Alcamentos

Departamento de Estadística, Estructura y O.E.I.

0801

THE IMPACT OF E-COMMERCE ON THE TOURIST PURCHASE DECISION: AN EMPIRICAL ANALYSIS

Juan Muro, Cristina Suárez y María del Mar Zamora

Universidad de Alcalá y Alcamétrica

DEPARTAMENTO DE
ESTADÍSTICA, ESTRUCTURA ECONÓMICA Y O.E.I.
Plaza de la Victoria, 2
28802 Alcalá de Henares (Madrid)
Teléfono: 91 885 42 01
http://www.uah.es/centros_departamentos/departamentos

Universidad de Alcalá

Alcamentos

Nº: 0801

THE IMPACT OF E-COMMERCE ON THE TOURIST PURCHASE DECISION: AN EMPIRICAL ANALYSIS

Juan Muro, Cristina Suárez y María del Mar Zamora

Universidad de Alcalá y Alcamétrica

D E P A R T A M E N T O D E
ESTADÍSTICA, ESTRUCTURA ECONÓMICA Y O.E.I.
Plaza de la Victoria, 2
28802 Alcalá de Henares (Madrid)
Teléfonos: 91 885 42 01
http://www.uah.es/centros_departamentos/departamentos

THE IMPACT OF E-COMMERCE ON THE TOURIST PURCHASE DECISION: AN EMPIRICAL ANALYSIS

Juan Muro
Cristina Suárez*
María del Mar Zamora[1]

Universidad de Alcalá and Alcamétrica

January, 2007

Abstract

The development in the tourist industry linked with the rapid growth in e-commerce has put in evidence the existence of a new customer. We empirically investigate the microeconomic determinants of the internet purchased tourist goods. We adopt a reduced form demand for online goods model, extended to incorporate possible selectivity biases stemming from interactions between unobserved individual heterogeneity associated with specific internet use choice. The model is estimated using a very rich dataset from EGATUR (*Encuesta de Gasto Turístico*), the Spanish Foreign Tourist Expenditure Survey. The sample allows us to explore the influence of price and income related variables as well as personal characteristics on internet purchased goods. Price and income results are consistent with theory. Unobserved individual heterogeneity linked with the use of the internet is significantly correlated to unobserved individual heterogeneity related to online purchases.

Keywords: e-Commerce, Tourism, Binary choice model with selectivity
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)
e-mail: cristina.suarez@uah.es

[1] We would like to thank participants at ATMC held in Valencia (2007) for the highly constructive comments received. Errors remain our sole responsibility.

THE IMPACT OF E-COMMERCE ON THE TOURIST PURCHASE DECISION: AN EMPIRICAL ANALYSIS

Juan Muro, Cristina Suárez and María del Mar Zamora

1. Introduction

The Internet phenomenon is changing people's habits in developed countries and many socio-economic studies are currently being developed related to this event, in particular with the on-line retail sales of goods and services (for example OECD (1999)). In general, the Internet could be deemed as an information system and also as an electronic market place, so, these characteristics allow the Internet to be considered as an intermediary between buyers and sellers to exchange information about prices and product offerings.

One of the Internet markets that has been developing towards higher levels of sales is the online travel-tourist market which has increased by as much as 34% from 2004 to 2005 (Marcussen (2006)). This market has special characteristics for both buyers and sellers. One of the main features of the tourist product is its intangible nature when purchased; it is merely a piece of information stored in a reservations system, subsequently buyers do not require sizable investments in organizational transformations. These electronic market systems reduce the search costs that buyers must pay to obtain information about the prices and product offerings available in the market, some authors, as Combes and Patel (1997), described the customer environment for Internet-based travel services as allocation where consumers could be compared with ease and, also, they can inquire about various aspects of a travel destination without having to speak to a travel agent or they can quickly and simply find the lowest fare anywhere. In addition

to their ability to reduce search costs, if we take into account the fact that the tourist is normally not able to try the product until the moment agreed, we can draw the conclusion that the tourist product fits perfectly with the new technologies.

The use of the Internet with a tourism purpose is twofold: it can act as a promotional tool or it can be used focusing on its capacity to do e-trade. At first, the promotional application was the reason why the tourist became so interested in the Internet. The second reason was e-commerce, which should be understood as the reservations and/or shopping for tourist products.

