Annals of Tourism Research Empirical Insights 2 (2021) 100014

Contents lists available at ScienceDirect

Annals of Tourism Research Empirical Insights

journal homepage: www.elsevier.com/locate/annale

Factors explaining length of stay: Lessons to be learnt from Madeira Island

António Almeida [a,1], Luiz Pinto Machado [a,1], Chen Xu [b,*]

[a] Center for Advanced Studies in Management and Economics, Universidade da Madeira, Portugal
[b] Welsh Economy Research Unit, Cardiff University, United Kingdom

ARTICLE INFO

Article history:
Received 4 September 2020
Received in revised form 10 January 2021
Accepted 21 February 2021

Keywords:
Length of stay
Duration models
Survival models
Madeira Island

ABSTRACT

The length of stay for tourists is shrinking for traditional tourism destinations, with tourists instead opting for short breaks to multiple destinations. The reasons for these changes include the increasing number of low cost airlines reduces the cost per journey, alongside heightened disposable income and strong marketing strategies by competing destinations. Madeira Island is well placed in this study as it faces a typical issue of declining length of stay, meanwhile acquires rich data in carrying out thorough analyses in the factors that explain length of stay in Madeira Island-Portugal by five different econometric approaches, further policy implications of the research findings, particularly those that could potentially prove useful to increase the length of stay, are also discussed.

1. Introduction

Tourism length of stay is a key issue in tourism demand management during a time of increasingly shorter stays and fierce competition to attract affluent visitors. Curtailment of the duration of stay could have a significantly detrimental impact upon the economic benefits generated by tourist activity and deter future investment interest in the sector. For example, data released by the Statistical Institute from Madeira Island in 2016 show a daily average spending of 124 euros, implying a loss of 124 euros per tourist for every tourism length of stay drop by one day. More precisely, in a destination with an average number of 1.2 million tourists per year, such as Madeira Island, an income loss of 124 euros per stay/tourist represents an annual loss in revenue of 148.8 million euros. Therefore, to increase or maintain current figures, tourism length of stay is essential for demand management in the current context of proliferating external shocks and difficulties in attaining expenditure targets (Yang & Liu, 2003; Martinéz-Garcia & Raya, 2008.

In addressing the existing tourism challenge, this study analyses the determinants of tourism length of stay in a tourism destination. The case study centres on Madeira Island, a popular and traditional tourism destination that forms an archipelago in an autonomous region of Portugal located off the northwest coast of Africa, and the study is based on a sample of tourists hosted in a type of specialist accommodation, boutique hotels on Madeira Island. Survival models have been adopted by several authors (see Barros, Correia, & Crouch, 2008; Barros & Machado, 2010; Gokovali, Bahar, & Kozak, 2006; Machado, 2010; Menezes et al., 2008; Rodríguez, Martínez-Roget, & González-Murias, 2018; Thrane, 2012) to study the determinants of length of stay.

However, recent research suggests that survival models are plagued with diverse theoretical weaknesses and technical complexities caused by the very nature of the dependent variables (Thrane, 2012). While survival models have proved to be particularly popular among researchers studying the determinants of tourism length of stay, owing to the wide range of different model specifications available, other models are also appropriate and more robust from a purely theoretical point of view (Santos, Ramos, & Rey-Maquieira, 2015; Thrane, 2012). For that reason, in addition to running survival models, we also employed different methods to identify commonalities in terms of the main determinants of tourism length of stay and a finite mixture of models to analyse the data owing to the presence of unobserved heterogeneity.

Hence, the aim of this paper, grounded on the assumption that tourism length of stay is of paramount importance in tourism management, lies in ascertaining which covariates best explain tourism length of stay in the destination under analysis (Gokovali et al., 2006). In this study, we take into consideration several variables underutilised in previous studies. While most authors consider a wide range of sociodemographic and travel related variables in their studies, in this study we also include issues such as the impact of time-dependent variables (e.g. number of activities pursued while

* Corresponding author at: T05, Aberconway Building, Colum Drive, Cardiff Business School, Cardiff CF10 3EU, United Kingdom.
E-mail address: gemmachenxu@gmai.com (C. Xu).
[1] Postal addresses for both António Almeida and Luiz Pinto Machado, Universidade da Madeira Faculdade de Ciencias Sociais –DGE Campus da Penteada Caminho da Achada n10 9000–208 Funchal.

http://dx.doi.org/10.1016/j.annale.2021.100014

abroad) on tourism length of stay. We also estimate the impact of travel arrangements mixing up access by air and visits by cruise. Additionally, we acknowledged two major trends in modern tourism besides declining tourism length of stay. By studying the determinants of tourism length of stay within the current drive to develop different types of specialist accommodation to target specific segments, this study specifically offers an assessment of several relevant variables impacting in a sample of tourists opting to stay in boutique hotels, which has rarely been done before. A further trend in the context of island destinations relates to repeat visits prompted by previous visits by cruise.

The comparative analysis of the results provided by five different econometric approaches offers a basis on which to establish the relationship between the variables under analysis and tourism length of stay in a "conclusive" manner. In terms of econometric specification, the application of a finite mixture approach to analyse the data, still rare in this strand of literature, allows the analyst to perform simultaneous clustering and regression analysis; such an approach recognizes the vital importance of segmentation analysis in the field of tourism studies.

Madeira Island, like other destinations, is presently facing difficulties in attracting tourists at the same rhythm as in the past, owing to increasing difficulties experienced in retaining their traditional customers, namely UK tourists. Several dramatic events, such as the collapse of Thomas Cook Group and the imminency of Brexit, together with devaluation of the British pound relative to the euro, and the relatively high number of airline bankruptcies (14 airlines in the last four years and others following in the wake of Covid-19) affecting other markets, such as Germany and France, (those three markets represent a share of 65% regarding international arrivals in the island). In addition, owing to the traditionally negligible share of the domestic market (resulting from lack of access by car) the impact of such events is more dramatic on islands during and post the coronavirus pandemic, where international visitors are severely limited. Thus, it is important from a policy-making point of view to identify critical factors that encourage visitors to stay longer (Aguiló et al., 2005), and which variables can be manipulated and micro-managed at local level, as major external shocks can only be deflected to a certain extent.

This study has discovered several unexpected results that can be explained in view of the nature of the accommodation services under analysis, in terms of a price premium and different setting of motivations at work. This study also highlights the importance of allowing for different econometric approaches to identify similarities and differences by additionally considering unobserved heterogeneity.

2. Madeira Island

Madeira Island, an old and traditional tourism destination, was discovered by the Portuguese in 1418. Tourism to Madeira Island commenced at the beginning of the nineteenth century, based on the attraction for wealthy British visitors travelling by sea. Visitors can choose the elegance of the island's time-honoured hotels, opt for one of the more modern forms of accommodation, or select smaller units, often located in rural areas, offering tourists an opportunity to connect with nature and local culture. Madeira's hotels are largely renowned for their hospitality, high quality and personalised service, refined taste and luxurious furniture and interior architecture. In addition, many recent establishments are well situated to highlight the richness of local architecture, culture and identity. This recognition has been repeatedly confirmed with presentation of the most prestigious awards in the world, both to Madeira hotels and to the island itself as a destination, being considered by World Travel Awards in recent years to be the best island destination in the world.

Most tourists continue to arrive by air. However, Funchal is a port of call throughout the year on the itineraries of several transatlantic cruise-ships. Data for 2019 identify 293 ships and 591,000 passengers, which is equivalent to 37% of arrivals by air. By taking advantage of the climate, its protective mountains and geographical location that adds an exotic, sub-tropical flavour to the destination, Madeira offers pleasant temperatures and a high degree of thermal comfort. It has been classified as a Biogenetic Reserve owing to the endemic flora and fauna unique in the world. Madeira offers excellent opportunities to trekkers and hikers, as well as to those interested in activities such as deep-sea fishing, sailing, extreme sports, and surfing. (Oliveira & Pereira, 2008).

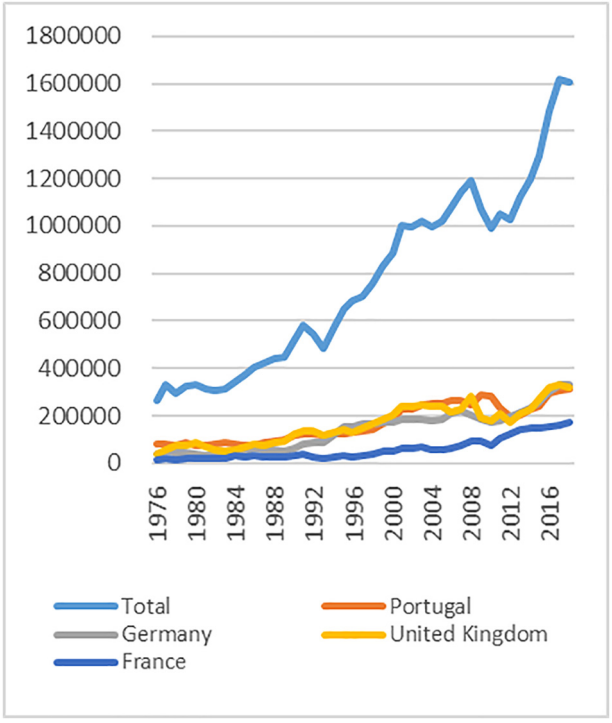
However, since 1976, the region has been losing ground as a winter destination, partly due to competition with other destinations in the Mediterranean Basin. Moreover, the tourism length of stay has tended to decline, raising concerns for the long-term prospects of the tourism sector in Madeira.

