USING SIMULATED MAXIMUM LIKELIHOOD (ML) TO ESTIMATE A PROBIT MODEL WITH SELECTIVITY AND ENDOGENOUS VARIABLES WITH APPLICATION TO THE ANALYSIS OF THE INFLUENCE OF TRAVEL ORGANIZATION MODE ON LOW COST CARRIER'S (LCC'S) TOURIST DEMAND

Juan Muro
Cristina Suárez
María del Mar Zamora*

Universidad de Alcalá and Alcamétrica

March, 2009

Abstract

Asymmetries of information and risk aversion play a relevant role in tourism demand. Insufficiently informed and high risk aversion tourists are very prone to organize their leisure travel through a package tour supplied by a tourist agency. Yet informed and low risk aversion tourists tend to organize their holidays by their own using more and more the new communication and information technologies. This translates to tourism microeconomic models a problem that has rarely been addressed, the endogeneity of the travel organization mode variables. In this paper we deal with this relatively unanswered question and show how to estimate a model of low cost carrier (LCC) demand in which endogeneity of the travel organization mode variable is controlled for in a fairly simple way.

Keywords: LCC's demand, Tourism, Binary choice model with selectivity and endogenous variables, Two-step estimation in qualitative models
JEL classification: C25, L83

*Corresponding autor.
Facultad de Ciencias Económicas y Empresariales
Universidad de Alcalá
Plaza de la Victoria, 2
28802 Alcalá de Henares, Madrid (SPAIN)
Tfno. y fax: 91.885.42.01
e-mail: juan.muro@uah.es
e-mail: cristina.suarez@uah.es
e-mail: mariam.zamora@uah.es

USING SIMULATED MAXIMUM LIKELIHOOD (ML) TO ESTIMATE A PROBIT MODEL WITH SELECTIVITY AND ENDOGENOUS VARIABLES WITH APPLICATION TO THE ANALYSIS OF THE INFLUENCE OF TRAVEL ORGANIZATION MODE ON LOW COST CARRIER'S (LCC'S) TOURIST DEMAND

1. Introduction

Asymmetries of information and risk aversion play a relevant role in tourism demand. Insufficiently informed and high risk aversion tourists are very prone to organize their leisure travel through a package tour supplied by a tourist agency. Yet informed and low risk aversion tourists tend to organize their holidays by their own using more and more the new communication and information technologies. This translates to tourism microeconomic models a problem that has rarely been addressed, the endogeneity of the travel organization mode variables. Tourism micro models are often based on survey data and have a qualitative nature. Organization mode variables are usually qualitative too. While estimating probit models with selectivity is now a standard in applied analysis examples of endogeneity treatment in this type of models are very scarce. In this paper we deal with this relatively unanswered question and show how to estimate a model of low cost carrier (LCC) demand in which endogeneity of the travel organization mode variable is controlled for in a fairly simple way.

Besides its empirical nature the paper has two main folds. Firstly we show how to make appropriate inference in a probit model with selectivity through a two-step method. Here the relevance of obtaining a robust variance estimator, Murphy-Topel (1985), Hardin (2002), is

stressed. Secondly we present a ML estimation of our selectivity probit model with endogeneity. In this case a fairly simple procedure is outlined that builds on research by Lahiri-Schmidt (1978) and Greene (1997, 1998).

We organize the paper in the following way. In section 2 we describe our microeconometric model. We deal with two different approaches to estimate a selectivity probit model with endogeneity in sections 3 and 4. We present our empirical results on the impact of the travel organization mode on LCC demand in section 5 while section 6 concludes.

