

# Tourists' Preferences for Hotel Booking

David Boto-García\*  
[botodavid@uniovi.es](mailto:botodavid@uniovi.es)

Marta Escalonilla  
[gonzalezemarta@uniovi.es](mailto:gonzalezemarta@uniovi.es)

Emma Zapico  
[zapicoemma@uniovi.es](mailto:zapicoemma@uniovi.es)

José Baños Pino  
[jbamos@uniovi.es](mailto:jbamos@uniovi.es)

\*Corresponding author

Department of Economics, University of Oviedo

Faculty of Economics and Business, University of Oviedo (Spain). Avenida del Cristo s/n 33006.

---

## Abstract:

This paper analyzes tourists' preferences for hotel booking mode using a sample of 17,921 tourists visiting a Northern Spanish region during 2005-2016. Four different booking modes are considered: telephone, the internet, travel agencies and other non-market-based intermediaries. We estimate a Finite Mixture Multinomial Logit Model that allows us to define three classes of tourists. Our results show that leisure tourists coming from distant locations and lodged at luxury hotels have a higher likelihood of online hotel booking in class 1. Travel agencies are preferred by offseason tourists with longer stays, while those travelling by public transit modes and staying at luxury hotels opt for non-market-based intermediaries in class 2. In class 3, first time tourists choose the internet, telephone is more prevalent among those staying at economy hotels and travel agencies are preferred among those travelling in the offseason and by public transit modes.

**Keywords:** *hotel booking; travel agencies; online booking; offline channels; Finite Mixture Modelling*

**JEL codes:** C25, Z30

## Acknowledgements

We thank the *Tourist Information System of Asturias* for the provision of the data.

David Boto-García acknowledges financial support from the Spanish Ministry of Education, Culture and Sport (FPU 16/00031).

Emma Zapico acknowledges financial support from the Regional Government of Asturias (PA-18-PF-BP17-069).

Marta Escalonilla acknowledges financial support from the University of Oviedo (PAPI-19-PF-02).

José F. Baños Pino also acknowledges financial support from the Project MINECO-18-ECO2017-86402-C2-1-R (Spanish Ministry of Economy and Competitiveness) and the Project PAPI-18-GR-2011-0026 (University of Oviedo).

## 1. INTRODUCTION

The emergence of online platforms, smartphones and apps have dramatically changed the way tourists make their accommodation reservations, since individuals today can directly interact with hospitality suppliers (Buhalis and Law 2008). Nowadays, more and more tourists book hotel rooms on the internet, which has become one of the main distribution channels. Notwithstanding this, a non-negligible share of tourists continues to make their reservations through offline channels (Murphy et al. 2016; Stangl et al. 2016).

Understanding tourists' preferences for the different booking modes has important implications for hotel managers. Cooperation with travel agencies or other intermediaries requires hotels to pay high commission fees for intermediation, which in some cases translates into lower revenues (Gazzoli et al., 2008; Thakran and Verma, 2013; Lei et al., 2019). Therefore, as highlighted by Law et al. (2015), the identification of the factors that make tourists more prone to book hotel rooms directly on the internet, through travel agencies or by telephone constitutes an issue of great relevance for hotel management.

Scholars have studied the linkages between accommodation booking mode choice and tourists sociodemographic and trip-related characteristics (Masiero and Law, 2016; Coenders et al., 2016). However, the tourism literature has shown that the underlying utility for each booking channel also depends on heuristics, attitudes and psychological traits (Fong et al., 2017), which are in most cases unobservable. For instance, there is evidence that the willingness to technology adoption and use depends on tourists' perceptions about the risk it entails (Luo et al., 2010; Park and Tussyadiah, 2017). This relates to the existence of unobservable heterogeneity in how tourists, conditional on their observable characteristics, perceive the usefulness of each reservation mode.

The main aim of this article is to explain the drivers of hotel booking mode choice. Specifically, we study the linkages between several trip-related characteristics like visiting the destination for the first time, travelling in the high season, trip purpose, hotel quality and distance to origin on the booking mode selected. Based on regression analysis, we assess how preferences for booking modes are associated with these trip characteristics. Therefore, the paper contributes to the literature on tourism distribution channels by shedding light on tourists' preferences for online and offline hotel booking modes. A second objective of the paper is the identification of different classes of tourists with different preferences. Since heterogeneity in preferences can be due to unobserved latent attitudes, we estimate a Finite Mixture Multinomial Logit Model (hereafter FMMNL) that allows us to estimate i) the effect of a set of sociodemographic characteristics on group (class) membership, and ii) the role of trip-related characteristics on the booking mode choice within each class.

The use of latent class analysis for distinguishing types of tourists is not new in tourism research (Alegre et al., 2011; Chen et al., 2019). However, to the best of our knowledge, this is the first empirical study that uses this methodology to disentangle heterogeneity in preferences for hotel booking channels. In this sense, our study mimics in some way the one by Masiero and Law (2016), who also examine customers' selection of sales

channels for booking hotel rooms. However, we depart from them in that rather than assuming preferences to be homogeneous across the sample, we allow different groups of tourists to have different tastes. By estimating the corresponding marginal effects for each class, we provide an illustration of how the effect of these variables on the booking mode choice varies by tourist segment.

The remainder of the paper is structured as follows. Section 2 reviews the related literature. Section 3 discusses the econometric modelling. Section 4 describes the dataset and the variables used. Section 5 presents the results and provides a discussion of the main findings. Section 6 concludes with some implications, limitations and lines for future research.

## 2. LITERATURE REVIEW

In this section, we provide an overview of the existing evidence on tourists' accommodation booking mode preferences. We first discuss the theoretical rationale and the empirical findings about how preferences for the booking channel relate to personal and trip-related characteristics. We then outline some recent evidence about the role of heuristics and subjective norms on booking mode patterns.

### *2.1. Booking mode patterns and personal characteristics*

Recent research shows that electronic booking modes are becoming increasingly popular for hotels, peer-to-peer accommodations and rural houses (Gössling and Lane, 2015). As a result, the traditional travel agent model has experienced a notable decline (Castillo-Manzano and López-Valpuesta, 2010). Nevertheless, a non-negligible share of tourists still prefers to make their trip reservations by traditional TAs or by telephone (Law et al., 2004; Pearce and Schott, 2011; Stangl et al., 2016; Murphy et al., 2016), mainly because of risk-related concerns (Golmohammadi et al., 2012). The fear that electronic providers collect and misuse personal information is a barrier that deters online booking adoption (Park and Tussyadiah, 2017; Talwar et al., 2020a), especially among elderly people (Talwar et al., 2020b). Given this heterogeneity in behaviour, a large body of literature has been concerned about studying consumers' booking preferences.

The theoretical background for booking mode preferences can be found in the Theory of Reasoned Action (TRA) (Fishbein and Ajzen, 1975) and the Technology Acceptance Model (TAM) developed by Davis (1989) and Davis et al. (1989). According to this framework, attitudes, subjective norms and intentions are antecedents of behaviour. In our context, booking online mainly depends on the perceived ease of use and usefulness compared to other options (Casaló et al., 2010; Amaro et al., 2015). Other relevant factors are perceived behavioural control (Park and Huang, 2017) and electronic trust (Wang et al., 2014; Li et al., 2017).

For explaining consumer behaviour, the representative agent assumption often fails since consumers are heterogeneous and therefore make different choices based on their characteristics (Pollak and Wales, 1981). Because of this, a great deal of attention has

been paid to how sociodemographic characteristics explain accommodation booking mode patterns. However, the empirical evidence on this is mixed. [Bonn et al. \(1998\)](#) show that the use of the internet for travel planning is more prevalent among high-educated and high-income tourists. Similar results are reported by [Wu et al. \(2013\)](#), who find that hotel website browsers tend to be young, well-educated, and particularly well-travelled. Nevertheless, [Law \(2009\)](#) shows that there are no differences in the propensity to book the hotel online based on gender, education and household income. Moreover, elderly people are found to be more likely to book online.

[Del Chiappa \(2013\)](#) explores Italian tourists' preferences for TAs versus the internet as booking modes. He finds that frequent buyers are more prone towards internet-based hotel reservations than occasional consumers. This suggests that a positive attitude towards technology could be developed through consumption, since the entry barriers dilute over time. Additionally, the frequency of online purchasing is positively related with age, education and income. Similarly, [Del Chiappa and Zara \(2015\)](#) document that young, employed and international tourists rely more on online intermediaries, whereas offline intermediaries are preferred by older and high-educated tourists who travel during the high season.

In the light of this evidence, it is unclear how the choice of online versus offline channels relates to personal characteristics. As we discuss later, this could stem from neglected unobserved preference heterogeneity.

## *2.2. Booking mode and trip characteristics*

The usefulness of online booking versus traditional offline modes might be context dependent so that the value of each alternative changes depending on trip circumstances ([Tversky and Simonson, 1993](#)). Accordingly, the choice of the accommodation booking channel is expected to be associated with trip-related characteristics such as travel purpose, party size or travel experience. For a sample of Spanish tourists, [Coenders et al. \(2016\)](#) find that the internet is more frequently adopted by tourists staying at low category hotels, trips planned long in advance, summer season trips and those involving friends and relatives. Conversely, the internet is less preferred by those with longer stays and who come from distant locations. [Masiero and Law \(2016\)](#) study customers' selection of hotel booking modes, considering both direct (hotel website) and indirect (online travel agencies, Destination Management Organization website and call centres) channels. They find that length of stay is positively associated with bookings made via hotel websites, while party size negatively influences reservations made by hotel and Destination Management Organization websites.

