AN ANALYSIS OF THE MICROECONOMIC DETERMINANTS OF TRAVEL FREQUENCY*

Joaquín Alegre

joaquin.alegre@uib.es

Llorenç Pou

llorens.pou@uib.es

Department of Applied Economics

Universitat de les Illes Balears

* The authors acknowledge financial support from SEC2002-15012 project. Corresponding author- Llorenç Pou: edifici Jovellanos-Campus UIB, C/Crta. Valldemossa km 7,5 s/n, 07122 Palma de Mallorca-Balearic Islands (Spain). Tlfn:+34971171319. Fax: +34971172389.

AN ANALYSIS OF THE MICROECONOMIC DETERMINANTS OF TRAVEL FREQUENCY

ABSTRACT

A critical factor in predicting the demand for tourism within a certain period of time is the number of trips individuals take. New tourists' behaviour shows a tendency toward more frequent travel. Nevertheless, the frequency of travel has received little attention in empirical literature. This paper uses household data to examine the determinants of the number of quarters with positive tourist expenditure within a year. The results highlight the relevance in travel frequency analyses of distinguishing between the participation decision and the frequency decision conditional on participation. Many socio-demographic variables only show explanatory power for the participation decision. The two most relevant factors by far in explaining each decision are the previous year tourism demand decisions (suggesting evidence of habit persistence in tourism decisions) and disposable income, although with an income elasticity below the unit.

Keywords: tourism demand, frequency of travel, habit persistence, household data.

JEL classification: C25, D12.

I. INTRODUCTION

Since the middle of the last century the tourist industry has undergone a sharp rise in growth. According to the World Tourism Organisation, the average annual revenue from international tourism during the 1980s and 1990s grew faster than revenue from both commercial services and exports of goods. Among the reasons that account for this trend since the 1950s, it is worth mentioning the extraordinarily high economic growth, general decrease in working hours, rise in the number of days' paid leave and high level of demographic expansion.

Although some authors predict that international tourism will maintain the same growth trend over the next few years (OECD, 2002; Papatheodorou and Song, 2005), the stagnating populations of developed countries could alter tourism flows. In this sense, it is particularly important to find out whether there are limits to the current tourism growth at a microeconomic level. At an individual level the demand for tourism can be broken down into three choices: the decision whether or not to travel (i.e. holiday participation), the number of selected trips (i.e. the frequency of travel) and tourist expenditure per trip, with the last two decisions being conditional on participation. Each decision might be affected by different sets of variables or indeed by the same set of variables but in a different way. For instance, Graham (2001) suggests that income and leisure availability might have a differing impact on holiday participation and the number of trips.

Analyses of holiday participation have been made by Hageman (1981), Van Soest and Koreman (1987), Melenberg and Van Soest (1996), Cai (1998), Hong, Kim and

Lee (1999), Fleischer and Pizam (2002), Mergoupis and Steuer (2003), Alegre and Pou (2004) and Toivonen (2004). Meanwhile Dardis et al., (1981), Hageman (1981), Van Soest and Kooreman (1987), Davies and Mangan (1992), Cai, Hong and Morrison (1995), Fish and Waggle (1996), Cai (1998), Hong, Kim and Lee (1999), and Coenen and van Eekeren (2003), among others, have also used household data to study the determinants of tourist expenditure. However, little attention has been given in literature to the determinants of the number of trips that are taken within a specific period of time, partly reflecting a lack of data. The empirical evidence is mainly descriptive, from studies of tourist profiles (European Commission, 1987; Romsa and Blenman, 1989; Bojanic, 1992; Opperman, 1995a and 1995b; Tourism Intelligence International, 2000a and 2000b). To the best of our knowledge, only studies by Hultkrantz (1995), Fish and Waggle (1996) and Hellström (2002) have estimated the determinants of the frequency of travel.

Vanhoe (2005) shows that the percentage of the population in certain European countries (Austria, Belgium, France, Germany, the Netherlands, Norway, Great Britain and Switzerland) that take at least one holiday a year (i.e. holiday participation) did not increase during the 1990s. Graham (2001) reaches the same conclusion for a longer period, spanning the 1970s to the early 1990s, for Great Britain, Germany, France and Holland. Interestingly, the values for holiday participation vary considerably among European countries, ranging from over 75% in Switzerland, Germany, Sweden and Norway to values of around 40% for Portugal and Ireland (European Commission, 1998). Leaving aside the effect of disposable income, the explanatory power of other variables, such as socio-demographic variables, labour-market participation or health status, help to explain the variance in participation among countries (European Commission, 1998; Mergoupis and Steuer, 2003).

On the other hand, the total number of per capita holidays (which takes into account travel participation and the frequency of travel) shows an increasing trend over the years in Great Britain, Germany, France and Holland (Graham, 2001; Tourism Intelligence International, 2000a, 2000b; Vanhoe, 2005). As with holiday participation, there is quite a big variance in the total number of per capita holidays among European citizens, ranging from an average frequency of 1.43 holidays in Finland to 0.40 in Portugal. Two conclusions can be reached: firstly, given the sluggish rise in holiday participation and the growth trend in the total number of holidays, the variable mainly responsible for explaining the increase in the total number of holidays is clearly the frequency of travel. Secondly, the varying frequency of travel among European countries calls for an analysis of its determinants.