2. A microeconometric model to measure the impact of the travel organization mode on LCC demand.

Our model is

$$TO_i^* = \theta' Z_{2i} + \upsilon_i \,, \qquad [1]$$

$$Y_i^* = \beta' X_i + \delta_1 TO_i + u_i \,, \qquad [2]$$

$$C_i^* = \gamma' Z_{1i} + v_i, \qquad [3]$$

with $C_i = 1[\gamma' Z_{1i} + v_i > 0]$; $Y_i = 1[\beta' X_i + \delta_1 TO_i \; u_i > 0]$ C_i , $\forall C_i = 1$; Y_i is unobserved , $\forall C_i = 0$; $TO_i = 1[\theta' Z_{2i} + \upsilon_i > 0]$. Where Y_i=1 stands for a tourist who travels in a LCC versus a full service airline; C_i =1 for a tourist who travels by air versus road; TO_i =1 for a tourist who travels with a package tour; X, Z_1, Z_2 are determinant variables matrices and θ, β, γ, δ_1 are parameters. Under joint normality υ_i, u_i , v_i , are distributed as a trivariate normal variable (TVN). We denote the relevant correlations by ρ_1, ρ_2, ρ_3 (correlations between υ_i and u_i; υ_i and v_i; u_i and v_i , respectively).

Model (1)-(3) above shows a recursive pattern. Equation (2) contains equation (1) endogenous variable but equations (1) and (3) do not include equation (2) endogenous variable. Unlike recursive models in a continuous setting the covariance structure of our model does not affect inference in our qualitative model. This is an extension of a well known result in the econometric literature first outlined by Lahiri and Schmidt (1978), see also Greene (1997, 1998).

The paper has two main folds, (a) and (b) below.

3. A two-step corrected estimation of our microeconometric model.

(a) *Inefficient, but consistent, estimation [Two step corrected estimation]*.

a-1 Estimate (1) by ML probit and obtain TO_i predictions, i.e., $\Phi(\hat{\theta}'Z_{2i})$.

a-2 Substitute predictions obtained in previous step in place of observed TO_i in (2) and estimate the system by ML probit with selectivity.

a-3 Calculate appropriate corrected variance-covariance estimations; Murphy-Topel (1985), see also Greene (1997, pp.141-142), Hardin (2002) and Hole (2006).

Building on previous work we use this procedure to obtain consistent but inefficient estimate of the parameter of interest, in particular δ_1, consistent estimation. In our case we consider equation in (1), a probit model of TO, as the model to be estimated in the first step and model in equations (2) and (3), a selectivity probit model of LCC demand, as the model to be estimated in the second step. So, we have to correct the estimated covariance matrix for

the selectivity probit model in equations (2) and (3), sometimes named the naïve covariance matrix, due to its conditional nature. As is well known, Murphy-Topel (1985), the estimate of the variance for a two-step model is

$$\hat{V}_2 + \hat{V}_2\left(\hat{C}\hat{V}_1\hat{C}' - \hat{R}\hat{V}_1\hat{C}' - \hat{C}\hat{V}_1\hat{R}'\right)\hat{V}_2. \qquad [4]$$

Where V stands for the covariance matrices for models in step 1 and step 2. In addition,

$$\hat{C} = \sum_{i=1}^{N}\left(\frac{\partial \ln f_{i2}}{\partial \hat{\theta}_2}\right)\left(\frac{\partial \ln f_{i2}}{\partial \hat{\theta}_1'}\right).$$

$$\hat{R} = \sum_{i=1}^{N}\left(\frac{\partial \ln f_{i2}}{\partial \hat{\theta}_2}\right)\left(\frac{\partial \ln f_{i1}}{\partial \hat{\theta}_1'}\right).$$

Where *f* stands for each observation i´s contribution to the likelihood function of the respective model.

A fairly easy way of calculating expression in (4) using Stata is described in detail in Hole (2006) and Muro, Suárez and Zamora (2008).

4. A ML estimation of our microeconometric model.

The procedure described in section 3 above facilitates consistent but inefficient estimates of our model. Now we present a ML estimation for our model of interest.

(b) *ML estimation.*

To establish the estimation method by ML, we show that the probabilities associated with each element of the set of all possible events

(1,1,1) LCC with package tour; [5]

(0,1,1) LCC without package tour; [6]

(1,0,1) Traditional carriage with package tour; [7]

(0,0,1) Traditional carriage without package tour; [8]

(1,-,0) Road with package tour; [9]

(0,-,0) Road with package tour; [10]

are trivariate normal probabilities that mimics likelihood contributions of the bivariate probit model with selectivity. Let us see likelihood contributions in each case.