Graphs 1 and 2 present the arrivals, overnight stays, and tourism length of stay of the four main markets in Madeira Island, namely Portugal, Germany, the United Kingdom and France, between 1976 and 2018.

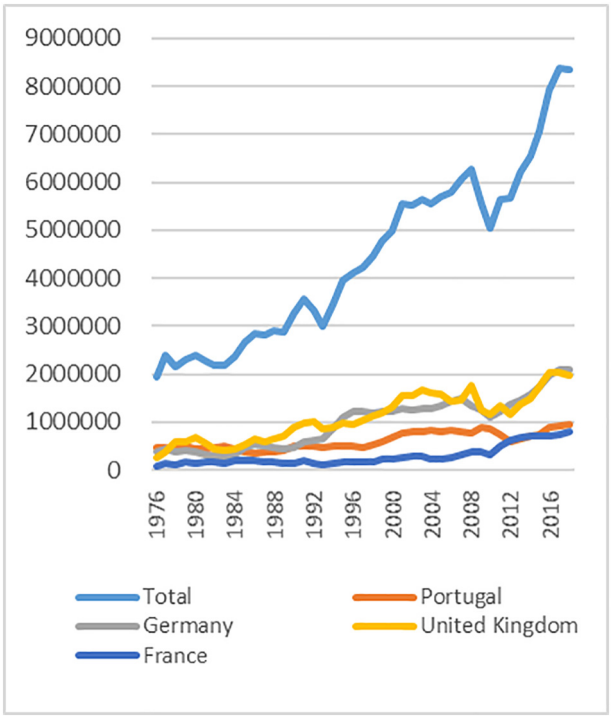
Graph 1 and Graph 2 show arrivals and overnights have risen since 1976, in comparison with other European tourism destinations. However, Graph 3 illustrates that the average tourism length of stay during the period in comparative markets follows a decreasing trend. This can be explained in part by the increasing availability of free time and extra disposable income among the population in general, which favours taking several vacation breaks during the year. The contemporary phenomenon of low-cost air travel compounded by strong marketing strategies determined to attract short-break visitors to alternative tourism destinations also provides a contributive factor, leading to shorter stays in traditional tourism destinations.

In recent years, the sector has witnessed different initiatives aimed at adapting dozens of derelict buildings and vacant premises into modern and stylish accommodation facilities, categorised as *boutique hotels*. Such establishments highlight their origins with prominent manor houses ideally positioned with magnificent panoramic views over the surrounding mountains and Funchal city centre and bay. Such establishments, grouped under the label *Quintas da Madeira*, share characteristics such as location close to the historical city; renovated buildings in view of their historic, economic and/or social use; botanical gardens; "elegantly furnished rooms"; and high service standards. In most cases, the main buildings are related to 18th century manor houses.

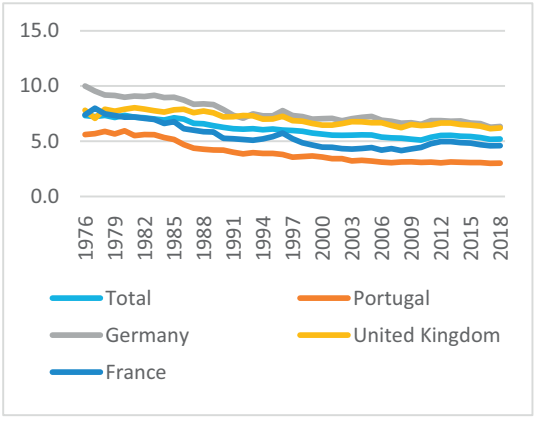
This case study illustrates the point that most destinations have experienced a downward trend in terms of LOS (Barros &; Martinéz-Garcia & Raya, 2008). In many cases, the declining trend in terms of average length of stay has been partially or completely offset by greater tourist volumes



Graph 1. Arrivals: 1976–2018. Source: Madeira Statistical Office and authors.



Graph 2. Overnigths: 1976–2018. Source: Madeira Statistical Office and authors.



Graph 3. Length of Stay: 1976–2018. Source: Madeira Statistical Office and authors.

and overnights and by a multiplication of the number of short stays, as evidenced by Gössling et al. (2018). However, since most countries around the globe have experienced a steady decline in LOS, this disputes whether it remains meaningful for destinations to seek to increase LOS. Gössling et al. (2018) question whether increasing LOS are worthwhile at a time when technological advances linked to global booking platforms, as well as sociological changes, contribute to a steady decline in LOS and greater volumes of tourists (Oklevik et al., 2018). Shorter stays and greater volume of visitors are a key feature of contemporary tourism. Overall, the negative impact of declining length of stay has been fully offset by the multiplication of market niches (e.g. short breaks, cultural tourism, etc.) and market developments such as the advent of low-cost airlines. The matter of declining length of stay has been continuously acknowledged as a major cause for concern and recent figures on the weight in percentage terms of shorter stays were noted. Encouraged by the constantly increasing numbers in terms of overnight and tourism receipts, DMOs are inclined to tacitly acknowledge that they have limited capacity to control the process. This can be observed in ongoing attempts to pursue a diversified and volume-focused strategy based on market segmentation and niche development. Such options are not available to the periphery. In such cases, the increase in terms of overnights have been fuelled by the overall performance of the industry, rather than by the development and generation of new sources of attraction/market niches. Therefore, such destinations are less able to withstand external shocks expressed in declining number of arrivals, because what matters is the consolidation of the main product. Under such circumstances, a volume based strategy is dependent on extended holidays. On the other hand, in several cases such regions have experienced an increase in average length of stay by means of an extra number of longer stays, which suggests that such approach is currently possible.

3. Literature review

Current studies confirm the statistical trend, indicating that people are travelling more frequently, yet spending less time at each destination (Soler, Gemar, & Correia, 2018). The short stays phenomenon received less attention from Destination Management Organisations while the number of arrivals and overnights continued to grow steadily. However, tourism length of stay began increasingly to be studied by researchers, as it was considered an important feature of demand with far reaching consequences in terms of daily and strategic planning of activities by operators. In effect, longer stays lead to higher levels of expenditure per stay and therefore to increased revenues and greater efficiency in operational terms from the point of view of the industry (Alegre, Mateo, & Pou, 2011; Thrane, 2012).

Several statistical models have been employed by researchers to study the determinants of tourism length of stay. By contrast, the regularity and constancy discovered in the literature, in terms of the explanatory variables employed to model tourism length of stay given, should be noted. Meta-analysis data show that it is relatively common to employ well-known socio-economic, travel-related and destination attributes variables. Survival analysis reigned for a long period, but other methods have becoming more common in recent times (Soler et al., 2018

An increasing “heterogeneity” of econometric methods has been used to examine the determinants of tourism length of stay and can be found in the literature (Brida, Meleddu, & Pulina, 2013; Gómez-Déniz & Pérez-Rodríguez, 2019; Thrane, 2012). The use of different econometric approaches in a single paper for comparative purposes is increasingly the norm. Studies based on survival analysis methods highlighted both parametric and Cox proportional hazard models as the most predominant methods to model the baseline (Thrane, 2012). Martinéz-Garcia and Raya (2008) examined the determinants of tourism length of stay of tourism travelling to Spain in low cost air carrier's tourism length of stay in Spain with log-logistics and Cox survival models. The main conclusion was that the Cox specification was unsuitable to model the data, which prompted the authors to run an accelerated survival model. A number of covariates of interest were statistically significant, namely socio-demographic, such as tourists' country of origin, age, choice of accommodation, time of year, and the geographical areas (urban vs rural) visited.

Hong and Jang (2005) employed survival analysis to model the duration of visits to a casino and Gokovali et al. (2006) employed survival model to analyse the tourism length of stay of tourists travelling to Turkey based on two traditional survival model specifications: the Cox model and the Weibull model. They found that the covariates of interest are the statistically significant variables, including the following: country of origin, age and professional occupation, in terms of socio-economic characteristics, and the package-vacation type being a first visit, quality, and visitors' perceptions on the level of hospitality and attractiveness of the destination, in addition to issues of night life, accommodation, image and promotion. Menezes et al. (2008a, 2008b) analysed tourism length of stay to study tourists hosted in the Azores islands based on a Cox proportional hazard model. The authors found that the variable first visit was of specific importance in explaining tourism length of stay.

Other traditional models have been extensively applied in this regard. Alegre et al. (2011) analysed tourists' tourism length of stay via latent class truncated Poisson regression to take into consideration the heterogeneity (segments) of the sample under analysis. They identified two segments containing individuals showing a preference for either short or longer stays. More recently, Yang and Zhang (2015) also applied a latent class duration model to analyse the determinants of tourism length of stay in different segments and reached similar conclusions to those of Alegre et al. (2011). Alén, Nicolau, Losada, and Domínguez (2014), supported by a negative binomial model, analysed the factors determining tourism length of stay of Spanish seniors. The variables of interest were respondent's age, the stated travel purpose, the type of accommodation establishment, the travel party size and nature and quantity of events attended, and activities pursued while abroad. Prebensen, Altin, and Uysal (2015) also applied a truncated negative binomial model to analyse tourism length of stay. Brida, Pereyra, and Scuderi (2014) based on quantile data regression analysed the determinants of tourism length of stay in Uruguay. The authors highlighted several variables of interest: namely, socioeconomic; travel-related; psychographic; and budgetary related characteristics. Rodríguez et al. (2018) same day excursionists and tourists to analyse the determinants of tourism length of stay in Santiago de Compostela based on a Heckman selection model the authors differentiated same-day visitors from traditional tourists.