TAs tend to be preferred for complex and planned trips ([Del Chiappa et al., 2015](#)). Additionally, traditional TAs are highly valued by those who appreciate personal services and who dislike engaging in large information search ([Bargeman and Van der Poel, 2006](#)). Social interactions, staff expertise and the possibility of saving time have been argued as the main reasons why some tourists opt for traditional TAs ([Buhalis and Licata, 2002](#); [Law and Lau, 2004](#)). This might explain why senior travellers mostly book through TAs ([Gronflaten, 2011](#)). Personal advice and friendliness are other important factors ([Wolfe et al., 2005](#)). In this vein, [Toh et al. \(2011\)](#) highlight that tourists search for

information online but then prefer to make phone calls to have personal contact and further information on hotel features.

For low-involvement and offhanded trips, tourists prefer to book their accommodation through the internet (Chu, 2001). The degree of involvement is strongly related with the distance between origin and destination, since long-haul travellers devote more effort to planning their trips. In this regard, Del Chiappa (2013) documents that long-haul tourists are more likely to book the accommodation through TAs, while Del Chiappa et al. (2015) show that the internet is the preferred mode for short-haul tourists. By contrast, Pearce and Schott (2011) report that domestic tourists book the accommodation by phone calls while outbound travellers prefer to do it online.

Accommodation booking usually takes place together with the booking of other services like transportation. In this way, travelling by public modes of transport makes transport and accommodation purchase to be strongly linked. Topolsek et al. (2014) show that TAs tend to cooperate with bus operators and airlines. Other factors that have been examined are travel experience and hotel category. Law (2009) and Jensen (2012) find that online travel shopping is positively related with travel experience and frequency. Qi et al. (2013) show that most five-star hotel guests make their reservations online.

Overall, the evidence on the role of trip characteristics points to electronic booking being associated with travel frequency, short stays, offhanded trips, travelling by public transit and opting for high quality accommodation. Notwithstanding this, the effect of trip features might not affect all consumers in the same direction.

### *2.3. Booking mode, heuristics and subjective perceptions*

Consistent with the TAM, several scholars have shown that online booking intentions strongly depend on website's aesthetic appeal, ease of use, usefulness and perceived risk (Jeong et al., 2003; Phelan et al., 2011; Bhatiasevi and Yoopetch, 2015; Amaro et al., 2015; Lu et al., 2015; Wang et al., 2015; Sun et al., 2020). Other important dimensions are information quality (Wang and Wang, 2010), haptic cues (Lv et al., 2020), choice set size and the existence of information filtering mechanisms (Guillet et al., 2020), the presence of photographs and their descriptions (Bufquin et al., 2020) and online reviews (Chan et al., 2017). A detailed review on the antecedents of online travel shopping can be found in Amaro and Duarte (2013).

Although the literature has proposed some multi-dimensional constructs to assess hotel website's quality (e.g. Huy Lee et al., 2020), in general website's usability and appealing cannot be directly observed and depends on user's subjective perceptions. This relates to a variety of heuristics including commitment, trust, attitude and perceived playfulness (Morosan and Jeong, 2008; Agag and El-Masry, 2016). In this regard, psychological factors like consumer's innovativeness (San Martin and Herrero, 2012) or locus of control (Fong et al., 2017) have been proposed as explanatory of online purchase intention. Moreover, consumer's cultural capital has been related to the likelihood of the online purchase of tourism services (Quaglione et al., 2020). These multidimensional constructs are difficult to measure and generally unobserved from the researcher viewpoint.

Heterogeneity stemming from both tourists' latent attitudes and the characteristics of the hotel booking channels at hand might cause the effect of trip characteristics on mode choice to differ by segment. That is, unobserved attitudes at the individual level might explain the inconclusive evidence presented above about the role of personal and trip characteristics on the booking mode choice. To fill this gap, we examine the relationship between a set of trip characteristics like travel purpose, distance to origin or length of stay and the booking mode choice while acknowledging the existence of groups of tourists with different preferences.

### 3. EMPIRICAL MODEL

#### 3.1. *Heterogeneity in tourism research*

One stylized finding in tourism research is that tourists' preferences differ by profile. Observable sources of preference heterogeneity such as motivations or sociodemographic features are acknowledged as explanatory of why different tourists make different choices (Heung et al. 2001). When the researcher has a priori theoretical reasons to consider that certain characteristic explains differences in choices, that covariate is included in the empirical model.

However, in most applications it is unlikely to have full information on all sources of heterogeneity. There are usually some unobserved factors that drive decisions. If not accounted for, parameter estimates can be severely biased (Jedidi et al., 1997). Although the estimates for models that treat all observations as being generated by the same data generation process can provide a good statistical fit, the implications and conclusions obtained from *average* results could be misleading. Assaf et al. (2016) discuss in detail the relevance and implications of unobserved heterogeneity in tourism research and practise, arguing that not accounting for it can bias the results and severely affect the reliability of the estimates. Because of these reasons, the literature is paying growing attention to how to account for unobserved heterogeneity.

Traditionally, scholars relied on two-stage procedures that combine cluster analysis in a first stage with multigroup analysis in the second stage. Examples of this include Chen and Lin (2012), Rid et al. (2014) and Srihadi et al. (2016), among others. This modelling framework exhibits at least two major caveats. First, the sample segmentation is done based on some arbitrary chosen distance measure that assigns individuals to groups based on similarity issues. As a result, different cluster methodologies will lead to different classifications. Therefore, results are heavily dependent on the method used (Eshghi and Haughton, 2011). Second, the two-staged nature of the segmentation makes the prior group classification and the subsequent regression analysis to be two separate processes that have been shown to produce biased results (Görz et al., 2000).

Due to these shortcomings, researchers have started to use latent class methodology (also referred as Finite Mixture Modelling, interchangeably). The intuition behind is that the data is assumed to be generated by a mixture of components (classes) with certain probability densities. Individuals are assigned to classes probabilistically. The mixture

distribution thus provides a natural representation of heterogeneity by dividing the sample into subpopulations and then modelling each one parametrically. Latent class membership is usually modelled as a function of some exogenous covariates, generally sociodemographic characteristics. Accordingly, the source of preference heterogeneity for the outcome of interest is explained by a set of factors, thereby providing an interpretation of why different groups exist in the data.

Latent class methodology has a long tradition in the recreational demand literature (e.g. [Boxall and Adamowicz, 2002](#); [Morey et al., 2006](#)). In the last decades, it has started to be also used in the tourism literature in different contexts, such as the modelling of tourists' length of stay ([Alegre et al., 2011](#)), preferences for packages tours ([Chen et al., 2019](#)), destination choice ([Crouch et al., 2016](#)) or cruisers' expenditure ([Baños and Tovar, 2019](#)).

### 3.2. Econometric Modelling

Based on Random Utility Maximization theory ([McFadden, 1974](#)), we assume that the latent utility that each individual  $i$  ( $i = 1, \dots, N$ ) obtains for booking the hotel through each available alternative  $j$  ( $j = 1, \dots, J$ ) is given by:

$$U_{ij}^* = V_{ij} + \varepsilon_{ij} \quad (1)$$

where  $i$  indexes individuals,  $j$  indexes the booking modes,  $V_{ij}$  is a function of individual observable characteristics and  $\varepsilon_{ij}$  is a random error term that reflects unobserved factors. The deterministic part of the latent utility is assumed to be a linear-in-parameters function of the tourist' characteristics so that:

$$V_{ij} = X_i' \beta_j \quad (2)$$

where  $\beta_j$  are parameters to be estimated for each mode.

Assuming that the error term follows a Type I Extreme Value (Gumbel) distribution, the probability that individual  $i$  chooses booking mode  $j$  ( $P_{ij}$ ) is:

$$P_{ij} = \frac{\exp(V_{ij})}{\sum_{j=1}^J \exp(V_{ij})} \quad (3)$$

This is the standard Multinomial Logit Model (hereafter MNL), widely used in tourism research (e.g. [Albadalejo and Díaz-Delfa, 2005](#)). One of its shortcomings is the Independence from Irrelevant Alternatives (henceforth IIA) property. This implies that the ratio of probabilities between two options does not change if a third (irrelevant) one is included or excluded from the choice set. Put another way, it means that the odds ratios are fixed independently of the availability of other alternatives. When this is not the case, coefficient estimates are consistent ([Train, 2009](#)), but biased.

A classical alternative to the MNL model is the Multinomial Probit Model (MNP). The MNP model assumes that the error terms follow a multivariate normal distribution and allows them to be correlated. The problem is that choice probabilities do not have a closed form and need to be estimated by simulation. Because the estimation of the MNP model involves as many integrals as one fewer than the number of alternatives, in some cases the model fails to converge. Even if so, some authors warn that the MNP is weakly identified and can lead to unreliable parameter estimates ([Dow and Endersby, 2004](#)).

In some settings, alternatives correlate with each other in different ways so that substitution between alternatives is not uniform. This can be accommodated using a Generalized Extreme Value (GEV) distribution. Among them, the Nested Logit Model (NL) is the most used because it allows for correlation between alternatives belonging to the same nest. In some contexts, the alternatives have a natural grouping that makes it easier for the researcher to define the nests. However, in most situations it is not clear which alternatives belong to the same nest. Additionally, this model is used when the researcher is interested in the effect of alternative-specific attributes on the probability of each option being chosen (i.e. the model is generally limited to alternative-specific characteristics). In our context, we aim to examine the effect of respondent-specific features, which are constant over alternatives.

These models (MNL, MNP and NL) assume homogeneity in preferences (i.e. they estimate *average* effects). However, as discussed in Section 2, it seems necessary to account for unobserved heterogeneity in tourists' preferences ([Assaf et al., 2016](#)). One way to do this probabilistically is the Finite Mixture Modelling approach. This procedure assigns individuals to classes based on observed characteristics and then estimates the structural parameters for each class. In doing so, preferences are heterogeneous across classes but homogeneous within classes. Hence, we propose a Finite Mixture Multinomial Logit Model (FMMNL).

