Revista de Economía Mundial
ISSN: 1576-0162
rem@uhu.es
Sociedad de Economía Mundial
España

Arbués, Pelayo; Baños, José; Mayor, Matías; Suárez, Patricia
Econometric Modelling of Long-Distance Domestic Travel
Revista de Economía Mundial, núm. 38, septiembre-diciembre, 2014, pp. 101-126
Sociedad de Economía Mundial
Madrid, España

ISSN: 1576-0162

Econometric Modelling of Long-Distance Domestic Travel

Modelización econométrica de los viajes nacionales de larga distancia

Pelayo Arbués
Universidad de Oviedo
gonzalezpelayo@gmail.com

José Baños
Universidad de Oviedo
jbanos@uniovi.es

Matías Mayor
Universidad de Oviedo
mmayorf@uniovi.es

Patricia Suárez
Universidad de Oviedo
suarezcano@gmail.com

Recibido: julio de 2014; aceptado: octubre de 2014

Abstract

The main goal of this study is to determine the effects of socio-economic, demographic and land-use factors on long-distance travel. Intercity trips are not very frequent activities, which caused a large share of the survey respondents to report zero kilometres travelled. The Double Hurdle Model, Infrequent Purchase Model and Heckman Selection Models are estimated. Distance travelled in trips longer than 50 km. is a function of gender, age, income, size of municipality, region of residence, mode of transport and trip purpose. Results indicate that zeroes reported in the survey are caused by corner solutions, which allow us to make recommendations about designing long-distance surveys.

Keywords: Long-Distance; Mobility; Econometrics.

Resumen

El objetivo principal de este estudio es analizar los efectos de los factores socioeconómicos, demográficos y de uso del suelo sobre los viajes de larga distancia. Este tipo de viajes son actividades que se realizan con una frecuencia baja lo que provoca que en los datos recogidos por las encuestas una amplia proporción de las respuestas sea de cero kilómetros viajados. Para tratar esta peculiaridad de los datos se estiman Modelos de Doble Valla, Modelos de Infrecuencia de Compra y Modelos de Selección de Heckman. La distancia recorrida en los viajes de más de 50 kilómetros es función del género, la edad, y la renta del encuestado, del tamaño del municipio y la región de residencia así como del modo de transporte y del propósito del viaje. Los resultados indican que los ceros recogidos en la encuesta son provocados por soluciones de esquina, lo cual permite hacer recomendaciones acerca del diseño de encuestas de viajes de larga distancia.

Palabras clave: Larga distancia; Movilidad; Econometría.

JEL Classification: C34, D12, R4.

1. Introduction

Long-distance travel is attracting the attention of scholars and policy makers due to its rapid increase in recent decades. Technological innovations in transportation modes and recent economic growth are among the main factors contributing to making intercity travel more frequent. The development of high speed railway infrastructure, the popularity of car ownership and greater accessibility to well-connected airports allow human activities to spread on larger spatial scales. In this sense, long-distance travel figures are important indicators of economic and social integration, creating stronger links between distant urban regions. These interregional interactions shape the future of urban regions through higher population mobility, greater productive resources and more companies.

Although intercity trips are less frequent activities than commuting and only represent a small share of all trips, a large share of total mileage travelled is long distance (Kuhnimhof *et al*. 2009). Increases in long-distance mobility have economic, social and environmental implications. Long-distance travel increases the opportunities for regions to be better connected to trade and economic networks, making industry clusters between distant areas more likely to occur (Limtanakool *et al*., 2006a). Rural areas benefit from private and public facilities located far away, allowing greater accessibility to services. Remote locations also seem more attractive in terms of recreational activities, raising demand for tourism in these areas. On the other hand, higher long-distance mobility also imposes negative externalities to society such as congestion and higher risks of traffic accidents. From an environmental perspective, long-distance travel negatively impacts society due to increased fuel consumption and contributions to global greenhouse emissions.

These implications make long-distance travel a significant concern to policymakers and companies in all industries, but especially in the tourism and transportation sectors. It would be beneficial to have a better understanding of the determinants of travel behaviour for several reasons. From a policy perspective, debate on proposed increases on fuel taxes is usually shaped by distributional issues, both regarding income groups and regions (Johansson-Stenman, 2002). Knowledge about the profiles of long-distance travellers can help transport companies with the design of price discrimination schemes. In this regard, the contribution of this study is twofold. First, it contributes to the scarce existing literature on long-distance travel by increasing the amount of empirical evidence on the factors behind the rise in long-distance travel demand. In particular, the main goal of this study is to examine the determinants of long-distance travel in Spain using a large micro data set, the 2007 Spanish Mobility Survey (Movilia).

The second contribution is to apply an alternative econometric approach to estimate distance travel determinants in intercity trips. As expected, a large share of the survey respondents report zero kilometres travelled. In order to reduce the number of zeroes, survey designers have usually increased the duration of reporting periods. However, this solution is complicated by the fact that respondents have more difficulty recalling events that occurred further in the past, and it requires a larger amount of resources to control and collect the data. Thus, reporting period is a fundamental characteristic of long-distance travel surveys. In this study, we estimate a set of models including the Double Hurdle Model, the Heckman Selection Model and the Infrequent Purchase Model. The application of these different techniques solves an econometric problem, but also allows us to gain insight into the origin of the zeroes.

The structure of the remainder of the paper is as follows. In the next section, the relevant literature is briefly reviewed to place the paper within the context of existing studies. Section 3 describes the information on long-distance travel in Spain contained in the Movilia 2007 database. The methodological part of the paper is presented in Section 4, where the econometric model and the results are discussed. Finally, some conclusions are drawn and directions for future research are outlined.

2. Previous studies on long-distance travel

National Mobility Surveys have usually collected data only on short-distance trips, with an emphasis on commuting to work or places of study, ignoring intercity transportation of passengers. This is probably caused by difficulties in measuring long-distance behaviour as noted by Axhausen (1999). As noted in Section 1, intercity trips are low-frequency events, which affects the design of long-distance surveys when selecting the duration of reporting periods. Moreover, long-distance travel is defined as trips longer than a distance threshold. Even determining the threshold definition is a common problem, and there is a lack of consensus in the literature with the lower limit ranging from 50 to 100 kilometres, according to Dargay and Clark (2012). These are some of the reasons why short-distance travel literature is much more extensive than scientific works on long-distance travel.

