C. (1994) "Developing and Using Psychographics in Tourism Research" in *Travel, Tourism and Hospitality Research*, John Wiley & Sons, New York.
- Riera, A. (2000) "Modelos de Elección Discreta y Coste del Viaje. Los Espacios Naturales Protegidos en Mallorca", *Revista de Economía Aplicada*, 8, 24, 181-201.
- Ryan, C. (1995) *Researching Tourist Satisfaction: Issues, Concepts and Problems*, Routledge, London.
- S.G.T. Informe de la Secretaría General de Turismo, (1989a) "Vacaciones de los Españoles en 1987", *Estudios Turísticos*, 102, 1-36.
- S.G.T. Informe de la Secretaría General de Turismo, (1989b) "Los Viajes de los Españoles al Extranjero en 1987", *Estudios Turísticos*, 102, 37-73.

- S.G.T. Informe de la Secretaría General de Turismo, (1989c) "El Gasto en Turismo de los Españoles", *Estudios Turísticos*, 102, 113-138.
- S.G.T. Informe de la Secretaría General de Turismo, (1992) "Las Vacaciones de los Españoles en 1992", *Estudios Turísticos*, 116, 87-112.
- S.G.T. Informe de la Secretaría General de Turismo, (1993) "Las Vacaciones de los Españoles", *Estudios Turísticos*, 119-120, 153-177.
- Schroeder, H.W. and Louviere, J. (1999) "Stated Choice Models for Predicting the Impact of User Fees at Public Recreation Sites", *Journal of Leisure Research*, 31, 3, 300-324.
- Seaton, A.V. and Palmer, C. (1997) "Understanding VFR Tourism Behaviour: The First Five Years of the United Kingdom Tourism Survey", *Tourism Management*, 18, 6, 345-355.
- Seddighi, H.R. and Theocharous, A.L. (2002) "A Model of Tourism Destination Choice: A Theoretical and Empirical analysis", *Tourism Management*, 23, 5, 475-487.
- Sigelman, L. and Zeng, L. (1999) "Analyzing Censored and Sample-Selected Data with Tobit and Heckit Models", *Political Analysis*, 8:2, 167-182.
- Silberman, J. (1985) "A Demand Function for Length of Stay: The Evidence from Virginia Beach", *Journal of Travel Research*, spring, 16-23.
- Smith, S.L.J. (1995). *Tourism Analysis: A Handbook*, Longman Group Limited, London.
- Smith, V.K. and Munley, V.G. (1978) "The Relative Performance of Various Estimators of Recreation Participation Equations", *Journal of Leisure Research*, 10, 3, 165-176.
- Spotts, D.M. and Mahoney, E.,M (1991) "Segmenting Visitors to a Destinations Region based on the Volume of their Expenditures", *Journal of Travel Research*, 29, 4, 24-31.
- Stopher, P.R. and Ergün, G. (1979) "Population Segmentation in Urban Recreation Choices", *Transportation Research*, 59-65.
- Taylor, Ch. E. and Knudson, D.M. (1976) "Area Preferences of Midwestern Campers", *Journal of Leisure Research*, spring, 39-48.
- Taylor, D.T., Fletcher, R.,R. and Clabaugh, T. (1993) "A Comparison of Characteristics, Regional Expenditures, and Economic Impact of Visitors to Historical Sites with Other Recreational Visitors", *Journal of Travel Research*, 32, 1, 30-35.
- Von Limburg, B. (1997) "Overnight Tourism in Amsterdam 1982-1993 – A Forecasting Approach", *Tourism Management*, 18, 7, 465-468.
- Walsh, R.G., John, K.H., McKean, J.R. and Hof, J.G. (1992) "Effect of Price on Forecasts of Participation in Fish and Wildlife Recreation: An Aggregate Demand Model", *Journal of Leisure Research*, 24, 2, 140-156.
- Wennergren, E.B. and Nielsen, D.B. (1968) "A Probabilistic Approach to Estimating Demand for Outdoor Recreation", Working paper, Utah State University.

