ECONOMIC DISCUSSION PAPERS

Efficiency Series Paper 7/2013

Determinants of ground transport modal choice in long-distance trips in Spain

Pelayo González, José F. Baños, Matías Mayor, Patricia Suárez

Departamento de Economía

Universidad de Oviedo

Available online at: http://economia.uniovi.es/investigacion/papers

Determinants of ground transport modal choice in long-distance trips in Spain

Pelayo González[a,1,*], José F. Baños[b,1], Matías Mayor [b,2,**], Patricia Suárez[b,1]

[a]*Facultad de Comercio, Turismo y Ciencias Sociales Jovellanos, Calle Luis Moya Blanco 261 33203 Gijón - Spain*

[b]*Facultad de Economía y Empresa, Avda del Cristo s/n. 33006 Oviedo - Spain*

Abstract

Information on the profiles of potential users is a major concern for producers and public agencies in every market. In the case of transport economics, modeling modal choice is a fundamental key for policy makers trying to improve the sustainability of transportation systems. However, existing empirical literature has focused on short- distance travel within urban systems. This paper contributes to the limited number of investigations on mode choice in medium- and long-distance travel. We employ data from the 2007 Spanish National Mobility Survey to shed light on how socioeconomic factors and trip attributes affect the selection of a primary mode of transportation. Our results confirm the impact of sociodemographic and land use variables on travel behavior. We have also tested the impact of a trip attribute not extensively used in the existing literature: stay duration. This attribute is found to be significant in explaining mode choice. A multilevel multinomial logit model accounts for spatial heterogeneity by including information about where an individual makes this travel decision.

Keywords:
Modal choice, Transportation demand, Multinomial logit, Multilevel

[*]Principal corresponding author. Telephone: 0034645463120

[**]Corresponding author

Email addresses: gonzalezpelayo@uniovi.es (Pelayo González), jbanos@uniovi.es (José F. Baños), mmayorf@uniovi.es (Matías Mayor), suarezcpatricia@uniovi.es (Patricia Suárez)

[1]Department of Economics

[2]Department of Applied Economics

analysis, Spatial heterogeneity

1. Introduction

Long-distance travel has rapidly increased in recent decades. Technological innovation, car ownership and economic growth are the major factors behind the rising demand for inter-city mobility. Improvements in highway, train and airport infrastructure have reduced travel costs and times while increasing safety. Schafer (1998) noted that travel time budgets have remained relatively constant, thus allowing people to travel further. Economic growth has also led to higher average disposable income; this favorable economic environment, along with less expensive car ownership, has contributed to the increase in long-distance business and personal trips.

Increased mobility implies economic, social and environmental consequences. According to Limtanakool et al. (2006), it permits a higher integration among regions, provides better accessibility to public services and social networks and extends the potential market for tourism activities. However, it can also have a negative environment effect, as longer trips involve increased energy consumption and greater emissions of pollutants. Furthermore, the investments required for enhancing long distance travel are substantial, implying high opportunity costs to the economy. In this sense, a deeper understanding of travel behavior in long-distance travel may ease transport policy challenges such as making mobility more sustainable and reducing negative externalities (Bhat, 1998). The demand for transportation services is the result of interactions between short- and long-term individual decisions. *Long term* travel planning typically involves making decisions concerning car ownership and residential or work locations. The final outcome of these decisions impact trip behavior, affecting short-term aspects such as modal choice, departure time and choice of route (Ben-Akiva and Bierlaire, 1999). To date, transportation and geographic researchers have typically studied the impact of these factors on daily and short-distance trips. Nonetheless, factors determining daily travel decisions and their interdependencies may have different impacts on less frequent events such as medium- and long-distance trips (Limtanakool et al., 2006).

This paper contributes to the overall discourse by increasing the knowledge in the determinants of modal choice in medium- and long-distance trips in Spain. More specifically, our primary goal is determining the influence of socioeconomic, land use and trip attributes on the selection of principal modes of transport among three possible choices: private car, bus and train. For this purpose, we use a database obtained from a survey carried out by the Statistical Office of Ministry of Public Works on the mobility patterns of Spanish residents. The data has a hierarchical structure, with travelers nested in their provinces of origin. This feature of the data permits us to incorporate unobservable variables through the application of multilevel analysis, which also enriches the existence knowledge in long-distance trips. In particular, we estimate a multilevel multinomial logit model with random intercepts to study the determinants of mode choice.

In line with previous research, our findings indicate that long distance modal choices are influenced by a combination of traveler socioeconomic characteristics, trip attributes and geographical factors. A key result is that the geographical context where the traveler begins the trip affects the alternative mode of transportation utilities in different ways. In addition, our analysis provides evidence of the positive effect of trip duration, measured as overnight stays, on railway demand; this has not been previously documented.

We initially review the existing literature on long-distance travel behavior, with an emphasis on modal choice studies. In Section 3, we explain the database and present some descriptive results of the included variables. In Section 4, we explain the foundations of rational choice relying on random utility theory in a multilevel framework. Section 5 displays the results of the estimated microeconometric models. Finally, Section 6 contains a summary of the major results and conclusions.

2. Previous research

Modal choice is the result of a complex process that includes objective and subjective determinants stemming from different disciplines and interrelated to a greater or lesser extent. In a recent survey, modal choice was defined as "the decision process to choose between different transport alternatives, which is determined by a combination of individual socio-demographic factor and spatial characteristics, and influenced by socio-psychological factors" (De Witte et al., 2013). In another survey, Buehler (2011) compared transport mode choice determinants in the case of Germany and the USA. Although both surveys were comprehensive examinations of the literature, they did not clearly distinguish between short- and long-distance studies.

Long-distance or inter-city trips are usually differentiated from short trips through the use of a distance threshold. Although there is no standard definition, trips are usually defined as long distance if they are longer than a threshold between 50 to 100 Km. (Axhausen et al., 2003). Long distance trips involve more time and out-of-pocket cost, so the traveler facing the modal choice decision is in a different situation than an individual making a short-distance trip. In addition, modal availability and travel purposes are also different. While the principal motive of short-distance travel is commuting, long distance transit is dominated by pleasure and business pursuits. Therefore, long-distance trips are less frequent, making travelers less familiar with available transportation alternatives. Decisions on mode choice can be affected by similar variables in short and long-distance trips, but the impact of the same variables can be conditional to the distance travelled. The purpose of this literature review is not to find these differences but to focus our attention on the determinants typically found in long-distance travel studies.

Sociodemographic factors play a significant role in transport mode decisions. Bhat (1997) applied an endogenous segmentation model to the estimation of inter-city travel mode choices in the Toronto-Montreal Corridor. The author found that women were more responsive than men to rail frequency improvements in Canadian inter-city travel. Limtanakool et al. (2006) estimated binary logit models that distinguished between private cars and trains, finding that women were more likely than men to use trains. Georggi and Pendyala (2001) and Mallett (1999) also found that women were slightly less car dependent in long-distance trips. The effect of the travelers age on modal choice was not as clear. Limtanakool et al. (2006) indicated that senior commuters were more likely to use private cars than middle aged and young travelers. In an analysis of elderly and low-income mobility, Georggi and Pendyala (2001) found that the elderly were more bus dependent. As for the impact of education on modal choice, only Limtanakool et al. (2006) included it as an explanatory variable for long-distance models. They found that highly educated commuters tended to use public transportation more often. It is important to note that age, education and occupation are related to income and car ownership. High income travelers have higher opportunity costs and values of time, which implies the selection of faster transportation modes such as airplanes because they are more sensitive to travel time improvements (Bhat, 1997). In another study by Mallett (2001), lower-income individuals were slightly more dependent on buses and other public transport modes.