Given that the relatively new phenomenon of e-commerce has important repercussions for tourist travel decisions, the aim of this paper is to analyse the microeconomic determinants of the tourists' purchase choices for foreign tourism arriving in Spain. The empirical literature on this subject is very scarce. It is limited to some mainly descriptive papers, for Spain consult, for example, IBIT (2001), or for others on very specific questions not related to the subject of tourism consult, for example, Goolsbee (2000) and Alm and Melnik (2005).

The purpose of the article focuses on the influence that e-commerce is having on the tourist sector by assuming that they compare the stochastic utility of several alternatives and select the one that maximizes their utility. Also, this paper demonstrates how access to the Internet with a tourism purpose is important in order to buy online tourist travel products or services. To do so, we use a probit model with sample selection in which the probability of the tourist e-commerce choice is conditional to the access to the Internet with a tourism purpose.

The model has been estimated with Spanish data on foreign tourism. We utilize the 2004 wave of a very rich database coming from EGATUR (*Encuesta de Gasto Turístico*) the Spanish Foreign Tourist Expenditure Survey. The survey is a questionnaire answered by more than a sixty thousand foreign tourists visiting Spain and it requests information on tourists' socioeconomic characteristics, attributes of the trip and other relevant variables including the e-commerce choice. The Egatur sample does not have problems of selection bias because the data include all types of tourists arriving in Spain and not only those that use the Internet.

The goal in this article is to analyze a new phenomenon, the impact of e-commerce on the tourist purchase decision for foreign tourists visiting Spain. The next section reviews a conceptual framework that explains the tourists' choice of commerce mode in section 2. This is followed by a description of the database and we also present, in section 3, the empirical results and analyse the main determinants of e-commerce choice. Finally, in the last section, we sum up with our main conclusions of the impact of e-commerce on the tourist purchase decision experience in Spain.

2. The Model

Online commerce presents Internet users with another method for purchasing goods. Almost all goods traded online can also be purchased in traditional commerce. In this respect, the Internet presents simply another venue for purchasing the same goods, and hence Internet-purchased goods can be considered as perfect substitutes to some goods purchased in traditional commerce.

We can therefore structure the consumer decision to purchase goods online in the following way. First, we assume that the utility function of the representative tourist is

$$U=U(q_1, \ldots, q_k, z_1, \ldots, z_n, d_1, \ldots, d_r),$$

where $q=(q_1, \ldots, q_k)$ represents the vector of goods that can be purchased preferably in traditional commerce, for example restaurant meals; $z=(z_1, \ldots, z_n)$ denotes consumer goods that can be purchased in online commerce and in traditional commerce, where they are perfect substitutes, for example, hotel beds; and finally, $d = (d_1, \ldots, d_r)$ represents a good that can be purchased preferably in online commerce, for example, low-cost airlines.

The consumer balance will be reduced to:

$$Max\ U= U(q_1, \ldots, q_k, z_1, \ldots, z_n, d_1, \ldots, d_r)$$

Subject to:

$$p_q\, q + p_z\, z + p_d\, d = Y$$

where p_q, p_z *and* p_d are the vectors of prices, and Y represents the income level.

In this setting each tourist is assumed to have to choose between tourist goods that can be purchased in online commerce and in traditional commerce. Due to the cross-sectional nature of our database we assume a myopic behaviour. For any given tourist, defined by means of individual observed characteristics, his/her utility is derived from a number of observed goods attributes and travel features and a set of unobservable ones.

The probability that a tourist *i* will choose to buy online equals the probability associated with a positive difference in the comparisons between the utility derived from buying online and the utility related to traditional commerce. The difference between the online commerce and the traditional commerce can be represented as an unobserved latent variable Y_i^*. So

$$Y_i^* = X_i'\beta + u_{i,} \qquad [1]$$

such that one observes only the binary outcome,

$Y_i = 1$ if $Y_i^* > 0$ and

$Y_i = 0$ if $Y_i^* \leq 0$.

However, one only observes Y_i for observation *i* if the tourist has decided to obtain access to the Internet ($C_i = 1$), where C_i^* follows

$$C_i^* = Z_i'\gamma + \varepsilon, \qquad [2]$$

where

$C_i = 1$ if $C_i^* > 0$ and

$C_i = 0$ if $C_i^* \leq 0$.