The aim of this article is to study the microeconomic determinants of households' frequency of travel. For this purpose a national survey, the Spanish Family Expenditure Survey (similar to the American Consumer Expenditure and British Family Expenditure Surveys) is used. The Spanish Family Expenditure Survey (*Encuesta Continua de Presupuestos Familiares*, henceforth the ECPF) is a nationally representative survey that monitors the same households for two years. The ECPF collects disaggregated data on household expenditure and income, along with socio-demographic and labour-related information about the household members. The survey does not provide information on the number of trips taken by households, but on quarterly tourist expenditure. So our variable of the frequency of travel is measured by the number of quarters with positive expenditure within a year. Thanks to the availability of a rich set of information on household characteristics, an analysis can be made of the effects that work/leisure decisions, preferences, demographics and income all have on the number of quarters with positive tourist expenditure. The survey was available for the period 1987 to 1996, thus covering a whole business cycle.

Previous papers that have studied the determinants of the frequency of travel have not taken into account the integer value characteristic of the travel frequency variable (Hultkranz, 1995; Fish and Waggle, 1996). The exception is a study by Hellström (2002). This paper differs from previous literature on several grounds. Firstly, it analyses the determinants of the frequency of travel taking into account the fact that it is a discrete variable that can only take nonnegative integer values. Secondly, with the econometric model that is applied, a distinction can be made between the determinants of the decision whether or not to travel and the determinants of the number of quarters with positive tourist expenditure. Thirdly, this paper uses household data to test for the existence of habit persistence in tourism decisions. Finally, thanks to the availability of a dataset that offers a rich source of household information, the explanatory power of preferences and budget and time constraints on travel frequency can be tested.

The findings of this paper highlight how important it is in travel frequency analyses to distinguish between the participation decision and the frequency decision. In fact, the explanatory power of many socio-demographic variables is limited to the decision to participate. The two most relevant factors by far in explaining each decision are the previous year tourism demand decisions (suggesting evidence of habit persistence in tourism decisions) and disposable income, although with an income elasticity below the unit.

The paper is organised as follows. Section 2 presents the database, motivation and estimation issues of this study. Section 3 outlines the results. Finally, the last section contains the concluding remarks and policy-making implications.

II. THE DATABASE, MOTIVATION AND ESTIMATION ISSUES

The Database

As commented previously, the database used in this paper is the Spanish Family Expenditure Survey for the period 1987-1996. The ECPF, conducted by the Spanish Bureau of Statistics, is a rotating quarterly panel survey representative of the Spanish population. The survey provides detailed information on consumer expenditure, income, and the socioeconomic and demographic characteristics of Spanish households.

Each quarter 3,200 households are interviewed. From these, 12.5% are randomly replaced each quarter, so that each household is monitored for up to eight consecutive quarters. In this paper, only those households that answered the survey for the whole eight quarters were considered, leading to a sample of 8,318 households. There were two reasons for this filter: firstly, working with four-quarter periods (i.e. one year) avoids the distorting effects of seasonality and, secondly, by focusing on those households that answered the survey for eight quarters, it was possible to examine whether tourism decisions taken the previous year affect current year decisions. In other words, monitoring the same households for two years makes it possible to test for the existence of habit persistence in the demand for tourism, as shown by Dynan (2000) for a general model of consumption.

Information from each household was summarized as follows. For ease of understanding, let us suppose that a household was interviewed in the years 1987 and 1988. Information was used from the last four quarters (i.e. the year 1988) relating to the explained variable (the number of quarters with tourist expenditure) and all the explanatory variables. The only exception was the variable for the number of quarters with tourist expenditure the previous year, which was constructed using the information from the first four quarters, i.e. those corresponding to the year 1987 in our example.

The ECPF records quarterly household expenditure on hotel stays and package holidays. Although it does not include all leisure-related tourist travel (for example, it excludes stays at second homes or homes owned by friends or relatives), a household is considered to have travelled if positive expenditure is recorded for either of those two categories during that quarter. Because of the ECPF's quarterly structure and the fact that our reference period covers one year (four quarters), our dependent variable (the frequency of travel) measures the number of quarters per year with positive tourist expenditure. Consequently, it ranges from zero (if no tourist trip is made throughout the entire year) to a maximum of four (if tourist expenditure is recorded in each quarter). As for the independent variables, they are defined in Table 1, while Table 2 shows summary statistics for the variables used in this study.

[INSERT TABLE 1 ABOUT HERE]

[INSERT TABLE 2 ABOUT HERE]

Motivation

From the 8,318 observations (households) available in our sample, 6,113 (73.49% of the sample) involved no travel during the year under analysis. The remaining 2,205 observations corresponded to households that travelled at least once, showing the expected decreasing pattern in the number of quarters with positive tourist expenditure (see Table 3).

[INSERT TABLE 3 ABOUT HERE]

When the sample is disaggregated according to certain household characteristics, there is considerable heterogeneity in the frequency of household travel. Table 4 and Figure 1 show the frequency distribution of the number of quarters in which travel occurred by income and age groups, respectively. As for the descriptive explanatory power of the disposable income variable (see Table 4), when a comparison of the percentage of households that do not travel is made by income quartiles, it is seen that the higher the income bracket, the lower the percentage of households that do not travel. This evidence is consistent with previous surveys, which point to a lack of money as being the main motive for not going on holiday (European Commission, 1987, 1998; Tourism Intelligence Information, 2000a, 2000b). Conditional on travel participation, the descriptive evidence in Table 4 also points to a positive association between the number of quarters with tourist expenditure and income.

[INSERT TABLE 4 ABOUT HERE]

Another interesting comparison is an analysis of the frequency of travel by age intervals. A substantial amount of literature relates many tourism decisions with the family life-cycle (Zimmermann, 1982; Lawson, 1989, 1991; Romsa y Blenman, 1989; Bojanic, 1992; Oppermann, 1995 a, 1995b; Collins and Tisdell, 2001, among others). Interestingly, the percentage of households that do not travel by age intervals describes a U-shape, where the lowest percentage corresponds to households aged between 35 and 44 (see Figure 1). In contrast, the distribution by age intervals of those households that do travel follows a hump-shaped pattern, peaking for the 35-44 age group. Thus the descriptive evidence presented here suggests that the effect of the independent variables (in this case the age variable) might differ, depending on the type of decision that is made: i.e. whether or not to travel, and how many quarters in which to travel conditional on participation. Obviously other independent variables are also associated with age and therefore with the family life-cycle, such as the size of the family, labour-market participation, disposable income etc. An empirical estimation is therefore needed to disentangle the partial effect of each independent variable.