Among the studies dealing with intercity trips, it is worth noting that a great share of scientific work has paid special attention to a traditional topic in travel demand literature: mode choice decisions[1]. Other works have developed com-

[1]See for instance Bhat (1997), Mallett (1999), Georggi and Pendyala (2001), Limtanakool *et al*. (2006b), and Arbués *et al*. (2013).

plete travel demand models applied to intercity trips, including the generation, distribution, mode choice and assignment models (Erhardt *et al*., 2007; Baik *et al*., 2008).

In this literature review, we focus on the relatively few studies on the determinants of long-distance travel. We expand upon the work of Dargay and Clark (2012) where they provide an up-to-date review of literature concerning long-distance travel demand. Some of these papers concentrate on different population groups, differentiating by demographics or socio-economic status.

Relying on data from the 1995 American Travel Survey about trips longer than 100 miles, Mallett (1999) studies gender differences in long-distance travel. Women made fewer inter-city trips than men. This was mainly because female respondents did less business travel, and also because women had lower incomes and lower employment rates. Men also reported more trips for outdoor recreation, which was more limited in the case of women as they were more likely to be caregivers of children. Georggi and Pendyala (2001) provide an analysis of long-distance travel behaviour of two key segments of market, the elderly and the low-income, also using data from the 1995 American Travel Survey. As expected, the authors found less long-distance mobility in terms of travel frequency in those individuals aged above 65 years, and specially in those above 75 years. Surprisingly, average trip length seems to increase with age. The comparison between income groups shows that there are clear differences in long-distance trips regarding the frequency and the length of the trip. On average, lower-income groups make half the number of trips (4 or 5) of those made by higher income households. Trips of the higher-income population are, on average, 80% longer than those of the low-income group. These differences are explained in part by the fact that low-income group does less business travel and less recreational travel, and in part because they are less likely to use air transportation than people with higher disposable incomes. Similar results were found in a parallel analysis of the same survey conducted by Mallett (2001).

Collia *et al*. (2003) analyse trips with a length over 50 miles using data from the 2001 National Household Travel Survey. More than two-thirds of the total intercity trips were taken by adults under 65, while only 8% of long-distance trips were made by the elderly. Cross-tabulation of the data also indicates gender differences in travel for all age groups, with men enjoying more long-distance trips than women. When comparing trip length by age groups, these authors found that elderly tend to take longer trips.

Using multivariate analysis, Mokhtarian *et al*. (2001) explore factors other than demographics or socioeconomic background that may explain travel distance. They measure some characteristics of the population such as personality, lifestyle or travel-related attitudes through different variables that are found to have an impact on distance travelled. The authors conclude that some segments of the population have a desire for travel for its own sake, instead of the traditional approach which views travel as a derived demand. Income is found to have a positive and significant impact, while increasing age seems to reduce the demand for airplane travel. As usual, differences in travel distance are found by gender, with women traveling less than men for work-related purposes, but farther when traveling on airplanes for long-distance entertainment.

In Limtanakool *et al*. (2006a) the authors study the propensity of people to participate in medium and long-distance travel in United Kingdom and the Netherlands. Relying on binary logit models applied to different segments of the sample, the overall results show how in leisure travel gender and household composition hardly affect decisions about long-distance travel as males and females within a household usually undertake this trips jointly. In business travel and commuting, gender and income differences have been found, where males, full-time workers and those with higher incomes are more likely to travel longer. Land use related factors also play a significant role, with citizens of metropolitan and high-density populated areas traveling shorter distances as they can find a wider range of services within a relatively shorter distance from home than their counterparts residing in low-density rural areas. Frändberg and Vilhelmson (2003) offer an account of Swedish long-distance travel, with a focus on international mobility, estimating regression models on trip frequency to other countries as a function of sex, region of origin, income and other personal variables.

Dargay and Clark (2012) study the socio-economic, demographic and geographic determinants influencing long distance travel in Great Britain in the period 1995-2006. Econometric results of models estimated by weighted least squares indicate that distance travelled is lower for women, and for people over the age of 60. Labour situations also affect distance travelled, with students and workers traveling longer than other groups. The authors also included length of residence, and the estimates show that the longer an individual lived at the same address, the less distance they travelled. Household composition also affects distance, and seems to support the idea that adults living on their own travel longer distances than adults living with other adults and children. Regarding geographic factors, Londoners travelled less (in terms of distance) than any other British region, while people living in rural areas travelled greater distances.

In Holz-Rau *et al*. (2014), the research question is focused on spatial differences in travel behaviour and a comparison of factors affecting distance travelled in daily and long-distance trips. Socio-economic and demographic variables are introduced in the mode using interactions between age, labour situation and gen- der. Using a Heckman model, explanatory variables are found to similarly affect distance travelled in both sorts of trips, although income and level of education affects the distance decision more in long distance trips than daily trips. Concerning geographic factors, participation and distance travelled increase with municipality size and population density in long distance trips, while opposite signs were found for daily commuting.

Overall, this literature review has shown that socio-economic, demographic and geographic factors have been previously found to have significant impacts on distance travelled in intercity trips. With regard to socio-economic and demographic factors, some consensus has been found in all studies reviewed. Conflicting results have appeared when studying geographic determinants. While in Dargay and Clark (2012) and Limtanakool *et al*. (2006b) residents in high- density and high-populated metropolitan areas travelled smaller distances than those from rural areas, while the opposite result is found in Holz-Rau *et al*. (2014).

3. Long-distance travel in Spain: Movilia 2007

In order to study the determinants of long distance travel in Spain a mobility survey (Movilia 2007) from Statistical Office of the Ministry of Public Works[2] is used for the analysis. This survey provides information that increases knowledge about the mobility patterns of the resident population in private households in Spain, their characteristics, and their determinants. The basis for the selection of households was the Municipal Register Long-distance trips are defined as those with a length over 50 km. It is important to note that some groups are not included in this survey because they do not live in private households in Spain. This is the case for tourists, illegal immigrants and people who reside in institutions or collective households. Information on long distance trips was collected by telephone interview questionnaire for a 4-week period between February 2007 and January 2008. Survey respondents reported information about their long-distance mobility in the previous month. The size of the sample and its distribution was designed to guarantee the data was representative at the provincial level. Consequently, the figures about the number of trips and the distribution by mode and purpose are statistically representative at the provincial and metropolitan levels.

The dataset employed consists of a continuous variable *distance* of kilometres travelled in intercity trips (dependent variable), and two dummy variables: *travel* that takes value 1 if the person has made a long-distance trip in the past four weeks and 0 otherwise, and *male* that takes value 1 when the respondent's gender is male and 0 when female. The remaining information comes in the form of categorical variables related to age, income, type of area, size of the municipality and region where the person lives.