X_i and Z_i are variable vectors of individual characteristics that can be common or not in both specifications [1] and [2]. u_i and ε_i are the error terms for equations [1] and [2], respectively, distributed as bivariate normal with mean zero, unit variance, and $\rho = Corr(u_i, \varepsilon_i)$. After controlling by observables our model allows for correlation between unobservables in equations [1] and [2].

As is well known, when $\rho \neq 0$, standard probit techniques applied to equation [1] yield biased results, and the probit model with sample selection provides consistent, asymptotically efficient estimates for all the parameters in such models.

3. Empirical analysis

In this section we present the empirical results of the analysis proposed in the last section that can be summarised with the following equations: the selection equation which is related to internet access with a tourism purpose (equation [2]) and the main equation which is related to online commerce and is only observed if internet access exists (equation [1]). These equations are estimated simultaneously according to maximum likelihood and the method is adapted from the article by Van de Ven and Van Pragg (1981), in which both equations have binary dependent variables.

All specifications incorporate a group of common variables included in X_i and Z_i which are related to characteristics that can influence tourist purchase choices and the possibility to undertake certain activities and they are common in both decisions (Internet access versus no Internet access, and online commerce versus traditional commerce). These variables are related

to the tourists' characteristics, trip attributes and other control variables, for example: age, level of education, level of income, length of stay, purpose and organization of the trip, size of travel group, type of accommodation, seasonality, fidelity, country of residence and main destination.

Some of these common variables are not included with the same specification in the main and selection equations because we assume that there are different ways of influencing, for example, age and purpose of the trip are sorted into different categories. Also, a group of variables, show specific characteristics related to the main decision and not with the selection decision, for example, travel to Spain with a low cost company because these types of companies do not frequently work with traditional commerce.

The data has been collected from the 2004 wave of EGATUR (*Encuesta de Gasto Turístico)* the Spanish Foreign Tourism Expenditure Survey, whose main objective is the 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 it collects information about the e-commerce for tourism activities without problems of selection bias because the data include all types of tourists arriving in Spain and not only those that use the Internet with a tourism purpose.

To highlight the importance of tourist access to the Internet and tourist use of online shopping, Figure 1 presents a structured view of these variables. It is important to remark the

importance of Internet use in general as a way of obtaining information to visit Spain, with forty percent of tourists, if we analysed these tourists who use the Internet, we show that the greater number of those correspond to tourists who purchase with e-commerce (24.5%).

[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[2]. Table 1 presents the percentage of tourists with purchases online analysed in the whole sample and in the access to Internet 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 e-commerce in products related to tourism in both samples (the Internet and in the total sample). In general, the tourist that is more likely to use e-commerce has the following characteristics: country of residence United Kingdom, main destination the Community of Valencia, travel in low cost company, without package tour, does not use tourism resort and with university education. Also, a peculiar fact is observed, on average, the younger tourist presents a strong presence in the use of the Internet, as it is expected, but when we analysed the sample of tourists with access to the Internet, the behaviour changed and older people preferred the e-commerce. This suggests the presence of a threshold age level under which older people do not find profitable the use of Internet profitable but when they use these new technologies they are more likely to use of e-commerce, especially if the purpose of the trip is to relax (sun and beach).

[Insert Table 1]

The results of the estimations of a probit model with sample selection, equation [1] and [2], are reported in Table 2.

[Insert Table 2]

Furthermore, we want to evaluate the marginal effects. Let us suppose W_i is a variable that belongs to both equations [1] and [2] and, additionally, W_i is a dummy variable. The total marginal effect of a change in W_i is possible to decompose in the sum of two terms, one will account for the "direct effect" on the probability that $Y_i = 1$ conditioned on $C = 1$, and the other will measure the "indirect effect" on this conditional probability. Thus,

$$\frac{\partial E[Y_i|C_i=1]}{\partial W_i} = \frac{\partial}{\partial W_i}\left[\frac{BVN(Y_i=1,C_i=1)}{Prob(C_i=1)}\right] = \frac{BVN(Y_i=1,C_i=1,W_i=1)}{Prob(C_i=1,W_i=1)} - \frac{BVN(Y_i=1,C_i=1,W_i=0)}{Prob(C_i=1,W_i=0)} =$$

$$= \Phi(X'_i\beta, W=1) - \Phi(X'_i\beta, W=0) + \qquad \text{(direct effect)}$$

$$+ \Xi_i(W=1) - \Xi_i(W=0). \qquad \text{(indirect effect)}$$

where $BVN(\bullet)$ is used to denote the Cumulative Distribution Function of the Bivariate Normal Distribution, $\Phi(\bullet)$ is the Distribution Function of the Standard Normal Distribution and $\Xi(\bullet)$ is:

