UNIVERSITAT DE LES ILLES BALEARS

FACULTAT D'ECONOMIA I EMPRESA

DEPARTAMENT D'ECONOMIA APLICADA

MULTIDESTINATION TRAVEL:
A CRITICAL THEORETICAL APPROACH AND EMPIRICAL EVIDENCE FOR THE CASE OF BRAZIL

Doctoral Thesis

Glauber Eduardo de Oliveira Santos

Directors:

Vicente Ramos Mir

Javier Rey-Maquieira Palmer

Programa Oficial: Economía del Turismo y del Medio Ambiente

Universitat de les Illes Balears

2012

Tesis Doctoral

Título:

MULTIDESTINATION TRAVEL: A CRITICAL THEORETICAL APPROACH AND EMPIRICAL EVIDENCE FOR THE CASE OF BRAZIL

Autor:

Glauber Eduardo de Oliveira Santos

Directores:

Vicente Ramos Mir

Javier Rey-Maquieira Palmer

Programa Oficial: Economía del Turismo y del Medio Ambiente

Universitat de les Illes Balears

2012

Acknowledgements

I would like to express my gratitude to my supervisors, Vicente Ramos Mir and Javier Rey-Maquieira Palmer, whose support, insights, criticisms and guidance were essential to the development of this thesis.

I also recognize the important role played by the professors of the Master in Tourism and Environmental Economics of the University of the Balearic Islands. Their lectures and ideas were of great value to my professional and personal development.

I acknowledge the financial support I received from the "Càtedra Sol Meliá d'Estudis Turistics" and the leave of absence conceded by the Federal Institute of Education, Science and Technology of São Paulo. I also acknowledge the provision of data from the Foundation Institute of Economic Research and the Tourism Ministry of Brazil.

I am profoundly indebted of Professor Wilson Abrahão Rabahy, who introduced me to the study of tourism economics and provided me enormous support since then.

Finally, my special thanks to my wife, parents and friend for their patience and encouragement.

Abstract

Multidestination tourism trips (MTTs) can be defined as a single tourism trip in which more than one destination is visited. Regarding the academic literature on MTTs, three main lacks are studied in this thesis. First, economic theoretical models of tourist choices developed hitherto are insufficient to explain MTTs consumption. On one hand, traditional microeconomic models of tourist choices, such as Morley (1992) and Stabler, Papatheodorou and Sinclair (2010), were developed within the single destination paradigm, overlooking the possibility of tourists choosing to visit multiple destinations in a single trip. On the other hand, available economic theoretical models of consumer choices within the MTT paradigm, such as Rugg (1973) and Tussyadiah, Kono and Morisugi (2006), were based on the Lancaster (1966) characteristics theory. However, two important criticisms can be made to this approach. Lancaster's theory does not allow for the consideration of singularities, overlooking the importance of uniqueness of tourist destinations. Besides, Lancaster's theory seems to add unnecessary issues to the discussion. Characteristics are argued to be an unrequired concept to explain why people have interest in visiting more than one destination in a single trip. Therefore, an alternative theoretical model for consumers' choices in the MTT paradigm based on the traditional economic theory is presented in this thesis. Some particularities of the traditional model regarding MTT are discussed, such as the proper shape of utility functions, special transport costs and the existence of two budget constraints. The model shows that the MTT paradigm may unfold special demand effects as compared with the single destination paradigm. Negative income effects may be experienced by destinations considered as normal goods, and positive price effects may be experience by non-Giffen destinations.

Second, despite of the relatively large theoretical literature on the determinants of MTTs consumption, a few studies had analysed this topic from an empirical perspective. The most relevant studies in this field are those conducted by Tideswell and Faulkner (1999, 2003) and Nicolau and Más (2005). This thesis develops an empirical analysis of the individual determinants of MTTs consumption, assessing theoretical propositions and empirical findings of previous studies. Inbound tourism in Brazil is

empirically studied. The earlier set of explanatory variables is extended to include some additional determinants, such as education level, type of accommodation and season. The time evolution of the dependent variable is also assessed. A censored zero-inflated negative binomial model is employed in order to overcome some econometric deficiencies of previous studies. The findings enlighten theoretical conflicting arguments proposed in the literature, especially those regarding the effect of monetary and time constraints. A qualitative difference between single and multiple destination trips was found. Hence, tourists' decisions to take single or multiple destination trips are shown to be different and somehow detached from the decision of how many destinations to visit.