[INSERT FIGURE 1 ABOUT HERE]

Figure 2 shows the evolution of the frequency of travel for the period 1987 to 1996, distinguishing between participation, the mean number of quarters with tourist expenditure per household (taking into account the whole sample), the mean number of quarters with tourist expenditure conditional on participation, and the percentage of households with recorded tourist expenditure during more than one quarter a year. Taking a base value of 100 for the year 1987, the mean number of quarters with tourist expenditure per household had increased by 30% by the end of the period for the whole sample group. The percentage of households with recorded tourist expenditure during at least one quarter and the percentage with recorded tourist expenditure during more than one quarter also followed this trend. That is, both series explain the increase in the average number of quarters with tourist expenditure by those households that did travel. Interestingly, the series that showed the highest growth rate corresponded to households with recorded tourist expenditure during more than one quarter per year. Lastly, Figure 2 shows that all the series follow the same evolution as average household income.

[INSERT FIGURE 2 ABOUT HERE]

Estimation Issues

As commented above, the dependent variable (the number of quarters with tourist expenditure) only takes non-negative integer values, y=0,1,…, 4, where y is measured in natural units. Its distribution is right skewed because it comprises a large proportion of zeros and a small proportion of households that travel during several quarters. This distribution implies that conventional OLS estimation techniques are inappropriate

(Long, 1997). In this context, count data models are a natural starting point for estimating the frequency of travel and, consequently, for explaining household's variability in terms of a set of explanatory variables. Although count data models have not been extensively applied to the demand for tourism, some authors (Ozuna and Gomez (1995), Gurmu and Trivedi (1996), Haab and McConnell (1996), among others) have applied them to the demand for recreation.

The simplest count data model is based on a Poisson distribution. In a basic Poisson regression model (PM), the number of events *y* (such as the number of quarters when at least a trip is made) corresponding to household *i* follows a Poisson distribution, with a conditional mean λ that is dependent on household characteristics, x_i :

$$\lambda_i = E(y_i | x_i) = e^{x_i \beta} \qquad [1]$$

And the probability that household *i* travels *y* times, given *x,* is:

$$\Pr(y_i | x_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} \qquad [2]$$

One particular feature of a Poisson distribution is the fact that its mean and its variance are both equal to its one parameter λ, i.e. $E(y_i | x_i) = Var(y_i | x_i) = \lambda_i$. However, because of its skewed distribution, count data very often displays "overdispersion", meaning that the conditional variance is larger than the conditional mean. This is the case with our sample (see Table 2). Overdispersion has similar qualitative consequences to heterocedasticity in linear regression models: the standard errors of β are biased (Cameron and Trivedi, 1998).

Overdispersion can be caused by unobserved heterogeneity, a high percentage of zeros or both. Unobserved heterogeneity can be handled by either enhancing the set of regressors in the mean function or by allowing the variance term to depend on further parameters. This second possibility is the underlying proposal behind Negative Binomial regression models (NBM), where parameter α is added so that the conditional variance can now exceed the conditional mean (Cameron and Trivedi, 1986). The increased variance in NBM results in substantially larger probabilities for small counts. Now the variance will be $Var(y_i|x_i) = \lambda_i + \alpha\lambda_i^2$, where α can be tested using the conventional *t*-test. If α is not significantly different from zero, the Negative Binomial model is reduced to a Poisson regression model.

As commented above, one of the usual characteristics of count data is the presence of two broad groups of observations: zero counts and positive counts. In a decision process, zero counts correspond to those households that decide not to travel and positive counts to those households that do actually travel. Poisson and Negative Binomial models assume that zeros are generated by the same process as positive observations and that, consequently, they share the same set of parameters. In other words, neither extracts information about the participation decision from the zeros in the data. As pointed out by Jones (2000), zero counts frequently have a special significance: they tell us about the participation decision in the underlying economic model.

The count data hurdle model introduced by Mullahy (1986) is a suitable model for analysing "excess zeros". Unlike Poisson and Negative models, which assume that all individuals are positively likely to travel, the hurdle model (HM) assumes that the statistical process governing households with zero counts and households with positive counts might be different. In other words, the set of variables that affects the decision whether or not to travel (the participation decision) might be different from the set that affects the decision how often to travel (the frequency decision). In addition, the same variables might affect the two decisions in different ways. The hurdle model can therefore be construed as a two-step approach (splitting mechanism) to analysing the decision-making process behind the choice to make a certain number of trips in a specific period of time:

(1) The first step is modelled using a binary choice model, which estimates the probability that an individual does not travel within the observed period,

$$\Pr(y_i = 0 | x_i) = f_1(0) \qquad [3]$$

where f_1 is the probability distribution function of not travelling and $1\text{-}f_1(0)$ is the probability of crossing the hurdle.

(2) The second step is modelled as a truncated-at-zero count model for positive observations,

$$\Pr(y_i > 0 | x_i) = f_2 \frac{1 - f_1(0)}{1 - f_2(0)} \qquad \text{for } y_i > 0 \qquad [4]$$

where $f_2(y_i)$ is the probability distribution function that governs the process once the hurdle has been passed, $1-f_1(0)$ gives the probability of crossing the hurdle, and $1-f_2(0)$ is the truncation normalization for f_2 so that the probabilities sum to one.