Survival analysis remains the prevalent econometric method employed to analyse tourism length of stay (Aguilar & Díaz, 2019). However, Santos et al. (2015) considers that survival models imply an unnecessarily high degree of technical complexity in estimating the determinants of tourism length of stay, despite their inherent advantage in terms of distributional flexibility. Santos et al. (2015) applied a generalised linear model, concluding that the log-gamma distribution exhibits better statistical proprieties compared to the log-linear Ordinary Least Squares. Unlike Thrane (2012), Santos et al. (2015) assumes that the log-linear Ordinary Least Squares is rarely the best alternative because it may result in misleading qualitative conclusions.

An increasing number of authors report on the analysis of the results based on two competing models. For example, Mortazavi and Cialani (2017) analyse the determinants of tourism length of stay in Venice based on a double approach: Ordinary Least Squares and zero-truncated negative binomial. The authors found that age, the nature of the return journey, to be a repeat visitor and travelling to the city in the summer positively influence tourism length of stay. By contrast, visits to other cities besides Venice and visitor's level of expenditure had a negative impact on tourism length of stay. It is worth mentioning that the Ordinary Least Squares method remains popular and several authors run an Ordinary Least Squares model for comparison purposes. For example, Esiyok, Kurtulmuşoğlu, and Özdemir (2018) analysed the determinants of tourism length of stay of older thermal tourists to discover that age, economics constraints, distance to the destination and seasonal effects significantly affect the tourism length of stay.

Thrane (2012) considers that “it makes little sense” to conceptualise tourists' tourism length of stay “as a positive random variable denoting survival times” (Box-Steffensmeier & Jones, 2004), because in most instances decisions concerning duration of stay need to be taken in advance. For example, large numbers of tourists opt for a week-long stay owing to practical issues such as flight availability, time constraints and accepted practice in this regard (e.g. weekly packages available in brochures). In such cases, for each of the first six days on vacation, “the probability of departure” (i.e. risk of the event happening) is zero percent”, except in special cases (familiar's sudden illness or an accident), “whereas it is 100 percent on the seventh day”. Thrane (2012) insists that in most instances the arrival and departure date will not (and cannot for all practical reasons) be changed, which results in length of journey being defined in advance. With regard to the region under analysis, the more obvious, characteristic and “realistic scenario” when planning tourism length of stay is to decide (and to book and pay) before the date of departure and arrival, in line with decisions taken in terms of airlines, airports of departure and flight timetables (Thrane, 2015; Thrane & Farstad, 2012).

From an even more comprehensive point of view, Gómez-Déniz and Pérez-Rodríguez (2019) argue that commonly employed models may be unappropriated. The authors refer that data on tourism length of stay exhibits both overdispersion (Alén et al., 2014; Boto-García, Baños-Pino, & Álvarez, 2018; Brida et al., 2013; Nicolau, Zach, & Tussyadiah, 2016) and bimodality or multimodality (Alegre et al., 2011; Salmasi, Celidoni, & Procidano, 2012). For that reason, traditional econometric methods, such as Ordinary Least Squares, count data or duration models may be misleading. Based on evidence pointing to bimodal data, the authors apply an infinite mixture model to consider the heterogeneous preferences of tourists in terms of the duration of stay. Therefore, in this study, we study the determinants of tourism length of stay supported on double approach: survival analysis and finite mixture models. In particular, we opted for a finite mixture model within the context of traditional survival analysis.

4. Data

The data source for this analysis emerges from a research project conducted to further understanding of the key factors attracting visitors to this type of specialist accommodation and to elucidate various other relevant issues. The study was conducted for a 12-month period between January 2014 and January 2015 to overcome selection bias relating to seasonality. For the same reason, i.e., avoiding selection bias pertaining to specific hotel locations, guests were approached in a dozen different establishments. In total, the research team collected 415 questionnaires. The distribution of the questionnaires, according to the timing of data collection process, is the following: 89% in the first semester; 11% in the second semester. The fact that the majority of questionnaires were collected in the first semester can be explained by taking into account the timing of data collection. The process started in January and as the days passed, owners' willingness and commitment to participate simply diminished.”

The hotels/owners approached belong to a network of boutique hotels called “Quintas da Madeira”. The hotels belonging to the network are scattered around the island, even though most are located in Funchal and surrounding areas”. In order to approach as many tourists as possible, the questionnaire was translated into three languages: English, French and German. As frequently the case, to reach a sufficient minimum size, the collection data process must continue over the course of time. While the objective was to complete the data recollection process in a six-month period, which would have enabled the research team to cover both the low and the high season, the process was extended for a further few months to facilitate the collection data process. In total, around 89% of the questionnaires were collected in the first semester and the remaining in the second semester.

The data derives from a self-administered questionnaire that requested visitors to rate the overall importance of 10 general motives to choose Madeira and 18 specific motives to choose one of the “Quintus” based on a 1-to-5 Likert scale. Respondents were also asked about personal characteristics and their duration of stay. Based on the data, we defined tourism length of stay as a strictly positive numeric discrete variable representing the number of nights spent at the destination.

The socio-demographic characteristics of the sample under analysis, and an overall profile of the tourists interviewed, are displayed in Table 1. Overall, the sample is evenly distributed in terms of gender, with the number of female respondents (50.8%) almost equal to the number of male respondents. The average age is 53.4 years old, with most respondents in the 50–59 (31.6%), 60 and plus (37.6%) age groups. Around 47% of the respondents report at least a college degree. The average monthly income is 3074.7 euros, with almost 79.5% of the respondents earning less than 3500 euros. A closer look at the results of the demographic analysis suggests, apart from figures in terms of nationality, that respondents' profile closely matches similar results provided in other studies concerning Madeira (Almeida & Garrod, 2018). The mean and standard deviation of the variable tourism length of stay is 10.34 days and 4.513 days, respectively. The results provided in Table 1 highlight the fact that visitors expressed a

Table 1
Basic statistics.

Variable	Description	Min	Max	Perc./Mean	Std. Dev
Length of Stay	length of stay in days	2	30	10.34	4.513
	Socio-economic characteristics				
Age	Age of the individual	…	…	55.3	…
	below 18			1.0%	
	18–24 years old			2.4%	
	25–29 years old			2.7%	
	30–39 years old			9.2%	
	40–49 years old			15.7%	
	50–59 years old			31.6%	
	60 or more			37.6%	
Gender	Female = 0, Male = 1	0	1	50.8%/49.2%	…
Civil status	Single = 1, other = 0	0	1	16.9%	
	Married = 1, other = 0	0	1	73.7%	
	Other = 1, other = 0	0	1	9.9%	
Education	*Level of education*	…	…		…
Primary	Primary = 1, other = 0 (reference category)			8.4%	
Secondary	Secondary = 1, other = 0			31.6%	
Undergraduate diploma	Undergraduate diploma = 1; other = 0			26.0%	
Master/PhD	Master/PhD = 1, other = 0			21.0%	
Other	Other = 1, other = 0			13.0%	
Nationality					
British	British tourist = 1, other nationality = 0 (reference category)	…	…	14.50%	…
German	German tourist = 1, other nationality = 0 (reference category)	…	…	39.80%	…
Portuguese	Port. tourist = 1, other nationality = 0 (reference category)	…	…	4.60%	…
French	French tourist = 1 other nationality = 0 (reference category)	…	…	11.60%	…
Other	"Other" = 1, other nationality = 0 (reference category)	0	1	29.60%	
Work status	Retired person: Yes = 1, No = 0 (reference category)	0	1	42.4%	
	Income and economic constraints				
Income	*Monthly net household income*			3074.7€	1895.6
Accommodation costs	Accommodation costs per stay			1239.6€	776.94
Daily prices	Daily room rates			117.19€	17.56
	Travel arrangements and decision making process				
Access to information	Number of sources of information consulted	1	6	1.45	0.83
Repeat Visit	Yes = 1, No = 0(reference category)	0	1	28.9%	
Previous Experiences	Experience in similar accommodation: Yes = 1, No = 2	1	2	78.60%	
Plane + cruise	Travelling by plane and then by cruise; Yes = 1, No = 0 (reference category)	0	1	19.5%	
Low cost	Travelling on a low cost carrier, Yes = 1, No = 0 (reference category)	0	1	24.3%	
	Motivations and behaviour				
Motivation	Sum of the score of 10 different motivations to visit the island; (1-low importance … 5-high importance)	10	50	39.34	6.839
Price/Quality ratio	Importance Price/Quality ratio (1-low importance … 5-high importance)	1	5	4.14	0.824
Satisfaction	Assessment of perceived satisfaction (1-low satisfaction … 5-high satisfaction)	1	5	4.4	0.79
Travelling in the Summer	Summer-season = 1, elsewhere = 0 (reference category)	0	1	24.8	
Key attribute	Importance of the "Weather" factor (1-low importance … 5-high importance)	1	5	4.37	0.835

rather positive interest in exploring the historical related resources of Quintas.

The sample does not follow the pattern identified in other studies in terms of the weight in percentage of each of the key countries of origin. Compared to the standard percentage distribution, an excess of German nationals and a reduced number of Portuguese and British nationals is well evident in this study.

5. Methodology, research design and hypothesis

The models employed to analyse the determinants of tourism length of stay are grounded on traditional consumer behaviour theory and assume that a consumer, facing a consumption decision subject to budgetary restrictions and aiming at maximizing the overall level utility, decides based on prices and income (Varian, 1987). Most variables are binary (dummy) or nominal in nature, but accommodation costs, motivation and importance can be understood as numeric. For estimation purposes, age is defined as a continuous variable. We applied survival analysis and finite mixture models to identify the determinants of tourism length of stay, involving a preliminary stage of identification of the most relevant variables for econometric analysis. Since the literature fails to reach a consensus on the most suitable econometric approach to deal with the characteristics of dependent variable, we tested the most employed econometric specifications reported in the literature.