Case (5):

Prob[TO=1, Y=1, C=1] = TVN($\theta'Z_2$, $\beta'X + \delta_1$, $\gamma'Z_1$, ρ_1, ρ_2, ρ_3).

It is a TVN pdf with a complete set of arguments: determinants and parameters of the organization mode; determinants and parameters of the LCC's demand (including δ_1); determinants and parameters of travel by air choice; finally correlation structure.

Case (6)

Prob[TO=0, Y=1, C=1] = TVN($-\theta'Z_2$, $\beta'X$, $\gamma'Z_1$, $-\rho_1$, $-\rho_2$, ρ_3).

It is a TVN pdf with an incomplete set of arguments (and sign changes): determinants and parameters of the organization mode; determinants and parameters of the LCC's demand; determinants and parameters of travel by air choice; finally correlation structure.

Case (7)

Prob[TO=1, Y=0, C=1] = TVN($\theta'Z_2$, -($\beta'X + \delta_1$), $\gamma'Z_1$, -ρ_1, ρ_2, -ρ_3).

It is a TVN pdf with a complete set of arguments (and sign changes): determinants and parameters of the organization mode; determinants and parameters of the LCC's demand (including δ_1); determinants and parameters of travel by air choice; finally correlation structure.

Case (8)

Prob[TO=0, Y=0, C=1] = TVN(-$\theta'Z_2$, -$\beta'X$, $\gamma'Z_1$,ρ_1, -ρ_2, -ρ_3).

It is a TVN pdf with an incomplete set of arguments (and sign changes): determinants and parameters of the organization mode; determinants and parameters of the LCC's demand; determinants and parameters of travel by air choice; finally correlation structure.

Case (9)

Prob[TO=1, C=0] = BVN($\theta'Z_2$, -$\gamma'Z_1$, -ρ_2).

It is a BVN pdf with an incomplete set of arguments (and sign changes): determinants and parameters of the organization mode; determinants and parameters of travel by air choice; finally specific correlation structure.

Case (10)

Prob[TO=0, C=0] = BVN($-\theta' Z_2$, $-\gamma' Z_1$, ρ_2).

It is a BVN pdf with an incomplete set of arguments (and sign changes): determinants and parameters of the organization mode; determinants and parameters of travel by air choice; finally specific correlation structure.

The above set of likelihood contributions mimics the likelihood contributions of a bivariate model with selectivity (in which only one equation is incidentally truncated and the other is fully observed) and then we can estimate our model as a bivariate probit model with selectivity regardless the presence of an endogenous qualitative variable in the right hand side of equations (2) and (3).

Bivariate probit models with selectivity can be consistently estimated by ML using numerical integration but it is our choice using simulated ML in order to achieve our results.

5. The impact of the travel organization mode on LCC demand.

The data used in this work come from the 2004 wave of EGATUR the Spanish Foreign Tourism Expenditure Survey, built by the Tourism Studies Institute. It is an annual survey which looks at a representative quantification of non-resident visitors coming to Spain and of their travel expenditure. The survey is a questionnaire answered by 60,011 foreign tourists visiting Spain and collected on a monthly basis in the frontiers. The EGATUR sample provides a very rich data set related to the tourists' behaviour, socioeconomic categories, attributes of the trip and other relevant variables and it allows the study of diverse questions.

The most important characteristic of the sample for this paper is that the sample doesn't have problems of selection bias because the data include all type of tourist arriving to Spain.

To highlight the importance of the means of entering the country (road or airport) Figure 1 shows the percentage of the tourist by the mode of transport. It is important to remark the importance of the airport in general as a way to coming to Spain, with more than eighty per cent of tourist travelling by air, if we analysed these air tourists, we show that the greater number of those correspond to full service carrier, but also the number of LCC tourist are important with a percentage of 21.3%.