The hurdle model can be specified in several ways by choosing different probability distributions for f_1 and f_2. Usually f_1 is specified as a logistic distribution. For f_2, two options are contemplated in this paper: a Poisson distribution, which gives a Poisson Hurdle regression model (PHM), and a Negative Binomial distribution, which gives a Negative Binomial Hurdle regression model (NBHM). As commented above for PM and NBM, the main difference between the PHM and the NBHM is the fact that the latter allows for unobserved heterogeneity through parameter α in the error term (Long, 1997; Cameron and Trivedi, 1998).

The count data models were estimated using the maximum likelihood method with robust standard errors by means of the STATA 8.0 Programme. To choose the model that best fitted the data, the values of the models' log-likelihood functions were compared. The Akaike Information Criterion (AIC) was also used, defined as AIC=-2LogL+2K, where LogL is the value of the model's log-likelihood function and K the number of estimated parameters. Models with higher log-likelihood values and smaller AIC values are preferable (Cameron and Trivedi, 1998). On the other hand, as commented above, overdispersion can arise from different sources: unobserved heterogeneity, excess zeros or both. Thus different sources of overdispersion must be tested for. A natural way to test for the effect of unobserved heterogeneity is to use a *t*-test for the significance of coefficient α in both the NBM and the NBHM. A statistically significant α implies that unobservable heterogeneity also accounts for overdispersion. In this case, the NBM (NBHM) is superior to the PM (PHM). In order to analyse whether a Hurdle regression model characterizes the data generation process better than a PM or a NBM, we used the following likelihood-ratio test, which in the case of a PM versus a PHM can be expressed as follows:

$$\rho = -2\left(LLF_{poisson} - LLF_{logit} - LLF_{\text{truncated poisson}}\right) \qquad [5]$$

where LLF represents the log-likelihood function value. Statistic ρ is chi-square distributed. Rejection of the null hypothesis supports the use of a Hurdle model.

III. RESULTS

This section presents the results of the estimates. The general model that was estimated took the form:

Number of quarters with tourism expenditure$_t$ = f (SD$_t$, labour$_t$, number of quarters with tourism expenditure$_{t-1}$, income$_t$)

where SD represents socio-demographic variables and household characteristics, labour represents labour-related variables and income represents disposable income.

Table 5 shows the results of the different tests that were used to choose the count regression model that best fitted the data. Based on the log-likelihood criterion, the Poisson regression model takes a lower value than the Negative Binomial regression model, indicating the presence of unobserved heterogeneity in the data. The statistical significance of parameter α at the 1% level also suggests that the NBM is superior to the PM. When P and NB models are compared with PH and NBH models, respectively, the two latter models clearly show a higher log-likelihood value. The Akaike Information criterion and likelihood ratio test for comparing Poisson/Negative Binomial models with hurdle models also rank them in a similar order. That is, the model-selection tests provide evidence that a splitting mechanism that distinguishes households that do not travel from ones that do best suits the data.

A final test in the model-selection process is to check whether unobserved heterogeneity still accounts for dispersion once we allow for the splitting mechanism. This was tested by checking the significance of parameter α in the NBH model. The results did not reject the null hypothesis of no significance of parameter α (*p-value* of 0.864). This explains why the log-likelihood value of the NBH model was equal to that achieved by the PH model. Interestingly, as expected, not including the number of quarters with tourist expenditure the previous year as an explanatory variable substantially reduced the estimates' goodness of fit and increased the unobserved heterogeneity factor (see row 2 of Table 5). Thus, the final model that was estimated was a Poisson Hurdle regression model, in which the number of quarters with tourist expenditure the previous year was included as a regressor. Its estimation results will be discussed in the remaining part of this section.

[INSERT TABLE 5 ABOUT HERE]

Tables 6 and 7 present the results of the most parsimonious specification of a Poisson Hurdle model for the number of quarters with tourist expenditure variable. Table 6 contains the estimation results of the participation decision, while Table 7 shows the results of the frequency decision. Due to the estimates' non-linearity, as well as showing the value of the coefficient and *t-statistic* for each variable, its marginal effect is also presented. The marginal effect is construed as the change in the dependent variable when the independent variable changes by one unit. In the case of the income variable, the income elasticity is reported instead.

Prior to commenting on each table separately, a comparison of the results of Tables 6 and 7 shows that fewer variables are significant for the travel frequency decision, suggesting that most of the explanatory power of the independent variables stems from its effect on zero counts. For the participation decision, the estimated coefficients show the expected signs. They highlight the trade-off from leisure/work decisions, the effect of time constraints and the consideration that tourism is a "normal" good. Interestingly, the variables that are statistically significant in the frequency decision were also significant in the participation decision, and with the same sign.

With regard to the participation decision, the estimation results show that income and the number of quarters with tourist expenditure the previous year are the most important factors in determining the probability of travelling (see Table 6). Consistent with the hypothesis that tourism is a "normal good", the coefficient on income is positive, meaning that the probability of travel increases as the level of income goes up. The income elasticity value is below the unit, 0.694, in consonance with previous literature that uses microdata (Hageman, 1981; Cai, 1998; Hong, Kim and Lee, 1999; Fleisher and Pizam, 2002; Mergoupis and Steuer, 2003; Alegre and Pou, 2004). The number of quarters with tourist expenditure the previous year is also highly significant. Its positive coefficient implies that travelling the previous year increases the probability of travelling this year, therefore suggesting the existence of habit persistence. The marginal effect of the lagged travel frequency is 0.1947. That is, an additional quarter with tourist expenditure the previous year increases the probability of travelling the following year by 19.47%.