Three concerns must be duly evaluated when analysing survival models: (1) identification of the type of data set under analysis (i.e., cross-section vs. time series/panel data); (2) issues of censoring and (3) identification of the most appropriate model specification. In common with most studies focused on the determinants of tourism length of stay, the data employed in this study reports to a cross-section study. No problems in terms of censored data are present in this study as respondents were interviewed at the end of their vacation.

Relative to the third issue, several solutions are worthy of consideration. The Cox proportional hazard model for single-event data is tested in most papers employing a survival analysis methodology. The distinctive characteristic of the Cox survival model relates to the absence of assumptions proposed about the shape of the underlying hazard function. However, the Cox proportional hazard model relies on the proportionality assumption. In practical terms, it is assumed that, given two observations associated to distinct values for the independent variables, the ratio of the hazard functions

for such observations does not depend on time, and therefore, each observation evolves during time independently and proportionally. An alternative approach lies in estimating parametric models such as the proportional hazard Weibull model (Gokovali et al., 2006). The choice of the best fitting survival model should be based on the shape of baseline vacation duration, which is the procedure adopted in the present research.

Heterogeneous models are employed to measure *overdispersion* by adding a latent multiplicative effect on the hazard function. The term frailty is defined as a "latent random effect that enters multiplicatively on the hazard function" being a "useful alternative to a standard survival model when the standard model fails to adequately account for all the variability in the observed failure times" (Gutierrez, 2002:23). In face of unobserved or unmeasured individual factors (unaccounted predictors), the results of the Cox proportional hazard model may be unreliable owing to sample selection bias. In such circumstances, it is advisable to compare the results of the standard proportional hazard model with frailty models that add a random component "designed to account for variability due to unobserved individual-level factors" (Faradmal, Talebi, Rezaianzadeh, et al., 2012:127). It is worth noting that a shared-frailty survival model is the counterpart in survival analysis of the random effects procedure in regression models. In this study, we consider that tourism length of stay is conditioned by tourists' degree of interest in opting for this type of specialist accommodation. Therefore, frailty relates to previous experiences in similar accommodation settings, leading visitors to report lower stays.

A common approach in statistical modelling of tourism length of stay is to adopt a "one size fits all" approach by running an econometric procedure on the sample as a whole, without sufficient regard to the possibility of the existence of relatively distinct segments, or by defining pre-defined segments of tourists based on variables of interest (e.g. nationality or first visitors vs repeat visitors). Finite mixture models permit identification of relevant segments only partially identified by standard analysis based on a-priori segments and are well suited to model unobserved heterogeneity and to "capture individual differences" (Georgea et al., 2013) by allowing the parameters to vary across observations (Theofilatosa & Yannis, 2019). Finite mixture models, as a "purely driven data method" (Georgea et al., 2013) have been increasingly used in a variety of fields, such as economics, biomedicine, and biology (Dyer, Pleck, & McBride, 2012; Georgea et al., 2013; Lanza, Kugler, & Mathur, 2011; Wong & Maffini, 2011). Finite mixture models in the context of this study can be viewed as an extension of standard survival model, as the identification of latent classes is based on a classic survival parametric model. Finite mixture models offer a flexible and attractive approach; however, dressed in a familiar tone, to distinguish between shorter and longer stays and to allow the determinants of tourism length of stay to be different (Rodríguez et al., 2018).

In this study, we employ a range of econometric methods to highlight the fact that such methods generally agree on the sign (strength and direction) of the impact/ association between a set of independent variables and length of stay. Due to the availability of several different assessment methods, one may be left in doubt as to whether a given method is appropriate. Based on the literature, and on the advantages and disadvantages put forward by Soler et al. (2018), one is almost tempted to propose a count data model (probably the negative binomial) as the best approach. Our approach highlights the fact that the methods under analysis lead to similar conclusions.

Based on the literature review, we analysed and tested the following hypotheses and assumptions.

Hypothesis 1. (socio-demographic characteristics). The tourism length of stay is positive related to a number of socio-demographic aspects, such as age, civil status, gender, and educational qualifications (Zimmer, Brayley, & Searle, 1995). In addition to those factors tested in the literature, it is hypothesed that tourism length of stay is a function of the nationality. This is a hypothesis proposed in most studies (Gokovali et al., 2006). We differentiate Portuguese nationals from foreigners, as nationals travelling from the Mainland have access to higher flight frequencies. Foreign tourists are usually dependent on weekly based frequencies. Another reason to differentiate nationalities is that nationals report lower levels of income, a matter of the utmost importance in this type of accommodation characterised by a price premium.

Hypothesis 2. (income constraints). Traditionally tourism length of stay is considered to be negatively related to individual's total expenditure and accommodation daily prices and a positive function of the level of income (Lancaster, 1966; Rosen, 1974). In general, the literature considers that prices, income, and the budget are earmarked in advance for the vacation, and other budget constraints define the budget constraint. In this study, the effect of economic constraints is assessed on the impact on tourism length of stay of the following variables: income; price/quality ratio; accommodation expenditure; and travelling in a low cost air carrier.

Hypothesis 3. (time constraints). Time constraints emerge as a key issue when selecting the number days abroad. In this paper, we hypothesise that retired individuals are allowed greater freedom to opt either for shorter or longer stays (Eugenio-Martin, 2003). Moreover, we also hypothesise that individuals benefiting from lighter time constraints and more upmarket than the average (proxied by the variable travelling by plane and cruise variable), enabling them to travel by plane to a port of call, can opt for longer stays. By contrast, individuals travelling in the Summer (variable Summer), when institutional factors are at stake, will opt for shorter stays also as a result of the increased daily room rates.

Hypothesis 4. (information and motivations). Tourism length of stay is, to a certain extent, influenced by the level of access to information about the destination. In the light of experience and the information possessed on the relevant destination, potential tourists both evaluate alternative destinations and define an adequate time horizon needed to visit all points of interest (Um & Crompton, 1990). Therefore, we analyse in this study a composite variable labelled "sources of information" defined as the sum of the different sources of information accessed (min = 0; max = 6). Furthermore, we also expect that tourists expressing a keen interest to participate fully in the activities available (proxied by the variable "Motivation") and more motivated to explore the resources available will stay for a prolonged period of time.

Hypothesis 5. (travel arrangments and tourist behaviour). Tourism length of stay is a positive function of tourist behaviours, such as being a repeat visitor, previous experience in this type of accommodation and degree of satisfacton reported (Carr, 2002; Kemperman, Borgers, & Timmermans, 2009). Moreover, as mentioned above, we also hypothesise that travelling in the high-season (Summer) leads to shorter stays. Furthermore, in this paper, we assume that satisfaction (as well as accommodation expenditure) is a time-varying variable.

Hypothesis 6. (destination attributes). Tourism length of stay is a positive function of respondent's attitude towards the destination's attractiveness and of the degree of importance attached to a number of key attributes (Barros et al., 2008; Barros, Butler, & Correia, 2009; Oliveira & Pereira, 2008). It is hypothesised the higher the score attached to a number of key attributes, the higher the duration of stay. In this study, we examine the impact of an expressed stronger preference for pleasant weather conditions (variable "Climate") on tourism length of stay.

The discussion about which econometric method is the most appropriate approach to uncover the determinants of LOS has been fuelled by several researchers, notably Thrane (2012), Santos et al. (2015) and Soler et al. (2018). In fact, at least eight competing approaches can be found in the literature (survival analysis, OLS, Poisson regression, negative binomial regression, ordered logit model, binomial logit model, latent class model, dynamic mixed multinomial logit model), with lack of adequate information about the most suitable approach. Researchers such as Wang, Little, and DelHomme-Little (2012) are convinced that survival models are the best option to handle time-related data such as LOS. Santos et al. (2015)

maintain that survival models offer a number of advantages in terms of flexibility of the distributions available to the dependent variable, such as exponential, Weibull and log-logistic, which offers opportunities to estimate better coefficit estimates. Thrane (2012) points out that the advantages of using more complex models, such as survival analysis, do not offset the traditional advantages linked to simpler methods, such as OLS. However, Prebensen et al. (2015) state that the appropriateness of OLS is limited by frequency the violation of the assumptions regarding the dependent variable (e.g. error term normality distributed), a key conclusion also supported by Santos et al. (2015), which prompted the authors to adopt a count model, not the Poisson model, because the assumption equal variance and the mean do not hold, but the less restrictive negative binomial model. In this study, in line with Thrane (2012) and Santos et al. (2015), we tested different models, not to reach a conclusion on which model is the most suitable in pure technical terms, but to highlight that such models are equivalent to a certain extent in qualitative terms.

In summary, survival analysis is well suited to analyse time dependent variables, based on the flexibility offered by the number of distributions available, but the dependent variable must imply the tourist is at risk of "experiencing a transition from one state (staying) to another (leaving)" (Thrane, 2012:127), which is not the case in most instances (e.g. decision taken on length of stay in advance). OLS methods offer easily interpretable results. As observed by Thrane (2012), and we share his point of view, "there appears to be little to gain by discarding the well-known OLS regression model with its intuitive coefficients for the benefit of the more complicated survival models accompanied by their less-than-transparent coefficients". However, as mentioned above, the violation of a number of key assumptions suggests that OLS is not appropriate (*). The characteristics of the dependent variable, a positive integer count variable opens up space to either employ survival methods or count models. Because the application of survival methods is not without criticism, the literature suggests several reasons to opt for count data models, namely the negative binomial regression.