Consider  $C$  classes of individuals ( $c = 1, \dots, C$ ) whose latent utility is given by:

$$U_{ij|c}^* = X_i' \beta_{jc} + \varepsilon_{ij|c} \quad (4)$$

where  $X_i$  is defined as before,  $\beta_{jc}$  is a set of parameters to be estimated for each class and each booking mode, and  $\varepsilon_{ij|c}$  are Gumbel random error terms that are independent across classes.

The probability that individual  $i$  belongs to class  $c$  is modelled as a function of a vector of  $K$  ( $k = 1, \dots, K$ ) individual characteristics ( $Z_{ikc}$ ) so that:

$$\Pi_{ic} = \frac{\exp(Z_{ikc}' \theta_c)}{\sum_{c=1}^C \exp(Z_{ikc}' \theta_c)} \quad (5)$$

Class membership is thus modelled as a MNL, being  $\theta_c$  a vector of parameters to be estimated for  $c - 1$  classes.

The probability that individual  $i$  belongs to class  $c$  and chooses alternative  $j$  is given by:



$$Prob(mode = j) = \sum_{c=1}^C \left[ \frac{\exp(Z_{ikc}'\theta_c)}{\sum_{c=1}^C \exp(Z_{ikc}'\theta_c)} \right] \left[ \frac{\exp(X_i'\beta_j)}{\sum_{j=1}^J \exp(X_i'\beta_j)} \right] \quad (6)$$

The parameters of the class membership and mode choice equations are jointly estimated by Maximum Likelihood using the expectation-maximization (EM) algorithm.

Several remarks are in order. First, the FMMNL model is a kind of Mixed Logit Model with a finite number of support points. Hence, it can represent any Random Utility Model (McFadden and Train, 2000). Second, under this model, the IIA property is not imposed across classes. Third, when  $\theta_c = \theta$  and  $\beta_c = \beta$  for each class, the model collapses to the baseline MNL model with homogeneous preferences.

## 4. DATA

### 4.1. Database

Our database is drawn from personal surveys directed to a representative sample of visitors to Asturias (Spain) over 18 provided by the *Tourist Information System of Asturias* (SITA). Data were collected both on the street and in collective establishments using a mixture of quota random sampling and pure random sampling. Data were obtained over a period of 12 years, between January 2005 and December 2016. Questionnaires were available in different languages (Spanish, German, English and French). The survey collects information about tourists' sociodemographic characteristics, travel purpose, length of stay, mode of transport and chosen type of accommodation, among others.

Asturias is a Northern Spanish region with 10,602 square kilometers for whom the tourism industry has recently become a major source of income. The tourism sector accounts for about a 10% of its Gross Domestic Product and 11% of its total employment. In our analysis, we restrict the sample to those tourists who spent at least one night in Asturias (i.e. same-day visitors are excluded), lodge at hotels and have made a previous reservation. During our study period, a total of 1,439 tourists (6.4%) declared that they did not book the hotel but merely asked for a room at the time of arriving (walk-in). Since these tourists represent a small share and we expect their trip decision-making process to be different from those who book in advance, they are excluded from the analysis. After having also dropped those with missing values in the variables of interest, our database comprises a total of 17,921 valid observations.

In the questionnaire, respondents are asked about the channel through which they booked the hotel. Four modes are considered: telephone (calling directly the hotel asking for a room), the internet (either directly through the hotel website, through the online DMO or through market-based online platforms), TA and by other channels (hereafter *intermed*). These four options are our dependent variables. The latter refers to declaring that the firm where the tourist works, friends or relatives did the booking on behalf of her. Hence, this category collapses non-market-based intermediaries, which we in short will refer to as 'intermediaries'.

#### 4.2. Evolution of booking mode choices over time

Figure 1 depicts the percentage of tourists in the sample that make the reservation by each of the four possible alternatives per year. As can be seen, there has been a considerable change in the preferences for hotel booking mode over time. In 2005 most hotel guests opted for booking the hotel by telephone (61.9%) and only a small fraction did it on the internet (6.6%). By contrast, in 2016 the telephone represented only 17.8% while the internet was chosen by 55.6% of the sample. Therefore, there has been a substitution of the telephone for the internet. On the other hand, the share of tourists who book through TAs has remained relatively unchanged around 8%. However, this channel has gained popularity from 2014 onwards. Finally, the share of hotel bookings through friends, relatives or the firm has slightly decreased over time, changing from 19.7% in 2005 to 8.5% in 2016.

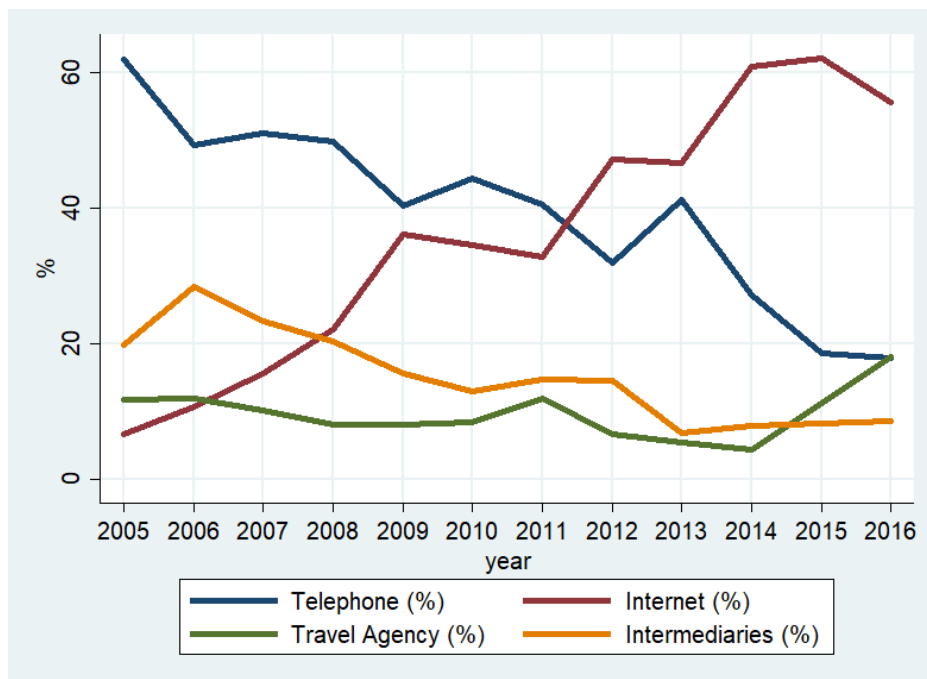


Figure 1.- Percentage of hotel booking mode choices per year

#### 4.3. Variable definition and model specification

Tourists' preferences for the booking channel can be affected by several factors. In line with the related literature, we block them into five groups: i) trip-related characteristics, ii) travel purpose, iii) distance to origin, iv) hotel type, and iv) time trend.

- Trip-related characteristics (Trip): to examine potential differences between first-time and repeat visitors, we define a dummy variable for whether it is the first time the tourist visits Asturias (*first-time*). Similarly, the mode of transport constitutes another relevant dimension that might influence the channel through which tourists book the hotel. We define a dummy variable for travelling in a public mode

of transport (bus, train or plane, denoted by *public\_trans*) as opposed to private car/motorbike. Additionally, booking mode preferences might be seasonal-dependent. Therefore, we also consider whether the tourist travels in the high season (July, August or September) through the dummy variable *high\_season*. Similarly, we include a binary indicator for whether it is a weekend trip or not (*weekend*). Finally, the length of the stay (*LOS*) at the destination is also considered as a determinant of the booking mode.

- Travel purpose (Purp): preferences for booking channels might also be influenced by travel purpose. We define a binary indicator for whether the tourist states that leisure and entertainment (*leisure*) is the main reason for travelling. The omitted category thus gathers all other purposes (work-related, visiting friends or relatives, religious peregrination, health-related treatments, etc.).
- Distance to origin (*distance*): consistent with the literature, we hypothesize that booking mode preferences correlate with the geographical distance to origin. To explore this, we compute Euclidean distance in kilometres from tourists' origin to Asturias using a GPS system.
- Hotel type: we consider two dummy variables for midscale and luxury hotels (*midscale* and *luxury*, respectively). The definition of the hotel type is based on hotel star rating and follows Bi et al. (2020), since there is fair evidence that they are a valid proxy of expected quality (e.g. Mohsin et al., 2019). We define 3-star hotels as *midscale* hotels while 4- and 5-star hotels are included in the *luxury* label. Hotels with 1 or 2 stars (*economy*) are the reference category.
- Time trend (trend): based on the evidence presented in Figure 1, we include a time trend to control for a potential change in preferences over time.

Accordingly, we model the latent utility for each possible booking mode  $j$  conditional on membership to class  $c$ , baseline equation (4), as follows:

$$U_{ij|c}^* = \alpha_{jc} + \beta_{1jc} \text{first-time}_i + \beta_{2jc} \text{public\_trans}_i + \beta_{3jc} \text{high\_season}_i + \beta_{4jc} \text{weekend}_i + \beta_{5jc} \text{LOS}_i + \beta_{6jc} \text{leisure}_i + \beta_{7jc} \text{distance}_i + \beta_{8jc} \text{midscale}_i + \beta_{9jc} \text{luxury}_i + \beta_{10jc} \text{trend}_i + \varepsilon_{ij|c} \quad (7)$$

for  $j = \text{telephone, the internet, TAs, and intermediaries}$ ,  $c = 1, \dots, C$ , and  $i = 1, \dots, N$ , where  $\alpha_{jc}$  is a mode-specific constant term.

Concerning class membership, we believe that sociodemographic characteristics are suitable candidates for segmenting tourists into classes. We specifically consider gender (*male*); age (in years); place of residence (a dummy variable denoted by *foreign* for whether the tourist lives abroad) and labour status (distinguishing among *retired*, *employee* and *self-employed*). The reference category gathers housewives, students and unemployed people. Since we lack information on income, this grouping is intended to also control for income differences. Additionally, the regional area within Asturias where the tourist stays might be another relevant source of preference heterogeneity. Therefore, the class allocation function further considers two dummy variables for staying in the central or the east area (*centre* and *east*, respectively), being the west the reference category. Finally, we consider a dummy for whether the tourist visits more regions apart from Asturias in the current trip (labelled as *multidest*). This relates to

recent evidence that stopover tourism represents a different market segment (Masiero et al., 2020).