[Insert Figure 1]

In general independent variables have been defined as dummy variables which take a value of 1 if the tourist belongs to the category specified and 0 otherwise. (See Appendix and Muro Suárez and Zamora (2009) for variables definition and more details of the sample). Table 1 presents the percentage of the LCC tourist analysed in the whole sample and in the air traveller sample and we distinguish three types of independent variables: tourist' characteristics, trip attributes and other control variables. While most of the characteristics are similar across the samples, it is important to remark the growing importance of the LCC tourist in the air sample versus the total sample of the Netherlands tourist, retired, with low level of income, without package tour, with a number of visits greater than ten and using Internet for booking the trip.

[Insert Table 1]

On Table 2 we report estimated parameters for our model under three different set of assumptions: selectivity probit model without endogeneity; two-step estimation of a selectivity probit model with endogeneity; ML estimation of a selectivity probit model with endogeneity.

[Insert Table 2]

The result showed for the rho coefficient, which corresponds to the correlation between u_i and v_i (ρ_3), display the adequacy of the probit model with sample selection. Also, the negative point estimate of ρ_3 implies that the unobserved factors affecting the probability of equation [2] and equation [3] are negatively correlated.

In addition, the endogeneity of the travel organization mode variables, as it is presented in equation [1], play an important role in the model. The results show in Table 2 confirms the adequacy of the two main folds proposed in this paper: the probit model with selectivity through a two-step method with a robust variance estimator and, also, the ML estimation of the selectivity probit model with endogeneity. The findings confirm the importance of incorporate the asymmetries of information and risk aversion in tourism microeconomic models.

The two ways proposed in section 3 and 4 do not offer divergent results in terms of statistical significance and signs observed. Also, the triprobit model has similar results with the different draws considered, confirming the consistence of the findings.

6. Conclusions.

Asymmetries of information and risk aversion play a relevant role in the choice of travel organization mode which in turn is one of the main determinants of tourism demand. In empirical analysis this means that inference in microeconometric models of tourism demand must deal with the endogeneity of travel organization mode variables. In this paper we show how to do that in a fairly simple way in the context of a selectivity probit model with endogeneity to analyse the impact of travel mode organization on LCC demand. We present two-step and ML procedures. In the first case we describe how to calculate corrected estimated covariance matrices, Murphy-Topel (1985) and Hardin (2002). In the second we show that in the context of qualitative models the issue of endogeneity can be considered in a recursive simultaneous equation setting without particular consideration of the covariance structure of our model.

Our empirical results show the importance of dealing with the endogenous character of travel organization mode variables in order to make appropriate evaluation of the impact of these variables on the tourist's demand. In particular with Spanish data on foreign tourism we find that treating the travel organization mode as endogenous modifies in percentage its contribution to the demand for LCC.

References.

Greene, W. (1997), *Econometric Analysis*. 3rd ed. Prentice Hall. Englewood Cliffs, NJ.

Greene, W. (1998), "Gender Economics Courses in Liberal Arts Colleges: Further Results", *Journal of Economic Education,* 29, pp. 291-300.

Hardin, J.W. (2002), "The robust variance estimator for two stage models. *The Stata Journal*, 2, pp. 253-266.

Hole, A.R. (2006), "Calculating Murphy-Topel Variance Estimates in Stata: A Simplified Procedure", *The Stata Journal*, 6, pp. 521-529.

Lahiri, K. and P. Schmidt (1978), "On the estimation of Triangular Structural Systems", *Econometrica*, 46, pp. 1217-1221.

Muro, J., C. Suárez and M.M. Zamora (2008), "A Note on Computing Murphy-Topel Corrected Variances in a Heckprobit Model with Endogeneity in Stata". Alcamentos 0804.

Muro, J., C. Suárez and M.M. Zamora (2009), "Access and Use of E-commerce in the Spanish Tourism Market" in *Advances in Tourism Destination Marketing*, ed. Kozak M., Gnoth, J., Andreu, L., Routledge

Murphy, K.M. and R.H. Topel (1985), "Estimation and Inference in Two-step Econometric Models", *Journal of Business and Economic Statistics*, 3, pp. 88-97.