Table 1 shows different information on the composition of the database and cross-tabulation between the decision to travel and the variables included in the survey. In 23,388 cases, the survey respondent did not make any long-distance trip in the previous four weeks, representing a large proportion of zeroes in the sample used to estimate the model. We can observe how there is a similar share of the total 50,730 observations of male respondents (48.6%) and females (51.4%). For low- and medium-income individuals, this represents

[2] Methodology and definitions applied in this survey comply with the requirements set by the European Commission for long-distance travel surveys. Movilia survey 2007 is the most recent version of the Movilia survey, published in 2008.

almost 83% of total observations. The majority of the surveyed population lives in not metropolitan areas (70%) and lives in municipalities with a standing population below 500000 in- habitants.

In our analysis of travel decisions it is important to note that the percentage of individuals who made long-distance trips (53%) is slightly higher than the individuals who did not make any trip with a length of over 50 kilometres in the previous three months (47%). When these global figures are analysed by gender, we observe that the percentage of men who did long-distance travel is greater than the percentage of women (57.8% and 51.5%, respectively).

Exploratory analysis also indicates important differences by age; we find that 74.6% of individuals 65 or older did not travel whereas in the rest of the age groups, percentage of non-travel individuals was lower than 50%. The percentage of non-participation was minimal for individuals ages 15-29. These values suggest an inverse relation between age and long-distance travel.

We constructed a proxy for a personal income variable, relying on education level and labour information. The low-income group was composed of the unemployed, housewives, retirees, students and unschooled children, and employed people with pre-primary education. The medium-income group was comprised of employed people with primary and secondary educations. The high-income group consisted of workers holding a university degree or vocational training. As we expected, individuals with low income mainly do not travel (58.8%) while in the medium- and high-income categories the proportion of travellers is higher, especially for individuals with high income (76.3%).

Regarding the geographical variables, cross-tabulation indicates that the percentage of individuals who decide to travel is slightly higher than 50% and the size of the municipality exerts positive effects in terms of participation. It is also possible to analyse the distribution of individuals who travelled and did not travel by region of residence of the respondent. In Castilla-La Mancha, Navarra and Extremadura, the percentage of individuals who travelled was equal to 61.6%, 61% and 60.9%, respectively. However, in Murcia on average 56.5% of the individuals did not travel in the reference period.

Table 1. Cross-tabulation of travel decisions

	Indicates if the person travels				
	No travel		Travel		Total
	Col. %	Row %	Col. %	Row %	Col %
Gender					
Female	56.3	51.5	47	48.5	51.4
Male	43.7	42.2	53	57.8	48.6
Age					
15 to 29	15.4	34.7	25.6	65.3	20.8
30 to 39	13.8	38.2	19.8	61.8	17
40 to 49	15.5	38.4	22	61.6	18.9

50 to 64	21.9	46.3	22.5	53.7	22.2
65 or more	33.4	74.6	10.1	25.4	21
Income					
Low	62.6	58.8	38.8	41.2	50
Medium	28.7	41.1	36.4	58.9	32.8
High	8.7	23.7	24.8	76.3	17.2
Area					
Not metropolitan	68.8	46.1	71.2	53.9	70.1
Metropolitan area	31.2	49	28.8	51	29.9
Municipality size					
Less than 10000	36.8	49	33.9	51	35.3
10000 to 50000	25.2	48.3	23.9	51.7	24.5
50000 to 500000	32.5	44.5	36	55.5	34.3
More than 500000	5.5	44	6.2	56	5.9
Region of residence					
Madrid	3.3	48.2	3.1	51.8	3.2
Galicia	9.7	51.9	8	48.1	8.8
Asturias	2.9	55.6	2.1	44.4	2.5
Cantabria	2.8	53.1	2.2	46.9	2.4
Basque Country	7.4	53.1	5.8	46.9	6.5
Rioja	1.7	46.7	1.7	53.3	1.7
Navarre	2.1	39	2.9	61	2.5
Aragon	5.7	42.7	6.8	57.3	6.3
Catalonia	8.7	46.4	8.9	53.6	8.8
Valencia	8.2	53.8	6.2	46.2	7.2
Murcia	2.7	56.5	1.8	43.5	2.2
Castile and León	15.9	42.1	19.4	57.9	17.7
Extremadura	3.1	39.1	4.2	60.9	3.7
Castile-La Mancha	7.7	38.4	10.9	61.6	9.4
Andalusia	18.2	50.2	16	49.8	17
Sample size	23,833		26,897		50,730
Total	100	47	100	53	100

Distance travelled in intercity trips, measured in kilometres, is the dependent variable of the model estimated in next section. Table 2 collects information on this variable, including different descriptive statistics for the sample, restricted to observations representing a long-distance trip. The mean distance is slightly higher for female travellers than men, while the median of the distance travelled is higher for male travellers, although these differences are very small. An analysis of distance travelled by age groups show that younger individuals travel shorter distances than older individuals. According to the statistics, the maximum distance is reached for individuals in the group between 50 and 64 years, while distance declines for seniors above 65. The residence of travellers also seem to affect distance travelled. In particular, the length of trips are higher on average for those individuals living in metropolitan areas. Residents of the Madrid Autonomous Community travel further than inhabitants of other regions, with a mean distance value of more than 200 kilometres, followed by travellers from Basque Country and La Rioja. On average, people living in Aragon are the travellers with the shortest long-distance trips (130.1 km).

TABLE 2. SUMMARY OF DISTANCE VARIABLE IN KM. FOR TRAVELLERS (N= 26,879)

	Mean	Sd	Median
Gender			
Female	151.2	191.2	89
Male	150.7	182.2	90
Age			
15 to 29	144	172.6	89
30 to 39	148.6	180	89
40 to 49	153	183.5	91
50 to 64	157.3	195.2	91
65 or more	154.6	216.7	82
Income			
Low	148.8	195.4	86
Medium	147.4	177.4	90
High	159.5	184.9	96
Area			
Not metropolitan	143.3	175.4	87
Metropolitan area	169.7	210.2	96
Municipality size			
Less than 10000	133.5	170.1	83

10000 to 50000	150.1	181.2	89
50000 to 500000	164	203.3	93
More than 500000	173.6	181.5	104
Region of residence			
Madrid	201.8	205.4	123
Galicia	141.9	203.7	78
Asturias	171	281.1	82
Cantabria	166.7	225.3	90.5
Basque Country	187.5	243.8	90
Rioja	182.9	258.9	102
Navarre	151.7	215.9	78
Aragon	130.1	126.2	93
Catalonia	153.6	223.4	90
Valencia	150.3	158.9	86
Murcia	145.7	168.9	73
Castile and León	150.6	177.3	90
Extremadura	138.4	136.6	89
Castile-La Mancha	138.3	147.8	90
Andalusia	144.5	165	90
Mode of transport			
Car	127.9	119.7	86
Bus	157.3	149.3	97
Train	200.8	172.9	126
Plane	875.7	570.4	661
Other	149.7	154.9	84
Trip purpose			
Pleasure	171.6	212.8	100
Business	152.2	158.1	90
Second residency	106.9	119.2	69
Other purpose	97	86.3	74
Total	150.9	186.5	89

This section has provided some insights into long-distance travel in Spain relying on an exploratory analysis. The estimation of econometric models presented in the next section allows us to make inferences on the effects of socio-economic, demographic and geographical variables on travel distance for intercity trips.