#### 4.4. Descriptive statistics

Table 1 provides summary statistics of all the variables defined above. Our sample is characterized by 56% of males with an average age of 40 years old. Most of them live in Spain, with only 6.5% coming from abroad. The majority participate in the labour market, either as employees (65%) or self-employed (18.2%), and travel for leisure purposes (84.4%). Average distance to origin is 670.9 kilometres, being private transport the most common travel mode (78.6%). The central area is the preferred one (55.4%), with 44% of the sample visiting Asturias for the first time. Almost half of the tourists travel in the high season (47.6%) with an average length of stay of 4.1 days. Midscale hotels are the most chosen (34.5%), closely followed by luxury (32.9%) and economy ones (32.6%).

<b>Continuous variables</b>	<b>Description</b>	<b>Mean</b>	<b>SD</b>	<b>Min</b>	<b>Max</b>
<i>age</i>	Age (years)	40.3	12.2	18	91
<i>LOS</i>	Length of stay (days)	4.111	3.1	1	60
<i>distance</i>	Distance to the place of origin (km)	670.9	1,116.0	0	19,800
<b>Categorical variables</b>	<b>Description</b>	<b>Percent (%)</b>			
<i>male</i>	=1 male	55.9			
<i>female</i>	=1 female	44.1			
<i>foreign</i>	=1 lives abroad	6.5			
<i>resident</i>	=1 lives in Spain	93.5			
<i>student</i>	=1 student	5.5			
<i>housewife</i>	=1 housewife/househusband	3.2			
<i>employee</i>	=1 employee	65.0			
<i>self_employed</i>	=1 self-employed	18.2			
<i>unemployed</i>	=1 unemployed	2.2			
<i>retired</i>	=1 retired	5.9			
<i>first-time</i>	=1 first-time visitor	43.5			
<i>repeat</i>	=1 repeat visitor	56.5			
<i>high_season</i>	=1 high season (June/July/August/September)	47.6			
<i>low_season</i>	=1 low season	52.4			
<i>weekend</i>	=1 weekend trip	54.1			
<i>private_trans</i>	=1 travels by car/motorbike	78.6			
<i>public_trans</i>	=1 travels by public transit (bus/train/plane)	21.4			
<i>multidest</i>	=1 visits other regions in the current trip	14.4			
<i>only_Asturias</i>	=1 only visits Asturias	85.6			
<i>leisure</i>	=1 leisure as main trip purpose	84.4			
<i>other_purpose</i>	=1 other trip purpose	15.6			
<i>centre</i>	=1 stays in the central area (Oviedo/Gijón/Avilés)	55.4			
<i>west</i>	=1 stays in the west area	16.6			
<i>east</i>	=1 stays in the east area	27.9			
<i>economy</i>	=1 stays at a 1- or 2-star hotel	32.6			
<i>midscale</i>	=1 stays at a 3-star hotel	34.5			
<i>luxury</i>	=1 stays at a 4- or 5-star hotel	32.9			
Sample size		<b>17,921</b>			

Table 1.- Descriptive statistics

## 5. RESULTS

### 5.1. Class membership

Prior to the estimation of the FMMNL model, we first examine the booking mode choice decision using a classical MNL model. The parameter estimates and the corresponding average marginal effects (henceforth AME) are presented in Tables A1-A2 in the Supplementary Material. As a diagnosis test, we run global Likelihood Ratio (LR) and Wald tests for joint statistical significance. These tests show that the explanatory variables are jointly statistically significant for explaining the booking mode choice, justifying our empirical specification. We conducted LR and Wald tests for whether any of the four modes should be combined. These tests clearly reject the null hypothesis that any pair of alternatives could be collapsed. Additionally, we computed the Hausman & McFadden test for the IIA assumption (Hausman and McFadden, 1984). This test strongly rejects the null hypothesis that the odds ratio between alternatives  $j$  and  $k$  are independent of other options, so we turn to the estimation of the FMMNL model (see Tables A3-A5 in the Supplementary Material).

For the choice of the number of classes, scholars typically consider a small number and select the preferred specification based on information criteria statistics like the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC). Table 2 presents AIC and BIC values for the case of two and three classes. As can be seen, the model with three segments seems to be the preferred one.

Model	Log L	AIC	BIC
MNL	-22,422.86	36,872.43	37,082.87
FMMNL 2 classes	-17,796.47	35,726.93	36,249.11
<b>FMMNL 3 classes</b>	<b>-17,437.6</b>	<b>35,089.2</b>	<b>35,923.13</b>

Table 2.- Information criteria statistics

To address concerns about potential multicollinearity, we performed the Variance Inflation Factor (VIF) diagnostic test in both the class membership and in the booking mode equations. For all the variables included in the model, the VIF is lower than 10, which is usually taken as the threshold value for potential multicollinearity problems.

Since the interpretation of the parameter estimates is not straightforward, we report the AME (in percentage) for the FMMNL model with three classes. The coefficient estimates can be found in Table A6 in the Supplementary Material.

Table 3 presents the AME for the tourists' class membership probabilities. Classes appear to be balanced in terms of gender, with females being slightly more likely to be in class 1 (at 10% significance level). Elderly people and tourists coming from other Spanish regions are significantly more likely to belong to class 1, whereas class 3 is mainly composed of foreign and young tourists. As for labor status, tourists who work (either as employees or self-employed) are more likely to be in class 3, while retired people are more represented in class 1. Class 2 is significantly composed of tourists

lodged at the west area of the region and who only visit Asturias in the current trip. Those who carry out a multi-destination trip are more likely to be in class 1. Overall, classes 1 and 3 gather significantly different tourist's profiles, while class 2 is more balanced. About 20% of the sample is classified as belonging to class 1 and 24% to class 2. The remaining 55% belong to class 3, which is the largest.

Variables	Class 1 AME (%)	Class 2 AME (%)	Class 3 AME (%)
male	-2.84*	1.25	1.59
age	0.92***	-0.08	-0.84***
foreigner	-29.23***	-1.05	30.28***
employee	-7.54***	-4.92**	12.46***
self_employed	-9.71***	3.06	6.65**
retired	9.60**	-13.02**	3.42
center	1.64	-21.76***	20.12***
east	6.22**	-14.72***	8.50***
multidest	6.73***	-8.31***	1.58
Share (%)	20.5	24	55.3

Table 3.- AME for class probabilities

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

## 5.2. Booking mode preferences by class

Table 4 shows the AME for the booking mode choice. Conditional on belonging to class  $c$ , the marginal effect for a generic variable  $X_k$  is computed as follows:

$$\frac{\partial \text{Prob}(\text{booking mode} = j|c)}{\partial X_k} = \text{Prob}(\text{booking mode} = j|X, c) [\beta_{jk} - \sum_{j=1}^J \beta_{jk} \text{Prob}(\text{booking mode} = j|X, c)] \quad (8)$$

It is important to highlight that, contrary to the binary logit case, both the sign and the statistical significance of the AME differ from the coefficient estimates because the formula of the marginal effects for alternative  $j$  involves the coefficient estimates for the remaining options.

Before moving to the discussion of the results, let us first characterize each class. Figures 2, 3 and 4 depict the estimated probabilities for each booking mode for classes 1, 2 and 3, respectively. As can be seen, the shares of the booking modes are balanced in class 1. However, tourists in class 2 mostly opt for the telephone (59.7%), while those in class 3 show a strong preference for the internet (58.5%). Based on this evidence, tourists belonging to class 1 are deemed *hybrids*, those in class 2 *traditionalists* and those in class 3 *techys*.