Faced with a number of competing models, it remains to be seen which model to adopt. The results of this study indicate that (sign of the coefficients) regarding the impact of the variables under analysis is almost identical among models. Of course, the methods presented in Table 3 are not "comparable in a strict statistical sense because they are fundamentally different in terms of parameterisation" (Thrane, 2015:182). The only approach is to use qualitative judgment based on the sign and significance of the coefficients. Moreover, based on the qualitative judgment, we can consider the approaches under scrutiny as "comparable".

6. Discussion and results

The empirical approach was retained to follow a three-step strategy. First, we fitted a standard Cox proportional hazard model to the data. The proportional hazards assumption was tested by means of the Schoenfeld residuals, but we failed to find support to validate the proportional hazards assumption. For comparative and illustrative purposes, the results of the Ordinary Least Squares method were also computed (see Table 2).

The proportional hazards assumption represents a major constraint in these types of models. The test based on the Schoenfeld residuals indicates that most variables met the proportional hazards assumption (global p =

Table 2
Cox model, Cox-model invariant variables, Cox stratified model, Gamma model and Ordinary Least Squares.

	Cox model		Cox-model tvc		Cox stratified		Gamma model		Ordinary Least Squares	
	coef.	p.	coef.	p.	coef.	p.	coef.	p.	coef.	p.
Socio-demographic										
Age	0.121**	0.018	0.118**	0.022	0.117**	0.022	−0.013**	0.050	−0.169***	0.006
Income	0.006	0.884	0.003	0.942	0.008	0.835	−0.006	0.256	−0.091*	0.063
Gender	−0.073	0.472	−0.077	0.449	−0.077	0.453	0.022*	0.086	0.024	0.849
Portuguese	0.607**	0.016	0.614**	0.015	0.581**	0.023	−0.104***	0.001	−1.110***	0.000
Civil Status	0.109	0.413	0.104	0.432	0.060	0.665	−0.022	0.195	−0.059	0.715
Work status	−0.321***	0.006	−0.326***	0.005	−0.338***	0.004	0.050***	0.001	0.385***	0.010
Travel Related variables										
Experience	0.165	0.117	0.168	0.110	0.172	0.104	−0.021	0.111	−0.212	0.109
Repeat Visit	0.324***	0.007	0.322***	0.007	…	…	−0.070***	0.000	−0.631***	0.000
Information	−0.151**	0.015	−0.147**	0.018	−0.162***	0.010	0.023***	0.004	0.275***	0.001
PlaneCruiser	−0.499***	0.001	−0.506***	0.001	−0.479***	0.002	0.088***	0.000	0.899***	0.000
Lowcost	−0.048	0.689	−0.039	0.744	−0.057	0.639	0.019	0.237	0.029	0.848
Motivations and Behaviour										
Motivations	−0.022**	0.027	−0.022**	0.028	−0.024**	0.018	0.005***	0.000	0.043***	0.002
Price/Quality	0.072	0.290	0.075	0.271	0.078	0.259	−0.017*	0.063	−0.164*	0.063
Attribute: weather	0.174**	0.039	0.179**	0.034	0.145*	0.090	−0.043***	0.000	−0.314***	0.003
Summer	0.219*	0.088	0.222*	0.092	0.197	0.125	−0.043***	0.008	−0.369**	0.020
Expenditure	−0.004***	0.000			−0.004***	0.000	0.001***	0.000	0.008***	0.000
Satisfaction	−0.183**	0.01			−0.170**	0.012	0.040***	0.000	0.367***	0.000
Constant							1.599***	0.000	2.824***	0.003
Time variant variables										
Satisfaction			−0.173***	0.010						
Expenditure			−0.003***	0.000						
Statistics										
obs	415		415		415		415		415	
log likelihood	−1991.393		−1998.346		−1745.586			230.459		
prob>ch2		0.000		0.000		0.000		0.000		0.000
proportional hazards test	84.45	0.000			82.51	0.000				
θ; LR test θ			0.04							
/lnsigma							−2.052	0.000		
/kappa							0.687	0.000		
Sigma							0.128			
R2									0.923	

Significance levels denotes as follows: ∗∗∗(1%), ∗∗(5%), and ∗(10%).

0,000). However, as we failed to validate the proportional hazards assumption for the model as a whole, we tested alternative modelling specifications, such as the stratified Cox model and time-variant covariates model. The stratified Cox model also failed to hold the proportional hazards assumption. As covariates violating the proportional hazards assumption must be adequately adjusted, we estimated a Cox regression model with time-varying covariates. This specification assumes that we take into account the effect of the impact of time on certain variables. For example, both satisfaction and expenditure can be regarded as time varying variables. We may assume that most visitors begin their stay with positive but cautious expectations in terms of satisfaction. If we assume an initial neutral stance (score 3), with the level of satisfaction increasing at an exponential rate as time goes on, we can estimate that the actual score reported by respondents is proportional to the initial score times exp. $^{0.1t}$, with t reflecting the impact of time. The results of this specification reflecting the impact of time are reported in Table 2. As the reported level of satisfaction increases, the impact of satisfaction in terminating the holiday increases.

Given the failure to hold the proportional hazards assumption, we estimated a number of parametric specifications. The results (data not shown) of five models in the Accelerated Time Failure form in terms of the log likelihood estimation, Akaike Information Criterion values as well as ancillary parameters and Wald tests in order to identify the best model suggests that the gamma model dominates the other models, both in terms of the log likelihood and Akaike information criterion. Therefore, the gamma model is retained in this study.

Given that we examined five econometric approaches "non comparable" "in a strict statistical sense" (Rodríguez et al., 2018; Thrane, 2015), it is important to establish whether they are qualitatively equivalent in terms of the main findings. The various models produce statistically significant results. Globally, the results suggest that the various econometric models are usually consistent with each other. Table 3 presents a summary of the main findings. For reasons of space, we do not show the coefficients and level of significance for the Tobit and Poisson model. However, the main findings are provided in Table 4. The signs of the regression coefficient differ only once across models. That is, with the exception of a single one (low cost), all variables report coefficients with identical signs and striking similarities in terms of statistical significance.

Table 4
Finite mixture model.

	fmm1		fmm2	
	coef.	p.	coef.	p.
Socio-demographic				
Age	0.487***	0.010	−0.243***	0.000
Income	−0.338**	0.026	0.142***	0.008
Gender	0.615	0.102	−0.502***	0.001
Portuguese national	0.436	0.603	0.611*	0.096
Civil Status	−0.020	0.973	−0.246	0.145
Work status	−0.317	0.474	−0.994***	0.000
Travel Related				
Previous experience	0.739**	0.045	−0.221	0.137
Repeat Visit	0.594	0.165	0.529***	0.002
Sources of information	−0.174	0.419	−0.071	0.474
PlaneCruiser	−0.081	0.910	0.109	0.604
Lowcost	1.589***	0.002	−1.042***	0.000
Motivations				
Motivations	−0.086*	0.058	−0.047***	0.005
Price-Quality ratio	0.323	0.136	0.037	0.693
attributes	−0.068	0.828	0.299***	0.008
Summer	0.807	0.110	0.409**	0.033
Accommodation expenditure	−0.042**	0.020	−0.019***	0.000
Satisfaction	−0.296	0.247	−0.346***	0.002
constant	−4.048	0.146	−0.804**	0.011
ln_p	1.129		1.358	

Significance levels denotes as follows: ∗∗∗(1%), ∗∗(5%), and ∗(10%).

Therefore, the discussion of main findings of this study can be based on any of the econometric specifications. However, in the next section we discuss the results of the Gamma model.

The regression results indicate that most variables are statistically significant, except for income, civil status, previous experience and travelling in a low cost carrier. It is, therefore, evident that most variables are significantly related (either positively or negatively) to the hazard rate of ending the journey. Nevertheless, the expected levels of significance, in several cases, are not substantiated and further information is needed to understand the results. For example, the results indicate that the variable income is not statistically significant. Usually a higher income tends to lead to

Table 3
Compared analysis.

	Cox		Cox strat.		Cox t.c.		Gamma		Ordinary Least Squares		Tobit		Poisson	
	sign	stat.	sign	stat.	sign	stat.	sign	stat.	sign	stat.	sign	stat.	sign	stat.
Socio-demographic														
Age	−	yes	−	yes	−	yes	−	yes	−	yes	−	yes	−	no
Income	−	no	−	no	−	no	−	no	−	yes	−	yes	−	no
Gender	+	no	+	no	+	no	+	yes	+	no	+	no	+	no
Portuguese	−	yes	−	yes	−	yes	−	yes	−	yes	−	yes	−	yes
Civil Status	−	no	−	no	−	no	−	no	−	no	−	no	−	no
Work status	+	yes	+	yes	+	Yes	+	yes	+	yes	+	yes	+	yes
Travel Related														
Experience	−	no	−	no	−	no	−	no	−	no	−	no	−	yes
Repeat Visit	−	yes	…	yes	−	yes	−	yes	−	yes	−	yes	−	yes
Information	+	yes	+	yes	+	yes	+	yes	+	yes	+	yes	+	yes
PlaneCruiser	+	yes	+	yes	+	yes	+	yes	+	yes	+	yes	+	yes
Lowcost	+	no	+	yes	+	no	+	no	+	no	+	no	−	no
Motivations and Behaviour														
Motivations	+	yes	+	yes	+	yes	+	yes	+	yes	+	yes	+	yes
Price/Quality	−	no	−	yes	−	yes	−	yes	−	yes	−	yes	−	yes
Attributes	−	yes	−	yes	−	yes	−	yes	−	yes	−	yes	−	yes
Summer	−	yes	−	yes	+	yes	−	yes	−	yes	−	yes	−	yes
Acc. Expenditure	+	yes	+	yes			+	yes	+	yes	+	yes	+	yes
Satisfaction	+	yes	+	yes			+	yes	+	yes	+	yes	+	yes
tvc														
Satisfaction					+	yes								
Expenditure					+	yes								

longer stays (Wang et al., 2012), although accommodation costs can be interpreted as a proxy for respondents´ level of income. From this point of view, our results corroborate past studies because the higher the accommodation costs (intrepreted as an indicador of economic capacity), the longer the duration of stay. Meanwhile, the impact of income may also be analysed with reference to the Portuguese nationality, reporting the lower levels of income. Portuguese nationality impacts negatively on tourism length of stay, which suggests that income constraints are of relevance to predict tourism length of stay. In contrast, being a female and a retired person leads to longer stays.

