Distance Traveled in Times of Pandemic:
An Endogenous Switching Regression Approach

David Boto-García
University of Oviedo
botodavid@uniovi.es

Veronica Leoni
Ca' Foscari University of Venice
veronica.leoni@unive.it

Abstract:

This paper studies the change in the distance traveled by domestic tourists considering the pre- and post-pandemic outbreak summer periods of 2019 and 2020. Using representative monthly microdata involving more than 31,000 trips conducted by Spanish residents, we examine the heterogeneity in behavioral adaptation to COVID-19 based on sociodemographic and trip-related characteristics. To account for selection effects and the potential change in the population composition of travelers between the two periods, we estimate an endogenous switching regression that conducts separate regressions for the pre- and post-pandemic periods in a unified econometric framework. Our results point to heterogeneous shifts in the distance traveled by domestic travelers after COVID-19 outbreak per sociodemographic group, with notable differences by travel purpose and lower relevance of traditional determinants like income.

Keywords: *COVID-19; domestic tourism; distance traveled; endogenous switching regression; crisis-resistant tourists*

Cite as:

Boto-García, D. and Leoni, V. (2021). Distance traveled in times of pandemic: a switching regression approach. *Tourism Economics*, forthcoming. https://doi.org/10.1177/13548166211059414

1. INTRODUCTION

Geographic distance plays an important role in human spatial behavior. When it comes to tourism, the distance between the origin and the destination is among the main determinants of tourism flows and travel destination choice. Existing studies point to the so-called 'antinomy of distance' (Cao et al., 2020), by which distance is associated with both frictional and catalytic effects on travel choice. On the one hand, distance acts as a dissuasive factor by increasing travel and time costs. A negative relationship between distance and aggregate tourism demand has been largely documented in studies using gravity models (e.g., Morley et al., 2014; McKercher, 2018). On the other hand, distant locations are usually associated with greater tourism product differentiation relative to the source market, making distant destinations more attractive (Nicolau and Mas, 2006). Despite being a well discussed topic in the tourism literature, the factors that explain the distance traveled are still far from being properly understood.

Tourists' travel habits and destination choice are influenced by macroeconomic shocks as well as by global health emergencies. The dependence of the tourism sector on human mobility makes it highly exposed to international crises. In this regard, COVID-19 disease has heavily hit the travel sector from both supply and demand sides, forcing tourists to adapt their travel patterns to the new circumstances and limitations. Recent evidence points to a shift in preferences towards local trips (e.g., Jin et al., 2020); however, little is known yet about the pandemic-induced effects in destination choice preferences. Since the pandemic is expected to produce long-lasting effects in the tourism industry, analyzing the change in the distance traveled becomes even more relevant, especially if made at the individual level. As opposed to aggregate data, individual-level data allow for a better understanding of the micro-level sources of heterogeneity in behavioral adaptation to COVID-19.

In light of the above, the objective of this paper is to examine the change in the distance traveled between pre- and post-pandemic periods and its underlying determinants. We analyze the effect of COVID-19 pandemic on the distance traveled by domestic tourists, offering empirical evidence on the effects of this macro-shock on travel preferences. Moreover, the paper provides insights on the changes in the population composition of domestic travelers, since health crises might not only affect the distance traveled, but also the sample profile of tourists in the post-crisis state relative to pre-crisis periods. Specifically, we seek to identify i) which tourist segments are more resilient to COVID-19 in terms of not modifying the distance traveled, and ii) how the pandemic has changed the importance of the underlying determinants.

Evidence from previous health crises shows that under infection risks people adopt precautionary behaviors, prioritizing local trips and outdoor activities (e.g., Wen et al., 2005). Our paper aims to examine if this pattern also holds in the case of COVID-19 pandemic. Since behavioral adaptation has been shown to vary with sociodemographic (e.g. Lepp and Gibson, 2003) and tripographic characteristics (e.g. Rittichainuwat and Chakraborty, 2009), we expect the plausible drop in the distance traveled after COVID-19 to be heterogeneous across tourist profiles. This relates to Hajibaba et al. (2015) framework of crisis-resistant tourists (i.e., those who sustain their planned travel patterns under external shocks).

We use detailed monthly microdata for domestic leisure trips from the Spanish Domestic Travel Survey for the summer periods of 2019 and 2020 (July, August and September). The chosen time frames allow for the comparison between pre-pandemic (2019) and post-pandemic outbreak (2020) travel behavior with the advantage that during the summer of 2020 Spanish tourists were free of mobility limitations. From a methodological standpoint, we estimate an endogenous switching regression model that considers separate regressions for the pre- and the post-pandemic outbreak periods in a unified framework, in which selection effects are explicitly considered. In this way, the analysis addresses potential sample selection bias arising from unmeasured heterogeneity affecting tourists' traveling decision in the summer of 2020 and the choice of a domestic (vs abroad) trip. Since we aim to assess potential slope heterogeneity, the proposed modelling approach allows to formally test whether the distance semi-elasticities of socio-demographic and trip-related characteristics are significantly different between the two periods.

The analysis focuses on domestic tourism (i.e., trips by Spanish residents within the country's borders). The reason behind this choice is twofold: i) the existence of inter-country travel restrictions and the contraction of overseas transportation supply, which precludes a proper comparison of travel preferences for international trips before and after the shock; and ii) the key role played by domestic tourism to bail out the travel sector. Arbulù et al. (2021) investigate the capacity of Spanish domestic demand to relieve the tourism industry, showing that the loss caused by the decline in international arrivals can be compensated by the reorientation of Spaniards to domestic travel. As a result, domestic tourism is nowadays regarded as the key factor to resume the tourism industry, at least until international travel can bounce back (Haywood, 2020). Therefore, since in times of heath crises tourists prefer local trips, studying domestic tourism gains more importance than usual.

2. LITERATURE REVIEW

2.1.The antinomy of distance

Given the spatial dimension of the tourism phenomenon, the distance between the area of residence and the possible destinations is a key factor for leisure travel choice. Although this topic has received substantial attention, there is no consensus about tourists' preference for nearby versus distant destinations. On the one hand, traveling to faraway destinations entails higher expenses and opportunity costs, which makes distance a dissuasive factor. In this regard, a large body of research has documented a pattern of distance decay by which tourism demand decreases as distance to the origin increases (Morley et al., 2014; Chandra et al., 2014; McKercher, 2018; Alvarez-Díaz et al., 2020).

On the other hand, tourists might be willing to travel far away if in exchange they discover new places, cultures and environments. This closely relates to Plog's travel personality model (Plog, 2001), by which allocentric (venturer) people are more open-minded and willing to travel abroad to live memorable experiences. Since traveling is also a way to escape from everyday life that has important restorative benefits (Packer, 2021), long-haul traveling becomes attractive if the utility obtained from discovering new places is greater than the extra costs it entails.

These two opposite effects, inhibition versus attraction, constitute the so-called 'antinomy of distance' (see Cao et al., 2020). In a recent paper, Wong et al. (2020) attempt to bridge a gap between these two streams of research by proposing a dual distance model where the economic distance moderates the relationship between geographic distance and tourism demand. According to their results, neither geographical nor economic distance can alone explain such a complex phenomenon, which, instead, results from an interaction of the two dimensions.

2.2.The drivers of the distance traveled

Several studies have analyzed the potential drivers of tourists' heterogeneous preferences regarding distance. Nicolau (2008) examines how tourists' sensitivity to distance can be explained by sociodemographic characteristics. The author shows that the disutility of distance decreases with income, organizing the trip through intermediaries and the size of the city of residence. Conversely, nearby destinations are preferred when traveling with children. Nicolau (2010) explores how the effect of distance on destination choice depends on tourist's motivations. His results show that *variety-seeking behavior* decreases the disutility of distance whereas *inertial behavior* increases it. Similarly, Nicolau and Más (2006) document that motivations like the search for a better climate or discovering new places make people willing to cover longer distances. Wynen (2013) studies the effects of several sociodemographic and trip-related characteristics on the distance traveled during same-day trips. The author shows that the distance traveled exhibits an inverted U-shaped relationship with age and is positively related with education and the hours spent at the destination. More recently, Boto-García et al. (2021) investigate heterogeneous preferences for destination hedonic attributes in the context of nature-based tourism in Spain. They show that the disutility of distance is moderated by income and the willingness to perform aquatic or adventure activities but reinforced by travel party size. Overall, these studies document a high degree of heterogeneity in the taste for distance that depends on trip purposes and sociodemographics.

2.3.The change in travel preferences after macroeconomic shocks and health crises

Another stream of research has examined temporal changes in travel habits and the distance traveled following economic and health crises. Using longitudinal data, Wong et al. (2017) analyze how travel frequencies among in-state, out-of-state and international destinations are affected by the unemployment rate. They report that under economic downturns more people prefer traveling less frequently and, if so, prioritize local trips. Cafiso et al. (2018) study how the 2008 economic crisis affected domestic tourism flows in Italy. They provide evidence that the dissuasive effect of distance becomes greater during the 2008-2012 period, suggesting that tourists choose closer destinations in times of crisis. Similar findings are reported by Eugenio-Martin and Campos-Soria (2014). Wong et al. (2016) document that tourists' outbound travel behavior is not only affected by the business cycle but also by the tourist's life cycle so that preferences for distant destinations are time-variant. In a recent paper, Weigert et al. (2021) exploit repeated cross-sectional survey data for German tourists for the period 1971-2018 and show that the distance traveled mainly depend on the life cycle and macro-level factors, with cohort differences being less pronounced.

Health risks represent a major threat for tourist, with profound implications on their behavior and expenditure patterns. Existing evidence from previous health crises shows that greater perceived risks generally translate into either travel avoidance (e.g., Cahyanto et al., 2016) or protective behavior at the destination (e.g., Wen et al., 2005), leading to several behavioral changes (Li et al., 2020). The recent outbreak of COVID-19 has heavily hit the tourism and travel sectors. Lockdown policies and stay-at-home orders, the mandatory closure of tourism establishments, and the drop in transportation supply have limited national and international mobility. Furthermore, personal fears and uncertainties related to the pandemic together with perceived severity have reduced people's willingness to travel (Neuburger and Egger, 2021; Das and Tiwari, 2021), leading to an overall decrease in tourism participation. Preliminary evidence about COVID-19 suggests that destination and accommodation risks would deter holiday intentions, especially among elderly people (Pappas, 2021). In general, travel fears induce people to engage into more cautious and self-protective travel behaviors (Zhang et al., 2021).