4. Econometric Analysis

4.1. Model Specification

In the empirical part of the study we aim to estimate the contribution of different factors to long-distance travel by the Spanish population. The dependent variable, length of the trip measured in kilometres, is truncated since a large fraction of the observations are zeroes as it was reported in the previous section. Hence, a basic ordinary least square (OLS) regression would yield biased parameter estimates. In this situation, different methodologies are usually applied to address this problem.

Zero observations may appear in any microeconomic data for different reasons as explained by Gibson and Kim (2011): an individual is a genuine non-consumer of the transport services, due to either abstention (for any sort of preferences) or to a corner solution where the individual is a potential consumer of the service but cannot afford it at the current levels of income and prices. The case of abstention can be modelled using Cragg's model, which is compounded by a Probit equation that accounts for the participation decision in the market and a Tobit model that explains the quantity of the service consumed or the expenditures. Conventionally, a Tobit model is applied to corner solutions where the demand for the service is censored from below. In addition to this, even when the individual usually consumes the analysed (goods or) services, the existence of zeroes may be due to infrequently purchased goods if the survey period is shorter than the times between consumption of the good. In this context, the design of the survey becomes an important theme since several waves of the survey may reduce the second source of zeroes. A more detailed explanation can be found in Newman *et al*. (2003) and in the references given therein. In this situation the appropriate specification would be the Infrequency of Purchase Model.

The standard Tobit model (Tobin, 1958) has been frequently used to estimate demand relationships with limited dependent variables. However, the Tobit model can be very restrictive because, apart from econometric reasons, it assumes that zero consumption observations only arise from corner solutions generated by a constrained budget. To overcome this restrictive assumption, different double-hurdle models, as used in the present analysis, have been proposed in the economic literature.

An alternative commonly known as the Double-Hurdle Model (DHM), originally formulated by Cragg (1971), supposes that individuals make their consumption decisions in two steps, each of which is determined by a different set of explanatory variables. Accordingly, two separate hurdles (the decision about whether or not to consume and about the quantity consumed) must be passed before observing a positive level of consumption. Thus, the model contains a participation equation,

$$d^* = \alpha' z + u \qquad u \sim N(0,1) \tag{1}$$

where the latent variable d^* represents, in our case, the decision to make long-distance trips as a linear function of first-hurdle regressors (z), and a consumption equation,

$$y^* = \beta' x + v \qquad v \sim N(0,1) \tag{2}$$

characterised by the latent variable y^*, the distance travelled, as a linear function of second-hurdle regressors (x). The observed distance travel, y, is such that

$$y = \begin{cases} y^* & \text{if } d^* > 0 \text{ and } y^* > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Assuming that the error terms u and v can be correlated allows for the possibility that the decision to travel and the decision about travel distance are made simultaneously. Then, it is considered that u and v are distributed as a bivariate normal,

$$(u,v) \sim BVN(0,\Sigma), \quad \Sigma = \begin{bmatrix} 1 & \sigma\rho \\ \sigma\rho & \sigma^2 \end{bmatrix} \tag{4}$$

where ρ is the correlation coefficient.
Estimating this bivariate model requires the maximisation of the following likelihood equation,

$$L = \prod_0 \left[1 - \Phi\left(\alpha' z, \frac{\beta' x}{\sigma}, \rho\right)\right] \prod_+ \left[\Phi\left(\frac{\alpha' z + \frac{\rho}{\sigma}(y - \beta' x)}{\sqrt{1-\rho^2}}\right) \frac{1}{\sigma} \phi\left(\frac{y - \beta' x}{\sigma}\right)\right] \tag{5}$$

denoting zero distance travelled as 0 and positive distance as +, and where Φ and φ are the standard normal cumulative and density functions, respectively.

If considered with the independent double-hurdle model (independence between u and v), the likelihood reduces to

$$L = \prod_0 \left[1 - \Phi\left(\alpha' z, \frac{\beta' x}{\sigma}\right)\right] \prod_+ \left[\Phi(\alpha' z) \frac{1}{\sigma} \phi\left(\frac{y - \beta' x}{\sigma}\right)\right] \quad (6)$$

It should be noted that the standard Tobit model is a nested version of this Cragg model, when $\Phi(\alpha' z) = 1$.

An alternative to the double-hurdle model is to consider that the participation decision dominates the quantity decision, as in the Heckman selection model (Heckman, 1979). In this case, the corresponding likelihood function can be written as

$$L = \prod_0 \left[1 - \Phi(\alpha' z)\right] \prod_+ \left[\Phi\left(\frac{\alpha' z + \frac{\rho}{\sigma}(y - \beta' x)}{\sqrt{1-\rho^2}}\right) \frac{1}{\sigma} \phi\left(\frac{y - \beta' x}{\sigma}\right)\right] \quad (7)$$

The Heckman specification reduces to a probit model for participation and an ordinary least squares model for the consumption if both equations are independent, that is $\rho = 0$.

In addition to these models, we have considered the infrequency of purchases model (Deaton and Irish, 1984; Blundell and Meghir, 1987), also based on a bivariate structure. In this model, zero observations may be due either to the survey period being too short for positive purchases to be observed, or to a standard corner solution. In fact, the likelihood of the infrequency of purchases model with independent error terms is similar to the double-hurdle model and can be expressed as

$$L = \prod_0 \left[1 - \Phi\left(\alpha' z, \frac{\beta' x}{\sigma}\right)\right] \prod_+ \left[[\Phi(\alpha' z)]^2 \frac{1}{\sigma} \phi\left(\frac{\Phi(\alpha' z) y - \beta' x}{\sigma}\right)\right] \quad (8)$$

although now the purchase decision is related to the actual costs through the expression $\Phi(\alpha' z)y$. Indeed, the Tobit model is nested within the above model when $\Phi(\alpha' z) = 1$.