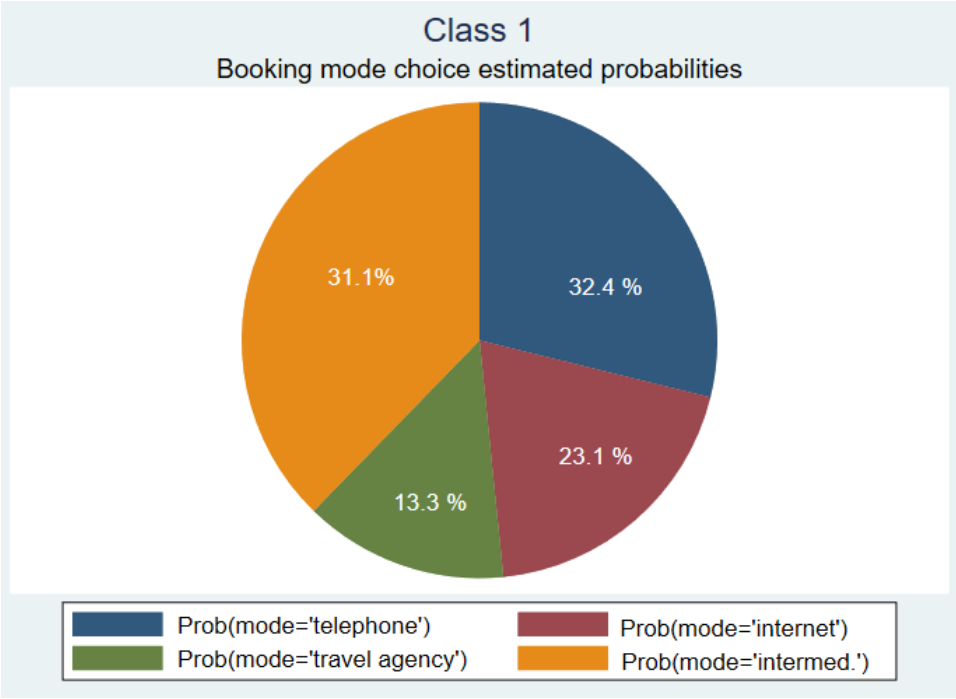


Figure 2.- Booking mode choice estimated probabilities for class 1

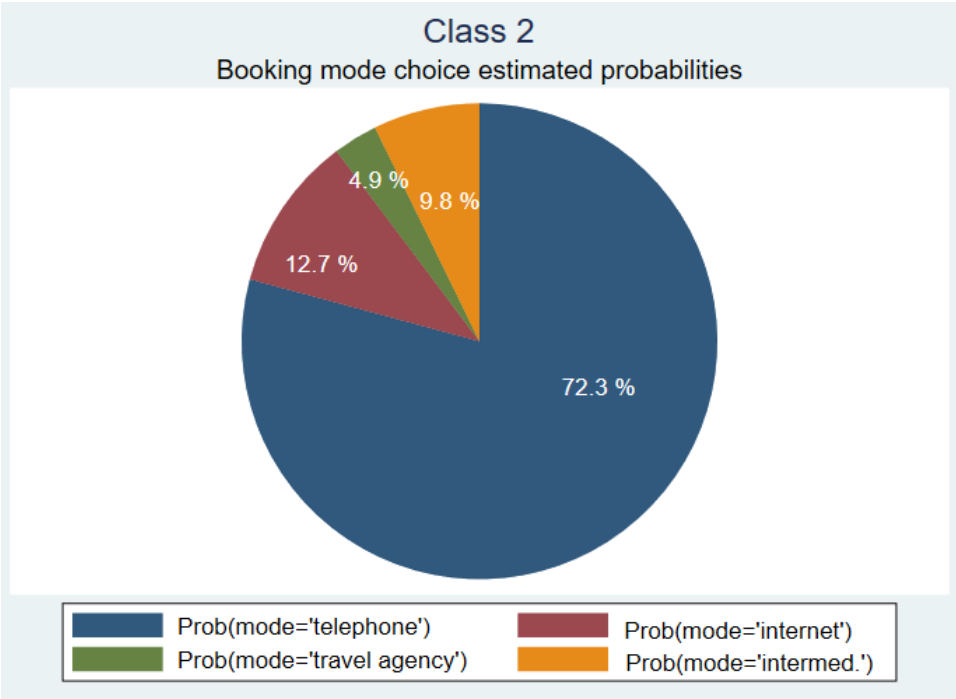


Figure 3.- Booking mode choice estimated probabilities for class 2

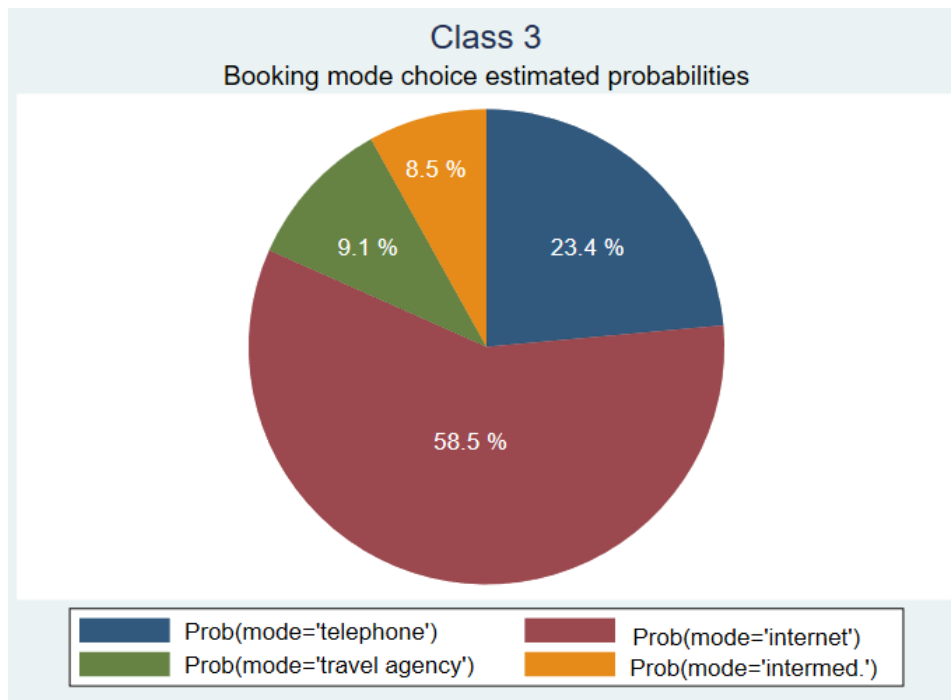


Figure 4.- Booking mode choice estimated probabilities for class 3

We now proceed to discuss the AME for each class.

*Class 1 (hybrids):*

First-time visitors in class 1 are less likely to book the hotel by telephone and TAs and prefer instead to do it through non-market-based intermediaries. Tourists travelling by public modes of transport are associated with higher probability of room booking by TAs and intermediaries. This is consistent with [Topolsek et al. \(2014\)](#), who show that public transport operators cooperate with TAs. However, those travelling in the high season seem to dislike TAs as the booking channel. Interestingly, TAs are preferred for weekend trips, whereas those who stay for longer favour the telephone. The latter contradicts [Wu et al. \(2013\)](#), who find that longer trips are associated with online hotel room bookings. Compared to other purposes, leisure tourists like better the internet and intermediaries. This is in line with [Gössling and Lane \(2015\)](#) and [Gronflaten \(2011\)](#). Concerning the role of distance, we document that telephone is more prevalent among those coming from nearby origins, with those coming from distant locations preferring other channels. In comparison to economy hotels, tourists lodged at midscale and luxury hotels prioritize booking the room through intermediaries, and strongly dislike making it by telephone. Finally, the trend term indicates that there has been an increase in online booking over time at the cost of TAs and other intermediaries.

*Class 2 (traditionalists):*

Tourists who have never been to Asturias do not show a significant preference for any booking mode. Similar to class 1, those travelling by public modes of transport tend to opt for non-market-based intermediaries, although here the marginal probabilities are of lower magnitude (11.7% versus 39.2%). Interestingly, those travelling by private



transport seem to book the hotel by telephone. Those who come to Asturias in the high season dislike booking the hotel by TAs, being the effect here notably larger (-4.28% vs -2.3%). As for weekend trips, there is a preference for the telephone (although at 10% significance level). Contrary to class 1, longer stays are associated with a significantly higher probability of booking the room by TAs. By contrast, leisure tourists display a 6.6% lower likelihood of hiring hotel services through TAs. The booking mode choice is not related to either distance or opting for midscale hotels. However, tourists staying in a luxury hotel have a 12% higher probability of booking the room by non-market intermediaries and a 7.5% higher probability of doing it on the internet. This is in line with [Qi et al. \(2013\)](#) and [Masiero and Law \(2016\)](#). Lastly, reservations made on the internet and TAs have increased over the study period in this class.

#### *Class 3 (techys):*

Unlike the other two classes, first-time visitors in class 3 show a significantly higher probability of booking the room online. Regarding the mode of transport, public transit increases the likelihood of booking through TAs by 2%. TAs are also marginally preferred as LOS increases (0.25%) but are notably disregarded for leisure tourists and those coming from distant locations. Like class 1, TAs and intermediaries are preferred for midscale and luxury guests. Tourists opting for economy hotels are by contrast more likely to choose the telephone. Similar to class 2, luxury hotel guests are about 7.5% more likely to book the room on the internet. Furthermore, there has been a shift towards the internet over time at the expense of the other channels.

Overall, our findings can be summarized as follows. Considering the existence of three classes of tourists, the preference for the internet for first-time visitors only holds for those in class 3 (*techys*). By contrast, first-time tourists in class 1 (*hybrids*) tend to book hotel rooms through other non-market intermediaries. Travelling by public transit is positively related to other intermediaries for *traditionalists* and to TAs for *techys*. *Traditionalists* travelling by car/motorbike prefer booking the room by telephone. The three segments like TAs better in the low season, although the magnitude of the effects differs across classes. Longer stays are positively associated with the telephone for *hybrid* tourists and with TAs for *techys* and *traditionalists*.

Regarding travel purpose, the preference for the internet for leisure tourists is significantly higher for *hybrids*. *Hybrids* coming from distant locations prefer either the internet or other intermediaries, while those from distant locations do not exhibit a significant preference order for the different channels in either *techys* or *traditionalists*. Concerning hotel quality, economy hotels appear to be mainly booked by telephone whereas reservations for midscale ones are made by other intermediaries. Luxury hotels are also more likely to be booked through intermediaries in the three classes but also on the internet and TAs for the case of *techys*. Consistent with Figure 1, there has been i) a rise in online bookings over time at the expense of the telephone and other intermediaries for the three classes, and ii) an increase in bookings through TAs among *traditionalists*.