The estimated effects of a family's composition and labour-market participation tally with the expected time constraint effect. That is, the bigger the family or the higher the number of earners, the lower the probability of taking a trip. For instance, compared with a one-person household, being a childless couple decreases the probability of travelling by 11.49%. For the number of earners, the marginal effects of this variable, -0.0298, show that an additional earner increases the leisure constraints.

As for household preferences, the results are also consistent with previous literature. The estimates reject the null hypothesis of no relationship between age and the probability of travel, as also detected in Cai (1998), Mergoupis and Steuer (2003), Alegre and Pou (2004) and Toivonen (2004). The estimation results corroborate the descriptive evidence from Figure 1, obtaining a non-linear relationship between age and the probability of travel that takes an inverted-U shape with a maximum probability at the age of 40. The dummy variables for the level of education are also statistically significant. Compared with the reference group (household heads with less than a primary school education), all the education levels show a higher probability of travel: the marginal effects are 0.0729 for the primary education level, 0.1281 for the secondary education level and 0.0967 for higher education levels. Living in a big city and owning at least one car also imply a positive effect on the probability of travel, with marginal effects of 0.0420 and 0.0457, respectively. On the other hand, the unemployment and home tenure variables, which can be associated with a precautionary motive (Deaton, 1992), show the expected negative coefficient, reducing the probability of travel by 2.09% and 3.06%, respectively. The remaining variables, i.e. living in municipalities with between 10,000 and 500,000 inhabitants, being a female, having a mortgage, and a non-linear causal relationship for income did not show statistically significant effects.

[INSERT TABLE 6 ABOUT HERE]

As for the frequency of travel conditional on participation (see Table 7), the estimation results once again show that the lagged frequency of travel and disposable income are the most relevant factors. Both are statistically significant at the 0.1% confidence level. In the case of the lagged number of quarters with tourist expenditure, the results indicate the existence of habit persistence in this second step too. One extra quarter with tourist expenditure the previous year increases the number of current year quarters with recorded tourist expenditure by 42.37%. For disposable income, the positive coefficient points to the consideration of travel frequency as a normal “good”. Interestingly, its income elasticity is again below the unit, 0.152. Thus big increases in total tourist expenditure should not be expected to be caused by an increase in the frequency of travel in the context of moderate income increases. As with the participation decision, the estimates rejected a non-linear causal effect by income on the number of quarters with tourist expenditure.

For the remaining variables, time constraints associated with family size affect the frequency decision in the expected manner: couples without children, couples with children, and couples with adults all have fewer quarters with positive tourist expenditure than one-person households. Their marginal effects were -0.1927, -0.2327, -0.1725, respectively. This effect is also obtained for the number of earners: having an additional earner reduces the frequency of travel by 9.65%. The age variable is also statistically significant, with a positive sign. Unlike the participation decision, however, the null hypothesis for the absence of a non-linear age effect is not rejected. Interestingly, as commented above, the coefficients on family size, the number of earners and age show the same sign for both the frequency of travel and participation decisions. The remaining sociodemographic variables, however, were not statistically significant. The estimation results therefore show that the set of variables that determines each decision is not the same, corroborating the validity of the splitting mechanism for our sample. In this sense, our results suggest that applying the same process to the participation decision and frequency of travel decision might lead to inconsistent estimates and to economic misinterpretations.

[INSERT TABLE 7 ABOUT HERE]

IV. CONCLUSIONS

Statistics from developed countries show that the component of the average number of trips per individual that has grown the most steadily over the last decades is the number of trips conditional on travel, while the percentage of the population that travels seems to be reaching a threshold. Part of the increase in the number of trips per individual is explained in literature through the habits of new tourists. Despite its relevance for tourism demand, the frequency of travel has received little attention in empirical literature, partly due to lack of databases with information in the tourists' countries of origin.

Using a Spanish national survey, this paper has analysed the microeconomic determinants of the number of yearly quarters with positive tourist expenditure. By applying different count data models, it was tested whether the participation decision and the frequency of travel decision conditional on travel follow the same process. Furthermore, the availability of a national survey with both a high number of observations and information on the same household for a two- year period made it possible to examine how preferences, time and budget constraints, and habit persistence affect the frequency of travel by households.

The results of this paper show the relevance in tourism demand analyses of distinguishing between the travel participation decision and the frequency of travel decision conditional on participation. In fact, most socio-demographic variables only have explanatory power in the participation decision. Interestingly, however, all the variables that affect the frequency of travel decision also explain the participation decision. The two most relevant factors by far in explaining each decision for Spanish households are the number of quarters with tourist expenditure the previous year and disposable income. In this sense, the paper has found evidence that habit plays a role in determining both travel decisions. For disposable income, the results corroborate its expected positive coefficient. Moreover, as well as being considered a “normal” good, the estimation results show income elasticity values below the unit for both decisions. The income elasticity values are robust to the count data model and to the financial measure that is applied.

Several implications can be drawn from this study. Firstly, as commented above, in many developed countries the percentage of the population who travel has remained nearly constant. Consequently the future evolution of the population that travel will be more dependent on population growth, and the future trend in the total number of trips will mainly be explained by the frequency of travel by those individuals that already travel. Policy decisions aimed at promoting tourism should therefore mainly focus on the frequency of travel. Secondly, the below-unit values for income elasticity that were obtained for both participation and frequency decision show tourism to be a “necessity” for Spanish households. Consequently, big increases in the number of quarters with tourist expenditure cannot be expected as a result of moderate changes in income. Thirdly, the detection of habit persistence in the demand for tourism suggests the latter’s stability over time and a tendency for the frequency of travel to grow over time. Fourthly, as long as the frequency of travel is endogenous in relation to other tourism demand variables, such as the length of stay at destinations and daily expenditure per trip, steady changes in the frequency of travel should also permanently affect these other tourism variables. Fifthly, the independent variables’ differing explanatory power in the travel participation and frequency decision highlights the need to pinpoint different marketing targets depending on the chosen decision.