Spatial configuration indicators such as population density, diversity of land use and accessibility to transportation infrastructure are typically included in short-distance mobility studies (De Witte et al., 2013). Few papers studying inter-city modal choice include land use factors. An exception is Limtanakool et al. (2006), which found that higher population densities and higher degrees of mixed land use around public transport stations make these modes more attractive in long-distance travel. In a recent descriptive study, Garmendia et al. (2011) found that travelers from cities less than 10 Km. away from a high-speed rail station more frequently chose trains than cars for their trips to metropolitan areas. However, it is possible that there may be a self-selection effect; individuals with a preference for public transport may move to areas with an abundance of these services (Buehler, 2011).

Journey characteristic indicators such as purpose, distance, frequency and travel time also impact the mode selection in medium- and long-distance travel. Moeckel et al. (2013) crafted an exhaustive description of long-distance mode choice studies focusing on trip attributes. Among the principal travel motives, existing literature usually distinguishes between commuting, business and leisure. The results in Limtanakool et al. (2006) indicated that private car use was very prominent for business trips, while commuters relied to a greater extent on trains. Georggi and Pendyala (2001) compiled a descriptive cross tabulation analysis of the 1995 American Travel Survey, showing that modal distribution changed across trip purposes. In a paper that attempted to assess the impact of high-speed train investments on the mobility of Spanish residents, Martín and Nombela (2007) found that public transport modes were more attractive in commuting to work. Faster travel modes are usually preferred in the case of longer distances. Martín and Nombela (2007) found distance to be a positive effect in selecting trains and a negative for buses. Koppelman and Sethi (2005) estimated models with different methodologies and concluded that distance discouraged travel by automobile. The empirical analysis carried out by Bel (1997) indicated the importance of travel times in modeling modal choice. Longer railway travel times negatively affect rail demand, while travel times by road have a positive relationship on rail demand.

Overall, the existing empirical evidence confirms that socioeconomic indicators, trip characteristics and land use factors affect mode choice decisions. Another source of variation in long-distance travel behavior arises from traveler location. Bricka (2001) showed that traveler compositions are different among states in terms of household income and race. While trip purposes seem to be quite stable for the analyzed states, dissimilar mode choices are explained by availability of modes and the urban form in the location of origin. The results from Lapparent et al. (2013) suggested disparities between European countries when heterogeneous preferences were taken into account. Spatial features seem to play an important role in modal decisions, not only through the impact of land use configurations but also through the spatial heterogeneity in traveler composition and preferences.

Many statistical methods assume that relationships are constant over the space of the sample, i.e., all coefficients are forced to be identical (or stationary) for all individuals, locations or zones. This hypothesis is likely to be violated in the case of mode choice due to the influence of the geographical and socioeconomic context, as mentioned above. Páez (2006) offered a review of the different alternatives proposed in existing literature to overcome this problem: market segmentation, the introduction of dummy variables in the model, the Casetti's expansion method and multilevel models. After evaluating the database information in Section 3 and the theoretical implications discussed in Section 4, we selected a multilevel multinomial model.

3. Database description

In this section, we discuss the data employed in the estimation of the model explained in section 4. The primary source of information is a mobility survey (Movilia 2007) carried out by the Statistical Office of the Ministry of Public Works [3]. The objective of the survey was to study the basic characteristics of Spanish resident travel to better understand Spanish population mobility habits. The sample unit only considered the movements of Spanish residents; it did not include trips by tourists and non-resident immigrants that might have affected certain locations. In this survey, long-distance trips were defined as those longer than 50 Km.. They collected the microdata information on long-distance trips longer than 50 Km. through quarterly telephone surveys between February 2007 and January 2008. The dataset revealed preferences about the individual characteristics of the trip maker, land use factors and attributes of the trip. For each trip, we know the gender of the interviewed person, their age group, employment status and educational level. The trip characteristics included the province of origin, the trip

[3]Methodology and definitions applied in this survey comply with the requirements set by the European Commission for long-distance travel surveys.

purpose, the selected mode of transportation, the distance and its duration. Despite the robust information provided by the survey, we did not have information on individual income, household characteristics such as number of children, car ownership, disposable income and time or cost of the trip. Moreover, the geographical information was insufficient for computing accessibility to public transportation infrastructure. While there was information on the province where the trip began, there was no detailed information on the specific town of origin. As for trip destinations, information was available on whether the destination province was adjacent or of the same region as the province of origin.

We overcame these caveats by using the available information and including unmeasured and unobservable characteristics. In our analysis of the variability of ground mode choice, we not only controlled for the effect of the individual and trip characteristics but also for the spatial context where they belong. This feature of the multilevel model (that will be explained in Section 4) allowed us to include random intercepts containing specific information on the trips province of origin. This information could be related to specific area characteristics that are not specifically included as explanatory variables but may be relevant for modal choice. The individuals in the same geographical unit are likely to be similar in some ways due to these unobserved characteristics (Hong et al., 2013). For instance, the spatial environment where the traveler makes the choice is bounded by the borders of the trip province of origin, which is linked to the access to transportation infrastructure and the public services provides within this area.

The empirical model to be estimated requires some database modifications [4]. We constructed a cost variable by multiplying the distance reported in the survey by the average price per kilometer for each mean of transportation [5]. The motivation to use a cost (rather than distance) variable can be found in Whalen et al. (2013), where the authors criticized the use of a dis-

[4]We have disregarded observations on plane trips and other means of transport because our objective is the study of ground transportation modes and their interactions.

[5]The average prices used in these computations are 0.19 /Km., 0.0877 /Km. and 0.09202 /Km. for car, bus and train, respectively. These prices were obtained from different sources. In the case of cars, the source is the law that sets the trip compensation for public workers in case of travel. The average bus price comes from a report of the National Competence Commission in 2006, while the train price per kilometer was obtained from the Railway Yearly Report by the Spanish Railway Foundation.

tance variable because it was an individual-specific attribute and equivalent all modes. In contrast, the cost variable is treated as a mode-specific variable and (according to these authors) superior to distance because it is more easily interpreted and more accurately represents the characteristics of the trip.