Variables	Class 1 AME (%)				Class 2 AME (%)				Class 3 AME (%)			
	Telephone	Internet	TA	Intermed.	Telephone	Internet	TA	Intermed.	Telephone	Internet	TA	Intermed.
<i>first-time</i>	-5.13*	0.74	-4.45***	8.85***	-2.70	2.08	0.41	0.21	-5.95***	4.06***	0.20	1.69
<i>pub_transit</i>	-4.59	-51.17	16.57***	39.20**	-15.44***	6.25*	-2.58*	11.76***	-2.51	-0.71	1.49**	1.72
<i>high_season</i>	-0.94	1.89	-2.31**	1.40	0.10	5.67	-4.28***	-1.49	-0.42	2.04	-2.92***	1.30
<i>weekend</i>	2.55	1.09	2.93**	-6.58***	5.23*	-2.12	0.62	-3.74*	-0.14	1.06	-2.96***	2.03
<i>LOS</i>	1.62***	-1.63**	-0.33	0.34	-0.37	-0.03	0.78***	-0.37	-0.11	-0.57**	0.25***	0.43**
<i>leisure</i>	-9.02**	17.18***	-16.87***	8.72***	-18.84	24.71*	-6.60***	0.74	5.45	2.85	-20.38***	12.08
<i>distance</i>	-0.04***	0.02***	6.3e-03***	0.02***	1.1e-03	1.2e-03	-5.69e-06	4.63e-06	-3.20e-06	1.65e-06	-3.77e-06*	5.32e-06
<i>midscale</i>	-6.65**	-1.95	1.54	7.05***	-6.63	0.14	1.35	5.14*	-13.97***	1.64	1.53**	10.80***
<i>luxury</i>	-11.56***	-0.87	1.05	11.37***	-20.84***	7.51**	1.34	11.99***	-16.68***	7.56***	3.24***	5.89**
<i>trend</i>	-0.86	3.27***	-1.07***	-1.34***	-2.82***	3.14***	3.22***	-3.54***	-5.45***	6.54***	-5.81e-03	-1.08***
Marginal means	<b>0.32</b>	<b>0.23</b>	<b>0.13</b>	<b>0.31</b>	<b>0.72</b>	<b>0.12</b>	<b>0.05</b>	<b>0.10</b>	<b>0.23</b>	<b>0.58</b>	<b>0.09</b>	<b>0.08</b>

Table 4.- FMMNL model Average Marginal Effects  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

## 6. CONCLUSIONS

### 6.1. Summary of findings

In this paper we estimate a Finite Mixture Multinomial Logit to model the channel through which tourists book hotel rooms. Four options are considered: telephone, the internet, TAs and other non-market intermediaries (friends, relatives, etc.). Using a large dataset of tourists visiting Asturias during the period 2005-2016, we estimate a FMMNL model with three latent classes. In doing so, we show that preferences vary across classes. In this way, models based on a homogeneity assumption, although informative, may mask the existence of different segments of tourists.

Our results show that there are three different groups of tourists. Whereas one class (*techys*) exhibits a strong preference for booking hotel rooms on the internet, other class (*traditionalists*) prefers to do it through traditional booking modes, mainly by telephone. The group labelled *hybrids* is by contrast more heterogeneous, being the predicted mode choice probabilities more balanced. Tourist segmentation into these three classes has been done based on sociodemographic characteristics. Contrary to other methodologies, the class allocation and the modelling of the booking choice has been performed jointly in a single step. We find that *techys* are composed of young and foreign individuals who participate in the labour market. Conversely, *hybrids* are older, travel domestically and visit other regions in the current trip. *Traditionalists* are those who only visit Asturias and lodge in the west area.

Overall, we conclude that: i) tourists are not homogeneous in terms of their preferences for hotel booking mode but segmented into classes according to some observed and unobserved factors; ii) there is a clear distinction between tourists who are into new technologies and book online (*techys*) and tourists who still prefer to make the reservation through traditional offline channels (*traditionalist*); and iii) within classes, preferences for one booking mode or another change depending on trip-related features.

### 6.2. Practical implications

Our results have important implications for hospitality managers. At the hotel level, our findings allow managers to identify to whom it is better to sell the rooms online and to whom it is still necessary to continue providing offline channels. Apart from customizing sales channels, hoteliers should not focus on one specific booking channel but be present in different ones to accommodate different segments desires. This evidence is illustrated in the case of first-time visitors and the length of the stay. First-time tourists in the *techys* group prefer the internet whereas for the case of *hybrids* other intermediaries are the preferred option. *Traditionalists* with longer stays prefer TAs while *hybrids* counterparts are more likely to opt for the telephone. Therefore, despite the growing popularity of the internet as a distribution channel, hoteliers need to continue marketing their services through both online and offline channels.

At the destination level, the identification of the online and offline hotel guest profile can help public authorities to develop better promotional campaigns. In general terms, digital

platforms seem to be more relevant for those classified as *techys*. Since foreign tourists mainly belong to this segment, it seems that public authorities should develop different promotion strategies depending on the intended customer base. If the destination aims to increase the number of foreign visitors, it appears that the online channels is the best way to promote the destination among this segment.

Even though hoteliers market their rooms both online and offline, they usually focus on one channel. Whereas some hotels sell their rooms mainly through third parties or intermediaries, others interact with their customers directly. To maximize revenues, hotel managers thus need to pay greater attention to the booking mode preferences of the type of tourists they intend to attract. For instance, hotels in the main cities whose customer base is mainly composed of non-leisure tourists need to be aware that they exhibit a strong preference for TAs. Consequently, these hotels need to work with TAs to offer their rooms through them. However, repeat visitors and those who stay longer prefer booking the room by telephone, both in the case of *hybrids* and *techys*. This suggests that the availability of a hotel call centre is still relevant. Hoteliers interested in attracting a specific market segment must then allow customers the possibility of making the reservation by their preferred mode. We believe that hospitality and tourism research needs more theoretical and empirical investigations into the underlying preference heterogeneity.

### 6.3. *Limitations and future research*

This study has some limitations. First, we lack information on how tourists usually book hotel rooms in other trips. In this sense, booking mode preferences might differ depending on whether it is a nature-based or a coastal destination. In case longitudinal data were available, future studies should examine tourists' choice of the booking channel in different types of trips. This would address whether mode preferences are context specific. Similarly, we focus on the mode choice for hotels. However, booking channel preferences might be different for other types of accommodation like rural houses or private apartment. Future studies should also address whether the heterogeneity in preferences analysed here also applies to other lodgings.

Second, this study does not explore how the mode choice relates to the advance of the booking. We might expect channels like travel agencies to be more likely for trips planned long in advance, whereas the internet might be preferred mode for unplanned last-minute trips. Future research should therefore study in detail the linkages between the mode choice and the degree of ahead trip planning.

In the context of the pandemic situation caused by COVID-19, future research should address the expected changes in tourists' preferences for hotel booking. The increase in the use of online platforms for teleworking and social interactions might reduce the distrust in the online channel due to higher exposure. At the same time, social distancing rules might disincentive tourists from resorting on physical travel agencies.

## REFERENCES

- Agag, G. and A.A. El-Masry. 2016. "Understanding the determinants of hotel booking intentions and moderating role of habit." *International Journal of Hospitality Management* 54: 52-67.
- Albadalejo-Pina, I.P. and M.T. Diaz-Delfa. 2005. "Rural tourism demand by type of accommodation." *Tourism Management* 26: 951-959.
- Alegre, J., S. Mateo, and L. Pou. 2011. "A latent class approach to tourists' length of stay." *Tourism Management* 32: 555-563.
- Amaro, S. and P. Duarte. 2013. "Online travel purchasing: a literature review". *Journal of Travel & Tourism Marketing* 30(8): 755-785.
- Amaro, S. and P. Duarte. 2015. "An integrative model of consumers' intentions to purchase travel online". *Tourism Management* 46: 64-79.
- Assaf, A.G., H. Oh, and M.G. Tsionas. 2016. "Unobserved Heterogeneity in Hospitality and Tourism Research." *Journal of Travel Research* 55(6): 774-788.
- Baños, J.F. and B. Tovar. 2019. "Explaining cruisers' shore expenditure through a latent class tobit model: Evidence from the Canary Islands." *Tourism Economics* 25(7): 1105-113.
- Bargeman, B. and H. Van der Poel. 2006. "The role of routines in the vacation decision-making process of Dutch vacationers." *Tourism Management* 27: 707-720.
- Bhatiasevi, V. and C. Yoopecth. 2015. "The determinants of intention to use electronic booking among young users in Thailand." *Journal of Hospitality and Tourism Management* 23, 1-11.
- Bi, J.W., Y. Liu, Z.P. Fan, and J. Zhang. (2020). "Exploring asymmetric effects of attribute performance on customer satisfaction in the hotel industry". *Tourism Management*, forthcoming. <https://doi.org/10.1016/j.tourman.2019.104006>
- Bonn, M.A., H.L. Furr, and A.M. Susskind. 1998. "Using the internet as a pleasure travel tool: an examination of the sociodemographic and behavioral characteristics among internet users and nonusers." *Journal of Hospitality and Tourism Research* 22(3): 303-317.
- Boxall, P.C. and W.L. Adamowicz. 2002. "Understanding heterogeneous preferences in Random Utility Models: a Latent Class Approach." *Environmental and Resource Economics* 23: 421-446.
- Bufquin, D., J.Y. Park, R.M. Back, M.W.W. Nutta, and T. Zhang. 2020. "Effects of hotel website photographs and length of textual descriptions on viewers' emotions and behavioral intentions." *International Journal of Hospitality Management*, forthcoming <https://doi.org/10.1016/j.ijhm.2019.102378>
- Buhalis, D. and M.C. Licata. 2002. "The future eTourism intermediaries." *Tourism Management* 23(3): 207-220.
- Buhalis, D. and R. Law. 2008. "Progress in information technology and tourism management: 20 years on and 10 years after the internet – the state of eTourism research." *Tourism Management* 29(4): 609-623.
- Casaló, L.V., C. Flavián and M. Guinalíu. 2010. "Determinants of the intention to participate in firm-hosted online travel communities and effects on consumer behavioral intentions". *Tourism Management* 31: 898-911.
- Castillo-Manzano, J.I. and L. López-Valpuesta. 2010. "The decline of the traditional travel agent model." *Transportation Research Part E* 46: 639-649.
- Chan, I.C.C., L.W. Lam, C.W.C. Chow, L.H. Nang Fong, and R. Law. 2017. "The effect of online reviews on hotel booking intention: the role of reader-reviewer similarity." *International Journal of Hospitality Management* 66: 54-65.
- Chen, N., L. Masiero and C.H.C. Hsu. 2019. "Chinese Outbound Tourist Preferences for All-Inclusive Group Package Tours: A Latent Class Choice Model." *Journal of Travel Research* 58(6): 916-931.
- Chen, C. and Y.H. Lin. 2012. "Segmenting Mainland Chinese Tourists to Taiwan by Destination Familiarity: a Factor-cluster Approach." *International Journal of Tourism Research* 14: 339-352.
- Chu, R. 2001. "What online Hong Kong travellers look for on airline/travel websites?" *International Journal of Hospitality Management* 20(1): 95-100.