To sum up, the results of this study highlight the fact that unlike what is usually assumed, particularly with aggregate data, the frequency of travel is not an exogenous variable. Just as tourism literature has shown in the case of tourist expenditure per trip, this paper has demonstrated that the frequency of travel is also influenced by household preferences and time and budget constraints. Overall, the results of this paper point to the need for a more detailed analysis of the demand for tourism, where the frequency of travel is included as a key factor in facilitating a more accurate explanation of variability in the demand. From an empirical viewpoint, these challenges call for considerable efforts gathering data in the countries of origin.

REFERENCES

Alegre, J., and Ll. Pou (2004). "Microeconometric determinants of the probability of tourism consumption," Tourism Economics, 10: 125-144.

Bojanic, D.C. (1992). "A Look at a Modernized Family Life Cycle and Overseas," Travel. Journal of Travel & Tourism Marketing, 1: 61-79.

Cai, L. A. (1998). "Analyzing Household Food Expenditure Patterns on Trip and Vacations: a Tobit Model," Journal of Hospitality & Tourism Research, 22: 338-358.

Cai, L. A., G.S. Hong, and A.M. Morrison (1995). "Household Expenditure Patterns for Tourism Product and Services," Journal of Travel & Tourism Marketing, 4: 15-40.

Cameron, A.C., and P.K. Trivedi (1998). *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.

Cameron, A.C., and P.K. Trivedi (1986). "Econometric Models Based on Count Data: Comparisons and applications of Some Estimators," Journal of Applied Econometrics, 1: 29-53.

Coenen, M., and L. van Eekeren (2003). "A Study of the Demand for Domestic Tourism by Swedish Households using a two-staged Budgeting Model," Scandinavian Journal of Hospitality and Tourism, 3: 114-133.

Collins, D., and C. Tisdell (2001). "Age-Related Lifecycles. Purpose Variations," Annals of Tourism Research, 29: 801-833.

Dardis, R., F. Derrick, A. Lehfeld, and K.E. Wolfe (1981). "Cross-Section Studies of Recreation Expenditures in the United States," Journal of Leisure Research, 13: 181-194.

Davies, B., and J. Mangan (1992). "Family Expenditure on Hotels and Holidays," Annals of Tourism Research, 19: 691-699.

Dynan, K.E. (2000). "Habit Formation in Consumer Preferences: evidence from Panel Data," American Economic Review, 90: 391-406.

European Commission (1998). Facts and Figures on the Europeans on Holidays, 1997-1998. Executive Summary; A Eurobarometer survey carried out on behalf of the European Commission, Directorate General XXIII.

European Commission (1987). The European and their Holidays. Brussels.

Fish, M., and D. Waggle (1996). "Current Income versus Total Expenditure Measures in Regression Models of Vacations and Pleasure Travel," Journal of Travel Research, 35: 70-74.

Fleisher, A., and A. Pizam (2002). "Tourism Constraints among Israeli Seniors," Annals of Tourism Research, 29: 106-123.

Graham, A. (2001). "Using Tourism Statistics to Measure Demand Maturity." In Tourism Statistics. International Perspectives and Current Issues, edited by J.J. Lennon, London: Continuum. Pp. 199-214

Gurmu, S., and P.K. Trivedi (1996). "Excess Zeros in Count Models for recreational Trips," Journal of Business & Economic Statistics, 14: 469-477.

Haab, T.C., and K.E. McConnell (1996). "Count Data Models and the Problem of Zeros in Recreation Demand Analysis," American Journal of Agricultural Economics, 78: 89-102.

Hageman, R.P. (1981). "The Determinants of Household Vacation Travel: Some Empirical Evidence," Applied Economics, 13: 225-234.

Hellström, J. (2002). "An Endogenously Stratified Bivariate Count Data Model for Household Tourism Demand," Umea Economic Studies nº 583.

Hong, G-S., S. Kim, and J. Lee (1999). "Travel Expenditure Patterns of Elderly Households in the US," Tourism Recreation Research, 41: 43-52.

Hultkranz, L. (1995). "On the Determinants of Swedish Recreational Domestic and Outbound Travel, 1989-93," Tourism Economics, 1: 119-145.

Jones, A.M. (2000). "Health Econometrics." In Handbook of Health Economics, edited by J.P. Newhouse and A.C. Culyer, Amsterdam: North-Holland ch. 7.

Lawson, R. (1991). "Patterns of Tourist Expenditure and Types of Vacation Across the Family Life Cycle," Journal of Travel Research, 29: 12-18.

Lawson, R. (1989). "Family Life Cycle." In Tourism Marketing and Management Handbook, edited by S.F. Witt and L. Moutinho, Hertfordshire: Prentice Hall, U.K. Pp. 147-151.

Long, J. S. (1997). "Regression Models for Categorical and Limited Dependent Variables." In Advanced Quantitative Techniques in the Social Sciences, Vol. 7, edited by Melenberg, B., and A. Van Soest, Sage Publications.

Long, J. S. (1996). "Parametric and Semi-Parametric Modelling of vacation Expenditures," Journal of Applied Econometrics, 11: 59-76.

Mergoupis, T., and M. Steuer (2003). "Holiday Taking and Income," Applied Economics, 35: 269-284.

Mullahy, J. (1986). "Specification and Testing of Some Modified Count data Models," Journal of Econometrics, 33: 341-365.

OECD (2002). "Household Tourism Travel: Trends, Environmental Impacts and Policy Responses." OECD Sector Case Studies Series, Paris.

Oppermann, M. (1995a). "Family Life Cycle and Cohort Effects: A Study of Travel Patterns of German Residents," Journal of Travel & Tourism Marketing, 4: 23-44.