Rivers, D. and Q.H. Vuong (1988), "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models", *Journal of Econometrics,* 39, pp. 347-366.

Figure 1: Tourist by mode of entering the country

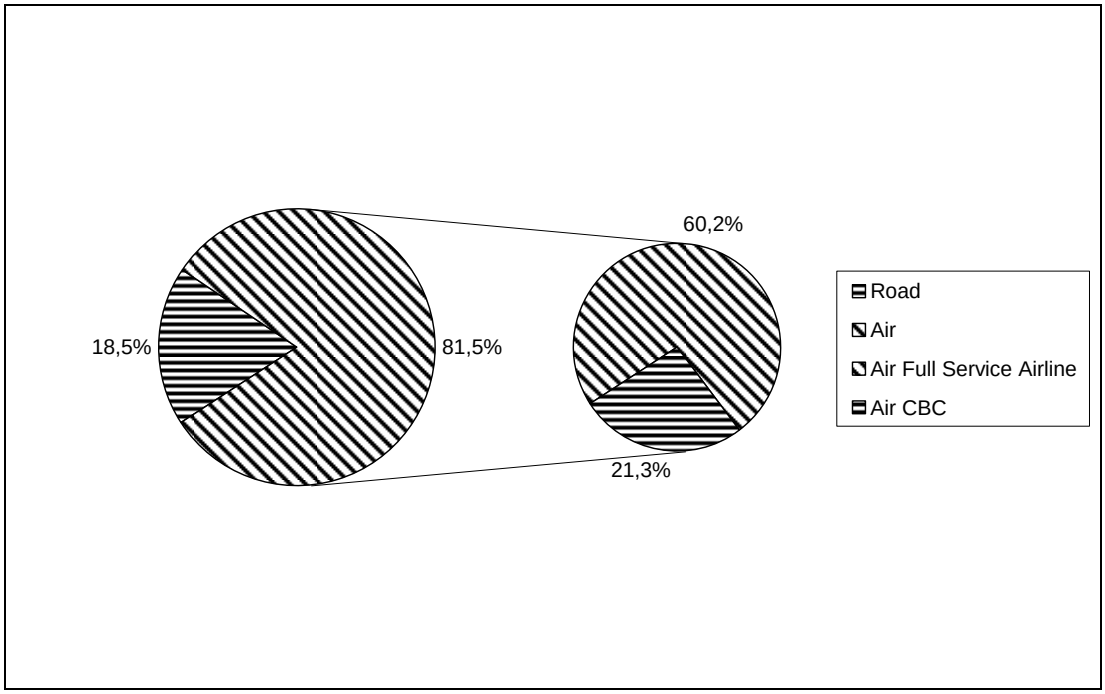

Table 1: Percentage of LCC tourist by tourist' characteristics, trip attributes and other control variables