However, the literature has documented important heterogeneity in travelers' behavioral responses and perceived behavioral control under health risks based on both sociodemographic and psychological factors (Karl et al., 2020). In this sense, some tourist segments have traditionally exhibited great resilience against external shocks and crises (i.e., crisis-resistant tourists), moving ahead with their travel plans even if unexpected events occur (Hajibaba et al., 2015). For instance, Kement et al. (2020) report that the travel intentions of Turkish tourists are not affected by the outbreak of COVID-19. Similarly, Boto-García and Leoni (2021) find that people living in areas that were more affected by the disease during the first wave exhibit a positive attitude towards traveling in the summer of 2020. As several studies have shown, tourism can generate substantial and widespread psychological benefits (e.g., Packer, 2021); in particular, outdoor and nature-based activities have important restorative benefits for mental health after strain (Buckley and Westamay, 2020).

An emerging body of literature documents a reshape in tourists' preferences towards local tours and proximity areas. In this regard, Mackenzie and Goodnow (2021) discuss how COVID-19 has forced many to embrace 'locavism' practices (i.e., short distance local trips). This also happened among Chinese tourists after SARS (Wen et al., 2005). Jin et al. (2020) report that Chinese tourists would resume travel domestically and plan to start traveling to nearby regions. Using network analysis, Jeon and Yang (2021) examine tourists' travel routes during the same period before and after the outbreak of COVID-19. They find that the network connectivity between regions weakened and that movement to coastal regions increased to a greater extent than to inland regions. Relatedly, Nguyen et al. (2020) show economic uncertainty reduces outbound tourism but boosts domestic trips.

Tourists traveling after COVID-19 outbreak appear to be unwilling to travel to crowded destinations and prioritize free and independent trips (Wen et al., 2021). Ivanova et al. (2021) show that, for their first trip after COVID-19 outbreak, respondents will travel within the country, by car and with their family. Craig (2021) analyzes traveler camping and glamping decisions in the U.S. after COVID-19, finding that crowding avoidance is the main predictor of tent camping engagement. Osti and Nava (2020) show a shift in destination preferences and loyalty from seaside to mountain destinations, especially among risk-sensitive tourists. The latter are perceived as safer because they offer a "natural" distancing landscape.

Since COVID-19 and the travel risks and restrictions that surround the pandemic are predicted to have affected destination choice decisions heterogeneously, and in line with the literature on the effect of macro-shocks on travel behavior, the current work enriches the debate on the crisis-induced changes on the distance traveled by investigating the sociodemographic and trip-related factors that make tourists more resilient. Specifically, we evaluate how the sociodemographic profile, travel purpose and the characteristics of the place of residence affect i) the probability of engaging into domestic tourism, ii) the likelihood of traveling in the summer of 2020 (relative to 2019) and iii) the distance traveled before and after the pandemic shock. The comparison of pre- and post-pandemic outbreak summer periods therefore allows for a characterization of crisis-resistant travelers.

3. DATA AND METHODS

3.1.Database

We use survey microdata from the Spanish Domestic Travel Survey (ETR/Familitur), conducted on a monthly basis by the National Statistics Institute to a representative sample of the Spanish population. This survey is a subsample of the Continuous Household Survey, which is obtained by multistage sampling, stratified by clusters with proportional section of primary (cities) and secondary units (census sections). Each month, around 8,000 individuals are sampled and surveyed by telephone about their recent trips, for any kind of purpose. For the current study, we only consider trips for purposes other than business (including leisure & entertainment, visiting friends or relatives (hereafter VFR), health treatments, shopping, gastronomy and religious pilgrimage) taken during the third quarter of the years 2019 and 2020 (July, August and September). We choose this time span for two reasons. First, the summer period is when most people travel for leisure reasons (INE, 2020). Second, after the severe lockdown between middle March and middle June 2020, it was only during the summer period when Spanish residents were allowed to travel with no imposed movement restrictions. By the beginning of October, several Spanish regions dictated some local confinements due to the worsening of their epidemiological situation (RD 925/2020, of October 25). Therefore, we consider the period of 2020 straddling the first two COVID-19 waves.

Among other trip-related information, respondents are asked about the (main) destination of their trips. There are three possibilities: 1) international (outside the country), 2) domestic outside the region of residence, and 3) domestic within the region of residence. Table 1 presents descriptive statistics of the number of individuals in the sample, disaggregated by month and destination choice, separately for before (2019) and during the pandemic (2020). Heat maps with the share of international and domestic trips per autonomous community of residence are presented in Appendix A.

TABLE 1 HERE

According to Table 1, there has been a substantial drop in the share of international trips in the summer of 2020 compared to the same period in 2019 (11% versus 3%). Although the total number of trips is lower than in the pre-pandemic period (18,393 versus 13,835), there has been a redistribution of flows in favour of domestic trips, with a notable increase in the share of trips that take place within the region of residence (51% versus 45%). This tentatively suggests that, in the first summer after the pandemic outbreak, Spanish residents have reduced the distance traveled and opted for more local trips.

Compared to the year 2019, the drop in the number of trips abroad in the summer of 2020 can be explained by both demand and supply reasons. On the one hand, international trips might entail greater risks than domestic destinations because the former normally requires traveling by mass transit, in which it could be more difficult to keep the necessary social distancing. Furthermore, amid COVID-19 pandemic, traveling abroad might require hiring a specific health insurance, which supposes additional costs and might thus constitute an important barrier. On the other hand, mobility limitations have reduced the supply of international flights, thereby exposing customers to higher fares. Overall, the greater difficulties faced to organize international trips might have fuelled a higher participation in domestic trips after the pandemic outbreak. Due to the low number of outbound trips in 2020, the analysis that follows only concentrates on the distance traveled within the country borders.

3.2.Dependent variable: distance traveled

To examine preferences for nearby versus distant destinations, we first compute the Euclidean distance (in kilometers) between each respondent's place of residence and the chosen destination. In the survey, respondents indicate the regional area where they stay at the NUTS 3 regional disaggregation level (Spanish provinces, equal to 50). However, for confidentiality reasons, the information on their place of residence (origin) is only provided at the NUTS 2 level (Spanish autonomous communities, equal to 17). This hinders an exact calculation of distances because if we computed the distance between NUTS2 regional centroids, that would set to zero the distances for all intra-regional trips. In this respect, whereas seven autonomous communities consist of a single province, others like Castilla and Leon or Andalusia involve nine and eight provinces each. Given this heterogeneity, doing so would substantially reduce the variability of the distance variable and ignore the distance covered to travel to other provinces within the same autonomous community.

To address this issue, for each tourist in the sample, we calculate a Euclidean weighted measure of distance that considers tourists' place of origin probabilistically. Specifically, we first compute the distance between the centroids of all Spanish provinces (NUTS 3). This is given by the lower triangular of a 50x50 matrix. Next, we compute bilateral distances between each province (*p*) and each autonomous community (*c*) in the following way:

$$d_{c,p} = \sum_{p \in c} \frac{pop_p}{pop_c} * d_{p,p'} \qquad \text{for } p = 1, \dots, 50 \text{ and } c = 1, \dots, 17 \tag{1}$$

where $d_{c,p}$ is the distance between each province destination p and each autonomous community c, $d_{p,p'}$ is the distance between the province destination and each potential province of origin and pop_p and pop_c are the population in each province and autonomous community of origin. Accordingly, distances between origin and destinations consider the likelihood of the individuals living in each province based on population weights. In other words, since we know the province of destination but only the autonomous community of origin, we consider the probability that an individual traveling from a given autonomous community lives in each of the provinces that integrate it. The resulting weighted distance ($d_{c,p}$) is labelled *distance* and is similar to the one implemented in Chandra et al. (2014) for modelling cross-border traveling.

To better illustrate how the distance measure in (1) is computed, let us provide the following example. Suppose a tourist who travels from Catalonia (CAT, NUTS 3) to Zamora (ZAM, NUTS 2). In that case, we would have:

$$d_{CAT,ZAM} = \frac{pop_{LLEIDA}}{pop_{CAT}} d_{LLEIDA,ZAM} + \frac{pop_{GIR}}{pop_{CAT}} d_{GIR,ZAM} + \frac{pop_{BARC}}{pop_{CAT}} d_{BARC,ZAM} + \frac{pop_{TARR}}{pop_{AND}} d_{TARR,ZAM} \tag{2}$$

Figure 1 plots a histogram of the calculated distance traveled by Spanish residents within the country before (2019) and after the pandemic outbreak (2020). As can be seen, the distributions are fairly similar, although the one for the summer of 2020 seems to be slightly shifted to the left. Table 2 presents the mean distances before and after the pandemic outbreak by autonomous community of residence together with the number of observations in each case. The last column reports a t-test for mean equality in the two periods. Figures 2 and 3 map these mean values for the summer periods of 2019 and 2020, respectively. As shown in the first column of Table 2 and in Figure 2, people from the central areas of Spain (particularly Madrid and Castilla-laMancha) travel the greatest distances in the summer of 2019, while those in Mediterranean regions seem to prefer traveling within their autonomous communities or to nearby locations. Residents at the Canary Islands also travel large distances, but this could be attributable to their geographic position. Focusing on the summer of 2020, we document an overall decrease in the mean distance traveled in all the regions. Nevertheless, these results are merely descriptive and only provide an overall picture. Because the mean values comparison might be affected by differences in the sample composition across regions, in the following subsections we develop a formal econometric analysis.