The negative coefficient associated with age may be explained on work responsibilities grounds, if we consider that, in line with expectations, retired people stay longer (Alén et al., 2014). Until retirement age, increasing work responsibilities at the managerial level may force respondents to micro-manage their stays abroad by staying just one week (on average), although they take several breaks per year. As a consequence, older (middle aged) but not retired tourists, performing more demanding functions in terms of management tasks and leadership until they reach retirement age, are forced to opt for shorter stays compared to younger tourists performing less demanding functions. That is, older respondents in their late forties and fifties undertaking managerial and directorship roles may deem it to be inappropriate to take a two week-long break. Wang et al. (2012) found that highly qualified tourists (with postgraduate education qualifications) reported shorter stays. As shown in Table 1, age defined as a categorical variable is characterised by a mean of 55.3 years old, with two thirds of the sample in the 40–59 year old cohort.

In all, our results concerning age do not appear to corroborate previous findings, because age impacts negatively on length of stay. Alegre and Pou (2003), Fleischer and Pizam (2002) and Barros et al. (2009) found a positive impact of age on LOS. Issues, such as lack of time constraints, were proposed to explain the positive relationship identified. However, a close analysis of results suggests that the main findings reported in other studies are also of relevance in the context of this study. Gender appears to influence tourism length of stay, with females staying, on average, longer than males, which corroborates previous studies, such as Barros and Machado (2010), but contradicts the majority of the studies available. As mentioned above, we find that education is not associated with tourism length of stay, which contradicts past studies, notably Barros and Machaod, Santos et al. (2015) and Rodríguez et al. (2018). The same applies to being married (*). Neverthless, to be a Portuguese national impacts length of stay negatively, in line with past evidence found by *. Therefore, Hypothesis 1 is supported, to a certain extent, by the results.

Moreover, booking a hotel room in the high-season (Summer), characterised by heavily overpriced daily rates, leads in the same diretion. Also, the higher the importance attached to the variable “price”, which points a price-conscious visitor, the lower is the tourism length of stay. In each case, the higher the levels of available income, the longer the duration of stay, which corroborates past studies, notably Barros and Machado (2010) and Martinéz-Garcia and Raya (2008). To have access to higher levels of discretional income offers the opportunity to enjoy more prolonged stays while housed in a luxury hotel. Therefore, Hypothesis 2 is supported by the data.

In terms of time constraints, we notice that travelling by plane plus cruise has a significant positive impact on tourism length of stay. Past research indicates that cruiser tourists report higher levels of income, and are more likely to be out of the labour market, compared to other tourists (Brida et al., 2014). Respondents who travel to Madeira in the high-season, which is more likely to be the case of individuals still in the labour market, are more likely to opt for shorter stays, while tourists who arrive in the the mid and low-season stay the longest. Our results contradict Thrane (2012), Rodríguez et al. (2018)and Martinéz-Garcia and Raya (2008) because travelling in the summer leads to shorter stays. The summer months offer the opportunity to undertake extended trips to tourist destinations. However, either time constraints or budgetary considerations are at work in this case. Hypothesis 3 is therefore supported by the results.

In line with expectations, respondents finding the destination more attractive in a number of items, are more likely to stay longer. The predictive mean of tourism length of stay for respondents in the 25th quartil in terms of motivation is 10.15. The similar figure for the 75 quartil is 10,53. In the same vein, tourists reporting higher levels of access to information are similarly more inclined to stay longer. In this vein, this study corroborates past studies reporting a positive relationship between the degree of importance attached to a number of attributes and length of stay; notably, Brida et al. (2013), Alén et al. (2014). Hypothesis 4 is therefore corroborated in this study.

Visitors reporting previous experience in this type of accommodation report lower levels of tourism length of stay but the coefficient is not statistically significant and visitors are more likely to stay longer. The results suggest that respondents categorised as repeat visits are similarly more likely to stay longer (10.8 vs 10.1 in terms of the predicted mean of the dependent variable), which can be explained on place-attachment grounds. It is worth mentioning again that this type of specialist accommodation is relatively expensive if compared with the average daily price charged by competitors. However, as the factor “motivation” was found to impact tourism length of stay positively, we may tentatively consider, for example, that high levels of reported interest in exploring the resources and experiences available are translated into longer stays. Finally, the factor satisfaction is also significant, as visitors reporting high levels of satisfaction stay longer. Therefore, this study corroborates past studies, notably Gokovali et al. (2006). Therefore, Hypothesis 5 is supported.

Regarding the destination attributes, past research indicates that some key aspects, such as weather, availability of cultural resources and sport or entertainment activities, are a significant predictor of tourism length of stay (See Barros & Machado, 2010; Brida et al., 2013; Alén et al., 2014; Peypoch et al., 2012; Rodríguez et al., 2018). The variable weather exibits a negative and significant relation with tourism length of stay, as discovered by Barros and Machado (2010). The appealing attributes of the destination were found by Alén et al. (2014) to impact the length of stay positively, and Rodríguez et al. (2018) discovered that the degree of attractiveness of Santiago of Compostela positively impacted the duration of stay. However, in this study, visitors travelling to Madeira to enjoy sunny days and mild temperatures are less likely to stay longer. This pattern might be related to time constraints, because sun-lovers may prefer to travel in Summer, which imposes travel restrictions due to institutional and economic factors. Hence, we fail to find evidence to support Hypothesis 6.

Both the availability of sufficient financial own resources and time availability issues appear to be the determinant issues at stake. Travelling in the summer leads to shorter stays; in all likelihood, as a result of higher prices charged by hotels and institutional constraints. Travelling both by cruiser and plane implies that respondents in such circumstances are not constrained either by economic resources or by time restrictions. Therefore, they can stay longer. Hypothesis saving time to access information and to enjoy the destination (translated in higher levels of satisfaction) also leads to longer stays. By the same token, reporting higher levels of motivation to explore the island's resources also leads to longer stays.

Given that we report unexpected outcomes regarding a number of variables, it may be worth considering an alternative route in terms of estimating the impact of the variables under analysis. It could be the case that several respondents behave in a certain way, while others behave in the opposite when confronted with similar stimulus or market conditions. The finite mixture model's procedure offers an opportunity to test whether a given variable produces comparable results in different segments of respondents. The finite mixture model's procedure leads to the identification of two meaningful classes in terms of tourism length of stay: class 1 with 20.98% of probability; class 2 with 79.02% of probability. The optimal number of latent classes was determined according to the Akaike Information Criterion and Bayesian Information Criterion, with lower Bayesian Information Criterion and Akaike Information Criterion values indicating better fit. Table 5 displays the results of the finite mixture models. These results indicate that the factors affecting tourism length of stay affect each segment differently, with even opposite signs.

Table 5
Basic *t*-tests and chi-squared test.

Variable	Class 1	Class2	Tests and statistics
Length of Stay	14.76 days	9.79 days	$T = 4.683$; sig = 0.000
Price-quality	4.11	4.14	T = −0.024;sig = 0.803
Income	3.72	4.39	T = −2.940;sig = 0.000
Age	5.46	5.86	T = −1.829;SIG = 0,073
Satisfaction	4.43	4.44	T = −0.044:sig = 0.965
Daily rate	116.5	117.237	$T = 1.639$;sig = 0.809
Sources of information	1.261	1.447	T = −1.332;sig = 0.189
Factor history	14.369	14.126	$T = 0.468$;sig = 0.641
Low levels of income	21.739%	10.298%	$\chi = 5.234$;sig = 0.022
High levels of income	8.696%	21.951%	$\chi = 4.413$;sig = 0.036

Most variables were found to be significant; nonetheless, a sizeable number of them fail to reach statistical significance. Far-reaching differences in terms of statistical significance are also evident. Firstly, it should be noted that we can clearly identify two groups (see Graph 4).

As highlighted in Graph 4, the "longer stays" group (class 1) exhibits more variability, while class 2 clusters around the mean. Class 1 exhibits an average tourism length of stay of 14.76 days and class 2 reports an average tourism length of stay of 9.79 days. Class 1 concentrates mainly constraint-free individuals staying 14 days or more (67.4% of the cases). The pike in terms of density is attained for 14 days. Class 2 concentrates individuals reporting a tourism length of stay clustered at 9–10 days, with 53.4% of the individuals reporting seven to eight days. Moreover, class 2 concentrates 80% of all cases reporting short stays (fewer than seven days). A closer look at the two classes indicated that around 93% of the respondents stayed for a one to two week period (between seven and 14 days with an average of placed 10 days).