After the data were prepared, the analyzed database includes 19,514 observations as displayed in Table 1. Each observation represents a trip and collected information on the individual characteristics of the traveler, variables describing the trip origin and specific attributes such as purpose, duration and distance to build the cost variable. Personal information showed a similar proportion of men and women and that the age of the respondents in almost 69% of the sample observations was below 50 years. Two thirds of the reported cases had completed secondary education. We constructed a proxy for a personal income variable relying on educational level and labor information, the latter being primarily composed of employed workers (66%). 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 explained in Section 2, among the common explaining variables of mode choice found in existing literature were geographical variables with information about land use, population density and accessibility to transportation infrastructure. Available information in this survey included the province of origin, a categorical variable for the city size and a variable to determine if the origin was located in a metropolitan area. More than two thirds of the trips began in a non-metropolitan area. The size of the city of origin, measured in population, was less than 50,000 habitants in 54% of the observations, while trips beginning in a city larger than 500,000 residents accounted for 7.1%. We expected that larger cities located in metropolitan areas would have a higher population density. Higher densities are related to improved public transportation in the sense of higher frequencies of public transport and better connections.

Finally, two variables containing information about attributes of the travel were available. The distance travelled during the trip indicated that almost half of the observations were trips shorter than 100 kilometers. The duration of the trip was measured as the number of overnight stays. In half of the sample observations, the traveler did not stay overnight; in 40% of the cases, overnight stays ranged from 1 to 6 nights.

Table 1: Descriptive Statistics

	Frequency	*Percentage*	*Cum. Perc.*
Gender			
Female	8,858	45.39	45.39
Male	10,656	54.61	100.00
Age group			
15 to 29	4,960	25.42	25.42
30 to 39	4,042	20.71	46.13
40 to 49	4,428	22.69	68.82
50 to 64	4,437	22.74	91.56
65 or more	1,647	8.44	100.00
Labour situation groups			
Employed	12,997	66.60	66.60
Housewife	1,977	10.13	76.73
Unemployed	635	3.25	79.99
Retired	2,116	10.84	90.83
Students	1,789	9.17	100.00
Educational level			
Pre-primary education	870	4.46	4.46
Primary education	5,223	26.77	31.22
Secondary education	6,910	35.41	66.63
Vocational training FP	1,655	8.48	75.12
University degree	4,856	24.88	100.00
Income level			
Low Income	6,764	34.66	34.66
Medium Income	7,410	37.97	72.64
High Income	5,340	27.36	100.00
Municipality size			
Less than 10000	6,081	31.16	31.16
10000 to 50000	4,590	23.52	54.68
50000 to 500000	7,445	38.15	92.84
500000 and more	1,398	7.16	100.00
Type of area			
Not metropolitan	13,470	69.03	69.03

Continued on next page

Table 1 – *Continued from previous page*

	Frequency	*Percentage*	*Cum. Perc.*
Metropolitan area	6,044	30.97	100.00
Purpose			
Pleasure	14,594	74.79	74.79
Business	2,706	13.87	88.65
Second residency	2,214	11.35	100.00
Distance			
50-99 Km.	9,244	47.37	47.37
100-249 Km.	6,887	35.29	82.66
250-499 Km.	2,650	13.58	96.24
500 Km. and more	733	3.76	100.00
Nights			
0	9,771	50.07	50.07
1 to 6	8,190	41.97	92.04
7 to 14	1,086	5.57	97.61
15 and more	467	2.39	100.00

In table 2, we conduct a descriptive cross tabulation analysis of mode choice by trip purpose (pleasure, business and second residence) with some of the variables included in the survey. Pleasure includes holiday trips, visits to relatives and leisure trips. Business comprises those trips made for professional reasons. Trips to a second residence are in a different category. Although different motives such as holidays or leisure might be motivating travel to a second residence, these are folded into this category. Table 2 shows the predominant role of private car usage, with a minimum modal share of almost 84% in the case of pleasure trips, 89% of business trips and a similar share in those heading to a second residence. The remaining two modes of transport were chosen differently, depending on the trip motivation. While trains were hardly used for second residence trips, they were used slightly more than buses for business trips. Bus shares for pleasure purposes were almost twice the share for business trips and slightly larger than the shares in case of visiting a second residence.

A descriptive analysis of sociodemographic variables provides a useful insight in mode choice. Male travelers appear to be more dependent than females on private car usage, as noted in previous studies. The largest difference in gender mode choice appears in the case of pleasure trips, where 87% of male travelers select cars while only 79% of females do. The effect of age on mode choice is a little more difficult to distinguish through descriptive analysis. In general, car usage increases until middle age and then decreases in the senior years. Buses and trains reach their maximum market shares in people aged over 65 for pleasure and business purposes, respectively, although the modal split is totally different. In trips motivated by pleasure, senior respondents rely heavily on buses, while in business trips the final choice is the train. As for income, car usage is the highest in medium income individuals, closely followed by high-income travelers. Descriptive data on bus usage conditional to income appear to support the idea of a bus as an inferior good; for all trip purposes, the market share decreases when income rises.

Another expected fact found in the data is that public transport was more intensively used in larger cities. Particularly in the case of rail, there was a strong positive correlation between city size and rail demand. This was likely caused by improved access to principal railway stations with connections to a wider range of destinations and more frequent services. The bus mode shares were not as closely related to city size as train demand. Travelers with origins in a metropolitan area also depended more on railway transport than those living in non-metropolitan zones. Private car usage did not differ much between metropolitan and non-metropolitan cities. Public buses were more important for non-metropolitan travelers. Trip characteristics such as distance and overnight stays also appear to have had an impact on primary transport mode choice. While car usage decreased with distance travelled, bus and train importance grew in longer distance trips. Additional nights spent travelling hindered car usage and favored the use of trains, particularly for pleasure purposes, with no clear pattern emerging for the other purposes.

Exploratory analysis using descriptive statistics is a helpful tool for identifying the stylized facts and characteristics of the data. This information provides important insights in the description of mode choice, but the estimation of a multivariate model allows us to assess the impacts of the different variables on probabilities of choosing a particular mode of transport.

Table 2: Selected mode of transport by purpose of the trip

	Pleasure			Business			Second residency		
	Car	Bus	Train	Car	Bus	Train	Car	Bus	Train
Gender									
Female	79.22	14.12	6.66	83.09	9.43	7.48	85.60	9.75	4.65
Male	87.36	8.48	4.17	91.01	3.97	5.02	91.72	5.67	2.60
Age group									
15 to 29	78.63	14.07	7.29	85.82	8.08	6.10	77.53	14.90	7.58
30 to 39	92.90	3.88	3.22	90.03	3.79	6.18	94.93	2.90	2.17
40 to 49	90.75	5.15	4.09	90.71	4.04	5.25	93.60	3.93	2.48
50 to 64	82.66	12.03	5.31	91.67	4.96	3.37	90.94	5.94	3.12
65 or more	61.29	31.04	7.67	61.29	9.68	29.03	79.35	17.00	3.64
Income									
Low Income	71.35	21.02	7.63	82.00	9.00	9.00	81.76	13.13	5.12
Medium Income	91.36	5.10	3.54	91.19	5.29	3.52	93.00	5.18	1.83
High Income	90.56	4.86	4.58	88.11	4.46	7.43	93.47	3.04	3.50
Municipality size									
Less than 10000	84.10	12.28	3.62	92.12	4.82	3.06	88.02	9.58	2.40
10000 to 50000	84.08	11.01	4.92	91.59	4.59	3.82	87.96	8.10	3.94
50000 to 500000	82.84	10.56	6.60	86.34	6.11	7.55	90.48	6.35	3.17
More than 500000	80.58	10.52	8.91	82.58	3.87	13.55	84.57	8.68	6.75
Metropolitan area or not									
Not metropolitan	83.21	12.11	4.68	90.11	5.64	4.25	88.52	8.44	3.05
Metropolitan area	83.84	9.14	7.02	86.87	4.11	9.02	88.65	6.85	4.50
Distance									