4.2 MODEL SELECTION AND ESTIMATION RESULTS

Before discussing these results in detail, we make some comments about the consideration of alternative models. In this exercise, we consider some exclusion restrictions to identify the participation effect although there is not a common strategy in the literature about their need and selection. Moreover, the first stage participation hurdle is likely the result of psychological and cultural factors aside from the levels of prices and income, which should be included with any other economic factors in the second hurdle of the model according to the basic assumptions of the models (Newman and Matthews, 2001). In the second equation, we have added the income variables to explain the length of the trip, allowing for the existence of corner solutions. From an econometric point of view the use of different sets of regressors, i.e. imposing some exclusion restrictions in each hurdle, can help with parameter identification (Jones, 1992). Similarly to Humphreys *et al*. (2011) the convergence is not achieved without these restrictions when the double-hurdle model is estimated, allowing for correlation between the errors of the participation equation and quantity (distance trip) equation.

In order to select the most suitable model for this particular application we estimated different models and ran appropriate statistical tests, which are shown in Table 3. The first step is the comparison between the Tobit model and the independent double-hurdle model using a likelihood ratio test (LR) because Tobit model is nested in Cragg's Double Hurdle Model (1971) under the hypothesis that the parameters on both equations are identical. This is the expression of the LR test

$$LR = -2[L(\hat{\beta}_1,\hat{\sigma}_1^2) - L(\hat{\beta}_2,\hat{\sigma}_2^2) \sim \chi_m^2 \qquad (9)$$

where m is the number of exclusion restrictions, $L(\hat{b}_1,\hat{s}_1^2)$ is the maximum value of the double hurdle log-likelihood function and $L(\hat{b}_2,\hat{s}_2^2)$ is the maximum value of the Tobit log-likelihood function. The null hypothesis establishes the equivalence between the double hurdle model and the Tobit model, which implies that the more parsimonious Tobit model fits the data better. Consequently, the rejection of the null hypothesis means that the double hurdle model is a better alternative to fit the data. In this application, a likelihood ratio test with value of 56486.18 clearly rejects the null hypothesis of model equivalence between the Tobit model and Double Hurdle Model, favouring the selection of the latter. Similarly, instead of assuming independence of error terms between the two parts of the model, we estimated a Dependent Double Hurdle Model finding a ρ parameter statistically insignificant, so we stick with the Independent Double Hurdle Model.

In the second step, we also estimated another common model used to address zeroes in the dependent variable, the Heckman Selection Model (HSM). Parameter ρ, which indicates dependence between the participation equation and the second hurdle, was not significant. Next we have to choose between the independent double-hurdle model and HSM, which is not nested in the double hurdle model. The LR test is not feasible in this case and the best model specification is determined using the Voung test for non-nested models (Vuong, 1989).

Suppose that if $f(y_i|X_i,\boldsymbol{\theta})$and $g(y_i|X_i,\boldsymbol{\gamma})$ are two alternative specifications for the density of the random variable y, the test requires the computation of both sets of predicted probabilities to obtain

$$m_i = \log \frac{f\left(y_i \middle| X_i, q\right)}{g\left(y_i \middle| X_i, g\right)}$$

This value is the difference (for each observation) between their contributions to the likelihood function. The Voung's statistic is

$$V = \frac{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} m_i\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_i - \bar{m})^2}} \qquad (10)$$

The distribution of this standard statistic is a limiting normal standard distribution and it is bidirectional[3]. If the absolute value of the standard statistic is less than two, the two competing models are considered equivalents. In our results, the value of the standard statistic is 4.513 so DHM is preferred instead of the HSM.

Finally, as it was described above, the IPM model is estimated and evaluated in comparison to the independent double-hurdle model. In this case, the Voung test statistic is 11.05 so the hypothesis of model equivalence is not rejected and both models are not significantly different. The large value of this statistic indicates that the DHM is the preferred specification alternative.

Overall, test results indicate that Cragg's Double Hurdle Model is superior to the rest of the specifications. This result is consistent with our previous hypothesis because DHM allows us to introduce flexibility into the model, contemplating the possibility of budget restrictions of individuals making them choose not to travel. Consequently, zeroes found in the survey seem to respond to corner solutions instead of the infrequency of long distance trips[4].

[3] Test is completely described in Greene (2003).

[4] Estimation of the DHM is performed using the Craggit routine by Burke (2009). IPM is estimated using code based on the work provided by Gibson and Kim (2011).

Table 3. Model Selection tests

Models	Test	Value	Conclusion
IDHM vs Tobit	LR test (m = 2)	56486.180***	Reject H_0
IDHM vs HSM	Non-nested Vuong	4.513***	Reject H_0
HSM vs IPM	Non-nested Vuong	11.016***	Reject H_0
IDHM vs IPM	Non-nested Vuong	11.051***	Reject H_0

Significance code: *p < .1, **p < .05, ***p < .01

IDHM: Independence double hurdle model (Cragg, 1971); HSM: Heckman sample selection model (Heckman, 1979); IPM: Infrequency purchase model (Gibson and Kim, 2011).

Table 4 displays Cragg's Double Hurdle model (DHM), revealing interesting patterns in long-distance decisions[5]. For the most part, our results are consistent with those reported from the studies on long-distance travel reviewed in Section 2. The participation equation indicates the effect of the explanatory variables on the probability of being involved in long-distance travel, while the second equation tries to assess the impact of the covariates on distance travelled. This second equation follows a semi-log linear form by employing the logarithms of the dependent variable. The magnitude of the ML estimates in these models cannot be interpreted in the same fashion as, for example, OLS estimates, as they are based on latent equations. However, the sign of the parameters can be interpreted, presenting an intuitive interpretation of the factors determining long-distance travel.

Females are less likely than males to participate in long-distance travel (estimated coefficient of dummy variable Male is 0.190) and, if they participate at all, they are involved in shorter trips than men (the value of 0.039 indicates that men travel longer distances). Long-distance travel participation clearly declines with age, however the effect of age on distance travelled is not so clear. According to DHM estimates, distance travelled increases for aging adults, taking as a reference the base category of 15 to 29 years old. It seems that being a senior is not different, in terms of distance travelled, from being in the omitted category. These results are similar to those obtained in Dargay and Clark (2012) but contrary to the findings in Collia *et al*. (2003) and Georggi and Pendyala (2001), although the latter two studies only relied on cross-tabulation analysis. As stated above, income variables only appear in the second equation. In both models, results indicate that individuals in the higher income categories travel further than low-income citizens, that is, there is a positive relationship between income and distance travelled.