	Total Sample	By Air Sample		Total Sample	By Air Sample
Tourists' characteristics			***Trip attributes***		
Age			**Size of travel group**		
<= 24 years	28,95%	25,92%	Alone	37,80%	34,15%
24 < age < 45	25,77%	23,34%	Couple	33,52%	34,29%
44 < age < 65	23,66%	24,79%	More than two	28,68%	31,56%
age > 64	21,62%	25,94%	**Tourist main destination**		
Level of education			Andalusia	21,97%	20,60%
Basic education	30,41%	31,34%	Balearic Island	18,69%	15,12%
Secondary education	37,62%	37,19%	Canary Island	9,59%	7,68%
University education	31,97%	31,47%	Catalonia	10,44%	15,49%
Ocupational situation			Community of Valencia	21,61%	21,83%
Employed	19,58%	19,02%	Madrid	8,23%	7,23%
Student	24,02%	21,61%	Rest of the Autonomous Communities	9,46%	12,06%
Retired	17,62%	22,73%	**Length of stay**		
Housewife	20,26%	18,07%	Between 1 and 3 days	27,00%	32,67%
Other	18,52%	18,56%	Between 4 and 7 days	37,88%	33,07%
Country of residence			Over 8 days	35,12%	34,27%
France	1,05%	2,89%	**Type of accomodation**		
Germany	16,27%	14,50%	Home (Freee accomodation)	44,48%	42,63%
United Kingdom	26,37%	22,95%	Tourism resort	26,11%	22,05%
Italy	11,93%	12,19%	Other type of accomodation	29,40%	35,32%
Netherlands	33,26%	35,89%	***Other control variables***		
Rest of the world	11,13%	11,58%	**Seasonality**		
Level of Income			First Quarter	27,95%	26,19%
High	27,51%	26,98%	Second Quarter	24,17%	23,94%
Medium	33,18%	30,18%	Third Quarter	23,11%	25,46%
Low	39,30%	42,84%	Fourth Quarter	24,78%	24,41%
Purpose of the trip			**Use Internet for transport reservations (booking)**		
Leisure	35,63%	34,50%	Yes	78,59%	73,83%
Work & Business relations	25,46%	23,00%	No	21,41%	26,17%
Other purpose	38,91%	42,50%	**Periodicity in a year**		
Organization of the trip			Less than one visit	34,78%	33,97%
With package tour	38,25%	31,06%	More than one visit	33,92%	36,22%
Without package tour	61,75%	68,94%	One in a year	31,30%	29,81%

Table 2: Estimations

Number of obs: 60011 Censored obs: 11106 Uncensored obs: 48905	Probit model wiht sample selection						Triprobit model (GHK simulator)								
	without instrument			with instrument			(15 draws)			(25 draws)			(40 draws)		
					Mtopel										
	Coef.	Std.Err.		Coef.	Std.Err.		Coef.	Std. Err.		Coef.	Std.Err.		Coef.	Std.Err.	
Low Cost Carriers vs Full Service Airline															
Age															
<= 24	0.1204	(0.045)	***	0.1138	(0.045)	**	0.1278	(0.045)	***	0.1311	(0.045)	***	0.1274	(0.045)	***
24 < age < 45	0.0353	(0.037)		0.0328	(0.037)		0.0424	(0.037)		0.0486	(0.037)		0.0429	(0.037)	
44 < age < 65	0.0330	(0.034)		0.0296	(0.034)		0.0335	(0.035)		0.0354	(0.035)		0.0354	(0.035)	
Level of Education															
Basic education	0.0215	(0.023)		0.0206	(0.023)		0.0199	(0.023)		0.0196	(0.023)		0.0141	(0.023)	
University education	-0.1517	(0.016)	***	-0.1511	(0.016)	***	-0.1496	(0.016)	***	-0.1479	(0.016)	***	-0.1480	(0.016)	***
Occupational situation															
Employed	0.0394	(0.054)		0.0395	(0.054)		0.0444	(0.054)		0.0464	(0.055)		0.0435	(0.054)	
Student	-0.0405	(0.061)		-0.0386	(0.061)		-0.0310	(0.061)		-0.0242	(0.062)		-0.0312	(0.061)	
Retired	0.2122	(0.060)	***	0.2057	(0.060)	**	0.2047	(0.060)	***	0.2003	(0.061)	***	0.2055	(0.060)	***
Housewife	-0.0241	(0.066)		-0.0295	(0.066)		-0.0056	(0.066)		0.0002	(0.066)		-0.0077	(0.066)	
Country of residence															
France	-1.6113	(0.082)	***	-1.5762	(0.083)	***	-1.6876	(0.062)	***	-1.7502	(0.060)	***	-1.6895	(0.064)	***
Germany	-0.6807	(0.031)	***	-0.6669	(0.037)	***	-0.6832	(0.031)	***	-0.6857	(0.031)	***	-0.6829	(0.031)	***
United Kingdom	-0.5679	(0.030)	***	-0.5528	(0.030)	***	-0.5698	(0.030)	***	-0.5721	(0.030)	***	-0.5675	(0.030)	***
Italy	-1.0219	(0.041)	***	-1.0029	(0.041)	***	-1.0261	(0.041)	***	-1.0300	(0.041)	***	-1.0232	(0.041)	***
Rest of the world	-0.9433	(0.033)	***	-0.9222	(0.032)	***	-0.9509	(0.032)	***	-0.9557	(0.032)	***	-0.9483	(0.032)	***
Level of Income															
High	-0.2108	(0.064)	***	-0.2080	(0.064)	**	-0.2034	(0.064)	***	-0.1998	(0.065)	***	-0.2084	(0.064)	***
Medium	-0.1920	(0.063)	**	-0.1915	(0.063)	**	-0.1785	(0.063)	***	-0.1709	(0.063)	***	-0.1847	(0.063)	***
Size of Group															
Alone	-0.1708	(0.028)	***	-0.1721	(0.028)	***	-0.1448	(0.025)	***	-0.1289	(0.025)	***	-0.1424	(0.025)	***
Couple	0.0366	(0.020)	*	0.0357	(0.020)	*	0.0441	(0.019)	***	0.0488	(0.019)	***	0.0448	(0.019)	***