Continued on next page

Table 2 – *Continued from previous page*

50-99 Km.	87.39	9.03	3.58	91.01	5.18	3.81	89.73	7.60	2.66
100-249 Km.	83.53	11.51	4.96	91.35	3.73	4.93	89.09	7.43	3.48
250-499 Km.	74.75	14.92	10.33	76.35	8.45	15.20	85.55	8.37	6.08
500 Km. and more	68.40	20.03	11.56	75.93	11.11	12.96	75.93	11.11	12.96
Nights									
0	86.48	10.18	3.34	91.15	4.72	4.13	89.74	7.69	2.56
1 to 6	81.44	11.49	7.07	79.96	7.59	12.45	89.52	6.96	3.51
7 to 14	73.57	17.78	8.65	95.45	4.55	0.00	81.82	11.48	6.70
15 and more	76.73	12.36	10.91	85.71	0.00	14.29	86.49	10.27	3.24
No. of Observations	12,171	1,638	785	2,414	141	151	1,961	172	81
Total Percentage by purpose	83.40	11.22	5.38	89.21	5.21	5.58	88.57	7.77	3.66

4. Utility framework and multilevel analysis

The mainstream approach to study modal choice assumes that travelers make rational decisions by selecting the alternative that maximizes their utility. From the optics of a rational choice perspective, subject n would choose the alternative with the highest utility. In our exercise, travel mode m is chosen by individual n if the utility of this alternative is greater than the utility of any other transportation mode t:

$$U_n^{(m)} > U_n^{(t)} \text{ for all } m \neq t \tag{1}$$

The workhorse tool for travel behavior analysis relies on the random utility framework. In these models, the decision rule is deterministic but the utility function includes a random component. The deterministic component includes information on the transport mode attributes, socioeconomic characteristics of the traveler, land use factors and other variables as in Section 2. The error term is included because the analyst assumes that there is incomplete information on the selection process faced by the individual (Manski, 1977).

In a multilevel framework, the utility $U_{nj}^{(m)}$ of an alternative m for individual n nested in cluster j is assumed to consist of a deterministic part $V_{nj}^{(m)}$, and a random component called the error term $\epsilon_{nj}^{(m)}$ (Skrondal and Rabe-Hesketh, 2003; Grilli and Rampichini, 2007) as follows:

$$U_{nj}^{(m)} = V_{nj}^{(m)} + \epsilon_{nj}^{(m)} \tag{2}$$

where $m = 1, 2, ..., M$ denotes the response category (mode of transport), $j = 1, 2, ..., J$ denotes the cluster (province of origin) and $n = 1, 2, ..., n_j$ denotes the traveler of the j-th province of origin. The deterministic part $V_{nj}^{(m)}$ represents the fixed part of the utility and is linearly related to the linear predictors of the model.

$$V_{nj}^{(m)} = \alpha^{(m)} + \beta^{(m)} X_{nj} + \beta X_{nj}^{(m)} + \zeta_j^{(m)} \tag{3}$$

where $\alpha^{(m)}$ is a fixed alternative-specific intercept, X_{nj} is a set of explanatory covariates that vary over travelers and $\beta^{(m)}$ is the set of associated coefficients to be estimated. Alternative specific covariate $X_{nj}^{(m)}$ is the attribute that varies between response categories m and travelers n nested in j and has a coefficient β that does not vary over alternatives m. The single level multinomial logit model would be solely composed of these components. The multilevel version of the model, where travelers are nested in the province of origin, also includes random alternative-specific intercepts $\zeta_j^{(m)}$ to account for unobserved heterogeneity at the province level. This setting allows for relaxing the multinomial logit assumption of independence of the irrelevant alternatives (IIA), which might be inappropriate in some choice situations, as discussed by Hausman and McFadden (1984)[6]. In the case of the multilevel multinomial logit model, the error terms $\epsilon_{nj}^{(m)}$ of the utility functions have Gumbel distributions and are independent over transportations modes, travellers and provinces (Rabe-Hesketh and Skrondal, 2012).

Discrete choice models estimate the probability that an individual will select a mode of transport from a given set of alternatives, based on the attributes of the alternatives and on his preferences (Ben-Akiva and Bierlaire, 1999). The number of model equations equals the number of mode choice alternatives (three in our analysis), while the utility maximization rule specified in Equation 1 helps in computing the probability of choosing an alternative, e.g., Alternative 2:

$$\begin{aligned} Pr(Y_{nj} = 2 \mid X_{nj}, X_{nj}^{(m)}, \zeta_j^{(m)}) &= Pr(U_{nj}^{(2)} > U_{nj}^{(1)}, U_{nj}^{(2)} > U_{nj}^{(3)}) \\ &= Pr(U_{nj}^{(2)} - U_{nj}^{(1)} > 0, U_{nj}^{(2)} - U_{nj}^{(3)} > 0) \qquad (5) \\ &= Pr(\epsilon_{nj}^{(2)} - \epsilon_{nj}^{(1)} > V_{nj}^{(2)} - V_{nj}^{(1)}, \epsilon_{nj}^{(2)} - \epsilon_{nj}^{(3)} > V_{nj}^{(2)} - V_{nj}^{(3)}) \end{aligned}$$

[6]Rabe-Hesketh and Skrondal (2012) showed the distribution of the vector of random intercepts and derived the correlations between utility differences where ψ denoted the covariances between them

$$\begin{bmatrix} \zeta_j^{(2)} \\ \zeta_j^{(3)} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \psi_2 & \\ \psi_{32} & \psi_3 \end{bmatrix} \right) \qquad (4)$$

where

$$V_{nj}^{(2)} - V_{nj}^{(1)} = \alpha^{(2)} - \alpha^{(1)} + \zeta_j^{(2)} - \zeta_j^{(1)} + (\beta^{(2)} - \beta^{(1)})X_{nj} + \beta(X_{nj}^{(2)} - X_{nj}^{(1)}) \quad (6)$$

and

$$V_{nj}^{(2)} - V_{nj}^{(3)} = \alpha^{(2)} - \alpha^{(3)} + \zeta_j^{(2)} - \zeta_j^{(3)} + (\beta^{(2)} - \beta^{(3)})X_{nj} + \beta(X_{nj}^{(2)} - X_{nj}^{(3)}) \quad (7)$$

Analogously similar expressions can be obtained for alternative *1* and *3*. To obtain an identifiable model, we must select an alternative as a base category whose fixed and random parameters are set to 0. Utilities and their differences are unobservable, but we do have information on the mode choice Y_{nj}^*. It can be shown that the resulting choice probability, when the base category is alternative *1*, is given by the multinomial logit model in a Generalized Linear Model formulation McFadden (1974).

$$Pr(Y_{nj} = m \mid X_{nj}, X_{nj}^{(m)}, \zeta_j^{(m)}) = \frac{exp[V_{nj}^{(m)}]}{1 + \sum_{l=2}^{M} exp[V_{nj}^{(l)}]} \quad (8)$$

where the denominator is the sum of the numerators of the probabilities of the three alternatives, guaranteeing that the probabilities sum to one. The model estimates presented in Section 5 were obtained through the Stata program gllamm [7].