[5] Estimates of the Heckman Selection Model and the Infrequent Purchases Model can be found in the Appendix.

Table 4. Independent Double Hurdle Model

	Participation Equation		ln(Distance)	
	Coeff.	S.E.	Coeff.	S.E.
Gender: Male	0.190***	-0.01	0.039***	-0.01
Age: 30 to 39	-0.0765***	-0.02	0.076***	-0.02
Age: 40 to 49	-0.0949***	-0.02	0.121***	-0.01
Age: 50 to 64	-0.298***	-0.02	0.129***	-0.01
Age: 65 or more	-1.059***	-0.02	0.006	-0.02
Income: Medium	-	-	0.046***	-0.01
Income: High	-	-	0.101***	-0.01
Municipality: 10,000-50,000	0.045***	-0.02	0.070***	-0.01
Municipality: 50,000-500,000	0.159***	-0.02	0.119***	-0.01
Municipality: $ > $500,000	0.235***	-0.03	0.200***	-0.03
Area: Metropolitan	-0.067***	-0.02	-0.027*	-0.01
Region: Galicia	0.031	-0.04	-0.319***	-0.03
Region: Asturias	-0.058	-0.05	-0.240***	-0.04
Region: Cantabria	-0.059	-0.05	-0.169***	-0.04
Region: Basque Country	-0.060	-0.04	-0.072**	-0.04
Region: Rioja	0.054	-0.06	-0.083*	-0.05
Region: Navarre	0.349***	-0.05	-0.192***	-0.04
Region: Aragon	0.264***	-0.04	-0.087***	-0.03
Region: Catalonia	0.125***	-0.04	-0.216***	-0.03
Region: Valencia	-0.113***	-0.04	-0.204***	-0.03
Region: Murcia	-0.184***	-0.05	-0.236***	-0.05

Region: Castile and León	0.286***	-0.04	-0.070**	-0.03
Region: Extremadura	0.300***	-0.05	-0.070*	-0.04
Region: Castile-La Mancha	0.302***	-0.04	-0.091***	-0.03
Region: Andalusia	-0.0457	-0.04	-0.204***	-0.03
Purpose: Business	-	-	-0.073***	-0.02
Purpose: Second residency	-	-	-0.497***	-0.01
Purpose: Other purpose	-	-	-0.273***	-0.02
Mode: Bus	-	-	0.196***	-0.02
Mode: Train	-	-	0.395***	-0.02
Mode: Plane	-	-	1.976***	-0.03
Mode: Other	-	-	0.155***	-0.04
Constant	0.136***	-0.04	4.570***	-0.03
σ	-	-	0.790***	0.00
Log likelihood	63,994.6			
AIC	128,105.1			
BIC	128,617.5			

Significance code: *p<.1, **p<.05, ***p<.01

Geographical, or land-use, variables also play a significant role explaining the participation and travel distance decisions. Residents of larger cities are more likely to be involved in long-distance travel (as it can be seen in the coefficients 0.045, 0.159 and 0.235), and moreover, these segments of the population travel further than those living in smaller municipalities (as indicated by estimates of the second equation with values of 0.070, 0119 and 0.200). A plausible explanation for this result might be that longer distances travelled by individuals in higher populated areas might be due to greater levels of accessibility to transportation services, especially those designed for long-distance travel, such as high capacity roads, well-connected airports and high-speed rail. Additionally, greater participation of residents of large cities might also be related to this sort of reasoning, although in some studies which found a negative relationship between city size and long-distance travel participation may have motivated this result differently. More precisely, authors like Limtanakool *et al*. (2006a) stated that individuals who have a better access to services, i.e., those living in areas with higher population density, travel shorter distances than rural populations. That might be the case for the estimates related to the variable that indicates if a person lives in a metropolitan area or not. Our results show that living in a metropolitan area reduces the chances of engaging in long-distance travel (-0.067) and lowers the trip distance for those who travel (-0.027). A similar result has been found in the study of Mallet (2001) who found that the most immobile households were those located in large metropolitan areas. Dargay and Clark (2012) also found that those residents of rural areas were more prone to travel long distances than metropolitan citizens.

The region of residence of the respondent is introduced as a control variable and the omitted region is the Autonomous Community of Madrid. Results of IPM show that residents in any other region are more, or at least equally, likely to make a long-distance trip than those living in Madrid. This result does not hold for the DHM where people living the Valencian Autonomous Community and the Region of Murcia are less likely to travel more than 50 km. than the Madrid population. DHM estimates of the second equation show that populations living in any region outside of the Madrid Autonomous Community make shorter trips. More variability is found when estimating the IPM, where depending on the region impact on distance travelled is positive, negative or null.

For the purpose of the trip the omitted category is the one collecting holidays and leisure-motivated trips, while private vehicle is the base category in the mode of transport used. Variables related to purpose and mode only appear in the second equation, as only the individuals who actually travel report information on these. Results confirm that when the purpose of the trip is leisure and holidays, long- distance users travel much further than those on business trips or those on trips to second residences. Mode of transport used is also introduced in the model as a control variable, and as expected, trips by car reduce the distance travelled compared to any other mode of transport, but especially when compared to train and plane.

Table 5. Discrete effects of the Double Hurdle Model

Variable	Prob.	Uncond.
Gender: Male	0.068***	0.337***
Age: 30 to 39	-0.027***	-0.088***
Age: 40 to 49	-0.034***	-0.094***
Age: 50 to 64	-0.107***	-0.428***
Age: 65 or more	-0.383***	-1.763***

City size: 10,000 to 50,000	0.016***	0.112***
City size: 50,000 to 500,000	0.057***	0.327***
City size: More than 500,000	0.085***	0.498***
Area: Metropolitan	-0.024***	-0.097***

Significance code: *p<.1, **p<.05, ***p<.01

Bootstrapped standard errors.

As previously indicated, the interpretation of the estimates must be taken with caution. In order to get a sensible interpretation of the results, marginal effects are computed and presented in Table 5. For categorical explanatory variables, marginal effects are used to compute percentage change in probability and unconditional level when the value of the variable shifts from zero to one, holding the rest of the variables constant.