FIGURE 1 HERE
TABLE 2 HERE
FIGURE 2 HERE
FIGURE 3 HERE

3.3.Econometric modelling

We initially propose the following canonical regression approach to explore the differences in the distance traveled before and after the pandemic outbreak, controlling for the influence of other factors:

$$\ln distance_i = \beta X_i + \theta R_j + \delta T_t + \tau y2020_i + \epsilon_i \tag{3}$$

where $\ln distance_{it}$ is the log of distance (in kilometers) between tourists' origin and the chosen destination, i denotes individuals in the sample, j stands for the region of origin and t refers to the time period; X_i is a set of sociodemographic and tripographic characteristics; R_j are region of origin fixed effects[1]; T_t are month fixed effects capturing seasonal variability in travel preferences within the summer; $y2020_i$ is a dummy variable that takes value 1 for the summer of 2020 and 0 otherwise, β, θ and δ are vectors of parameters to be estimated, τ is the coefficient for *y2020* and ϵ_{it} is a normally distributed error term.

Equation (3) can be easily estimated by Ordinary Least Squares. In principle, the estimate of τ would indicate the difference in traveled distance before and after the pandemic outbreak, everything else being equal. However, it is highly likely that the pandemic has not only produced a shift in the mean distance but also a change in the influence of personal characteristics on the distance travelled (i.e., the slope of the explanatory variables in (3)). One way to explore this would be by running separate regressions for 2019 and 2020 (i.e., *y2020*=0 and *y2020*=1) and compare the coefficient estimates. However, since we work with a sample of actual travelers that have decided to travel domestically, the results from separate regressions are likely to be affected by a composition effect stemming from the different characteristics of the subsamples. First, the population composition of travelers might be different in the summer of 2020 relative to the pre-pandemic periods, as shown by some studies (Ivanova et al., 2021; Jin et al., 2020). Second, among those who travel, domestic (versus international) travelers are likely to be self-selected based on both observable and unobservable factors (Eugenio-Martín and Campos-Soria, 2010). As a result, the data might suffer from sample selection bias because unmeasured heterogeneity affecting observability during the pandemic period correlates with

[1] The inclusion of region of origin fixed effects controls for heterogeneity in transport infrastructure or climate conditions, among others, that might affect travel preferences (Eugenio-Martín and Campos-Soria, 2010).

unobserved factors driving the destination choice (and therefore the distance covered). As such, we face a double selection problem that might cause endogeneity and inconsistent estimates.[2]

A more suitable modelling approach is the following switching regression:

$$\begin{cases} \ln distance_{0i} = X_i{}'\beta_0 + \theta_0 R_{0j} + \delta_0 T_{0t} + \kappa_0 \lambda_{0i} + \epsilon_{0i} & if\ y2020_i = 0 \\ \ln distance_{1i} = X_i{}'\beta_1 + \theta_1 R_{1j} + \delta_1 T_{1t} + \kappa_1 \lambda_{1i} + \epsilon_{1i} & if\ y2020_i = 1 \end{cases} \quad (4)$$

where $\lambda_{it}=\frac{\phi(W_{it}{}'\hat{\gamma})}{\Phi(W_{it}{}'\hat{\gamma})}$ is the inverse Mills ratio (IMR) that controls for domestic traveling selection bias (Heckman, 1979), being $\phi(.)$ and $\Phi(.)$ the probability density and cumulative distribution functions of the standard normal distribution. The IMR is derived based on the estimates from the following auxiliary probit regression:

$$domestic_i{}^* = W_i{}'\gamma + \xi_i$$
$$\begin{cases} domestic_i = 0\ if \quad domestic^*{}_i \leq 0 \\ domestic_i = 1\ if \quad domestic^*{}_i > 0 \end{cases} \quad (5)$$

where $domestic_i{}^*$ is a latent variable for the utility of traveling domestically (vs internationally), W_i are a set of explanatory variables including regional and month fixed effects, γ is a vector of parameters to be estimated and ξ_i is a normally distributed error term.

The model in (4) estimates different slopes depending on the value of the switching regime (here *y2020*) once accounting for selection, with the additional advantage that the two regressions are jointly estimated and the equality of the slopes can be formally tested.[3] In other words, we allow the effect of the explanatory variables to have shifted after the pandemic outbreak and therefore to differ depending on the observation period. To consider the differences in the likelihoods of observability (tourism participation) between the two periods, the switching regime (i.e., *y2020*) is further treated as endogenous and modelled as follows:

$$y2020^*{}_i = \alpha + \pi Z_i + \vartheta R_j + \mu T_t + \eta_i$$

$$\begin{cases} y2020_i = 0\ if \quad y2020^*{}_i \leq 0 \\ y2020_i = 1\ if \quad y2020^*{}_i > 0 \end{cases} \quad (6)$$

where α is a constant term, Z_i is a set of explanatory variables explaining being observed in the summer of 2020, R_j and T_t are regional and month fixed effects, π, γ and μ are parameters to be estimated and η_i is a random error term.

[2] Please note we tackle selection effects *within* the population of leisure tourists. The decision to travel and sorting into a leisure trip over other purposes is not modeled. Therefore, we draw inference only about the population of leisure travelers.

[3] The reader is referred to Hotchkiss and Pitts (2005) and Di Falco et al. (2011) for technical details of the model formulation and empirical applications in different settings.

In this way, the proposed endogenous switching linear regression approach estimates different slopes for pre- and post-pandemic domestic travelers, conditional on the endogenous observation period (regime) and controlling for domestic traveling selection bias. To tackle correlation in observables, the error terms in (4) and (6) are allowed to be correlated. The model is jointly estimated using a Full Information Maximum Likelihood estimator, which is more efficient than two-stage procedures.

3.4. Descriptive statistics, model specification and exclusion restrictions

Detailed sociodemographic data including gender, age, education level, household income or nationality is also collected in the survey. Together with the region of residence, information on the population size and density of the municipality of residence is also gathered. Table 3 reports summary statistics of the variables available in the survey along with their definition.

TABLE 3 HERE

The following variables are considered in the regressions: gender (a dummy for being a female), age (in years), education level (a dummy for university education), household income (in intervals), nationality (a dummy for being foreigner), labour status (employed and unemployed, with inactive or retired collapsed in the reference category), household size (number of people living in the household), travel companions (traveling alone or in a couple, with the rest of options collapsed in the reference category), population size and density of the municipality of residence (in intervals), travel purpose (sun & beach, cultural, nature, sports and VFR, with the rest of motivations gathered in the omitted category), region of residence and monthly fixed effects.[4]

The appropriate identification of the model proposed in (4-6) requires both the selection equation in (5) and the endogenous switching equation in (6) to include distinct exclusion restrictions. That is, there must be at least one variable entering in equations (5) and (6) that is excluded from (4).[5] Finding suitable exclusion restrictions is not easy in this context, since they must be correlated with domestic traveling likelihood and observability in 2020 relative to 2019 but uncorrelated with the distance traveled.

For the selection equation in (5), we use the monthly Consumer Confidence Index (henceforth CCI) per region. This index captures citizens' optimism and perceptions about the current economic situation and their expectations for the incoming six months. We assume individuals living in areas with relatively greater optimism are more likely to travel internationally. The

[4] The inclusion of travel purposes as a control follows Nicolau (2008; 2010), Nicolau and Más (2006) and Boto-García et al. (2021), who document that trip motivations exert a moderating effect on the willingness to cover long distances.

[5] Although identification could be achieved without exclusion restrictions through the non-linearity of the model, having at least one valid exclusion restriction provides an additional source of variability to the inverse Mills ratio and the endogenous switching regime that alleviates collinearity problems in the former case (Puhani, 2000), and improves the estimator precision of the correlation between the error terms in the latter (Monfardini and Radice, 2008).

CCI index has been widely applied in the economic literature as a predictor of household consumption (Allenby et al., 1996; Gausden and Hasan, 2018). We compute it using survey microdata from the Consumer Confidence Barometer, conducted on a monthly basis by the Spanish Institute for Sociological Research (CIS in Spanish) to a representative sample of the Spanish population. Appendix B provides details about its construction and summary statistics. This variable varies across regions and months. Appendix B Figure A1 plots the distribution of regional CCI per period. To inspect whether it is a valid exclusion restriction, Appendix B Figure A2 presents a scatterplot of the distance traveled against the CCI. As can be seen, there is no association between the two. An auxiliary OLS regression as in (3) including CCI corroborates CCI is conditionally uncorrelated with the distance travelled (t-statistic=0.56, p-value=0.582). Therefore, it can be considered as a valid exclusion restriction.

For the switching regime likelihood in equation (6), we select the regional unemployment rate at the region of residence per period. Here we assume the likelihood of traveling in the summer of 2020 relative to 2019 could be affected by the situation of the labor market at the place of residence through precautionary savings motive (Malley and Moutos, 1996; Papatheodorou and Pappas, 2017). Since its effect might differ depending on the respondents' labor status, we also consider an interaction term between unemployment rate and the indicator for being unemployed. Rather than using the generic regional unemployment rate, we impute to each respondent the corresponding one to his/her age range (less than 25 years old, between 25 and 55 years old, and more than 55 years old), gender and region of residence. In this way, each individual is linked with the relevant unemployment rate, with the additional advantage that this increases its variability. Unemployment rate varies from 4.8% (males with more than 55 years old in Cantabria in 2019Q3) to 67.2% (males under 25 in the Canary Islands in 2020Q3), with a mean value equal to 19.8% and a standard deviation equal to 12.9. Figures A3 and A4 in Appendix C plot the unemployment rates per region and age ranges for the third quarters of 2019 and 2020. As shown there, there is large variability in unemployment rates across regions and age ranges. To be a valid exclusion restriction, this variable must be uncorrelated with the distance traveled. Figure A5 in Appendix C presents a scatterplot of distance against unemployment rate. As can be seen, there is no relationship between the distance traveled and the associated unemployment rate, thereby working as a valid exclusion restriction.[6] An auxiliary OLS regression as in (3) including the unemployment rate further corroborates this; the unemployment rate is conditionally uncorrelated with the distance travelled (t-statistic=1.35, p-value=0.197).