In this context, a multilevel model has a number of advantages over a traditional single-level multinomial model. Bhat (2000) proposed a multilevel analysis as a useful tool for incorporating a spatial context where individuals must make decisions while allowing simultaneous consideration of the spatial heterogeneity between higher level units (provinces). This spatial heterogeneity can be explained through different reasons. First, the existence of spatial autocorrelation is typical in any regional analysis Tobler (1970). In this case, individuals belonging to the same higher level units exhibited similar behavior (modal choice), but there is no complete individual information to explain this pattern. Second, it is possible to denote important differences in terms of structural behavior and relationships between higher level units. Both phenomena must be included in the modeling issue to avoid estimation and testing errors and parameter instability. Bhat (2000) and Jones and

[7] See Rabe-Hesketh et al. (2005) and Skrondal and Rabe-Hesketh (2003).

Duncan (1996) claimed to have adequately differentiated the heterogeneity among higher level units and individual heterogeneity, which may be related to socio-psychological factors. All of these issues were satisfactorily addressed through the multilevel framework.

5. Model Results

In this section, we present the empirical results obtained from applying the multilevel multinomial model to the mode choice sample [8]. We also estimate a standard multinomial model as a benchmark. Table 3 presents the results of these two models. The baseline mode of transport is the private car, and the results show odds ratios. Odds ratios, also referred to as relative risk ratios, indicate the ratio of the probability of choosing one outcome category (train or bus) over the probability of choosing the baseline category (private car) [9]. The results were derived by Maximum Likelihood Estimation using gllamm program in Stata.

The effects of all explanatory variables are very similar in the standard multinomial model and in its multilevel counterpart, although significant testing of the multilevel model based on the likelihood ratio test, Bayesian information criterion (BIC) and Akaike information criterion (AIC) shows that the model has significant spatial heterogeneity; therefore, this model is preferred to the conventional regression model. As expected, socio-demographic variables are important explaining the mode choice outcomes. Being male, *ceteris paribus*, decreases the odds of selecting a bus over a car by 40%, and the odds of taking the train compared with driving are reduced by an estimated 36%. According to this result, men are more car dependent than women in long-distance trips. Limtanakool et al. (2006) appointed different factors in explaining these gender differences found in the literature such as inequality in monetary rewards and different household task allocations. Categorical variables referred to age indicated that elderly trip makers were more likely to choose a bus over a car. The relative risk ratio associated with the age category 30-39 years implies that their odds of choosing a bus over

[8] Similar results were obtained using a constrained sample were trips shorter than 100 Km. were disregarded.

[9] Odds ratios are the ratio of probabilities of events and take the form: $\frac{Pr(Y_{nj}=m|X_{nj},X_{nj}^{(m)},\zeta_j^{(m)})}{Pr(Y_{nj}=1|X_{nj},X_{nj}^{(m)},\zeta_j^{(m)})}$.

a car relative to the base age of 15-29 years is multiplied by a factor of 0.332 (considerably less than 1), which sharply reduces their chances of selecting a bus over a car. This reduction in the probability lessens with older travelers, as shown by the estimated odds ratios for the middle aged, which are still below 1 but larger than 0.332. In the case of travelers between 40-49 years, the reduction would be 60%, while travelers aged between 50 to 64 years would reduce the odds to 30%. When switching from the base age category to the one including 65 years or older, travelers would increase their chances of selecting a bus over a car by 64%. This result concurs with the findings in Georggi and Pendyala (2001), as shown in Section 2.

A similar pattern can be observed in the case of studying train and car mode choice relationships where the factor that multiplies the odds of choosing a train over a car is below 1 but increases in the subsequent categories of older people. It is important to note that the odds ratio associated with the category of 65 years or more reaches a value that is not significantly different from 1, meaning that the probability of train selection is similar in the youngest and oldest participants. A likely explanation of these results relates to the impact of age on public transportation demand; students and the elderly have reduced access to car ownership and enjoy similar discounts in travel fares in both train and bus services, thus favoring the use of these modes.

In addition, income variables also show expected effects. Switching from the low-income group to the medium- or high-income groups decreases the odds of the traveler choosing a bus over a car. A higher disposable income also discourages the use of trains, though the decrease of the odds when switching from low to high income is smaller than in the case of buses. The train mode of transport also includes high-speed trains that can reduce travel times for some destinations and become more attractive for higher income travelers.

Land use variables included in the models consist of the size of the city of origin and an indicator of whether this location is a metropolitan area. Travelers beginning their inter-city trips in non-metropolitan areas are more likely to select a bus over a car than those users in metropolitan locations. The model results also show that, when all else is equal, the impact of the size of the city is insignificant in determining the odds of bus vs. car usage. The influence of land use indicators is different when confronting train and private car usage. In the multilevel version of the model, the metropolitan characteristic of a city of origin has no effect on the mode choice decision.

However, city size has a significant impact. The odds of choosing a train over a car when departing from a large city are almost twice the odds of a city of origin of less than 10,000 habitants. Larger cities are expected to have greater accessibility to high-speed rail and other important rail stations. Travelers from these cities enjoy a better supply of public transport services, including more frequent departures and a larger variety of destinations.

Table 3: Multinomial logit model estimates

	One level		Two levels	
	Odds Ratio	z- Statistic	Odds Ratio	z-Statistic
Fixed parameters				
Cost	0.978***	(-14.66)	0.978***	(-14.19)
Bus				
Gender: Male	0.596***	(-9.72)	0.594***	(-9.70)
Age: 30-39 years	0.333***	(-11.40)	0.332***	(-11.37)
Age: 40-49 years	0.405***	(-10.58)	0.406***	(-10.48)
Age: 50-64 years	0.700***	(-5.26)	0.705***	(-5.10)
Age: 65 years or more	1.644***	(6.53)	1.642***	(6.41)
Income: Medium	0.358***	(-14.67)	0.363***	(-14.36)
Income: High	0.316***	(-14.26)	0.316***	(-14.13)
Metropolitan Area: Yes	0.711***	(-4.94)	0.750***	(-3.56)
City size: 10001-50000	0.930	(-1.06)	0.963	(-0.53)
City size: 50001-500000	0.957	(-0.68)	0.940	(-0.91)
City size: 500000 or more	1.262	(1.86)	1.253	(1.63)
Purpose: Business	0.994	(-0.06)	1.010	(0.10)
Purpose: Secondary residence	0.674***	(-4.36)	0.666***	(-4.44)
Overnight stays: 1 to 6	0.987	(-0.23)	0.982	(-0.31)
Overnight stays: 7 to 14	0.948	(-0.51)	0.940	(-0.58)
Overnight stays: 15 and more	0.557***	(-3.51)	0.554***	(-3.51)
Intercept	0.296***	(-16.95)	0.287***	(-15.12)
Train				
Gender: Male	0.643***	(-6.36)	0.640***	(-6.28)
Age: 30-39 years	0.530***	(-6.04)	0.546***	(-5.64)
Age: 40-49 years	0.608***	(-5.08)	0.658***	(-4.17)
Age: 50-64 years	0.648***	(-4.71)	0.663***	(-4.33)
Age: 65 years or more	1.053	(0.44)	1.060	(0.48)
Income: Medium	0.465***	(-8.35)	0.441***	(-8.74)
Income: High	0.577***	(-6.03)	0.557***	(-6.23)