Oppermann, M. (1995b). "Travel Life Cycle," Annals of Tourism Research, 22: 535–552.

Ozuna, T., and I.A. Gomez (1995). "Specification and Testing of Count Data Recreation Demand Functions," Empirical Economics, 20: 543-550.

Papatheodorou, A., and H. Song (2005). "International Tourism Forecatss: Time-series Analysis of World and Regional Data," Tourism Economics, 11: 11-23.

Romsa, G., and M. Blenman (1989). "Vacation Patterns of the Elderly German," Annals of Tourism Research, 16: 178-188.

Toivonen, T. (2004). "Changes in the Propensity to Take Holiday Trips Abroad in EU Countries between 1985 and 1997," Tourism Economics, 10: 403-417.

Tourism Intelligence International (2000 a). How Germans Will Travel 2005. Bielefeld.

Tourism Intelligence International (2000 b). How British Will Travel 2005, Bielefeld.

Van Soest, A., and P. Kooreman (1987). "A Micro-Econometric Análisis of Vacation Behaviour," Journal of Applied Econometrics, 2: 215-226.

Vanhoe, N. (2005). The Economics of Tourism Destinations. Elsevier.

Zimmermann, C. A. (1982). "The Life Cycle Concept as a Tool for Travel Research," Transportation, 11: 51-69.

Table 1. Definition and Measurement of Independent Variables

VARIABLE	DEFINITION AND MEASUREMENT
Income	Log of real after-tax income of all household members
Earners	Number of earners
Unemployed	Household head unemployed (unemployed=1, else=0)
Retired	Household head retired (retired=1, else=0)
Age	Household head's age
Age squared	Household head's age squared
Place of residence	
Small (reference)	The municipality in which the household lives has fewer than 10,000 inhabitants (fewer than 10,000 inhabitants=1, else=0)
Medium	A municipality with more than 10,000 inhabitants and fewer than 500,000 inhabitants (between 10,000 and 500,000 inhabitants=1, else=0)
Big	A municipality with over 500,000 inhabitants (over 500,000 inhabitants=1, else=0)
Car	The household owns at least one car (owns a car=1, else=0)
Trfreq1	The number of quarters with positive tourist expenditure the previous year
Gender	(If household head is female=1, else=0)
Tenant	(If the house is rented=1, else=0)
Education	
Illiterate or with no education (reference)	(Household head with less than a primary school education=1, else=0)
Primary school education	(Household head with a primary school education=1, else=0)
Secondary school education	(Household head with a secondary school education=1, else=0)
Higher education (university)	(Household head with more than a secondary school education=1, else=0)
Family size	
One-person household (reference)	(Single household head=1, else=0)
Childless couples	(Married couples without children=1, else=0)
Couples with children	(Married couples with children up to 14=1, else=0)
Couples with adults	(Married couples with children over 14=1, else=0)
Year 1988 (reference)	Year of interview (if household is interviewed in year 1988=1, else=0)

Table 2. Summary Statistics

VARIABLE	MEAN	STANDARD DEVIATION	MAXIMUM	MINIMUM
Trfreq	0.347	0.727	4	0
Income	13.854	0.589	16.355	11.396
Earners	1.754	0.839	7	1
Unemployed	0.083	0.276	1	0
Retired	0.416	0.492	1	1
Age	53.55	14.44	85	25
Place of residence				
Small	0.210	0.407	1	0
Medium	0.426	0.494	1	0
Big	0.363	0.480	1	0
Car	0.723	0.447	1	0
Gender	0.162	0.369	1	0
Tenant	0.117	0.322	1	0
Education				
Illiterate or with no education	0.273	0.445	1	0
Primary school education	0.566	0.495	1	0
Secondary school education	0.091	0.288	1	0
Higher education (university)	0.068	0.252	1	0
Family size				
One-person household	0.079	0.269	1	0
Childless couples	0.159	0.365	1	0
Couples with children	0.192	0.393	1	0
Couples with adults	0.568	0.495	1	0

Table 3. Frequency Distribution of the Number of Quarters with Tourist Expenditure

Number of quarters	Frequency	Percentage	Cumulative percentage
0	6,113	73.49%	73.49%
1	1,512	18.17%	91.66%
2	499	5.99%	97.66%
3	147	1.76%	99.43%
4	47	0.56%	100.00%

Table 4. Frequency Distribution by Income Quartiles

Number of quarters with tourist expenditure	Q1	Q2	Q3	Q4
0	88.0	81.2	70.6	57.6
1	8.8	14.1	21.0	26.4
2	2.3	3.2	6.4	10.9
3	0.6	1.0	1.4	3.6
4	0.1	0.3	0.4	1.2

Note: Q1 represents 25% of the sample households with a lower income, while Q4 represents 25% of the households with higher financial resources.

Table 5. Model Selection Tests

	Poisson model	Negative Binomial Model	Truncated Poisson model	Truncated NB model
Log-likelihood	-5,786	-5,780	-5,646	-5,646
Log-likelihood (without $Tfreq_{t-1}$)	-6,303	-6,234	-6,045	-6,045
Akaike Information Criterion	-5,767	-5,761	-5,608	-5,608
Ln(α) [overdispersion test]		0.000		0.864
ρ test for truncation			ρ=268 (*p-value*=0.000)	ρ=280 (*p-value*=0.000)

Table 6. Travel Participation Decision
(the dependent variable takes a value of 1 if the household travelled and 0 otherwise).