[6] We tested the possibility of including unemployment rate as a second exclusion restriction in the selection equation in (5), since authors like Wong et al. (2017) used this situational factor as explanatory of the destination choice. However, it resulted non-significant, thereby not satisfying the first requirement for being a valid instrument. Similarly, we also examined whether CCI could be used as a second exclusion restriction in the switching regime in (6). This is not possible because CCI perfectly predicts the regime status. This produces a problem of quasi-complete separation that precludes identification.

4. RESULTS

4.1.Main findings

Columns 1 and 2 in Table 4 present the coefficient estimates of the endogenous switching regression depending on the regime (pre- and post-pandemic, respectively). These estimates can be interpreted as semi-elasticities (i.e., percentage increase in the distance traveled if there is a marginal increase in the explanatory variable).[7] In column 3, we report chi-squared tests of parameter equality, which inform about whether the semi-elasticities are statistically different between the two periods. Finally, columns 4-5 and 6-7 present the coefficient estimates and average marginal effects (AME) for the endogenous switching and the selection equations, respectively. Due to space constraints, the regional fixed effects are omitted but are available from the authors upon request. To avoid aggregation bias due to the inclusion of covariates defined at the regional level (Moulton, 1990), all the standard errors are clustered at the autonomous community of residence level.

TABLE 4 HERE

To facilitate the interpretation of the results, we provide a horizontal reading of Table 4 for each variable. Whereas there are no gender differences in the distance traveled neither before nor after the pandemic outbreak, females are less likely to be observed traveling in the summer of 2020 (-9%). This relates to some studies that document females are more risk averse against health crises (Karl et al., 2020; Neuburger and Egger, 2021). The distance traveled is negatively associated with age, being the effect about the same in the two periods (around -1% per year, on average). This result is consistent with evidence presented in Wynen (2013). Interestingly, elderly people are also more prone to domestic travel (+0.1%) and less likely to be observed traveling in the summer of 2020 (-0.5%). The latter result corroborates findings by Das and Tiwari (2021) and Pappas (2021). Regarding educational background, we find that respondents with university education i) are more likely to travel in the summer of 2020 (+6.4%) and ii) prefer to travel abroad (+0.9%). Nonetheless, there are no differences in the distance traveled by educational level.

Moving to the role of income, the distance traveled is positively related to income in both the pre-pandemic and the post-pandemic outbreak summer periods. As shown by the related literature (Van Nostrand et al., 2013; Nicolau, 2010), high-income people are less deterred by distance due to facing lower budget constraints. Strikingly, in the post-pandemic summer the effect of income differences is quantitatively lower: the semi-elasticity of distance with respect of the medium-high income category is 21.0% in 2020 whereas it is 36.8% in 2019 (relative to the low-income category). Moreover, there are no differences in the distance traveled in 2020 for the high-income segment. We also document that wealthier people exhibit a lower likelihood of traveling domestically (-2.2%) but are relatively more likely to travel after the

[7] For the binary indicators, the semi-elasticities are computed as follows (Halvorsen and Palmquist, 1980): $\frac{\partial \ln Distance_{ijt}}{\Delta D} = \exp(\beta) - 1$

pandemic outbreak (+3.3%). This pattern is consistent with evidence by Boto-García (2020) showing a larger persistence in tourism participation among high-income people.

Spanish residents with a foreign nationality used to travel more distance before the pandemic outbreak (at 90% confidence level); however, in the summer of 2020 there are no differences based on nationality. This makes the effect to be significantly different between the two periods. The underlying explanation could be that foreigners are less likely to travel domestically (-14.6%) and more likely to travel in the summer of 2020 (+7.0%). As a result, the combination of the two effects might make no relevant differences to be detected. As for the labor status and relative to inactive and retired respondents (excluded category), employed people are less likely to travel within the country (-1.6%) but more likely to be observed in 2020 (+14.7%). These individuals also travel farther away; nonetheless, the distance gap has significantly decreased in 2020 relative to 2019 (+30.7% vs. 61.7%). Individuals with children are more likely to be represented in 2020 (+2.4%) but they travel significantly less distant then than before (-14.6%). By contrast, households with more members are less probable to travel in 2020 (-1.9%), with no differences in the distance traveled based on household size in the post-pandemic outbreak summer.

Concerning the influence of the population density of the municipality of residence, respondents from low-density areas are significantly less likely to travel in 2020 (-6.2%) relative to those in medium-density municipalities, with no significant differences in the distance traveled. By contrast, those living in highly populated areas are more deterred to cover long distances, both in 2019 (-18.5%) and in 2020 (-22.8%). Interestingly, people from small municipalities are relatively more likely to be observed traveling in 2020 (+4.3%).

As for the travel purpose, tourists seeking coastal destinations travel significantly farther away in 2020 (+138.7%), which matches the results by Jeon and Yang (2021). Recall here that the reference group are trips for shopping, gastronomy, well-being and religious peregrination, which in general involve shorter distances. With regards to cultural tourists, this segment is less likely to travel domestically (-26.7%), less likely to be observed traveling in 2020 (-11.0%) and, if so, they travel longer distances in 2020 than the reference category (+224.1%) but less than in 2019. This finding is in line with Nicolau (2008; 2010) and Nicolau and Mas (2006), who report that tourists who aim to discover new places are significantly more willing to travel farther away. Therefore, domestic cultural tourists seem to be who travel to more distant locations. Those seeking nature-based activities tend to cover longer distances in 2020 (+53.6%). Interestingly, this trip purpose is negatively associated with the likelihood of traveling domestically (-4.3%) but is more prevalent in 2020 (+7.4%). This is consistent with several studies documenting an increase in the taste for outdoor activities after the lockdown period (Osti and Nava, 2020). The practice of sports is less associated with domestic traveling (-3.9%), whereas VFR has increased its share over total trips in 2020 (+3.1%). Interestingly, although the distance traveled associated with these two latter travel motivations seems to have increased in 2020 relative to 2019, we cannot reject the null hypothesis that the distance semi-elasticities are equal in the two periods.

We also find that individuals traveling alone cover shorter distances before and after the pandemic outbreak (-35.4%). Interestingly, solo trips are more prevalent in 2020 (+4.1%). Weekend trips are positively associated with domestic traveling (+8.0%) and are less likely in the post-pandemic outbreak period (-4.4%). Weekend tourists travel shorter distances (-78.3% and -68.1%, respectively) both before and after the pandemic began, which is consistent with these individuals facing tighter time constraints.

The estimation results indicate that the regional unemployment rate for the tourist's reference group is positively associated with observability in 2020 in general, but it exerts a negative effect for unemployed individuals. The later finding makes sense: a worsening in labor market prospects deters tourism traveling in 2020 relative to the pre-pandemic period for unemployed individuals. However, the positive effect for the whole sample is less self-apparent. In principle, we would expect a negative relationship too, since unemployment in the reference group should also decrease observability in the summer of 2020 through precautionary saving motives (e.g. Malley and Moutos, 1996). Nonetheless, whether households change their consumption patterns under job-loss risks is still inconclusive (Carroll et al., 2003). Beyond this, there are two channels that explain this result. On the one hand, the fear of unemployment has been shown to deteriorate people's well-being (Bunnings et al., 2017). Some studies document that taking a tourist trip constitutes a way to reduce stress and burnout (Etzion, 2003) that facilitates recovery from mental strain (Chen et al., 2016). As such, those more exposed to lose their jobs could be in more need to take a leisure trip. On the other hand, after the pandemic outbreak, many workers in Spain temporarily lost their jobs due to the implementation of social distancing measures that forbade them to carry out their activities. Due to the temporal character of the shock, many employers use temporary dismissals (*Expediente de Regulación Temporal de Empleo* in Spanish, ERTE) rather than layoffs; many people kept their jobs (and a share of their salaries, partially subsidized by Spanish government) even though they were virtually 'unemployed'. This safeguard could have made the precautionary motive less important than before, thereby making tourists in 2020 less sensitive to unemployment loss risk than in 2019.

Finally, we find the CCI index is negatively associated with domestic traveling. This indicates that when economic prospects and expectations for the near future worsen, people prefer to travel within the country borders, in line with Smeral (2009). Strikingly, the selection effects from the IMR have opposite effects in the 2019 and 2020 subsamples, being not significant in both cases. This suggests that there is no selection bias in domestic tourism based on unobserved factors so that the set of controls does a good job in capturing all the sources of sample heterogeneity. We also find that the correlation between the error terms of the regime and the outcome equations is positive but not statistically different from zero; this might be attributable to the computation of clustered standard errors. While heteroskedasticity-robust standard errors render a significant correlation parameter (available upon request), the clustering adjustment partially embeds shared unobservables between the chance of being observed traveling in the summer of 2020 and the distance traveled, thereby making the remaining correlation not different from zero.[8] A similar result is also found in other studies adopting endogenous switching regression modelling (e.g., Di Falco et al., 2011).

The absence of shared unobservables could lead us to conclude that the switching regime modelling is not needed and that a more parsimonious OLS regression is valid in this case. However, note that the results from the outcome equation point to structural differences in the magnitude of the coefficient estimates between regimes. A pooled regression would ignore such slope heterogeneity, as discussed in Di Falco et al. (2011). Moreover, even if the switching regime could be taken as exogenous, the regime equation informs us about sample compositional effects that need to be considered when evaluating the differences in the role of sociodemographic and trip-related characteristics on the distance traveled.

TABLE 5 HERE

4.2. Robustness checks

We have performed a battery of sensitivity checks to our findings. First, because the dependent variable construction could be affected by measurement error, we repeated the analysis by grouping the computed distance traveled into ten discrete intervals based on the distribution deciles. In this way, this auxiliary interval regression acknowledges the imprecision in the specific kilometers covered by the individual. The parameter estimates from an interval regression remain largely unchanged (Appendix D, Table A3), which indicates our main findings are robust to measurement error.

Second, we re-estimated the model excluding travelers whose origin is the Balearic or the Canary Islands. The reason is that residents in these islands benefit from airfare subsides imposed through Public Service Obligations to airlines in the form of reduced rates and increased frequencies (Fageda et al., 2017). Nevertheless, the results are consistent with the main analysis (see Appendix D, Table A4).