[2] See Appendix for variables definition.

$$\Xi_i(X'_i\beta, Z_i'\gamma, \rho) = \frac{BVN(Y_i = 1, C_i = 1) - \Phi(X'_i\beta)\Phi(Z_i'\gamma)}{\Phi(Z_i'\gamma)}.$$

Table 3 shows the marginal effects estimate for the model.

[Insert Table 3]

The statistical significance of the correlation coefficient (see Table 2) suggests that controlling for the likelihood of the use of Internet is critical in determining the e-commerce effects. The positive point estimate of ρ implies that the unobserved factors affecting the probability of internet access or online shopping are positively correlated. In other words, the two outcomes are positively correlated after controlling for the tourists' characteristics and the attributes of the destination itself.

The characteristics of tourists were introduced to explore differences in sensitivity to different aspects of e-commerce. Differences in age show coefficients that are significant and with predictable signs. The results show that being younger makes a tourist more likely to have access to Internet and, also, more likely to buy online when the purpose of the trip is sun and beach. Also, it is important to remark that the existence of a threshold age, as suggested in Table 1, is rejected when we control for the sample selection and the other tourist characteristics. On the other hand, if the purpose of the trip is work and business relations, the tourist of principal working age (between 24 and 64 years old) has less probability of buying the tourist product online. Business tourists behave differently than leisure tourists, the reason could be found in the higher price sensitivity of leisure travellers, and, as a consequence, e-commerce acts as a determinant in the

decision-making processes of this type of tourist. The same result is found if we look for the probability of access to Internet and the purpose of the trip, leisure tourists look for tourist information on the Internet more than business travellers. Table 3 does not show the indirect marginal effects due to the different variable definition for this concept in the selection and the main equations, for this reason, the direct marginal effects are equal to the total marginal effects.

Education has a consistently positive and statistically significant impact on Internet access rate and on the probability of buying online. A higher level of education implies more knowledge about the Internet and its utilities for looking up information and, in that sense, makes a tourist more likely to buy online. Another interesting fact is the magnitude of the total marginal effects. If the estimation does not take into account the selection equation, the marginal effects (direct marginal effects) are in absolute values highest and lead to a mistake. In general the results show that the indirect marginal effects (although not always significant) are changing the values of the total marginal effects and confirm the significance of controlling Internet access with a tourism purpose.

Income level does not have an impact on the decision to buy online, conditional on Internet access and relative to the omitted income category[3]. But it is important to remark the different behaviour of the income level on the decision regarding Internet access, these variables have a statistically significant and positive impact, with the magnitude of the coefficient on income increasing as income goes up. This result suggests that the wealthy segments of tourist are more likely to benefit from access to the Internet for tourist information but not for online purchases.

Differences by origin market show that the tourist who comes from United Kingdom presents the greatest probability of using the Internet and of buying online, as Marcussen (2006) points out the UK became the largest online travel market in Europe in the last years. At the other side, the tourist which comes from France has the lowest. British tourists are about 9.9% more likely to buy online, than tourist from the Netherlands (the reference category) nearly all of the positive effect of the British tourist is the direct effect on the probability itself, as opposed to the indirect effect with a very scarce effect and not significant.

The Community of Valencia is the main destination for tourists which access and purchase using the Internet while the Canary Islands show the lowest probability for tourism e-commerce. The percentage change in probability to buy online for destination Valencia destination in relation to the reference category (Balearic Island) is 6.5% more likely and the most substantial effect is the direct effect.

The trade-off among package holidays has a negative effect on the probability of Internet access, with a marginal effect of -23.8%. The explanation of this result can be found in the different ways of planning travel, directly or indirectly thorough distribution channels such as travel agents. A new tourism distribution system has emerged that requires a better knowledge of environments and the Internet can act as a promotional tool and/or it can be used focusing on its capacity to do e-commerce for individual tourists that planned their tourism without a package.