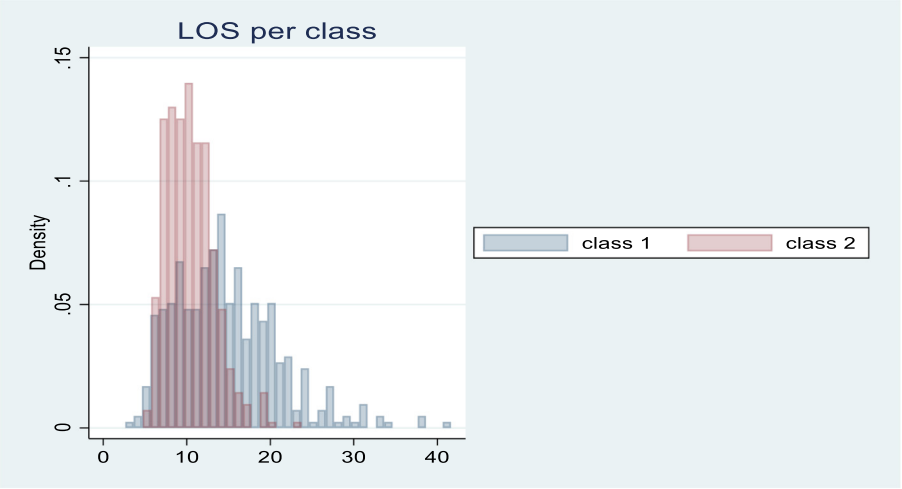
Regarding class 1, higher levels of income lead to a lower hazard rate and, therefore, to longer stays. However, travelling in a low cost carrier, interpreted in this study as a proxy for heavy economic constraints, leads to shorter stays. Reporting higher levels of motivation (and accommodation costs) travel lead to longer stays, while reporting previous experience leads in the opposite direction. In summary, the key dominant factor in class 1 appears to be economic issues. As depicted in Table 5, class 1 is characterised by a high number of low-income individuals and by income levels above average.

The results provided so far are reasonably in line with expectations. It is noteworthy that a number of authors link higher levels of income to longer stays (Fleischer, Peleg, & Byk, 2011; Mortazavi & Cialani, 2017; Nicolau & Más, 2009). However, the fundamental problem with class 1 is that longer stays are associated with levels of income below average. This is an issue for Destination Management Organisations eager to welcome highly motivated individuals (which is the case); however, those individuals are more likely to be "below average income" visitors. From the findings, we can also identify a number of statistically insignificant variables; for example, being a Portuguese national, married and retired, as well as being a repeat visitor, and benefiting from previous experience, importance, sources of information, travelling in a plane plus cruise. Class 2 appears to be driven by a different approach. Higher levels of income (in all likelihood associated to high status occupations) lead to shorter stays, for the reasons explained above. In contrast, being older or retired leads to longer stays. Reporting higher levels of access to information and satisfaction leads to longer stays. Being a female and travelling in a low cost air carrier also leads to longer stays, whereas being a Portuguese national leads to shorter stays.

In common between the two segments, several differences are strikingly evident in terms of coefficients displaying opposite signs, which warrant attention. For example, travelling in a low cost carrier leads to shorter stays in class 1, but to longer stays in class 2. In this instance, two separate groups are clearly identifiable. The short stays group exhibits more variability, as it comprises both on-average and longer stays. By contrast, observations included in segment 2 cluster around the mean. The analysis provided thus far is very valuable in providing assurance on how to interpret the determinants of stay and on how to define priorities. In essence, class 2 rallies around 80% of the respondents, while class 1 includes 20% of the sample, which suggests that most attention must be mainly channelled towards class 2. This may appear counterintuitive from the point of view of Destination Management Organisation officials, eager to contribute to more inflated figures on tourism length of stay.

As it stands, based on the impact of motivation and access to information, the local Destination Management Organisations and the sector have considerable scope for manoeuvrability and improvement in the medium and long term. However, in the short term, the only actions likely to positively impact tourism length of stay lie in the field of production and dissemination of useful and attractive information, in order to stimulate the desire "motivation" to visit the island and to encourage visitors to stay longer and to visit and experience more, and in ensuring high levels of quality and satisfaction among visitors based on the current advantages. The hotel's management is therefore required to excel in terms of well-staged experiences to guarantee high levels of satisfaction. Moreover, as female retired persons and older individuals are more inclined to stay longer, operators need to pay attention to such individuals.

Possibly, that is what can be done because several variables are not manageable to any substantial extent. One could be inclined to say that the identification of two segments of visitors increases our knowledge and ability to intervene to increase tourism length of stay. In effect, the finite mixture models approach offers the opportunity to handle heterogeneous data characterised by a finite number of unobserved segments,



Graph 4. Predicted tourism length of stay (days) per group.

which is the case. The finite mixture models procedure also offers ground to identify patterns based on classification and clustering analysis. However, an in-depth analysis of each segment, based on *t*-tests and chi-tests, shows that their share similar figures for a number of variables of interest.

Although satisfaction, access to information, importance attached to "price/quality ratio" etc. are undifferentiated, class 1 includes a high proportion of older individuals earning below average income. In contrast, class 2 includes a large number of middle-aged individuals who earn above average income. Nevertheless, as class 2 outnumbers class 1 by a 4 to1 ratio we take a closer look at class 2. Age positively impacts tourism length of stay, which means that ageing contributes to longer stays. Based on the demographic projections currently available, within the foreseeable future, old people will be the most representative segment of tourists and longer stays will predominate. Developing packages and services designed for the most demanding customers in the older age cohorts will offer scope to gain valuable experience in this market segment. The development, management and upgrading of programs aimed at the elderly people segment must be one of the priorities. The production of good quality information needs to be be another priority, as access to information increases tourism length of stay. This circumstance may arise because information creates an unfulfilled curiosity on the part of the respondents as to its transformation into actual activities and experiences.

7. Conclusions and implications

In this study, we carried out several econometric estimations to identify the determinants of tourism length of stay of visitors opting for a type of specialist accommodation, boutique accommodation. The focus on a specialist market niche is important, as understanding the factors that drive tourists to stay longer can help Destination Management Organisations and operators to devise strategies aimed at increasing tourism length of stay, especially when the decline of tourism length of stay at the destination has become the new norm. The estimation of the Cox proportional model was followed by the test of the basic assumption of proportional hazard. However, the results showed that the assumption was not confirmed, which led us to test a number of parametric models in addition to various special specifications within the scope of the Cox model. The gamma distribution was retained as the most suitable in line with the statistical criterion of "best-fit". Different approaches produced similar results in terms of the relationship between the covariates under analysis and tourism length of stay.

The results are, to a certain extent, in line with previous research. However, in several instances, the results contradict a considerable body of research, which can be explained on the grounds of dealing with a specific market segment. This market segment, characterised by a price premium, attracts an affluent clientele with sufficient resources to indulge in more leisure time and longer holidays. Unsurprisingly, above average daily room rates are not a relevant constraint in such circumstances. However, high levels of income are also associated with high-status occupations which limit the availability to stay on slightly longer. It was demonstrated that the operators need to target highly motivated tourists, retired persons and visitors travelling by plane and cruise. Moreover, the availability of reliable and precise information should be regarded as a winning strategy because the higher the level of information accessible, the longer the stay.

We also tested the covariates under analysis via finite mixture model approach. In reality, the results point to two different segments, with several factors holding a different relation with tourism length of stay. The results allow us to profile the main characteristics of each segment class 1, compared to class 2, and incorporate individuals reporting higher levels of tourism length of stay, but below average income. By contrast, tourists included in class 2 exhibit higher score in terms of income, but shorter stays. In any case, it should be considered that one segment (class 2) accounts for 80% of the sample.

According to the results, it is possible to devise recommendations for initiatives to increase tourism length of stay. As the results indicate that age positively affects tourism length of stay, operators need to implement actions aimed at encouraging this market segment by highlighting several characteristics of the accommodation that might appeal to its members from the number of elderly people on the sample. The above average "mean age" indicates that such issues should be highlighted, such as the proximity between most hotels and many key points of interest in the island's main city.

As superior motivation to visit the island prompts visitors to stay longer, the promotion abroad needs to excel in highlighting the overall attraction of the island. A practical way to increase the region's attraction lies in using internet related sources of information and coordinated promotion activity centred on highlighting the main charateristics of this type of accommodation, in addition to other main motives to travel to the island.

Another important issue concerns satisfaction. For all purposes, satisfaction cannot be considered an antecedent of tourism length of stay; consequently, it can be difficult to discuss the impact of increasing levels of satisfaction on the probability of ending the journey. However, 46.3% of the respondents stay for periods longer than seven to eight days. Therefore, it cannot be disregarded that some respondents book a number of days in addition to those originally envisaged because they were informed that, if they feel relaxed and satisfied, they have the option to extend this type of specialist accommodation.

As our results corroborate our earlier diagnostic hypothesis of more than one segment included in the sample, actions should be pursued to target the most promising segment. Given the budget restrictions to which respondents are subject, it is strongly advisable to increase the quality price ratio of the accommodation services provided by the local industry. The most promising route in this regard is likely to be the organisation of thematic sessions aimed at highlighting the historical, architectonic, botanic, environmental and landscape aspects of the architecture and surrounding areas. The involvement of guests in cultural events, music festivals and gastronomic initiatives occurring locally can be promoted by granting easy access to such initiatives. Managers can also take advantage of other natural and cultural services available on the island based on tourist routes organised by third parties, such as the "Madeira Sugar Cane Route" or the "Madeira Wine Route".

As different models appear to produce apparently related (but not identical) results, the practice of estimating the determinant of tourism length of stay through the lens of two or more econometric approaches must be pursued to produce the minimum consensus in terms of the most relevant variables. Furthermore, it is highly likely that in most cases, the distribution of the variable of interest is a mixed one, generated by two or more distributions. Given the specific geographical setting and market niche under analysis of this study, caution must be taken when comparing our results with those provided in previous papers. Nevertheless, results provided so far do support widespread views regarding the relevance of several socio-demographic and travel related variables in analysis of the determinants of tourism length of stay.