Third, most of the trips in the sample are made by car (84% in 2019 and 90% in 2020). We do not include the chosen mode of transport as a regressor because we assume that transport mode choice is made once chosen the destination (distance to travel). Nevertheless, to inspect potential differences between 2019 and 2020 associated with the mode of transport, we conducted separate regressions for those who traveled by car (87%) and those who traveled by public modes of transport (4.3% by plane, 3.7% by bus, 3.3% by train and 1.3% by ship or ferry). The coefficient estimates are presented in Appendix D, Table A5. The estimation results for the two groups are very similar to the ones reported in Table 4, with the exception that the correlation between the error terms is large in magnitude and statistically significant for those traveling in public modes of transport.

[8] By clustering at the autonomous community of residence to deal with Moulton bias and the clustered sampling of the dataset (Abadie et al., 2017), the residuals from individuals living in the same region are not taken as independent but potentially correlated. As a result, it appears that common unobserved factors that affect traveling observability in 2020 (relative to 2019) and the distance traveled (i.e. endogenous switching) mainly emanate at the regional level.

Finally, to illustrate the importance of considering selection effects and changes in the sample composition between the pre- and post-pandemic periods, Table A6 in Appendix D presents the estimation results from a naïve OLS regression model as described in (3). Comparing the magnitude and significance of the estimates with that in Table 4, we see that the implications derived from this simple regression could be misleading.

5. CONCLUSIONS

5.1.Summary of findings

This article analyzes the change in the distance traveled by Spanish residents before and after the outbreak of COVID-19. Using representative microdata for domestic trips, we assess the change in the factors that explain the distance traveled during the summer of 2020 relative to the pre-pandemic period (summer of 2019). We find that traveling long distances is highly associated with income, although the income semi-elasticity has decreased in the summer of 2020 relative to the pre-pandemic period. This suggests that household budget constraints have become less relevant for explaining how far to travel. In addition, the pre-pandemic existing differences between foreign and Spanish people and household size seem to have disappeared. By contrast, we find significant differences in the distance traveled by travel purpose after the pandemic outbreak, with tourists seeking sun & beach and nature traveling to more distant locations than before. Similarly, tourists seem to be willing to cover longer distances when traveling for sports or VFR purposes. Furthermore, people with children and cultural tourists seem to be relatively more averse to long-distance travel than before, whereas the effects of age and solo traveling remain unchanged. Overall, our results suggest heterogeneous responses against COVID-19 in domestic traveling patterns by sociodemographic group, with reduced importance of traditional determinants like income or household size.

5.2.Contribution and implications

The current work contributes to the tourism literature in two different ways. First, it enriches the existing evidence on the distance traveled by examining the microlevel sources of preference heterogeneity. Given the inconclusive evidence about whether distance is a deterrent or an appealing factor, our results provide new evidence on the differences in preferences for distant versus nearby destinations by sociodemographic group and travel motivations. In this regard, our findings extend previous works by Nicolau (2008; 2010), Nicolau and Más (2006) and Boto-García et al. (2021). Second, the study adds to an emerging body of research on the effects of COVID-19 on travel habits and behavioural adaptation. We illustrate how the pandemic has changed the influence of some personal characteristics on the distance traveled.

From a methodological viewpoint, we implement a switching regression model that considers i) slope heterogeneity, ii) selection effects stemming from unobserved heterogeneity affecting both domestic traveling and the distance traveled, and iii) changes in the population composition of travelers before and after the pandemic. As illustrated by comparing our model with a naïve OLS regression, our results highlight the importance of rigorous econometric analysis that try to capture all the confounding elements that might influence the outcomes of interest to avoid biased estimates. In this respect, the study of the pandemic-induced impacts on the tourism industry is challenging because there are many factors operating at the same time. Whenever we aim to examine the effects of COVID-19 on a given outcome, we must consider all other potential changes in the population composition of leisure tourists. This calls for avoiding naïve mean comparisons that ignore composition effects, especially if coupled with reduced sample sizes that are not representative for inference. Indeed, our analysis illustrates how some sociodemographic and trip-related characteristics do not only affect the distance traveled but also the probability of being observed traveling in 2020 and doing it domestically.

Incoming studies on the pandemic-induced changes in tourism and travel patterns could benefit from applying the switching regression approach. This methodology allows the researcher to consider distinct data generation processes driving outcomes in pre- versus during/post-pandemic scenarios together with correlated unobservables affecting regime observability and the outcome of interest. However, practitioners must bear in mind it requires exclusion restrictions for appropriate identification. The validity of the selected instruments needs to be proved to offer reliable results. In this setting, simple falsification tests (checks of univariate and conditional correlations between the instruments and the outcomes) as the ones implemented in the paper and advocated by Di Falco et al. (2011) are recommended.

Understanding the impact of the pandemic on tourists' travel patterns is not only relevant for academic reasons but also for policy makers and destination managers. Our findings offer empirical support to the concept of 'crisis-resistant tourists' (Hajibaba et al., 2020). Tourists exhibit different reactions to shocks like COVID-19: while some segments are risk averse and stop traveling (pruning strategy) or change their travel habits (behavioral adaptation), others are more resilient and tend to continue with their plans. Since it will take some time to the industry to recover normality, identifying crisis-resistant tourists is paramount for developing marketing campaigns. Furthermore, domestic tourism has been pointed as the key element to restart the sector and alleviate the negative effects of COVID-19. If the shift in destination choice towards local destinations (reduced distance) were permanent, this would offer a great possibility for the tourism development of local areas, which could also help in the transition towards more sustainable tourism practices, reduced the dependency on outbound flows and increase tourism employment in the home country.

5.3. Limitations and avenues for future research

The study is not without limitations. First, we only have information about actual travelers. As such, we cannot model the participation decision (travel *vs* not to travel) and our analysis thus takes tourists as the population of interest. Even though we partially control for self-selection through the endogenous assignment to the switching regime (*y2020*), future research should expand our work by also modelling the participation decision. Second, the distance variable used in the analysis is an approximation of the true distance covered by the tourist. Mobile position data could be a valuable tool to track tourists' travel routes with more precision. Third, the study covers a large country with substantial spatial heterogeneity. Even though our model specification includes autonomous community of origin fixed effects, we advocate for more case studies that examine spatial travel patterns after COVID-19 with more granularity. Finally, we use a pooled cross-sectional database. Future research might take advantage of pseudo-panel methods to offer a longitudinal perspective to the study. For instance, it would be interesting for incoming studies to characterize the dynamics in tourists' travel patterns under different contexts and scenarios throughout the pandemic.

REFERENCES

Abadie, A., Athey, S., Imbens, G.W. and Wooldridge, J. (2017). When should you adjust standard errors for clustering? NBER Working Paper Series 24003. DOI 10.3386/w24003

Allenby, G.M., Jen, L. and Leone, R.P. (1996). Economic trends and being trendy: The influence of consumer confidence on retail fashion sales. *Journal of Business & Economic Statistics*, 14(1), 103-111. https://doi.org/10.2307/1392103

Alvarez-Diaz, M., D'Hombres, B., Ghisetti, C. and Pontarollo, N. (2020). Analysing domestic tourism flows at the provincial level in Spain by using spatial gravity models. *International Journal of Tourism Research*, 22(4), 403-415. https://doi.org/10.1002/jtr.2344

Arbulú, I., Razumova, M., Rey-Maquieira, J. and Sastre, F. (2021). Can domestic tourism relieve the COVID-19 tourist industry crisis? The case of Spain. *Journal of Destination Marketing & Management*, 20, 100568. https://doi.org/10.1016/j.jdmm.2021.100568

Boto-García, D. (2020). Habit formation in tourism travelling. *Journal of Travel Research*, forthcoming. https://doi.org/10.1177/0047287520964597

Boto-García, D., Alvarez, A. and Baños, J. (2021). Modelling heterogeneous preferences for nature-based tourism trips. *Papers in Regional Science*, forthcoming. https://doi.org/10.1111/pirs.12631

Boto-García, D. and Leoni, V. (2021). Exposure to COVID-19 and travel intentions: Evidence from Spain. *Tourism Economics*, forthcoming. https://doi.org/10.1177/1354816621996554

Bünnings, C., Kleibrink, J. and Webling, J. (2017). Fear of unemployment and its effect on the mental health of spouses. *Health Economics*, 26, 104-117. https://doi.org/10.1002/hec.3279

Buckley, R. and Westaway, D. (2020). Mental health rescue of women's outdoor tourism: a role in COVID-19 recovery. *Annals of Tourism Research*, 85, 103041. https://doi.org/10.1016/j.annals.2020.103041

Cafiso, G., Cellini, R. and Cuccia, T. (2018). Do economic crises lead tourists to closer destinations? Italy at the time of the Great Depression. *Papers in Regional Science*, 97(2), 369-386. https://doi.org/10.1111/pirs.12242

Cahyanto, I., Wiblishauser, M., Pennington-Gray, L. and Schroeder, A. (2016). The dynamics of travel avoidance: The case of Ebola in the U.S. *Tourism Management Perspectives*, 20, 195-203. https://doi.org/10.1016/j.tmp.2016.09.004

Cao, J., Zhang, J., Wang, C., Hu, H. and Yu, P. (2020). How far is the ideal destination? Distance desire, ways to explore the antinomy of distance effects in tourist destination choice. *Journal of Travel Research*, 59(4), 614-630. https://doi.org/10.1177/0047287519844832

Carroll, C.D., Dynan, K.E and Krane, S.D. (2003). Unemployment risk and precautionary wealth: Evidence from households' balance sheets. *Review of Economics and Statistics*, 85(3), 586-604. https://doi.org/10.1162/003465303322369740

Chandra, A., Head, K. and Tappa, M. (2014). The economics of cross-border travel. *The Review of Economics and Statistics*, 96(4), 648-661. https://doi.org/10.1162/REST_a_00404

Chen, C., Petrick, J.F. and Shahvali, M. (2016). Tourism experiences as a stress reliever: Examining the effects of tourism recovery experiences on life satisfaction. *Journal of Travel Research*, 55(2), 150–160. https://doi.org/10.1177/0047287514546223

Craig, C.A. (2021). Camping, glamping, and coronavirus in the United States. *Annals of Tourism Research*, 89, 103071. https://doi.org/10.1016/j.annals.2020.103071

Das, S.S. and Tiwari, A.K. (2021). Understanding international and domestic travel intention of Indian travellers during COVID-19 using a Bayesian approach. *Tourism Recreation Research*, 46(2), 228-244. https://doi.org/10.1080/02508281.2020.1830341

Di Falco, S., Veronesi, M. and Yesuf, M. (2011). Does adaptation to climate change provide food security? A micro-perspective from Ethiopia. *American Journal of Agricultural Economics*, 93(3), 829-846. https://doi.org/10.1093/ajae/aar006

Etzion, D. (2003). Annual vacation: Duration of relief from job stressors and burnout. *Anxiety, Stress and Coping*, 16(2), 213–226.