Continued on next page

Table 3 – *Continued from previous page*

Metropolitan Area: Yes	1.155*	(1.79)	1.040	(0.38)
City size: 10001-50000	1.251*	(2.20)	1.193	(1.63)
City size: 50001-500000	1.620***	(5.22)	1.683***	(5.38)
City size: 500000 or more	2.216***	(5.60)	2.059***	(4.39)
Purpose: Business	1.897***	(6.22)	1.979***	(6.48)
Purpose: Secondary residence	0.518***	(-5.33)	0.491***	(-5.66)
Overnight stays: 1 to 6	1.717***	(6.93)	1.820***	(7.51)
Overnight stays: 7 to 14	1.571***	(3.31)	1.641***	(3.54)
Overnight stays: 15 and more	1.394*	(1.71)	1.459*	(1.91)
Intercept	0.052***	(-27.35)	0.040***	(-23.17)
Random intercepts				
			Est.	S.E.
ψ_{Bus}			0.091	(0.026)
ψ_{Train}			0.477	(0.082)
$\psi_{BusTrain}$			-0.126	(0.035)
No. Of observations	19,514		19,514	
Loglikelihood at convergence	-9,127.70		-8,933.33	
LR test				
Chi^2			388.7	
p value			0.000	
$ICC_{bus}(\%)$			2.755	
$ICC_{train}(\%)$			14.482	
BIC	18,639.6		18,283.8	
AIC	18,325.4		17,942.7	

** Significant at 10%. ** Significant at 5%. *** Significant at 1%*

As for trip attributes, trip purpose and duration have a strong relationship with mode choice. If the purpose of the trip is visiting a second residence, a private car is preferred to both bus and train. For business travel, the odds of choosing a train over a car are almost twice that of leisure purposes. The effect of trip duration is quite different in the case of trains and buses. Overnight stays have no effect when the duration of the trip is below 14 nights, but in trips longer than 15 nights, the probability of choosing a bus over a car is considerably reduced. In the case of trains, increased nights spent away from home increase the odds of choosing a train over a car, although the effect seems to vanish in longer duration stays.

The only transport mode specific covariate is the cost variable; it has a unique estimated coefficient that does not vary over the alternatives. The odds ratio value for this variable is 0.978, which is below one and corresponds with a -0.02 estimated coefficient. A negative coefficient means that if the cost increases for one category, then the demand for that category decreases and increases for the other categories. The impact of variation in costs can be studied in more detail by computing the impacts on mode market shares caused by potential scenarios.

In this sense, we examine the results of four policy measures using the estimated model. The first policy is a 50 % increase in car usage cost that might be related to the evolution of gas prices or caused by the introduction of a congestion pricing measure. The second and the third are, respectively, a 25% decrease in public bus fares and railways costs. Finally, a fourth scenario includes a combination of the other three scenarios: a 50% increase in private car use costs jointly with 25% decreases in public transport fares. The effect of each policy measure is assessed by modifying the costs magnitudes to reflect a change, computing predicted outcome probabilities of mode choice using the estimated model, calculating predicted aggregate market shares of each mode and finally obtaining a percentage change from the baseline estimates.

Table 4 displays the results of the different scenarios. As expected, the model shows a decrease in private car market share in Scenario 1 when the cost associated with using the car rises. The bus market shares are more sensitive than railway shares, indicating a higher degree of substitutability between car and public bus. Scenario 2 and 3 show the effect of lower fares in public transport modes. In both situations, the market share of private cars decreases in a percentage close to 0.30%. The impact of reducing travel prices is larger in the case of railway than in the case of public buses, which seem to be have a very low sensitivity to changes in their own fares compared to changes in car usage costs. Finally, the last scenario shows the effect of a simultaneous change in the costs of the three modes, which would induce a larger change in the modal split favouring the bus demand in a larger extent.

Along with studying the effects of different independent variables on mode choice, multilevel analysis also allows researchers to analyze the effect of spatial heterogeneity. The estimated varying intercept can consider the correlation among people living in the same province. The correlation between

Table 4: Results of policy evaluation

Scenario	Policy	Change in market shares Car	Bus	Train
1	50% Increase car usage cost	-2.39%	+14.12%	+9.79%
2	25% decrease bus fare	-0.35%	+2.92%	-0.18%
3	25% decrease railway fare	-0.29%	-0.08%	+4.61%
4	Scenarios 1-3 simultaneously	-2.88%	+17.88%	+10.22%

individuals in the same area can be computed in terms of an intraclass correlation coefficient (ICC). This coefficient determines the proportion of variability that is accounted for by differences among areas [10]. The calculated ICC for buses (2.7%) indicates that there is a small correlation of travelers that choose a bus over a car. The associated ICC for trains, however, is significantly larger (14.5%). Relatively small ICCs indicate that individual factors included in the model can explain most variance within a group. In the case of trains, differences between provinces seem to be more important in explaining the variation in train choice.

Spatial heterogeneity can be caused by very different factors such as non-homogeneous preferences in modes of transportation. Differences in preferences can be determined through historical reasons or regional policies that are not easily observable. Other unmeasured characteristics such as the supply level of public services or the accessibility to public transport stations can also play an important role in these differences. The variance of estimated random intercepts collects the odds variations caused by departing from different provinces. We obtained 94 realizations of these random intercepts, 47 for each mode of public transport (bus and train). The plotted maps differentiate between provinces with odds ratios below 1 and provinces with odd ratios above 1, indicating the areas that decreased and increased the probability of choosing a public mode over a car, respectively.

Figure 1 and Figure 2 display the exponential transformation of random intercepts associated with bus and train modes for each province, respec-

[10]See Grilli and Rampichini (2007) Rabe-Hesketh and Skrondal (2012) and Mercado and Páez (2009).

tively. In the multinomial logit model, this transformation is interpreted as the odd ratio of choosing a public mode over a private car relative to the departure from a particular province. The plotted maps differentiate between provinces with odds ratios below 1 and provinces with odd ratios above 1, indicating the areas that decreased and increased the probability of choosing a public mode over a car.