[3] The test hypothesis jointly equal zero ($\chi^2(2) = 1.07$) is not rejected. Inclusion of Income level combined with Purpose of the trip instead of the direct form has no significant impact on any of the coefficients reported in Tables 2 and 3.

It may be assumed that the cost of searching for new alternatives is generally too high and the expected gains associated with new alternatives too uncertain. If we analyse tourists with ten or more visits to Spain, we can observe that they prefer to use e-commerce with a marginal effect of 8.1% on the probability of buying online.

Also evidenced in Table 2 is the fact that the use of a low cost airline to come to Spain increases the probability of e-commerce. This is one of the most important characteristics of this type of company, which prefers direct access to a consumer only through call centres and the Internet and also it is important to remark that the type of service most demanded in the online market is air travel with 56% of the demand.

The shorter the length of the stay the greater the probability of buying online, with positive means marginal effects ranging from 8.8% (between 1 and 3 days) to 5.3% (between 4 and 7 days).

4. Conclusions

The impact of the Internet in activities related to tourism has seen a significant growth in the last years. Given that online commerce has important repercussions for tourist decisions, the aim of this paper is to analyse the microeconomic determinants of the tourists' e-commerce choices for foreign tourism arriving to Spain in 2004.

The econometric analysis employed to obtain these results used a probit model with sample selection. This model is necessary for the aim of controlling e-commerce effects for the likelihood of Internet use. We came to this conclusion showing the statistical significance of the correlation coefficient. Using a probit model with sample selection we have estimated the probability of e-commerce choice, for the tourists that use Internet with a tourism purpose, as opposite to buying tourist products using traditional commerce.

Our results allow us to define characteristics influencing e-commerce. In general, the results show that tourist users of online shopping meet the following requirements. Younger people are more likely to buy via internet, tourists who come to Spain looking to relax or for beach and sun (leisure tourists) prefer to purchase the tourist product online, tourists without package holidays and who travel by low cost companies have more probability of buying online than tourists planning the travel with package holidays or travel by air, in a full service airline, or by road. Furthermore, geographical characteristics show that tourist coming from United Kingdom and going to the beach in Community of Valencia have the greatest probability of using Internet for looking up information and also of buying online.

5. References

Alm, J. and M. I. Melnik (2005) Sales Taxes and the Decision to Purchase Online, *Public Finance Review*, **33(2)**, 184-212.

Combes, G.C. and J.J. Patel (1997) Creating lifelong customer relation-ships: why the race for customer acquisition on the Internet is so strategically important, *Iword*, **2(4)**, Hambrecht & Quist.

Goolsbee, A. (2000) In a World without Borders: The Impact of Taxes on Internet Commerce, *Quarterly Journal of Economics*, **115(2)**, 561-76.

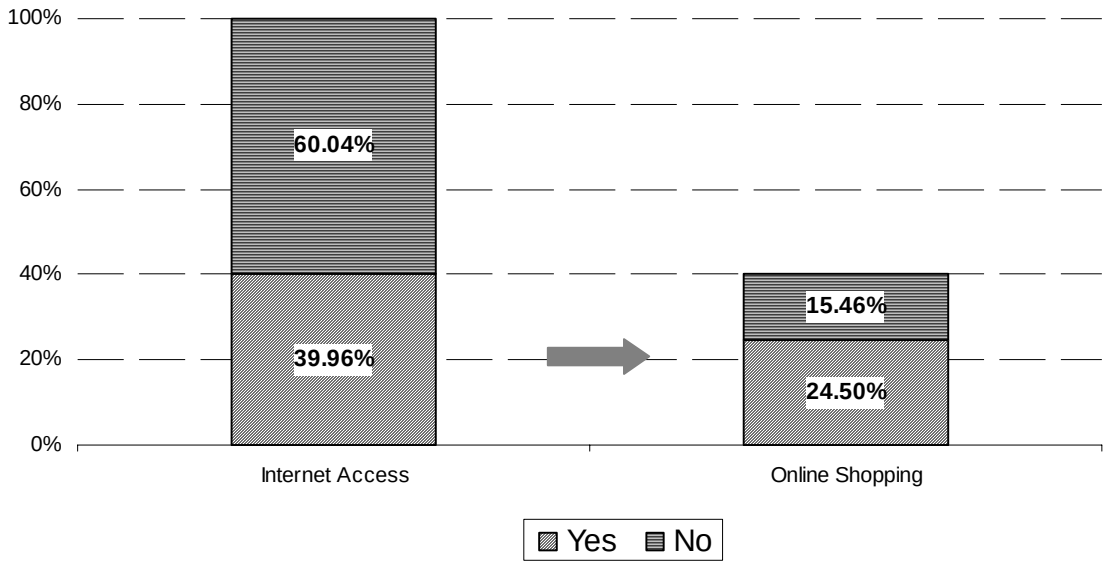
IBIT (2001) *Study on Electronic commerce in the value chain of the tourism sector*.