Eugenio-Martin, J.L. and Campos-Soria, J.A (2010). Climate in the region of origin and destination choice in outbound tourism demand. *Tourism Management*, 31, 744-753. https://doi.org/10.1016/j.tourman.2009.07.015

Eugenio-Martín, J.L. and Campos-Soria, J.A. (2014). Economic crisis and tourism expenditure cutback decision. *Annals of Tourism Research*, 44, 53-73. https://doi.org/10.1016/j.annals.2013.08.013

Fageda, X., Jiménez, J.L. and Valido, J. (2017). An empirical evaluation of the effect of European public policies on island airfares. *Transportation Research Part A*, 106, 288-299. https://doi.org/10.1016/j.tra.2017.09.018

Gausden, R. and Hasan, M.S. (2018). An assessment of the contribution of consumer confidence towards household spending decisions using UK data. *Applied Economics*, 50(12), 1395-1411. https://doi.org/10.1080/00036846.2017.1363859

Halvorsen, R. and Palmquist, R. (1980). The interpretation of dummy variables in semilogarithmic equations. *American Economic Review*, 70(3), 474-475. https://www.jstor.org/stable/1805237

Hajibaba, H., Gretzel, U., Leisch, F. and Dolnicar, S. (2015). Crisis-resistant tourists. *Annals of Tourism Research*, 53, 46-60. https://doi.org/10.1016/j.annals.2015.04.001

Haywood, K.M. (2020). A post COVID-19 future-tourism re-imagined and re-enabled. *Tourism Geographies*, 22(3), 599-609. https://doi.org/10.1080/14616688.2020.1762120

Heckman, J.J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153-161. https://doi.org/10.2307/1912352

Hotchkiss, J.L. and Pitts, M.M. (2005). Female labour force intermittency and current earnings: switching regression model with unknown sample selection. *Applied Economics*, 37, 545-560. https://doi.org/10.1080/0003684042000307003

Ivanova, M., Ivanov, I.K. and Ivanov, S. (2021). Travel behavior after the pandemic: the case of Bulgaria. *Anatolia*, 32(1), 1-11. https://doi.org/10.1080/13032917.2020.1818267

Jeon, C. and Yang, H. (2021). The structural changes of a local tourism network: comparison of before and after COVID-19. *Current Issues in Tourism*, forthcoming. https://doi.org/10.1080/13683500.2021.1874890

Jin, X., Bao, J. and Tang, C. (2021). Profiling and evaluating Chinese consumers regarding post-COVID-19 travel. *Current Issues in Tourism*, forthcoming. https://doi.org/10.1080/13683500.2021.1874313

Karl, M., Muskat, B. and Ritchie, B.W. (2020). Which travel risks are more salient for destination choice? An examination of the tourist's decision-making process. *Journal of Destination Marketing & Management*, 18, 100487. https://doi.org/10.1016/j.jdmm.2020.100487

Kement, U., Cavusoglu, S., Demirag, B., Durmaz, Y. and Bükey, A. (2020). Effect of perception of COVID-19 and nonpharmaceutical intervention on desire and behavioral intention in touristic travels in Turkey. *Journal of Hospitality and Tourism Insights*, forthcoming. https://doi.org/10.1108/JHTI-07-2020-0139

Lepp, A. and Gibson, H. (2003). Tourist roles, perceived risk and international tourism. *Annals of tourism research*, 30(3), 606-624. https://doi.org/10.1016/S0160-7383(03)00024-0

Li, Z., Zhang, S., Liu, X., Kozak, M. and Wen, J. (2020). Seeing the invisible hand: underlying effects of COVID-19 on tourists' behavioral patterns. *Journal of Destination Marketing & Management*, 18, 100502. https://doi.org/10.1016/j.jdmm.2020.100502

Mackenzie, S.H. and Goodnow, J. (2021). Adventure in the age of COVID-19: Embracing microadventures and Locavism in a post-pandemic world. *Leisure Sciences*, 43(1-2), 62-69. https://doi.org/10.1080/01490400.2020.1773984

Malley, J. and Moutos, T. (1996). Unemployment and consumption. *Oxford Economic Papers*, 48, 584-600. https://doi.org/10.1093/oxfordjournals.oep.a028586

McKercher, B. (2018). The impact of distance on tourism: a tourism geography law. *Tourism Geographies*, 3, 1-5. https://doi.org/10.1080/14616688.2018.1434813

Monfardini, C. and Radice, R. (2008). Testing exogeneity in the bivariate probit model: A Monte Carlo study. *Oxford Bulletin of Economics and Statistics*, 70(2), 271-282. https://doi.org/10.1111/j.1468-0084.2007.00486.x

Morley, C., Rosselló, J., and Santana-Gallego, M. (2014). Gravity models for tourism demand: theory and use. *Annals of Tourism Research*, 48, 1-10. https://doi.org/10.1016/j.annals.2014.05.008

Moulton, B.R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The Review of Economics and Statistics*, 334-338. https://doi.org/10.2307/2109724

Neuburger, L. and Egger, R. (2021). Travel risk perception and travel behavior during the COVID-19 pandemic 2020: a case study of the DACH region. *Current Issues in Tourism*, 24(7), 1003-1016. https://doi.org/10.1080/13683500.2020.1803807

Nguyen, C.P., Thanh, S.D. and Nguyen, B. (2020). Economic uncertainty and tourism consumption. *Tourism Economics*, forthcoming. https://doi.org/10.1177/1354816620981519

Nicolau, J.L. (2008). Characterizing tourist sensitivity to distance. *Journal of Travel Research*, 47, 43-52. https://doi.org/10.1177/0047287507312414

Nicolau, J.L. (2010). Variety-seeking and inertial behavior: the disutility of distance. *Tourism Economics*, 19(1), 251-264. https://doi.org/10.5367/000000010790871999

Nicolau, J.L. and Más, F.J. (2006). The influence of distance and prices on the choice of tourist destinations: the moderating role of motivations. *Tourism Management*, 27(5), 982-996. https://doi.org/10.1016/j.tourman.2005.09.009

Osti, L. and Nava, C.R. (2020). Loyal: to what extent? A shift in destination preference due to the COVID-19 pandemic. *Annals of Tourism Research: Empirical Insights*, 1(1), 100004. https://doi.org/10.1016/j.annale.2020.100004

Packer, J. (2021). Taking a break: exploring the restorative benefits of short breaks and vacations. *Annals of Tourism Research: Empirical Insights*, 2(1), 100006. https://doi.org/10.1016/j.annale.2020.100006

Papatheodorou, A. and Pappas, N. (2017). Economic recession, job vulnerability, and tourism decision making: A qualitative comparative analysis. *Journal of Travel Research*, 56(5), 663-677. https://doi.org/10.1177/0047287516651334

Pappas, N. (2021). COVID19: Holiday intentions during a pandemic. *Tourism Management*, 84, 104287. https://doi.org/10.1016/j.tourman.2021.104287

Plog, S. (2001). Why destination areas rise and fall in popularity. *Cornell Hotel & Restaurant Administration Quarterly*, 42(3): 13–24. https://doi.org/10.1177/0010880401423001

Puhani, P.A. (2000). The Heckman correction for sample selection and its critique. *Journal of Economic Surveys*, 14(1), 53-68. https://doi.org/10.1111/1467-6419.00104

Rittichainuwat, B.N. and Chakraborty, G. (2009). Perceived travel risks regarding terrorism and disease: The case of Thailand. *Tourism Management*, 30, 410-418. https://doi.org/10.1016/j.tourman.2008.08.001

Royal Decree 926/2020, of October 25, declaring the state of alarm for the management of the health crisis situation caused by COVID-19. Spanish State Official Newsletter (BOE). Available at: https://www.boe.es/buscar/doc.php?id=BOE-A-2020-12898.