Tourist main destinations															
Rest of the Autonomous Communities	-0.1540	(0.033)	***	-0.1701	(0.037)	***	-0.1638	(0.032)	***	-0.1748	(0.031)	***	-0.1465	(0.034)	***
Andalusia	0.2859	(0.030)	***	0.2867	(0.033)	***	0.2670	(0.028)	***	0.2543	(0.027)	***	0.2841	(0.030)	***
Canary Island	-0.2942	(0.021)	***	-0.2934	(0.021)	***	-0.2954	(0.021)	***	-0.2965	(0.021)	***	-0.3022	(0.021)	***
Catalonia	0.0055	(0.031)		-0.0221	(0.035)		-0.0134	(0.029)		-0.0271	(0.028)		0.0049	(0.031)	
Community of Valencia	0.1254	(0.026)	***	0.0996	(0.029)	**	0.1297	(0.027)	***	0.1307	(0.027)	***	0.1409	(0.027)	***
Madrid	-0.3813	(0.036)	***	-0.3930	(0.039)	***	-0.3831	(0.037)	***	-0.3855	(0.037)	***	-0.3665	(0.038)	***
Length of stay															
Between 1 and 3 days	0.2005	(0.027)	***	0.2062	(0.027)	***	0.1915	(0.026)	***	0.1826	(0.026)	***	0.1978	(0.027)	***
Between 4 and 7 days	0.0098	(0.016)		0.0087	(0.016)		0.0163	(0.016)		0.0202	(0.016)		0.0143	(0.016)	
Type of accomodation															
Home (Freee accomodation)	0.0034	(0.033)		-0.0041	(0.033)		0.0181	(0.032)		0.0281	(0.032)		0.0239	(0.032)	
Tourism resort	-0.1383	(0.034)	***	-0.1488	(0.036)	***	-0.1294	(0.033)	***	-0.1170	(0.032)	***	-0.1450	(0.035)	***
Seasonality															
Second	-0.1315	(0.020)	***	-0.1356	(0.020)	***	-0.1371	(0.020)	***	-0.1396	(0.020)	***	-0.1357	(0.020)	***
Third	-0.0764	(0.021)	***	-0.0781	(0.021)	***	-0.0896	(0.020)	***	-0.0951	(0.020)	***	-0.0880	(0.020)	***
Fourth	-0.1614	(0.020)	***	-0.1584	(0.020)	***	-0.1671	(0.020)	***	-0.1694	(0.020)	***	-0.1663	(0.020)	***
Fidelity															
First visit to Spain	-0.1083	(0.022)	***	-0.1063	(0.022)	***	-0.1023	(0.022)	***	-0.0962	(0.022)	***	-0.1043	(0.022)	***
Equal or Less than five visits	-0.0742	(0.017)	***	-0.0755	(0.017)	***	-0.0707	(0.017)	***	-0.0676	(0.017)	***	-0.0723	(0.017)	***
Periodicity in a year and trip's purpose															
Less than one visit and Sun & Beach	0.0925	(0.022)	***	0.0904	(0.021)	***	0.0923	(0.022)	***	0.0911	(0.022)	***	0.0908	(0.022)	***
More than one visit and Sun & Beach	0.0206	(0.021)		0.0215	(0.021)		0.0232	(0.021)		0.0252	(0.021)		0.0242	(0.021)	
Less than one visit and Work & Business relations	-0.2028	(0.044)	***	-0.1877	(0.044)	***	-0.1829	(0.043)	***	-0.1724	(0.043)	***	-0.1755	(0.043)	***
More than one visit and Work & Business relations	-0.1608	(0.032)	***	-0.1464	(0.033)	***	-0.1507	(0.032)	***	-0.1454	(0.032)	***	-0.1436	(0.032)	***
Use Internet for transport reservations (booking)															
Yes	0.7154	(0.016)	***	0.7668	(0.015)	***	0.7188	(0.015)	***	0.7223	(0.015)	***	0.7184	(0.015)	***
Organization of the trip															
With package tour	-0.4331	(0.022)	***	-0.4179	(0.050)	***	-0.4177	(0.029)	***	-0.4150	(0.025)	***	-0.3710	(0.036)	***
Constant	0.5215	(0.103)	***	0.5087	(0.104)	***	0.4632	(0.101)	***	0.4277	(0.101)	***	0.4510	(0.100)	***
rho	-0.3917	(0.056)					-0.2908	(0.034)	***	-0.2358	(0.033)	***	-0.3031	(0.037)	***