Marcussen, C.H. (2006) Trends in European Internet distribution of travel and tourism services, http://www.crt.dk/uk/staff/chm/trends.htm.

OECD (1999) *The economic and social impact of electronic commerce. Preliminary findings and research agenda*.

Van de Ven, W. P. M. M. and B. M. S. Van Pragg (1981) The Demand for Deductibles in Private Health Insurance: A Probit Model with Sample Selection, *Journal of Econometrics*, **17**, 229-252.

Figure 1: Percentage of Tourist with Internet Access and Online Shopping



Source: EGATUR and own elaboration.

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

Tourists' characteristics	Total tourists	Internet access	*Trip attributes*	Total tourists	Internet access
Age			**Size of travel group**		
<= 24 years	36.19%	23.91%	Alone	39.53%	38.32%
24 < age <= 44	29.69%	23.59%	Couple	29.64%	30.90%
44 < age <=64	21.60%	24.75%	More than two	30.83%	30.78%
>= 65 years	12.52%	27.76%	**Tourist main destination**		
Purpose of the trip			Rest of Spain	13.43%	13.18%
Work and Business	22.74%	30.66%	Andalusia	12.06%	16.25%
Sun and beach (or relax)	32.39%	30.17%	Balearic Island	16.44%	16.70%
Other motives	44.87%	39.17%	Canary Island	8.16%	7.76%
Age & Purpose of the trip			Catalonia	15.88%	14.64%
<= 24 years & Sun and beach (or relax)	32.56%	22.59%	Community of Valencia	22.46%	18.31%
24 < age <= 44 & Sun and beach	31.35%	23.18%	Madrid	11.57%	13.16%
44 < age <=64 & Sun and beach	23.16%	25.51%	**Length of stay**		
>= 65 years & Sun and beach	12.93%	28.71%	1 < days < 3	28.94%	35.46%
			4 < days < 7	37.96%	32.71%
<= 24 years & Work and Business	38.19%	26.15%	>= 8 days	33.10%	31.83%
24 < age <= 44 & Work and Business	23.68%	23.91%	**Type of accomodation**		
44 < age <=64 & Work and Business	20.32%	25.21%	Other type of accomodation	33.50%	33.27%
>= 65 years & Work and Business	17.81%	24.72%	Free accomodation	41.62%	40.43%
			Tourism resort	24.88%	26.30%
<= 24 years & Other motives	37.61%	23.57%	**Type of travel**		
24 < age <= 44 & Other motives	33.53%	25.56%	Full Service Airline	26.01%	31.09%
44 < age <=64 & Other motives	18.78%	24.19%	Low Cost Company	67.26%	51.64%
>= 65 years & Other motives	10.07%	26.68%	Road	6.73%	17.28%
Level of education			***Other control variables***		
Basic education	20.62%	27.16%	**Seasonality**		
Medium education	35.00%	35.99%	First Quarter	24.20%	26.16%
University education	44.38%	36.85%	Second Quarter	24.57%	24.63%
Country of residence			Third Quarter	23.90%	23.67%
France	6.84%	13.40%	Fourth Quarter	27.33%	25.54%
Germany	14.29%	14.58%	**Number of visits**		
United Kingdom	24.09%	20.90%	Number of visits >=10	47.58%	55.55%
Italy	17.98%	17.44%	Number of visits < 10	52.42%	44.45%
Netherlands	21.69%	18.30%			
Rest of the World	15.11%	15.38%			
Level of income					
High	30.07%	31.01%			
Médium	32.81%	31.99%			
Low	37.12%	36.99%			
Organization of the trip					
without package tour	70.19%	66.49%			
with package tour	29.81%	33.51%			