- Coenders, G., B. Ferrer-Rossell and E. Martínez-García. 2016. "Trip characteristics and dimensions of internet use for transportation, accommodation, and activities undertaken at destination." *Journal of Hospitality Marketing and Management* 25(4): 498-511.
- Crouch, G.I., T. Huybers, and H. Oppewal. 2016. "Inferring Future Vacation Experience Preference from Past Vacation Choice: A Latent Class Analysis." *Journal of Travel Research* 55(5): 574-587.
- Davis, F.D. 1989. "Perceived usefulness, perceived ease of use, and user acceptance of information technology". *MIS Quarterly* 13(3): 319-340.
- Davis, F.D., R.P. Bagozzi and P.R. Warshaw. 1989. "User acceptance of computer technology: a comparison of two theoretical models". *Management Science* 35(8): 982-1003.
- Del Chiappa, G. 2013. "Internet versus travel agencies: the perception of different groups of Italian online buyers." *Journal of Vacation Marketing* 19(1): 55-66.
- Del Chiappa, G., C. Lorenzo-Romero, and M.G. Gallarza. 2015. "Attitudes towards disintermediation in hotel reservations: Spanish travellers' Profile." *European Journal of Tourism Research* 9: 129-143.
- Del Chiappa, G. and A. Zara. 2015. *Offline versus online intermediation: a study of booking behavior of tourists travelling to Sardinia*. In: Tussyadiah I., Inversini A. (eds) *Information and Communication Technologies in Tourism 2015*. Springer.
- Dow, J.K. and J.W. Endersby. 2004. "Multinomial probit and multinomial logit: a comparison of choice models for voting research." *Electoral Studies* 23: 107-122.
- Eshghi, A. and D. Haughton. 2011. "Identifying groups: a comparison of methodologies." *Journal of Data Science* 9: 271-291.
- Fishbein, M. and I. Ajzen. 1975. *Belief, Attitude, Intention and Behaviour: An Introduction to Theory and Research*. Addison-Wesley, Reading, MA.
- Fong, L.H.N., L.W. Lam and R. Law. 2017. "How locus of control shapes intention to reuse mobile apps for making hotel reservations: Evidence from Chinese consumers." *Tourism Management* 61: 331-342.
- Gazzoli, G., W.G. Kim, and R. Palakurthi. 2008. "Online distribution strategies and competition: are the global hotel companies getting it right?" *International Journal of Contemporary Hospitality Management* 20: 375-387.
- Golmohammadi, A.R., B. Jahandideh, and K.D. O'Gorman. 2012. "Booking on-line or not: a decision rule approach." *Tourism Management Perspectives* 2-3: 85-88.
- Görz, N., L. Hildebrandt and D. Annacker. 2000. *Analyzing Multigroup Data with Structural Equation Models*. Classification and Information Processing at the Turn of the Millenium, edited by W. Gaul and R. Decker, 312-319. New York. Springer.
- Gössling, S. and B. Lane. 2015. "Rural tourism and the development of internet-based accommodation booking platforms: a study in the advantages, dangers and implications of innovations." *Journal of Sustainable Tourism* 23(8-9): 1386-1403.
- Gronflaten, O. 2011. "Predicting Travelers' Choice of Information Sources and Information Channels." *Journal of Travel Research* 48(2): 230-244.
- Guillet, B.D., A. Mattila and L. Gao. 2020. "The effects of choice set size and information filtering mechanisms on online hotel booking." *International Journal of Hospitality Management*, forthcoming. <https://doi.org/10.1016/j.ijhm.2019.102379>
- Hausman, J. and D. McFadden. 1984. "Specification Tests for the Multinomial Logit Model." *Econometrica* 52(5): 1219-1240.
- Heung, V.C.S., H. Qu, and R. Chu. 2001. "The relationship between vacation factors and socio-demographic and travelling characteristics: the case of Japanese leisure travellers." *Tourism Management* 22: 259-269.
- Huy Lee, V., H.T.T. Nguyen, N. Nguyen and S. Pervan. 2020. "Development and validation of a scale measuring hotel website service quality (HWebSQ)." *Tourism Management Perspectives*, forthcoming. <https://doi.org/10.1016/j.tmp.2020.100697>
- Jedidi, K., S.J. Harsharanjeet, and W.S. DeSarbo. 1997. "Finite-Mixture Structural Equation Models for Response-Based Segmentation and Unobserved Heterogeneity." *Marketing Science* 16(1): 39-59.

- Jeong, M., H. Oh and M. Gregoire. 2003. "Conceptualizing web site quality and its consequences in the lodging industry." *International Journal of Hospitality Management* 22(2): 161-175.
- Jensen, J.M. 2012. "Shopping orientation and online travel shopping: the role of travel experience." *International Journal of Tourism Research* 14(1): 56-70.
- Law, R. 2009. "Disintermediation of hotel reservations. The perception of different groups of online buyers in Hong Kong." *International Journal of Contemporary Hospitality Management* 21(6): 766-772.
- Law, R. L. Leung, and R.J. Wong. 2004. "The impact of the Internet on travel agencies." *International Journal of Contemporary Hospitality Management* 16(2): 100-107.
- Law, R. and W. Lau. 2004. "A study of the perceptions of Hong Kong hotel managers on the potential disintermediation of travel agencies." *Journal of Travel & Tourism Marketing* 17(2/3): 117-131.
- Law, R., R. Leung, A. Lo, D. Leung, and L.H.N. Fong. 2015. "Distribution channel in hospitality and tourism. Revisiting disintermediation from the perspective of hotels and travel agencies." *International Journal of Contemporary Hospitality Management* 27(3): 431-452.
- Lei, S.S.I., J.L. Nicolau and D. Wang. 2019. "The impact of distribution channels on budget hotel performance." *International Journal of Hospitality Management* 81:141-149.
- Li, L., M. Peng, N. Jiang and R. Law. 2017. "An empirical study on the influence of economy hotel website quality in online booking intentions." *International Journal of Hospitality Management* 63: 1-10.
- Lu, J., Z. Mao, M. Wang, and L. Hu. 2015. "Goodbye maps, hello apps? Exploring the influential determinants of travel app adoption." *Current Issues in Tourism* 18(11): 1059-1079.
- Luo, X., H. Li, J. Zhang and J.P. Shim. 2010. "Examining Multidimensional Trust and Multi-faceted Risk in Initial Acceptance of Emerging Technologies: An Empirical Study of Mobile Banking Services." *Decision Support Systems* 46: 222-234.
- Masiero, L. and R. Law. 2016. "Comparing reservation channels for hotel rooms: a behavioural perspective." *Journal of Travel & Tourism Marketing* 33(1): 1-13.
- Masiero, L., R.T.R. Qiu, and J. Zoltan. 2020. "Long-Haul Tourist Preferences for Stopover Destination Visits". *Journal of Travel Research* 59(5): 811-827.
- McFadden, D. 1974. "Conditional logit analysis of qualitative choice Behavior." In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105-142). New York: Academic Press.
- McFadden, D. and K. Train. 2000. "Mixed MNL models for discrete response." *Journal of Applied Econometrics* 15: 447-470.
- Mohsin, A., H. Rodrigues, and A. Brochado. 2019. "Shine bright like a star: Hotel performance and guests' expectations based on star ratings." *International Journal of Hospitality Management* 83: 103-114.
- Morey, E., J. Thatcher, and W. Breffle. 2006. "Using Angler characteristics and attitudinal data to identify environmental preference classes: a latent-class model." *Environmental and Resource Economics* 34: 91-115.
- Morosan, C. and M. Jeong. 2008. "Users' perceptions of two types of hotel reservation Web sites." *International Journal of Hospitality Management* 27(2): 284-292.
- Murphy, H.C., M. Chen and M. Cossutta, M. 2016. "An investigation of multiple devices and information sources used in the hotel booking process." *Tourism Management* 52: 44-51.
- Park, S. and Y. Huang. 2017. "Motivators and inhibitors in booking a hotel via smartphones." *International Journal of Contemporary Hospitality Management* 29(1): 161-178.
- Park, S. and I.P. Tussyadiah. 2017. "Multidimensional facets of perceived risk in mobile travel booking." *Journal of Travel Research* 56(7): 854-867.
- Pearce, D.G. and C. Schott. 2011. "Domestic vs outbound booking and channel choice behavior: evidence from New Zealand." *International Journal of Culture, Tourism and Hospitality Research* 5(2): 112-127.
- Phelan, K.V., N. Christodoulidou, C.C. Countryman, and L.J. Kistner. 2011. "To book or not to book: the role of hotel web site heuristics." *Journal of Services Marketing* 25(2): 134-148.
- Pollak, R.A. and T.J. Wales. 1981. "Demographic variables in demand analysis". *Econometrica* 49: 1533-1551.