Note: An asterisk after a number indicates that the underlying heckprobit model (regression) coefficient differs significantly from zero at the 10 percent level; two asterisk, 5% level; and three asterisk, 1%. Individual reference: Netherlands, Balearic Island, more than five visits, first quarter, without package tour, more than 64 years old, other type of accommodation, size of travel group over two, secondary education, low level of income, other labor market status and job category, no use internet for booking, length of stay over 8 days.

Appendix

Tourists' characteristics:

Age: We have established four categories related to the age of the tourist: Under 24, between 24 and 44, between 45 and 64 and over 64.

Level of Education: The educational level has been established in three different categories: Basic, Secondary and University Education.

Country of residence: We have considered six different origins: France, Germany, United Kingdom, Italy, Netherlands and the rest of the world.

Level of income: This variable considers different income levels which are placed into the following categories: High income level, Medium income level, and Low income level.

Occupational situation: This variable identifies what it is the tourist occupation: Employed, Student, Retired, Housewife and Others.

Purpose of the trip: These variables identify tourists whose principal motives of the Spain visit is Work and Business relations, Sun and Beach and other motives.

Organization of the trip: This variable recognizes if the tourists have visited Spain with a package tour or not.

Trip attributes:

Size of travel group: With this variable we identify if the tourist travels Alone, as a Couple or in a Group of more than two persons.

Tourist main destination: In order to collect the main tourism destinations in Spain we have defined seven dummy variables: Andalusia, Canary Islands, Balearic Island, Catalonia, Community of Valencia, Madrid and other destinations, respectively.

Length of stay: This variable recognizes three different categories: Between 1 and 3 days, Between 4 and 7 days and More than 7 days.

Type of accommodation: We use three different categories: Tourism resort, Free accommodation and other type of accommodation.

Type of travel: We use three different categories: Full Service Airline, Low Cost Company and Road.

Other variables:

Seasonality: these variables identify the quarter during which the trip is made.

Use Internet for transport reservations (booking): This variable recognizes if the tourists have used Internet for transport reservations.

Periodicity in a year: In order to identify the tourist's periodicity, we have considered the categories: More than once in a year, Once in a year and Less than once a year to refer the number of times that the tourists visit Spain in one year.