Smeral, E. (2009). The impact of the financial and economic crisis on European tourism. *Journal of Travel Research*, 48(1), 3-13. https://doi.org/10.1177/0047287509336332

Weigert, M., Bauer, A., Gernert, J., Karl, M., Nalmpatian, A., Küchenhoff, H. and Schmude, J. (2021). Semiparametric APC analysis of destination choice patterns: using generalized additive models to quantify the impact of age, period, and cohort on travel distances. *Tourism Economics*, forthcoming. https://doi.org/10.1177/1354816620987198

Wen, J., Kozak, M., Yang, S. and Liu, F. (2021). COVID-19: potential effects on Chinese citizens' lifestyle and travel. *Tourism Review*, 76(1), 74-87. https://doi.org/10.1108/TR-03-2020-0110

Wen, Z., Huimin, G. and Kavanaugh, R.R. (2005). The impacts of SARS on the consumer behavior of Chinese domestic tourists. *Current Issues in Tourism*, 8(1), 22-38. https://doi.org/10.1080/13683500508668203

Wong, I., Fong, L., and Law, R. (2016). A longitudinal multilevel model of outbound travel behaviour and the dual cycle model. *Journal of Travel Research*, 55(7), 957–970. https://doi.org/10.1177/0047287515601239

Wong, I., Law, R. and Zhao, X. (2017). When and where to travel? A longitudinal multilevel investigation on destination choice and demand. *Journal of Travel Research*, 56(7), 868-880. https://doi.org/10.1177/0047287516670269

Wong, I.A., Zhang, G., Zhang, Y. and Huang, G.I. (2020). The dual distance model of tourism movement in intra-regional travel. *Current Issues in Tourism*, 24(9), 1190-1198. https://doi.org/10.1080/13683500.2020.1738356

Wynen, J. (2013). Explaining travel distance during same-day visits. *Tourism Management*, 36, 133-140. https://doi.org/10.1016/j.tourman.2012.11.007

Zhang, D., Luo, Q. and Ritchie, B.W. (2021). Afraid to travel after COVID-19? Self-protection, coping and resilience against pandemic 'travel fear'. *Tourism Management*, 83, 104261. https://doi.org/10.1016/j.tourman.2020.104261

		Total Trips		Total Domestic		Domestic outside the region of residence		Domestic within the region of residence	
		N	%	N	% over total trips	N	% over dom. trips	N	% over dom. trips
Before the pandemic (2019)	July	6,008	18.64	5,403	89.93	2,839	52.54	2,564	47.46
	August	7,315	22.69	6,433	87.94	3,773	58.65	2,660	41.35
	September	5,070	15.73	4,546	89.66	2,347	51.63	2,199	48.37
	Summer	18,393	57.07	16,382	89.07	8,959	54.69	7,423	45.31
During the pandemic outbreak (2020)	July	4,098	12.71	4,098	97.76	1,954	47.64	2,144	52.32
	August	6,277	19.47	6,069	96.69	3,178	52.36	2,891	47.64
	September	3,259	10.11	3,259	96.82	1,437	44.09	1,822	55.91
	Summer	13,835	42.92	13,426	97.04	6,569	48.93	6,857	51.07
	Total	32,228	100	29,808	100	15,528	52.09	14,280	47.91

Table 1.- Descriptive statistics of destination choice by period

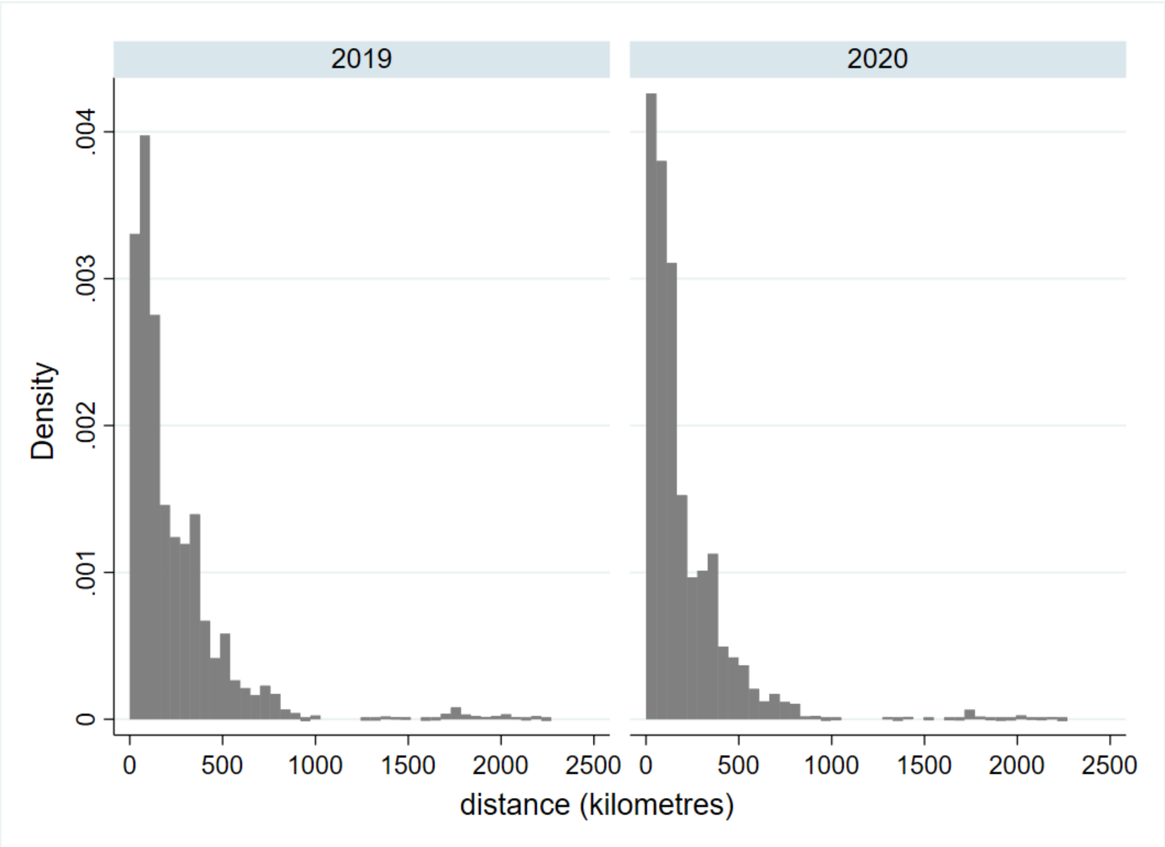

Figure 1.- Histogram of distance before (2019) and after (2020) the pandemic

Region of residence	2019 Distance travelled (km)	2019 Observations	2020 Distance travelled (km)	2020 Observations	t-test
Andalusia	229.74	1,897	199.08	1,451	5.050***
Aragon	211.14	805	181.12	643	2.703***
Asturias	198.29	642	115.71	533	5.179***
Balearic Islands	179.58	334	131.79	303	1.892*
Canary Islands	336.12	631	193.75	492	4.181***
Cantabria	235.25	415	188.87	287	2.004**
Castile and Leon	263.49	1,002	232.63	736	2.833***
Castilla-laMancha	275.42	830	255.65	584	1.979**
Catalonia	205.76	1,502	178.10	1,377	2.736***
Valencian Community	199.54	1,234	150.50	1,030	5.825***
Extremadura	278.23	537	178.87	423	7.032***
Galicia	185.36	802	121.87	565	4.551***
Community of Madrid	274.80	2,485	270.75	2,197	0.610
Murcia	158.14	677	104.15	577	4.101***
Navarre	224.77	717	215.20	564	0.638
Basque Country	257.78	1,049	243.49	959	1.159
La Rioja	239.31	522	179.61	498	3.208***
National mean	236.00	16,081	198.28	13,216	

Table 2.- Mean distance travelled per autonomous community of origin before (summer of 2019) and after (summer of 2020) the pandemic outbreak. *** p<0.01, ** p<0.05, * p<0.1

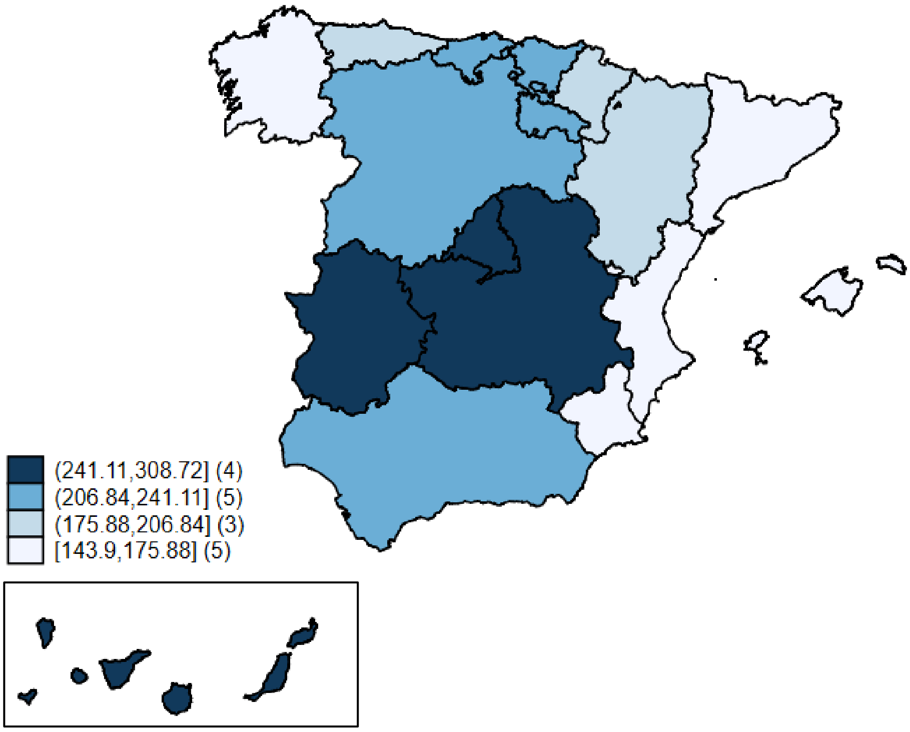

Figure 2.- Average distance travelled per autonomous community of residence (in kilometres) in the summer of 2019

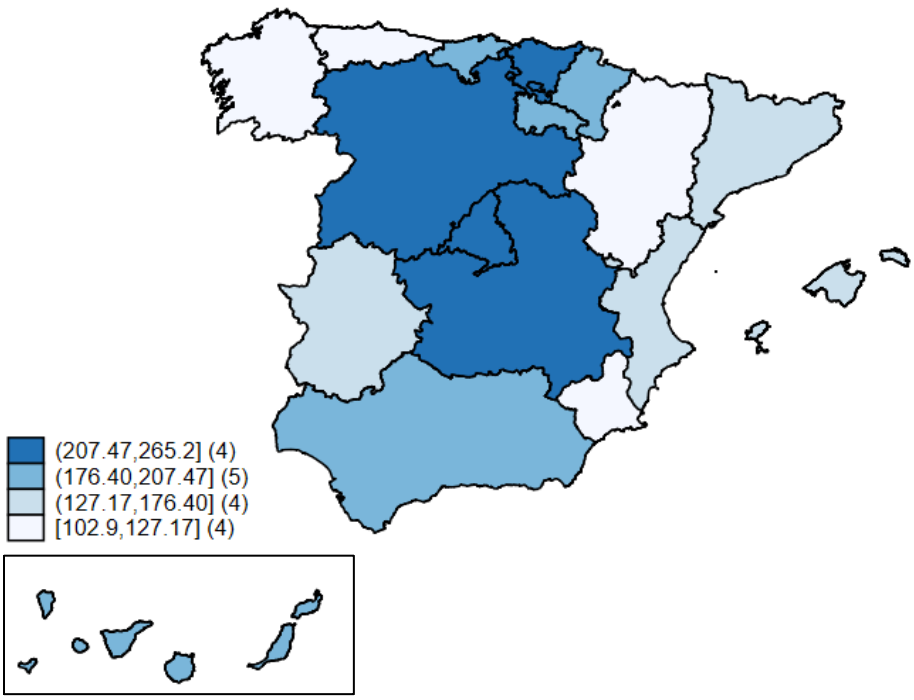

Figure 3.- Average distance travelled per autonomous community of residence (in kilometres) in the summer of 2020