- Qi, S., R. Law and D. Buhalis. 2013. "Who booked five-star hotels in Macau? A study of guests' online booking intention." *Journal of Hospitality and Tourism Management* 20: 76-83.
- Quaglione, D., A. Crociata, M. Agovino and L. Iaia. 2020. "Cultural capital and online purchase of tourism services." *Annals of Tourism Research*, forthcoming. <https://doi.org/10.1016/j.annals.2019.102797>
- Rid, W., I.O. Ezeuduji, and U. Pröbstl-Haider. 2014. "Segmentation by motivation for rural tourism activities in The Gambia." *Tourism Management* 40: 102-116.
- San Martín, H. and A. Herrero. 2012. "Influence of the user's psychological factors on the online purchase intention in rural tourism: integrating innovativeness to the UTAUT framework." *Tourism Management* 33: 341-350.
- Srihadi, T.F., Hartoyo, D. Sukandar and A.W. Soehadi. 2016. "Segmentation of the tourism market for Jakarta: Classification of foreign visitors' lifestyle typologies." *Tourism Management Perspectives* 19: 32-39.
- Stangl, B., A. Inversini and R. Schegg. 2016. "Hotels' dependency on online intermediaries and their chosen distribution channel portfolios: the country insights." *International Journal of Hospitality Management* 52: 87-96.
- Sun, S., R. Law, and M. Schuckert. 2020. "Mediating effects of attitude, subjective norms and perceived behavioral control for mobile payment-based hotel reservations." *International Journal of Hospitality Management*, forthcoming. <https://doi.org/10.1016/j.ijhm.2019.102331>
- Talwar, S., A. Dhir, P. Kaur and M. Mäntymäki. 2020a. "Barriers toward purchasing from online travel agencies." *International Journal of Hospitality Management*, forthcoming. <https://doi.org/10.1016/j.ijhm.2020.102593>
- Talwar, S., A. Dhir, P. Kaur and M. Mäntymäki. 2020b. "Why do people purchase from online travel agencies (OTAs)? A consumption values perspective." *International Journal of Hospitality Management*, forthcoming. <https://doi.org/10.1016/j.ijhm.2020.102534>
- Thakran, K. and R. Verma. 2013. "The emergence of hybrid online distribution channels in travel, tourism and hospitality." *Cornell Hospitality Quarterly* 54: 240-247.
- Toh, R.S., C.F. DeKay, and P. Raven. 2011. "Travel planning: searching for booking hotels on the internet." *Cornell Hospitality Quarterly* 52(4): 388-398.
- Topolsek, D., E. Mrnjavac, and N. Kovacic. 2014. "Integration of travel agencies with transport providers." *Tourism Management Perspectives* 9: 14-23.
- Train, K. 2009. *Discrete Choice Models with Simulation*. Second Edition. Cambridge University Press.
- Tversky, A. and I. Simonson. 1993. "Context-dependent preferences". *Management Science* 39(10): 1179-1189.
- Wang, H. and S. Wang. 2010. "Predicting mobile hotel reservation adoption: insight from a perceived value standpoint." *International Journal of Hospitality Management* 29: 598-608.
- Wang, L., R. Law, B.D. Guillet, K. Hung, and D.K.C. Fong. 2015. "Impact of hotel website quality on online booking intentions: eTrust as a mediator." *International Journal of Hospitality Management* 47: 108–115.
- Wang, L., R. Law, K. Hung, and B.D. Guillet. 2014. "Scale development of perceived trust in the hotel industry: the perspective of Internet users." *International Journal of Hospitality Management* 43: 35–46.
- Wolfe, K., C.H. Hsu, and S.K. Kang. 2005. "Buyer characteristics among users of various travel intermediaries." *Journal of Travel & Tourism Marketing* 17(2–3): 51–62.
- Wu, E.H., R. Law, and B. Jiang. 2013. "Predicting browsers and purchasers of hotel websites a weight-of-evidence grouping approach." *Cornell Hospitality Quarterly* 54(1): 38–48.



## **SUPPLEMENTARY MATERIAL**

### **Tourists' Preferences for Hotel Booking**

Variables	Telephone	Internet	Intermed.
first-time	-0.052 (0.069)	0.294*** (0.069)	0.360*** (0.074)
public_trans	-1.353*** (0.071)	-1.066*** (0.072)	0.060 (0.075)
high_season	0.565*** (0.072)	0.669*** (0.073)	0.563*** (0.078)
weekend	0.260*** (0.064)	0.148** (0.065)	-0.064 (0.071)
LOS	-0.036*** (0.008)	-0.083*** (0.010)	-0.015 (0.009)
leisure	2.417*** (0.070)	3.032*** (0.076)	3.037*** (0.090)
distance	-0.000 (0.000)	0.000*** (0.000)	0.000** (0.000)
midscale	-0.668*** (0.081)	-0.317*** (0.084)	0.379*** (0.093)
luxury	-0.897*** (0.082)	-0.133 (0.083)	0.320*** (0.094)
trend	-0.111*** (0.009)	0.153*** (0.010)	-0.122*** (0.011)
Constant	1.266*** (0.114)	-1.554*** (0.128)	-1.417*** (0.140)
N		17,921	
Log Likelihood		-18352.19	

Table A1.- MNL parameter estimates  
Standard errors in parentheses  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Variables	Telephone	Internet	TA	Intermed.
first	-6.57***	4.49***	-0.99**	3.08***
public_trans	-15.02***	-4.35***	6.05***	13.33***
high_season	0.30	3.28***	-3.79***	0.21
weekend	3.87***	0.00	-0.96**	-2.90***
LOS	0.35***	-1.08***	0.31***	0.42***
leisure	-4.25***	14.29***	-17.40***	7.37***
distance	-0.00***	0.00***	-0.00**	0.00
midscale	-11.87***	0.16	2.06***	9.65***
luxury	-18.37***	6.65***	2.33***	9.39***
trend	-3.37***	4.79***	0.10**	-1.52***

Table A2.- MNL marginal effects  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Variables	LR test		Wald test	
	Chi-squared	p-value	Chi-squared	p-value
first-time	114.67	<0.01	114.24	<0.01
public_trans	848.58	<0.01	847.60	<0.01
high_season	87.43	<0.01	84.98	<0.01
weekend	50.55	<0.01	50.54	<0.01
LOS	90.86	<0.01	82.18	<0.01
leisure	2225.96	<0.01	1988.93	<0.01
distance	56.60	<0.01	52.32	<0.01
midscale	344.70	<0.01	330.45	<0.01
luxury	512.37	<0.01	496.38	<0.01
trend	2633.53	<0.01	2253.20	<0.01
Observations	17,921			

Table A3.- LR and Wald tests for joint statistical significance

Alternatives combined	LR test		Wald test	
	Chi-squared	p-value	Chi-squared	p-value
Telephone-Internet	2909.11	<0.01	2360.64	<0.01
Telephone-TA	3110.68	<0.01	2274.88	<0.01
Telephone-intermed.	1525.15	<0.01	1329.37	<0.01
Internet-TA	3523.68	<0.01	2516.73	<0.01
Internet-intermed.	2287.62	<0.01	1894.22	<0.01
TA-intermed	2158.75	<0.01	1529.27	<0.01
Observations	17,921			

Table A4.- LR and Wald tests for combining alternatives

Omitted option	Hausman and McFadden test	
	Chi-squared	p-value
<b>Telephone</b>	89.00	<0.01
<b>Internet</b>	138.65	<0.01
<b>Intermed.</b>	113.13	<0.01

Table A5.- Hausman-McFadden test for the IIA property

Variables	Class 1			Class 2			Class 3		
	Telephone	Internet	Intermed.	Telephone	Internet	Intermed.	Telephone	Internet	Intermed.
first	0.213 (0.187)	0.436** (0.196)	0.739*** (0.135)	-0.236 (0.277)	0.060 (0.336)	-0.188 (0.448)	-0.431** (0.191)	0.057 (0.155)	0.109 (0.248)
public_trans	-2.899*** (0.288)	-5.483 (3.777)	-0.581*** (0.180)	0.586 (0.510)	1.356** (0.616)	2.652*** (0.768)	-0.497** (0.209)	-0.360** (0.166)	-0.159 (0.300)
high_season	0.223 (0.187)	0.360 (0.226)	0.284** (0.144)	1.229*** (0.335)	1.783*** (0.471)	0.987* (0.530)	0.703*** (0.205)	0.760*** (0.171)	0.911*** (0.256)
weekend	-0.148 (0.177)	-0.186 (0.212)	-0.504*** (0.140)	-0.112 (0.297)	-0.384 (0.358)	-0.769 (0.470)	0.751*** (0.183)	0.755*** (0.151)	1.032*** (0.244)
LOS	0.079** (0.034)	-0.057 (0.052)	0.038 (0.027)	-0.268*** (0.055)	-0.239*** (0.077)	-0.332*** (0.075)	-0.058** (0.023)	-0.069*** (0.018)	-0.002 (0.023)
leisure	1.594*** (0.297)	2.848*** (0.431)	2.042*** (0.247)	1.176*** (0.440)	4.162** (2.040)	1.408** (0.620)	5.573*** (0.671)	5.076*** (0.593)	6.913*** (1.334)
distance	-2.2e-03*** (3.2e-04)	-1.2e-04 (1.1e-04)	-1.7e-05 (5.9e-05)	1.4e-04 (1.5e-04)	2.8e-04 (1.8e-04)	2.2e-04 (2.0e-04)	8.0e-05 (5.1e-05)	9.9e-05*** (3.7e-05)	1.6e-04*** (6.4e-05)
midscale	-0.515** (0.211)	-0.407 (0.248)	0.068 (0.170)	-0.483 (0.335)	-0.381 (0.437)	0.413 (0.621)	-1.109*** (0.237)	-0.238 (0.196)	0.981** (0.393)
luxury	-0.683*** (0.236)	-0.357 (0.260)	0.262 (0.185)	-0.826** (0.333)	0.288 (0.448)	1.295* (0.669)	-1.826*** (0.244)	-0.559*** (0.196)	-0.195 (0.406)
trend	0.130*** (0.038)	0.333*** (0.040)	0.076*** (0.027)	-1.246*** (0.183)	-0.702*** (0.187)	-1.837*** (0.226)	-0.410*** (0.059)	0.139*** (0.025)	-0.233*** (0.056)
constant	0.788*** (0.402)	-2.860*** (0.572)	-1.251*** (0.389)	15.068*** (2.270)	4.882 (3.221)	14.114*** (2.396)	1.118*** (0.386)	-1.468*** (0.300)	-4.164*** (1.551)
N	17,921								
Log Likelihood	-17681.65								

Table A6.- FMMNL parameter estimates  
Standard errors in parentheses  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1