The map represented in Figure 1 suggests a dominance of buses over cars in provinces in the north and in the west of the country, a route traditionally dominated by the largest bus companies in Spain. Figure 2 shows that the odds of choosing a train over a car are increased in trips departing from provinces near the east coast of Spain. There also seems to be a strong effect in Madrid, Toledo, Ciudad Real, Córdoba and Sevilla; these five provinces enjoy a high-speed rail connection. However, provinces located along the southeast coast reduce the relative risk of choosing a train over a car. The radial structure of the railway network is easily recognizable, increasing the probability of choosing a train over a car in provinces covered by primary railway lines. The same map shows that certain areas in the north and south of the country not been favored by good access to train services reduce the use of trains.

In comparing both figures, we can see that odds ratios associated with buses on the northeast coast and some landlocked provinces in the south are below 1. The same provinces appear to increase their probabilities of choosing a train over a car. This might indicate that in certain areas, travelers tend to choose either buses or trains instead of cars. An insufficient provision of railway service might be appealing to bus companies, attracting them to operate in those markets where railway is less competitive. However, regions in the northwest (Galicia) of the country and in some landlocked provinces in the north middle east (Navarra, Alava, Guipuzcoa, Huesca, Teruel) and on the south coast (Mlaga, Almera, Murcia) show decreasing probabilities in both public transportation modes in relationship to cars. The dominance of cars in these areas might be related to different factors. The geographical remoteness of the provinces in the northwest and south of the country, along with a low supply of public transport in these provinces, might increase the need of private vehicle usage. The lack of a car ownership control variable might be a significant aspect affecting the probabilities of travelers from provinces in the Basque Country, one of the richest regions in the country, who are more likely to own a private vehicle than those in other areas.

Figure 1: Bus mode provincial random intercepts

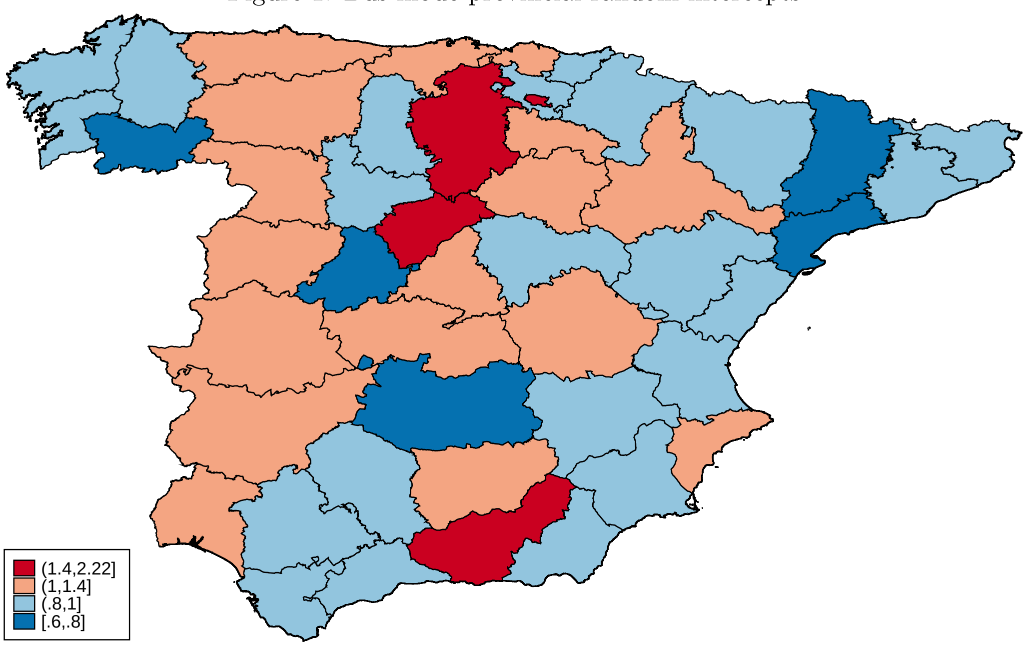
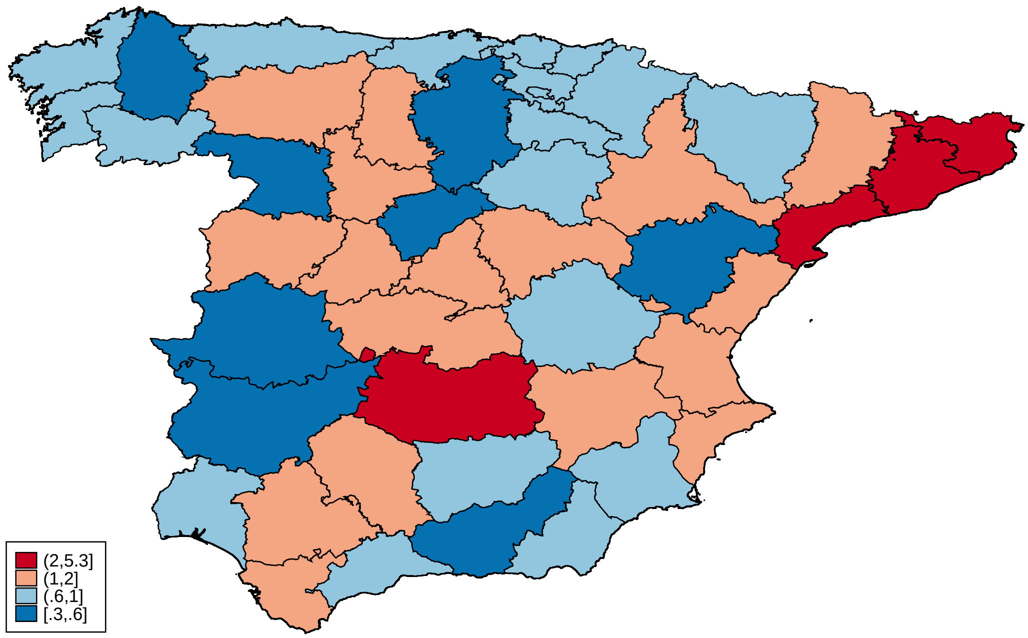


Figure 2: Train mode provincial random intercepts



6. Conclusions

A better understanding of modal choice is important to take adequate policy measures to guide mobility behavior towards more sustainable modes of transport. The aim of this paper has been to study the determinants of mode choice in long-distance trips in Spain. In particular, we focused our attention on three ground modes of transport: private car, bus and train. For this purpose, we applied discrete model choice techniques and tested the impact of several sociodemographic, land use and trip purpose independent variables.

To conduct the analysis, we employed data from a 2006 Spanish mobility survey that attempted to evaluate the travel behavior of Spanish residents. It is worth emphasizing two characteristics of the database that might be common to other long-distance databases in different parts of the world: missing variables and hierarchical structure. A literature review of the few papers dealing with inter-city travel helped in relating the most common explanatory variables in these previous studies. Comparing this information with the available indicators in the survey, we found some missing variables, including household income, car ownership and spatial indicators. Although all these variables may be important, our greatest concern was knowing the exact locations of origin and destination. Ignoring these geographical points prevented any attempt to include accessibility measures to transport infrastructure. As a solution, we proposed exploiting the hierarchical structure of the data. In this survey, trip observations might be nested in the provinces of origin for the trip. This characteristic feature of the data allowed us to estimate a multinomial multilevel model with random intercepts. Multilevel analysis permitted reducing the omitted variable bias and improving the estimation of standard errors through clustering of the observations.