Table 2: Estimation of the probit model with sample selection

E-commerce	**Coef.**	**Std. Err.**		**Internet access**	**Coef.**	**Std. Err.**	
Age & Purpose of the tripe				**Age**			
<= 24 years & Sun and beach (or relax)	0.1994	(0.048)	***	<= 24 years	1,1751	(0.029)	
24 < age <= 44 & Sun and beach	0.1427	(0.033)	***	24 < age <= 44	0.9609	(0.024)	***
44 < age <=64 & Sun and beach	0.0984	(0.035)	***	44 < age <=64	0.5836	(0.024)	***
<= 24 years & Work and Business relations	-0.2038	(0.125)		**Level of education**			
24 < age <= 44 & Work and Business relat.	-0.5062	(0.045)	***	Basic education	-0.3718	(0.020)	***
44 < age <=64 & Work and Business relat.	-0.4441	(0.073)	***	Medium-High education	-0.2868	(0.013)	***
Level of education				**Country of residence**			
Basic education	-0.2641	(0.042)	***	France	-0.8092	(0.030)	***
Medium education	-0.0646	(0.026)	**	Germany	-0.1312	(0.027)	***
Country of residence				United Kingdom	0.0945	(0.026)	***
France	-0.2326	(0.063)	***	Italy	-0.2623	(0.034)	***
Germany	0.0186	(0.042)		Rest of the world	-0.2643	(0.027)	***
United Kingdom	0.2602	(0.040)	***	**Level of income**			
Italy	-0.1976	(0.052)	***	High	0.2142	(0.052)	***
Rest of the world	-0.2270	(0.042)	***	Medium	0.1591	(0.050)	***
Level of income				**Organization with package tour**	-0.6599	(0.019)	***
High	-0.1021	(0.099)		**Purspose of the trip**			
Medium	-0.0959	(0.097)		Work and Business relations	-0.5606	(0.024)	***
Organization with package tour	-0.7892	(0.039)	***	Sun and beach	0.0761	(0.018)	***
Type of travel				**Size of travel group <=3**	0.0998	(0.015)	***
Full Service Airline	1.0015	(0.040)	***	**Tourist main destination**			
Low Cost Company	1.7268	(0.046)	***	Rest of Spain	0.1120	(0.025)	***
Size of travel group				Andalusia	-0.4450	(0.025)	***
Alone	0.0671	(0.032)	*	Canary Island	0.1409	(0.018)	***
Couple	0.0185	(0.025)		Catalonia	0.0959	(0.022)	***
Tourist main destination				Community of Valencia	0.1966	(0.024)	***
Rest of Spain	-0.2219	(0.042)	***	Madrid	-0.1903	(0.030)	***
Andalusia	-0.2107	(0.051)	***	**Length of stay**			
Canary Island	-0.6237	(0.036)	***	1 < days < 3	-0.1862	(0.019)	***
Catalonia	-0.0710	(0.037)	*	4 < days < 7	0.0610	(0.013)	***
Community of Valencia	0.1940	(0.042)	***	**Type of accomodation**			
Madrid	-0.3672	(0.048)	***	Free accomodation	-0.1250	(0.017)	***
Length of stay				**Seasonality**			
1 < days < 3	0.1806	(0.034)	***	Second Quarter	0.0696	(0.017)	***
4 < days < 7	0.1411	(0.021)	***	Third Quarter	0.0129	(0.016)	
Type of accomodation				Fourth Quarter	0.1613	(0.017)	***
Free accomodation	-0.0124	(0.041)		**Number of visits >=10**	-0.1709	(0.013)	***
Tourism resort	-0.1227	(0.039)	***	**Constant**	-0.6922	(0.069)	***
Seasonality							
Second Quarter	-0.0833	(0.030)	***	rho	0.2411	(0.077)	***
Third Quarter	-0.1356	(0.029)	***	Log pseudolikelihood	-477860.49		
Fourth Quarter	-0.0557	(0.030)	*	Number of obs. 60011			
Number of visits>=10	0.1661	(0.026)	***	Censored obs.	36031		
Constant	-0.4411	(0.131)	***	Uncensored obs.	23980		

[a]Individual reference: more than 64 years old, other motives of travel, University education, Netherlands, low level of income, without package tour, travel by road, size of travel group over two, Balearic Island, length of stay over 8 days, other type of accommodation, first quarter, less than ten visits.
***Level of significance 1%, **level of significance, 5%, *level of significance 10%.