An important limitation to this study concerns the treatment of variable motivation. Motivations are analysed via a proxy variable representing the sum of the scores attached to several attributes/reasons to visit the island. Although less than ideal from a theoretical point of view, it follows the practice in the literature and leads to the conclusion that higher levels of visitors' degree of interest or positive image results in extended stays.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Aguilar, M., & Díaz, B. (2019). Length of stay of international tourists in Spain: A parametric survival analysis. *Annals of Tourism Research, 79*, 102768.

Aguiló, E., Alegre, J., & Sard, M. (2005). The persistence of the sun and sea tourism model. *Tourism Management*, *26*, 219–231.
Alegre, J., Mateo, S., & Pou, L. (2011). A latent class approach to tourists' length of stay. *Tourism Management*, *32*(3), 555–563.
Alén, E., Nicolau, J., Losada, N., & Domínguez, T. (2014). Determinant factors of senior tourist's length of stay. *Annals of Tourism Research*, *49*, 19–32.
Almeida, A., & Garrod, B. (2018). A CATREG model of destination choice for a Mature Island destination. *Journal of Destination Marketing and Management*, *8*, 32–40.
Barros, C., & Machado, L. (2010). The length of stay in tourism. *Annals of Tourism Research*, *37*, 692–706.
Barros, C., Correia, A., & Crouch, G. (2008). Determinants of the length of stay in Latin American tourism destinations. *Tourism Analysis*, *13*, 329–340.
Barros, C., Butler, R., & Correia, A. (2009). The length of stay of golf tourism:A survival analysis. *Tourism Management*, *31*, 13–21.
Boto-García, D., Baños-Pino, J., & Álvarez, A. (2018). Determinants of tourists' length of stay: A hurdle count data approach. *Journal of Travel Research*, *58*(6), 977–994.
Box-Steffensmeier, J., & Jones, B. (2004). *Event history modeling. A guide for social scientists*. Cambridge: Cambridge University Press.
Brida, J., Meleddu, M., & Pulina, M. (2013). Factors influencing length of stay of cultural tourists. *Tourism Economics*, *19*, 1273–1292.
Brida, J., Pereyra, J., & Scuderi, R. (2014). Repeat tourism in Uruguay: Modelling truncated distributions of count data. *Quality and Quantity*, *48*, 475–491.
Carr, N. (2002). The tourism–leisure behavioural continuum. *Annals of Tourism Research*, *29*(4), 972–986.
Dyer, W., Pleck, J., & McBride, B. (2012). Using mixture regression to identify varying effects: A demonstration with parental incarceration. *Journal of Marriage and Family*, *74*, 1129–1148.
Esiyok, B., Kurtulmuşoğlu, F., & Özdemir, A. (2018). Hypothesis eterogeneity in the determinants of length of stay across middle age and senior age groups in thermal tourism. *Journal of Travel & Tourism Marketing*, *35*(4), 531–540.
Eugenio-Martin, J. (2003). Modelling determinants of tourism demand as a five-stage process: A discrete choice methodological approach. *Tourism and Hypothesis ospitality Research*, *3*, 341–354.
Faradmal, J., Talebi, A., Rezaianzadeh, A., et al. (2012). Survival analysis of breast cancer patients using cox and frailty models. *Journal of Research in Hypothesis ealth Sciences*, *12*, 127–130.
Fleischer, A., Peleg, G., & Byk, J. R. (2011). The impact of changes in household vacation expenditures on the travel and hospitality industries. *Tourism Management*, *32*(4), 815–821.
Georgea, M., Yang, N., Jaki, T., Feaster, D., Lamont, A., Wilson, D., & Hypothesis orn, M. (2013). Finite mixtures for simultaneously modelling differential effects and non-normal distributions, multivariate. *Behav Res*, *48*(6), 816–844 November.
Gokovali, U., Bahar, O., & Kozak, M. (2006). Determinants of length of stay: A practical use of survival analysis. *Tourism Management*, *28*, 736–746.
Gómez-Déniz, E., & Pérez-Rodríguez, J. (2019). Modelling bimodality of length of tourist stay. *Annals of Tourism Research*, *75*, 131–151.
Gutierrez, R. (2002). Parametric frailty and shared frailty survival models. *Stata Journal*, *2*, 22–44.
Hong, S., & Jang, H. (2005). Factors influencing purchasing time of a new casino product and its managerial implications: An exploratory study. *Journal of Travel Research*, *43*, 395–403.
Kemperman, A., Borgers, A., & Timmermans, H. (2009). Tourist shopping behavior in a historic downtown area. *Tourism Management*, *30*(2), 208–218.
Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*, 132–157.
Lanza, S., Kugler, K., & Mathur, C. (2011). Differential effects for sexual risk behavior: An application of finite mixture regression. *Open Family Studies Journal*, *4*, 81–88.
Machado, L. (2010). Does destination image influence the length of stay in a tourism destination? *Tourism Economics*, *16*, 443–456.
Martinéz-Garcia, E., & Raya, J. (2008). Length of stay for low cost tourism. *Tourism Management*, *29*(6), 1064–1075.
Menezes, A., Moniz, A., & Vieira, J. (2008a). The determinants of length of stay of tourists in the Azores. *Tourism Economics*, *14*, 205–222.
Menezes, A., Moniz, A., & Vieira, J. (2008b). The determinants of length of stay of tourists in the Azores. *Tourism Economics*, *14*, 205–222.
Mortazavi, R., & Cialani, C. (2017). International tourists' length of overnight stay in Venice. *Tourism Economics*, *23*(4), 882–889.
Nicolau, J., & Más, F. (2009). Simultaneous analysis of whether and how long to go on holidays. *The Service Industries Journal*, *29*(8), 1077–1092.
Nicolau, J., Zach, F., & Tussyadiah, I. (2016). Effects of distance and first-time visitation on tourists' length of stay. *Journal of Hypothesis Ospitality and Tourism Research*, *42*(7), 1023–1028.
Oliveira, P., & Pereira, P. (2008). Who values what in a tourism destination? The case of Madeira Island. *Tourism Economics*, *14*(1), 155–168.
Prebensen, N., Altin, M., & Uysal, M. (2015). Length of stay: A case of Northern Norway, Scandinavian. *Journal of Hypothesis Hospitality and Tourism*, *15*(sup1), 28–47.
Rodríguez, X., Martínez-Roget, F., & González-Murias, P. (2018). Length of stay: Evidence from Santiago de Compostela. *Annals of Tourism Research*, *68*, 9–19.
Rosen, S. (1974). Hypothesis edonic prices and implicit markets: Products differentiation in pure competition. *Journal of Political Economy*, *82*, 34–35.
Salmasi, L., Celidoni, M., & Procidano, I. (2012). Length of stay: Price and income semi-elasticities at different destinations in Italy. *International Journal of Tourism Research*, *14*(6), 515–530.
Santos, G., Ramos, V., & Rey-Maquieira, J. (2015). Length of stay at multiple destinations of tourism trips in Brazil. *Journal of Travel Research*, *54*, 788–800.
Soler, I., Gemar, G., & Correia, M. (2018). Length of stay for tourists' inland trips. *Journal of Destination Marketing & Management*, *10*, 49–60.
Theofilatosa, A., & Yannis, G. (2019). Exploring crash injury severity on urban motorways by applying finite mixture models. *Transportation Research Procedia*, *41*, 480–487.
Thrane, C. (2012). Analyzing tourists' length of stay at destinations with survival models: A constructive critique based on a case study. *Tourism Management*, *33*(1), 126–132.
Thrane, C. (2015). Research note: The determinants of tourists' length of stay: Some further modelling issues. *Tourism Economics*, *21*(5), 1087–1093.
Thrane, C., & Farstad, E. (2012). Tourists' length of stay: The case of international summer visitors to Norway. *Tourism Economics*, *18*(5), 1069–1082.
Um, S., & Crompton, L. (1990). Attitude determinants in tourism destination choice. *Annals of Tourism Research*, *17*, 432–448.
Varian, H. (1987). *Intermediate microeconomics: A modern approach*. New York and London: W. W. Norton & Co.
Wang, E., Little, B., & DelHomme-Little, B. (2012). Factors contributing to tourists' length of stay in Dalian northeastern China — A survival model analysis. *Tourism Management Perspectives*, *4*, 67–72.
Wong, Y., & Maffini, C. (2011). Predictors of Asian American adolescents' suicide attempts: A latent class regression analysis. *Youth Adolescence*, *40*, 1453–1464.
Yang, J., & Liu, A. (2003). Frequent flyer program: A case study of China airline'smarketing initiative—Dynasty flyer program. *Tourism Management*, *24*, 587–595.
Yang, Y., & Zhang, H. L. (2015). Modeling tourists' length of stay: Does one model fit all? *Tourism Analysis*, *20*, 13–23.
Zimmer, Z., Brayley, R., & Searle, M. (1995). Whether to go and where to go: Identification of important influences on seniors' decisions to travel. *Journal of Travel Research*, *33*(3), 3–10.

António Almeida works at the Management & Economics, Universidade da Madeira. His research interests include tourism length of stay, natural disasters and Destination Climate Change.

Luiz Pinto Machado works at the Management & Economics, Universidade da Madeira. Luis does research in Communication and Media, Tourism and Business Economics and Business Administration.

Chen Xu works in the Economics section of Cardiff University, her research interests include Tourism Satellite Account, tourism length of stay and tourism forecasting.