As expected, the analysis confirmed some empirical evidence found in previous papers and added new insights in the determinants of long-distance travel mode choice. Socio-demographic variables such as gender, age and income play a significant role on mode choice. While women are less car dependent that men, young and elderly travelers rely more on buses and trains. Higher income levels also reduce the odds of using public transport over private cars. Characteristics of the origin of trip locations such as city size and belonging to a metropolitan area were also found to be significant. Differences on mode choice also arose depending on the motive of the trip. Trips for business purposes are more likely to be made by train instead of car, while travel for leisure preferred car usage. In the case of comparing trips to a second residence with leisure trips, the former are more car dependent. Moreover, we tested the inclusion of a variable capturing the duration of the trip, overnight stays, that has not been used widely in previous literature on inter-city travel. A longer duration of inter-city trips appears to favor the use of railways, reducing the demand for buses. We also discussed some interesting significant relationships with the cost variable.

Random intercepts included in the model captured spatial differences on the probabilities of choosing transport modes, resulting in a more flexible specification. Individuals beginning their trip in the same province are affected by certain factors that might not be properly accounted for through explicative variables and may have important policy implications. Our results showed that, when controlled for individual characteristics (level 1), there is evidence of spatial differences (level 2). These differences can arise from different factors. The spatial distribution of preferences is not uniformly distributed, thus differently affecting the final decision depending on the departure province. The information plotted on the maps shows certain patterns that were explained in terms of the impact of important factors related to geography, institutions and transport policy.

This paper has clearly defined the dominance of cars over public transport modes in Spain. Regional heterogeneity in the design of transportation networks has yielded different spatial access to public transport modes. This factor, combined with inexpensive access to road networks by private vehicles, has promoted the predominant role of cars in mode choice decisions, especially in certain areas. While departing from certain provinces seem to increase the chances of using public transport modes (either bus or train), there are other areas where, according to our results, buses and trains are less likely to be used than private cars. Sustainability is an ineluctable challenge of any transportation system, and transport policy makers should consider if the current picture of the passenger transportation market in Spain, particularly in these areas, is on the path of reaching such a goal.

Our analysis provides interesting results and new insights that add to the existing knowledge on inter-city mode choice, and open new avenues for research on factors related to transport policy. Future research should further enrich this type of analysis by using more refined geolocated data on public transportation accessibility.

Axhausen, K., Madre, J.L., Toint, P. (Eds.), 2003. Capturing Long-Distance Travel. Research Studies Press Ltd., Hertfordshire, England, United Kingdom.

Bel, G., 1997. Changes in travel time across modes and its impact on the demand for inter-urban rail travel. Transportation Research Part E: Logistics and Transportation Review 33, 43–52.

Ben-Akiva, M., Bierlaire, M., 1999. Discrete choice methods and their applications to short term travel decisions, in: Handbook of transportation science. 1985.

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

Bhat, C.R., 1998. Accommodating variations in responsiveness to level-of-service measures in travel mode choice modeling. Transportation Research Part A: Policy and Practice 32, 495–507.

Bhat, C.R., 2000. A multi-level cross-classified model for discrete response variables. Transportation Research Part B: Methodological 34, 567–582.

Bricka, S., 2001. Variations in Long-Distance Travel. Transportation research circular , 197–206.

Buehler, R., 2011. Determinants of transport mode choice: a comparison of Germany and the USA. Journal of Transport Geography 19, 644–657.

De Witte, A., Hollevoet, J., Dobruszkes, F., Hubert, M., Macharis, C., 2013. Linking modal choice to motility: A comprehensive review. Transportation Research Part A: Policy and Practice 49, 329–341.

Garmendia, M., Ureña, J., Coronado, J., 2011. Long-distance trips in a sparsely populated region: The impact of high-speed infrastructures. Journal of Transport Geography 19, 537–551.

Georggi, N.L., Pendyala, R.M., 2001. Analysis of long-distance travel behavior of the elderly and low income. Transportation Research Board E-C026, 121–150.

Grilli, L., Rampichini, C., 2007. A multilevel multinomial logit model for the analysis of graduates skills. Statistical Methods and Applications 16, 381–393.

Hausman, J., McFadden, D., 1984. Specification Tests for the Multinomial Logit Model. Econometrica 52, 1219.

Hong, J., Shen, Q., Zhang, L., 2013. How do built-environment factors affect travel behavior? A spatial analysis at different geographic scales. Transportation .

Jones, K., Duncan, C., 1996. People and places: the multilevel model general framework for the quantitative analysis of geographical data, in: Longley, P., Batty, M. (Eds.), Spatial Analysis: Modelling in a GIS Environment. GeoInformation International, Cambridge. chapter 5, pp. 79–104.

Koppelman, F.S., Sethi, V., 2005. Incorporating variance and covariance heterogeneity in the Generalized Nested Logit model: an application to modeling long distance travel choice behavior. Transportation Research Part B: Methodological 39, 825–853.

Lapparent, M.D., Axhausen, K., Frei, A., 2013. Long distance mode choice and distributions of values of travel time savings in three European countries. European Transport \ Trasporti Europei , 1–7.

Limtanakool, N., Dijst, M., Schwanen, T., 2006. 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 Board 1693, 71–78.

Mallett, W., 2001. Long-distance travel by low-income households. Transportation Research Board E-C026, 169–177.

Manski, C.F., 1977. The Structure of Random Utility Models. Theory and Decision 8, 229–254.

Martín, J.C., Nombela, G., 2007. Microeconomic impacts of investments in high speed trains in Spain. The Annals of Regional Science 41, 715–733.

McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior, in: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, pp. 105–142.

Mercado, R., Páez, A., 2009. Determinants of distance traveled with a focus on the elderly: a multilevel analysis in the Hamilton CMA, Canada. Journal of Transport Geography 17, 65–76.

Moeckel, R., Fussell, R., Donnelly, R., 2013. Mode Choice Modeling for Long-Distance Travel. Transportation Research Board , 1–16.

Páez, A., 2006. Exploring contextual variations in land use and transport analysis using a probit model with geographical weights. Journal of Transport Geography 14, 167–176.

Rabe-Hesketh, S., Skrondal, A., 2012. Multilevel and Longitudinal Modeling Using Stata. Press, Stata, Lakeway Drive, TX. 3 edition edition.

Rabe-Hesketh, S., Skrondal, A., Pickles, A., 2005. Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. Journal of Econometrics 128, 301–323.

Schafer, A., 1998. The global demand for motorized mobility. Transportation Research Part A: Policy and Practice 32, 455–477.

Skrondal, A., Rabe-Hesketh, S., 2003. Multilevel logistic regression for polytomous data and rankings. Psychometrika 68, 267–287.

Tobler, W.R., 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography 46, 234–240.

Whalen, K.E., Páez, A., Carrasco, J.a., 2013. Mode choice of university students commuting to schooland the role of active travel. Journal of Transport Geography 31, 132–142.