Label	Description	Mean	SD	Min	Max
Continuous variables					
Distance	Weighted bilateral distance between the origin and destination (in kilometres)	218.99	267.64	0	2,269.97
Age	Age in years	48.52	15.28	15	85
Household size	Number of people living in the household	2.65	1.19	1	13
Binary variables		%			
Female	=1 if female	53.04			
Educ: primary	=1 if primary studies	4.36			
Educ: secondary	=1 if secondary studies	38.84			
Educ: university studies	=1 if university education	56.34			
Income: low	=1 if household monthly income is less than 1,000 €	8.31			
Income: low-medium	=1 if household monthly income is between 1,000 and 1,500 €	16.17			
Income: medium	=1 if household monthly income is between 1,500 and 2,500 €	33.68			
Income: medium-high	=1 if household monthly income is between 2,500 and 3,500 €	21.90			
Income: high	=1 if household monthly income is above 3,500 €	19.92			
Foreign	=1 if not Spanish	5.59			
Children	=1 if has children under 15	26.70			
Unemployed	=1 if currently unemployed	7.34			
Employed	=1 if employed	64.52			
Retired	=1 if retired	15.37			
Inactive	=1 if inactive (student, disabled, housekeeper)	12.31			
Mun density: Low	=1 if the municipality of residence is sparsely populated	10.59			
Mun density: Medium	=1 if the municipality of residence is moderately populated	27.35			
Mun density: High	=1 if the municipality of residence is highly populated	62.05			
Mun size: Small	=1 if the municipality of residence has more than 100,000 inhabitants	47.90			
Mun size: Medium	=1 if the municipality of residence has between 20,000 and 100,000 inhabitants	26.62			
Mun size: Big	=1 if the municipality of residence has less than 20,000 inhabitants	25.46			
Travel purpose: sun & beach	=1 if the trip purpose is sun & beach	26.07			
Travel purpose: cultural	=1 if the trip purpose is culture	7.66			
Travel purpose: nature	=1 if the trip purpose is to recreate in the nature	9.71			
Travel purpose: sports	=1 if the trip purpose is to practice sport	1.34			
Travel purpose: VFR	=1 if the trip purpose is to visit friends or relatives	34.74			
Travel purpose: other	=1 if the trip purpose is one of the following: religious peregrination, well-being, gastronomy, or shopping	20.45			
Alone	=1 if travels alone	16.71			
Couple	=1 if travels in a couple	59.32			
Other travel companions	=1 if travels in a group of friends or with relatives	60.00			
July	=1 if travels in July	31.68			
August	=1 if August	42.11			
September	=1 if September	26.20			
Weekend	=1 if travels during a weekend	40.20			
N	31,656				

Table 3.- Descriptive statistics of the sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Main Equation: ln Distance			Endogenous switching		Selection equation	
	If y2020=0	If y2020=1	Parameter equality	Prob (y2020=1)		Prob (domestic=1)	
Explanatory variables	Coeff. (SE)	Coeff. (SE)	Chi2(1) [p-value]	Coeff. (SE)	AME (SE)	Coeff. (SE)	AME (SE)
Female	-0.081	-0.060	0.12	-0.246***	-0.090***	-0.006	-0.001
	(0.054)	(0.050)	[0.724]	(0.069)	(0.024)	(0.027)	(0.003)
Age	-0.012***	-0.009***	0.46	-0.015***	-0.005***	0.006***	0.001***
	(0.005)	(0.003)	[0.497]	(0.002)	(7.4e-04)	(0.001)	(1.22e-04)
Educ: university studies	0.059	0.097	0.09	0.173***	0.064***	-0.094**	-0.010***
	(0.069)	(0.097)	[0.766]	(0.035)	(0.012)	(0.028)	(0.003)
Income: medium	0.250	0.079	1.28	0.023	0.008	-0.036	-0.004
	(0.155)	(0.063)	[0.257]	(0.030)	(0.011)	(0.040)	(0.004)
Income: medium-high	0.314**	0.191**	1.90	0.033	0.012	-0.126**	-0.014**
	(0.124)	(0.082)	[0.168]	(0.027)	(0.010)	(0.050)	(0.006)
Income: high	0.582***	0.218	3.80*	0.090**	0.033**	-0.189***	-0.022***
	(0.198)	(0.154)	[0.051]	(0.037)	(0.013)	(0.052)	(0.006)
Foreign	0.858*	0.008	4.75**	0.189***	0.070***	-0.908***	-0.146***
	(0.454)	(0.289)	[0.029]	(0.056)	(0.021)	(0.072)	(0.016)
Children	-0.072	-0.158**	1.17	0.067*	0.024*	0.089**	0.010**
	(0.050)	(0.061)	[0.279]	(0.036)	(0.013)	(0.045)	(0.005)
Unemployed	0.319	-0.026	2.32	0.668***	0.178***	-0.075	-0.009
	(0.215)	(0.099)	[0.128]	(0.134)	(0.024)	(0.081)	(0.010)
Employed	0.481***	0.268***	2.68	0.409***	0.147***	-0.147***	-0.016***
	(0.163)	(0.102)	[0.102]	(0.040)	(0.012)	(0.045)	(0.005)
Household size	-0.105**	-0.028	1.35	-0.052***	-0.019***	0.007	0.001
	(0.046)	(0.029)	[0.244]	(0.014)	(0.005)	(0.023)	(0.003)
Mun density: High	-0.205***	-0.259*	0.13	0.059	0.021	0.118**	0.013**
	(0.084)	(0.148)	[0.713]	(0.038)	(0.014)	(0.039)	(0.004)
Mun density: Low	0.229	-0.164	2.61	-0.171***	-0.062***	0.171***	0.017***
	(0.239)	(0.135)	[0.106]	(0.065)	(0.023)	(0.043)	(0.004)
Mun size: Small	-0.161	-0.232	0.11	0.116***	0.043***	-0.049	-0.005
	(0.119)	(0.195)	[0.744]	(0.034)	(0.012)	(0.050)	(0.006)
Travel purpose: sun &beach	0.575	0.870**	2.56	-0.056	-0.020	0.171	0.018
	(0.376)	(0.360)	[0.109]	(0.046)	(0.016)	(0.136)	(0.014)
Travel purpose: cultural	2.254***	1.176**	4.55**	-0.307***	-0.110***	-1.348***	-0.267***
	(0.769)	(0.579)	[0.033]	(0.056)	(0.019)	(0.094)	(0.025)
Travel purpose: nature	0.323	0.429**	0.17	0.200***	0.074***	-0.345***	-0.043***
	(0.246)	(0.178)	[0.681]	(0.062)	(0.023)	(0.040)	(0.006)
Travel purpose: sports	0.763*	1.223***	0.66	-0.056	-0.020	-0.302**	-0.039**
	(0.403)	(0.449)	[0.417]	(0.071)	(0.026)	(0.131)	(0.019)
Travel purpose: VFR	0.483***	0.504***	0.02	0.085**	0.031**	-0.103	-0.012
	(0.183)	(0.163)	[0.892]	(0.038)	(0.014)	(0.080)	(0.009)
Alone	-0.438***	-0.439***	0.01	0.110***	0.041***	0.136	0.014
	(0.129)	(0.141)	[0.991]	(0.042)	(0.015)	(0.090)	(0.009)
Couple	-0.040	-0.050	0.01	0.122***	0.045***	0.029	0.003
	(0.105)	(0.092)	[0.928]	(0.038)	(0.013)	(0.041)	(0.005)
Weekend	-1.528***	-1.144***	4.58**	-0.119***	-0.044***	0.830***	0.080***
	(0.400)	(0.272)	[0.032]	(0.041)	(0.015)	(0.126)	(0.009)
August	-0.065	-0.019	0.34	0.116***	0.043***	-0.143**	-0.016**
	(0.061)	(0.072)	[0.558]	(0.018)	(0.010)	(0.068)	(0.007)
September	-0.054	0.018	0.22	-0.012	-0.004	-0.296***	-0.035***
	(0.090)	(0.083)	[0.637]	(0.020)	(0.011)	(0.068)	(0.009)
IMR	-1.480	0.471	3.10*				
	(1.162)	(1.437)	[0.078]				
Unemployment rate				0.076***	0.027***		
				(0.009)	(0.003)		
Unemployment rate#Unemployed				-0.014**			
				(0.006)			
CCI						-0.018***	-0.002***
						(0.002)	(1.58e-04)
Constant	6.451***	5.497***		-2.920***		2.891***	
	(0.453)	(0.284)		(0.316)		(0.175)	
Origin Fixed Effects	YES	YES		YES		YES	
Corr (η_i,ϵ_i),			0.109				
			(0.088)				
Var (ϵ_i)			5.872***				
			(1.712)				
Observations			29,297			31,656	

Table 4.- Coefficient estimates for the endogenous switching regression model

Notes: Clustered standard errors at the Autonomous Community in parentheses. The third column reports the statistic of the parameter equality that follows a chi2(1) distribution. P-values in brackets. The reference categories are primary or secondary education, low or low-medium income, retired or inactive, living in a municipality that is moderately populated (medium density), living in a municipality that is big or medium sized, traveling in a family or with friends/workmates and July. *** p<0.01, ** p<0.05, * p<0.1