	Coefficient	Standard error	Marginal effect
Primary-school	0.445	0.080 *	0.0729
Secondary-school	0.725	0.120 *	0.1281
Higher education	0.571	0.136 *	0.0967
Medium	–		
Big	0.234	0.059 *	0.0420
Childless couple	-0.564	0.140 *	-0.1149
Couple with children	-0.558	0.154 *	-0.1134
Couple with adults	-0.677	0.141 *	-0.1345
Gender	–		
Car	0.267	0.083 **	0.0457
Mortgage	–		
Renter	-0.180	0.092 ***	-0.0306
Unemployed	-0.122	0.115	-0.0209
Earners	-0.169	0.044 *	-0.0298
Age	0.058	0.017 **	0.0102
Age squared	-0.0005	0.0001 **	-0.00009
Trfreq1	1.103	0.042 *	0.1947
Income	0.841	0.079 *	0.694 (†)
Income squared	–		
Constant	-14.576	1.082 *	
N	8,318		
Log-likelihood	-3,915		
LR chi2 (23)	1,790.2		
Prob > chi2	0.000		
Pseudo R2	0.186		

Note: The results for yearly dummy variables are not reported. (–) corresponds to those variables that were not statistically significant when all the independent variables were included. *, ** and *** indicate significance at the 0.1%, 1% and 5% levels, respectively. † refers to the elasticity value.

Table 7. Frequency of Travel Decision (y>0).

	Coefficient	Standard error	Marginal effect
Primary-school	–		
Secondary-school	–		
Higher education	–		
Medium	–		
Big	–		
Childless couple	-0.206	0.128 ****	-0.1927
Couple with children	-0.244	0.148 ****	-0.2327
Couple with adults	-0.192	0.122 ****	-0.1725
Gender	–		
Car	–		
Mortgage	–		
Renter	–		
Unemployed	–		
Earners	-0.092	0.038 **	-0.0965
Age	0.008	0.002 **	0.0090
Age squared	–		
Trfreq1	0.427	0.024 *	0.4237
Income	0.328	0.060 *	0.232 (†)
Income squared	–		
Constant	-5.545	0.848 *	
N	2,205		
Log-likelihood	-1,737		
LR chi2 (14)	335.7		
Pseudo R2	0.088		

Note: The results for yearly dummy variables are not reported. (–) corresponds to those variables that were not statistically significant when all the independent variables were included. *, **, *** and **** indicate significance at the 0.1%, 1%, 5% and 10% levels, respectively. † refers to the elasticity value.

Figure 1. Frequency Distribution of the Number of Quarters with Tourist Expenditure by Age Groups

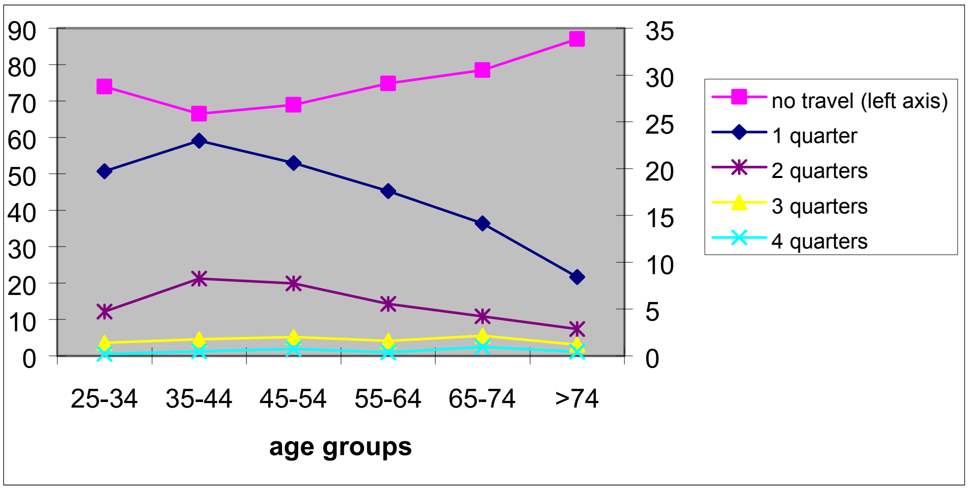
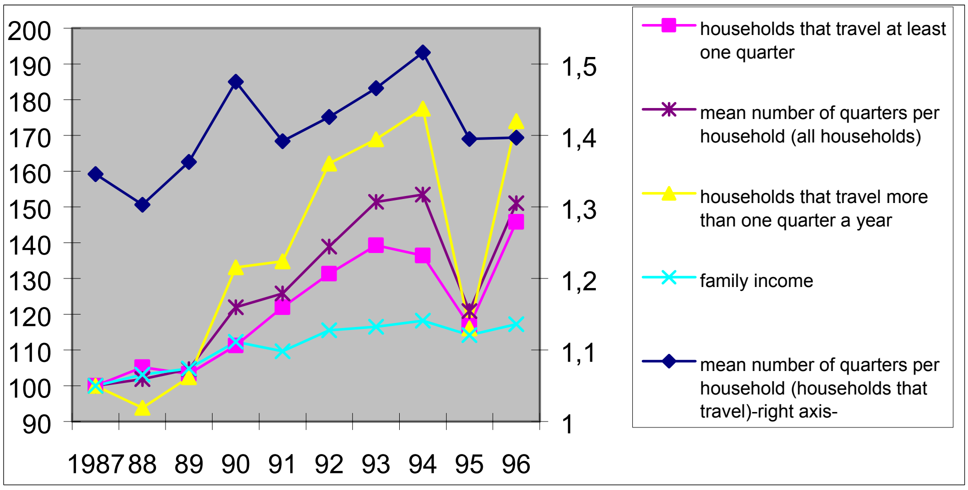


Figure 2. Travel Frequency Indicators, 1987-1996.



Note: Holiday participation, the mean number of quarters per household (whole sample), and the percentage of households that travelled more than one quarter a year and family income are all measured with a base value of 100 for the year 1987.