Table 3: Marginal effects and pseudo-elasticities of the probit model with sample selection

	Marginal Effects			Pseudo elasticity		
Use Internet for booking	**Direct effects**	**Indirect effects**	**Total**	**Direct effects**	**Indirect effects**	**Total**
Age & Purpose of the tripe						
<= 24 years & Sun and beach (or relax)	0.0749		0.075	22.7%		22.7%
24 < age <= 44 & Sun and beach	0.0531		0.053	16.1%		16.1%
44 < age <=64 & Sun and beach	0.0363		0.036	11.0%		11.0%
<= 24 years & Work and Business relations	-0.0701		-0.070	-21.3%		-21.3%
24<age<= 44 & Work and Business relations	-0.1578		-0.158	-47.9%		-47.9%
44<age<=64 & Work and Business relations	-0.1415		-0.142	-42.9%		-42.9%
Level of education						
Basic education	-0.0892	0.0123	-0.077	-27.1%	10.5%	-16.6%
Medium education	-0.0230	0.0187	-0.023	-7.0%	16.0%	9.0%
Country of residence						
France	-0.0793	0.0492	-0.030	-24.1%	42.1%	18.0%
Germany	0.0067	0.0107	0.017	2.0%	9.1%	11.2%
United Kingdom	0.0986	0.0002	0.099	29.9%	0.2%	30.1%
Italy	-0.0681	0.0088	-0.059	-20.7%	7.5%	-13.1%
Rest of the world	-0.0775	0.0070	-0.071	-23.5%	6.0%	-17.6%
Level of income						
High	-0.0361	-0.0199	-0.056	-10.9%	-17.1%	-28.0%
Medium	-0.0339	-0.0158	-0.050	-10.3%	-13.6%	-23.8%
Organization with package tour	-0.2203	-0.0181	-0.238	-66.8%	-15.5%	-82.3%
Type of travel						
Full Service Airline	0.3828		0.383	116.2%		116.2%
Low Cost Company	0.5712		0.571	173.3%		173.3%
Size of travel group						
Alone	0.0246		0.025	7.5%		7.5%
Couple	0.0067		0.007	2.0%		2.0%
Tourist main destination						
Rest of Spain	-0.0759	-0.0194	-0.095	-23.0%	-16.6%	-39.6%
Andalusia	-0.0723	0.0218	-0.050	-21.9%	18.7%	-3.3%
Canary Island	-0.1861	-0.0479	-0.234	-56.5%	-41.0%	-97.5%
Catalonia	-0.0253	-0.0102	-0.035	-7.7%	-8.7%	-16.4%
Community of Valencia	0.0728	-0.0083	0.065	22.1%	-7.1%	15.0%
Madrid	-0.1201	-0.0087	-0.129	-36.4%	-7.4%	-43.8%
Length of stay						
1 < days < 3	0.0676	0.0200	0.088	20.5%	17.2%	37.7%
4 < days < 7	0.0525	0.0004	0.053	15.9%	0.3%	16.2%
Type of accomodation						
Free accomodation	-0.0045		-0.004	-1.4%		-1.4%
Tourism resort	-0.0431		-0.043	-13.1%		-13.1%
Seasonality						
Second Quarter	-0.0296	-0.0089	-0.038	-9.0%	-7.6%	-16.6%
Third Quarter	-0.0475	-0.0076	-0.055	-14.4%	-6.5%	-20.9%
Fourth Quarter	-0.0199	-0.0141	-0.034	-6.0%	-12.1%	-18.1%
Number of visits>=10	0.0621	0.0185	0.081	18.8%	15.9%	34.7%

Appendix

Tourists' characteristics:

Age: The socio-demographic characteristics have been defined including information related to the age of the tourist. We have established four categories: 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.

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: In order to identify the tourist's fidelity, 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.

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 control variables:

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

Number of visits: In order to identify the tourist's fidelity, we have considered the categories: More than ten and Less than ten to refer the number of times that the tourists visit Spain in one year.

Relación de títulos publicados en la colección ALCAMENTOS.

Nº	Autor/es	Título
0801	Juan Muro Cristina Suárez Maria del Mar Zamora	The impact of e-commerce on the tourist purchase decision: An empirical micro analysis