For instance, the effect of 0.068 on the probability of travel implies that the probability of making a long-distance trip by a male (who is average in all other respects) is almost 7 percent higher than the probability when the individual is a female with similar characteristics. Of most interest is the overall effect on the dependent variable for values of the explanatory variables, known as unconditional expectation. The unconditional effect of 0.337 implies that the average individual distance is expected to increase by 34% when the person is male. More interestingly, when the unconditional effects of age are also analysed, it is found that growing older sharply reduces distance travelled, with a 176 percent reduction in the case of seniors with respect to the base category. The size of the city of residence also affects distance travelled, raising it by 50% for cities larger than 500,000 inhabitants in comparison to cities below 10,000 residents.

5. Summary and Conclusions

Long distance travel is increasing in frequency and especially in trip length. Although these sorts of trips are still outnumbered by daily commuting short trips, the distance travelled in intercity trips represents a growing share over the total distance. This leads to better integration between distant regions and greater accessibility to public and private services, but also causes congestion and pollution, among other important social, economic and environmental implications. A current topic of interest is studying factors affecting different aspects of long-distance travel demand such as mode choice, trip length and the travel-or-not decision.

In this study, factors determining long-distance travel are investigated. Using 2007 Spanish Ministry of Public Works' Mobility Survey, and different econometric modelling techniques, we explored the importance of different socioeconomic, demographic and land-use factors on the decision of making a long-distance trip and on the length of it. Other factors shown to be important for long-distance travel are the main motivation to travel and the mode of transport used.

The main findings indicate that males travel further than females, and younger segments of the population participate more in the long-distance travel market and also travel longer distances. Individuals who enjoy higher income levels travel more and to more distant places. Interesting findings related to geographical variables emerged. Citizens living in large cities travel further, but those whose residence is located in metropolitan areas travel less and make shorter trips. Airplane and railway passengers travel longer distances than those who use public bus services and especially those who drive private vehicles. The profile of the long-distance traveller is completed by the results showing that those who travel for leisure or holiday reasons travel more kilometres than those traveling for business or to visit a second residence.

Additionally, testing different model specifications allows us to make some comments about long-distance survey design. According to our results, the Double Hurdle Model is the preferred method over the Infrequent Purchase Model. The prevalence of DHM indicates that most zeroes in the survey are caused by corner solutions and the decision of abstention, i.e. with the different prices of transport and personal income given, individuals choose not to travel either because of economic reasons or preferences. This means that zeroes do not seem to be the result of the low frequency of a rare event such as an intercity trip. In this regard, it seems that the three-month period used in this survey is adequate to capture travel behaviour. However, our intuition is that this result may hold only if the current definition of long-distance travel is kept at 50 kilometres. We believe that if the threshold is increased to 100 kilometres, all other things being equal, it may require raising the duration of the reporting period if survey designers want to avoid infrequent zeroes. It is also important to note that although three months seems to be adequate in statistical terms for long-distance travel survey design, reducing the period by some weeks might increase the quality of survey responses. Designers must take into account the likely trade-off between reducing infrequency of purchase and the difficulty of recalling past events caused by long reporting periods. Additional research on these questions would be desirable in order to find the optimal reporting period for intercity trips.

References

Arbués, P., Baños, J.F., Mayor, M., Suárez, P. (20139: "Determinants of Ground Transport Modal Choice in Long-distance Trips in Spain", *Efficiency Series Paper* 7/2013, Departamento de Economía, Universidad de Oviedo.

Axhausen, K. (1999): "Preparing for the Long of It: Methodological Research for an European Survey of Long Distance Travel", Paper presented at *TRB Conference: Personal Travel: The Long and Short of It*, Washington DC, 28 June1 July .

Baik, H., Trani, A.a., Hinze, N., Swingle, H., Ashiabor, S., Seshadri, A. (2008): "Forecasting Model for Air Taxi, Commercial Airline, and Automobile Demand in the United States", *Transportation Research Record: Journal of the Transportation Research Board* 2052, 9–20.

Bhat, C.R. (1997): "An Endogenous Segmentation Mode Choice Model with an Application to Intercity Travel", *Transportation Science* 31, 34–48.

Blundell, R., Meghir, C. (1987): "Bivariate Alternatives to the Tobit Model", *Journal of Econometrics* 34, 179–200.

Burke, W. (2009): "Fitting and Interpreting Cragg's tobit Alternative Using Stata", *Stata Journal*, *9*(4), 584–592.

Collia, D.V., Sharp, J., Giesbrecht, L. (2003): "The 2001 National Household Travel Survey: A Look into the Travel Patterns of Older Americans", *Journal of Safety Research* 34, 461–470.

Cragg, J. (1971): "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods", *Econometrica: Journal of the Econometric Society* 39, 829–844.

Dargay, J.M., Clark, S. (2012): "The Determinants of Long Distance Travel in Great Britain", *Transportation Research Part A: Policy and Practice* 46, 576–587.

Deaton, A., Irish, M. (1984): "Statistical Models for Zero Expenditures in Household Budgets", *Journal of Public Economics* 23, 59–80.

Erhardt, G., Freedman, J., Stryker, A., Fujioka, H., Anderson, R. (2007): "Ohio Long-Distance Travel Model", *Transportation Research Record* 2003, 130–138.

Frändberg, L., Vilhelmson, B. (2003): "Personal Mobility: A Corporeal Dimension of Transnationalisation. The Case of Long-distance Travel from Sweden", *Environment and Planning A* 35, 1751–1768.

Georggi, N.L., Pendyala, R.M. (2001): "Analysis of Long-distance Travel Behavior of the Elderly and Low Income", *Transportation Research Circular E-C026*, 121–150.

Gibson, J., Kim, B. (2011): "Testing the Infrequent Purchases Model Using Direct Measurement of Hidden Consumption from Food Stocks", *American Journal of Agricultural Economics* 94, 257–270.

Greene, W.H. (2003): *Econometric Analysis*, Prentice Hall, New Jersey. 5th edition.

Heckman, J. (1979): "Sample Selection Bias as a Specification Error Author", *Econometrica: Journal of the Econometric Society* 47, 153–161.

Holz-Rau, C., Scheiner, J., Sicks, K. (2014): "Travel Distances in Daily Travel and Long-distance Travel: What Role is Played by Urban Form?" *Environment and Planning A* 46, 488–507.

Humphreys, B., Lee, Y., Soebbing, B. (2011): "Modeling Consumers' Participation in Gambling Markets and Frequency of Gambling", *Journal of Gambling and Business and Economics* 5, 8–14.

Johansson-Stenman, O. (2002): "Estimating Individual Driving Distance by Car and Public Transport Use in Sweden", *Applied Economics*, 34(8), 959–967.