**Table I. DETERMINANT FACTORS OF GOING ON HOLIDAY AND TOURISTS EXPENDITURES**  
(Standard errors in brackets)

Independent Variables	Heckit 1		Heckit 2		Heckit 3		Heckit 4	
	Eq. 1: <i>GH</i>	Eq. 2: <i>E</i>	Eq. 1: <i>GH</i>	Eq. 2: <i>E</i>	Eq. 1: <i>GH</i>	Eq. 2: <i>E</i>	Eq. 1: <i>GH</i>	Eq. 2: <i>E</i>
<b>Indiv-Dest. Charact.</b>								
Distance		0.022 <sup>a</sup> (0.004)		0.021 <sup>a</sup> (0.043)		0.023 <sup>a</sup> (0.050)		0.022 <sup>a</sup> (0.049)
Hotel				62.087 <sup>a</sup> (10.733)				69.861 <sup>a</sup> (11.617)
Campsite				0.035 (9.989)				2.422 (9.790)
Own apartment/villa				26.635 <sup>c</sup> (12.523)				42.772 <sup>b</sup> (13.217)
Rented apartment/villa				69.627 <sup>a</sup> (16.085)				82.213 <sup>a</sup> (16.210)
<b>Personal Restrictions</b>								
Income2	0.460 <sup>a</sup> (0.073)	9.577 (16.684)	0.460 <sup>a</sup> (0.073)	2.891 (17.462)				
Income3	1.005 <sup>a</sup> (0.109)	28.688 (24.362)	1.005 <sup>a</sup> (0.109)	18.736 (24.794)				
Income4	1.545 <sup>a</sup> (0.245)	48.551 (33.426)	1.545 <sup>a</sup> (0.245)	33.296 (34.188)				
Income5	1.392 <sup>c</sup> (0.590)	224.465 <sup>d</sup> (125.99)	1.392 <sup>c</sup> (0.590)	216.846 <sup>d</sup> (121.589)				
Household size	-0.071 <sup>b</sup> (0.023)	9.953 <sup>c</sup> (4.357)	-0.071 <sup>b</sup> (0.023)	9.14364 <sup>c</sup> (4.435)				
Children		6.353 (4.870)		7.557 (5.213)		7.331 (4.554)		8.472 <sup>d</sup> (5.006)
<b>Socio-demographic Characteristics</b>								
Age1					0.175 (0.114)	17.052 (19.306)	0.175 (0.114)	13.852 (18.925)
Age2					0.156 (0.095)	60.349 <sup>b</sup> (17.361)	0.156 (0.095)	47.839 <sup>b</sup> (15.615)
Age3					0.012 (0.096)	41.170 <sup>c</sup> (16.390)	0.012 (0.096)	31.289 <sup>c</sup> (15.226)
Marital status		51.642 <sup>a</sup> (9.298)		44.850 <sup>a</sup> (9.362)				
Education 2	0.415 <sup>a</sup> (0.079)	-6.877 (13.181)	0.415 <sup>a</sup> (0.079)	-11.835 (13.025)	0.553 <sup>a</sup> (0.081)	-11.081 (14.203)	0.553 <sup>a</sup> (0.081)	-16.922 (14.284)
Education 3	0.233 <sup>c</sup> (0.098)	-5.049 (13.890)	0.233 <sup>c</sup> (0.098)	-13.675 (13.896)	0.595 <sup>a</sup> (0.091)	-7.518 (14.272)	0.595 <sup>a</sup> (0.091)	-18.257 (14.516)
Size of city 2	0.036 (0.087)		0.036 (0.087)		0.098 (0.086)		0.098 (0.086)	
Size of city 3	0.244 <sup>b</sup> (0.087)		0.244 <sup>b</sup> (0.087)		0.307 <sup>a</sup> (0.085)		0.307 <sup>a</sup> (0.085)	
Size of city 4	0.373 <sup>a</sup> (0.104)		0.373 <sup>a</sup> (0.104)		0.484 <sup>a</sup> (0.101)		0.484 <sup>a</sup> (0.101)	
<b>Psychographic Factor</b>								
Favourable opinion	0.684 <sup>a</sup> (0.063)		0.684 <sup>a</sup> (0.063)		0.718 <sup>a</sup> (0.062)		0.718 <sup>a</sup> (0.062)	
Constant	-0.658 <sup>a</sup> (0.109)	24.559 (25.395)	-0.658 <sup>a</sup> (0.109)	29.377 (26.935)	-0.719 <sup>a</sup> (0.102)	72.594 <sup>b</sup> (26.931)	-0.719 <sup>a</sup> (0.102)	68.662 <sup>b</sup> (24.096)
Inverse Mills ratio		-46.736 <sup>d</sup> (24.297)		-52.950 <sup>c</sup> (24.480)		-56.569 <sup>c</sup> (23.437)		-60.883 <sup>b</sup> (23.218)
<b>Control Variable</b>								
Trip duration		0.639 <sup>c</sup> (0.270)				0.838 <sup>b</sup> (0.301)		
Adjusted R-squared		0.140		0.165		0.093		0.126
F-Snedecor		15.617 <sup>a</sup>		14.8207 <sup>a</sup>		13.310 <sup>a</sup>		13.614 <sup>a</sup>
McFadden $\rho$	0.206		0.206		0.155		0.155	
LR	418.355 <sup>a</sup>		418.355 <sup>a</sup>		311.879 <sup>a</sup>		311.879 <sup>a</sup>	
Schwarz B.I.C.	1162.03	6877.44	1162.03	6720.37	1207.71	6897.03	1207.71	6735.36
Log likelihood	-1116.55	-6832.07	-1116.55	-6664.72	-1169.79	-6862.13	-1169.79	-6690.15

a=prob<0.1%; b=prob<1%; c=prob<5%; d=prob<10%.

- [1] These authors test different models to analyse vacation expenditures. Among others, they propose a system of equation without imposing the not easy to justify *exclusion restrictions* of the Heckit model. We rely upon the Heckit model since our behavioural approach allows us to justify all the variables included in each equation, therefore the *exclusion restrictions* are treated naturally.
- [2] In both variables, “distance” and “number of days”, we assume a general effect, without discounting the existence of discount destinations or holidays with lower prices than those of other closer destinations (or with shorter journey times).
- [3] As can be seen,  $v_i$  is heteroskedastic, so to obtain consistent estimators of the standard errors we have to use a heteroskedasticity robust covariance-variance matrix.
- [4] With the exception of the variable *children*, which appears to be only significant at 10% in model 4.
- [5] In accordance with Davis & Mangan (1992), tourists expenditures rise steeply at first as income rises, but the rate of increase declines as saturation is approached.
- [6] These magnitudes correspond to model 2 (optimal according to Schwarz BIC). However, this order is maintained in all the models.
- [7] The complete results of this estimation are available from the authors.