Jones, A.M. (1992): "A Note on Computation of the Double-hurdle Model with Dependence with an Application to Tobacco Expenditure", *Bulletin of Economic Research* 44, 67–74.

Kuhnimhof, T., Collet, R., Armoogum, J., & Madre, J.-L. (2009): "Generating Internationally Comparable Figures on Long-Distance Travel for Europe", *Transportation Research Record: Journal of the Transportation Research Board*, 2105(-1), 18–27

Limtanakool, N., Dijst, M., Schwanen, T. (2006a): "On the Participation in Medium- and Long-distance Travel: A Decomposition Analysis for the UK and the Netherlands", *Tijdschrift voor Economische en Sociale Geografie* 97, 389– 404.

Limtanakool, N., Dijst, M., Schwanen, T. (2006b): "The Influence of Socioeconomic Characteristics, Land Use and Travel Time Considerations on Mode Choice for Medium- and Longer-distance Trips", *Journal of Transport Geography* 14, 327– 341.

Mallett, W. (1999): "Long-Distance Travel by Women: Results from the 1995 American Travel Survey", *Transportation Research Record* 1693, 71–78.

Mallett, W. (2001): "Long-distance Travel by Low-income Households". *Transportation Research Circular* E-C026, 169–177.

Mokhtarian, P.L., Salomon, I., Redmond, L.S. (2001): "Understanding the Demand for Travel: It's Not Purely 'Derived'. Innovation", *The European Journal of Social Science Research* 14, 355–380.

Newman, C., Henchion, M., Matthews, A. (2003): "A Double-hurdle Model of Irish Household Expenditure on Prepared Meals", *Applied Economics* 35, 1053–1061.

Newman, C., Matthews, A. (2001): "Infrequency of Purchase and Double-hurdle Models of Irish Households Meat Expenditure", *European Review of Agricultural Economics* 28, 393–412.

Tobin, J. (1958): "Estimation of Relationships for Limited Dependent Variables", *Econometrica: Journal of the Econometric Society* 26, 24–36.

Vuong, Q. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses", *Econometrica: Journal of the Econometric Society* 57, 307–333.

Appendix

	Heckman Selection Model				Infrequency of Purchase Model			
	Participation Equation		ln(Distance)		Participation Equation		ln(Distance)	
Variable	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
Gender: Male	0.190***	-0.01	0.038***	-0.01	0.135***	-0.01	0.258***	-0.02
Age: 30 to 39	-0.077***	-0.02	0.078***	-0.02	0.037***	-0.01	0.109***	-0.02
Age: 40 to 49	-0.095***	-0.02	0.123***	-0.01	0.033**	-0.01	0.127***	-0.02
Age: 50 to 64	-0.298***	-0.02	0.132***	-0.01	-0.102***	-0.01	-0.109***	-0.02
Age: 65 or more	-1.059***	-0.02	0.012	-0.02	-0.630***	-0.01	-1.107***	-0.03
Income: Medium	-	-	0.046***	-0.01	-	-	0.026***	-0.01
Income: High Income	-	-	0.100***	-0.01	-	-	0.063***	-0.01
Municipality size: 10,000 to 50,000	0.045***	-0.02	0.069***	-0.01	0.01	-0.01	0.0566***	-0.02
Municipality size: 50,000 to 500,000	0.159***	-0.02	0.118***	-0.01	0.080***	-0.01	0.207***	-0.02
Municipality size: > 500,000	0.235***	-0.03	0.199***	-0.03	0.200***	-0.02	0.475***	-0.04
Area: Metropolitan	-0.067***	-0.02	0.027*	-0.02	-0.081***	-0.01	-0.130***	-0.02
Region: Galicia	0.031	-0.04	-0.319***	-0.03	0.091***	-0.03	-0.001	-0.05
Region: Asturias	-0.058	-0.05	-0.240***	-0.04	0.056	-0.04	-0.025	-0.07
Region: Cantabria	-0.059	-0.05	-0.169***	-0.04	-0.018	-0.03	-0.121*	-0.07
Region: Basque Country	-0.060	-0.04	-0.074**	-0.04	-0.018	-0.03	-0.114**	-0.06
Region: Rioja	0.054	-0.06	-0.083*	-0.05	0.0303	-0.04	0.017	-0.08
Region: Navarre	0.349***	-0.05	-0.199***	-0.04	0.269***	-0.03	0.378***	-0.07

Region: Aragon	0.264***	-0.04	-0.087***	-0.03	0.347***	-0.03	0.567***	-0.06
Region: Catalonia	0.125***	-0.04	-0.217***	-0.03	0.135***	-0.03	0.127**	-0.05
Region: Valencia	-0.113***	-0.04	-0.204***	-0.03	-0.043	-0.03	-0.182***	-0.05
Region: Murcia	-0.184***	-0.05	-0.235***	-0.05	-0.023	-0.04	-0.160**	-0.07
Region: Castile and León	0.286***	-0.04	-0.070**	-0.03	0.268***	-0.03	0.447***	-0.05
Region: Extremadura	0.300***	-0.05	-0.070*	-0.04	0.244***	-0.03	0.399***	-0.06
Region: Castile-La Mancha	0.302***	-0.04	-0.093***	-0.03	0.307***	-0.03	0.498***	-0.05
Region: Andalusia	-0.0457	-0.04	-0.205***	-0.03	0.047*	-0.02	-0.019	-0.05
Purpose: Business	-	-	-0.072***	-0.02	-	-	-0.040***	-0.01
Purpose: Second residency	-	-	-0.498***	-0.01	-	-	-0.280***	-0.01
Purpose: Other purpose	-	-	-0.273***	-0.02	-	-	-0.152***	-0.01
Mode: Bus	-	-	0.196***	-0.02	-	-	0.108***	-0.01
Mode: Train	-	-	0.396***	-0.02	-	-	0.219***	-0.01
Mode: Plane	-	-	1.976***	-0.03	-	-	1.062***	-0.02
Mode: Other	-	-	0.155***	-0.04	-	-	0.112***	-0.02
Constant	0.136***	-0.04	4.570***	-0.03	0.022	-0.03	2.322***	-0.05
σ	-	-	0.790***	0.00	-	-	0.451***	0.00
λ	-	-	-0.009	-0.008	-	-	-	-
Log-likelihood	-64,042.8	-64,595.5						
AIC	128,203.5	129,306.9						
BIC	128,724.7	129819.3						

Significance code: *p<.1, **p<.05, ***p<.01