Third, tourists' lengths of stay at different locations of a multidestination trip are studied. The determinants of this variable was examined through a shared heterogeneity duration model where the set of independent variables included individuals' characteristics, travels' attributes, characteristics and identification of visited destinations and a set of dummy variables to capture variations over time. Five main conclusions were derived. First, several relevant empirical evidences about the effects of explanatory variables were provided. Some findings of earlier studies were confirmed and some were contradicted. Some examples of diverging results are those regarding the estimated effects of income, party size and previous visits to the destination. Income was not found to be a significant determinant of tourists' length of stay within the MTTs paradigm in the Brazilian case. The length of stay was found to follow a convex function of party size. Tourists who visit Brazil for the first time are expected to stay shorter than those who are repeating their visits, while additional previous trips for repeaters have a negative effect on the expected length of stay. Second, the increasing trend of international tourists' lengths of stay at Brazilian destinations was proved to be caused by changes in the composition of the inbound tourism flow. Third, the average length of stay was found to vary across regions and states, pointing out destinations that might be used as benchmarks for tourism management. Tourists' length of stay was also found to vary according to some destinations' characteristics. Fourth, positively skewed density distributions of tourists' length of stay were found to be appropriate. Finally, shared heterogeneity across observations statistically improves the explanatory capacity of duration models when multidestination tourism trips data is analysed.

Resumen

Los viajes de turismo a multidestinos (MTTs) pueden ser definidos como un viaje en que se visita más de un destino. La presente tesis estudia tres omisiones principales de la literatura académica sobre MTTs. En primer lugar, los modelos económicos teóricos de las elecciones de los turistas desarrollados hasta ahora son insuficientes para explicar el consumo de MTTs. De un lado, los modelos microeconómicos teóricos de las elecciones de los turistas, como Morley (1992) y Stabler, Papatheodorou y Sinclair (2010), se basan en el paradigma de viajes a destinos únicos, pasando por alto la posibilidad de que los turistas visiten múltiples destinos en un único viaje. De otro lado, los modelos económicos teóricos de las elecciones de los turistas basados en el paradigma de MTTs, como Rugg (1973) y Tussyadiah, Kono y Morisugi (2006), están basados en la teoría de las características de Lancaster (1966). Dos importantes críticas pueden ser hechas a ese enfoque. La teoría de Lancaster no permite la consideración de singularidades, pasando por alto la importancia de las peculiaridades de los destinos turísticos. Además, la teoría de Lancaster parece añadir problemas innecesarios a la discusión. Las características no son un concepto necesario para explicar por qué las personas tienen interés en visitar más de un destino en un viaje. Así, en esta tesis se desarrolla un modelo teórico alternativo de las elecciones de los turistas en el paradigma de MTTs y con base en la teoría económica tradicional. Algunas particularidades del enfoque económico tradicional aplicado a las MTTs son discutidas, como la forma adecuada de las funciones de utilidad, especificidades de los costes de transporte y la existencia de dos restricciones presupuestarias. El modelo revela que el paradigma de MTTs puede presentar efectos especiales en la demanda comparado al paradigma de viajes a destinos únicos. Efectos renta negativos pueden ocurrir para destinos considerados como bienes normales, y efectos positivos del precio pueden ocurrir para destinos que no son bienes de Giffen.

En segundo lugar, a pesar de que la literatura teórica sobre los determinantes del consumo de MTTs es relativamente extensa, pocos estudios examinaron ese tema desde una perspectiva empírica. Los estudios más relevantes en ese campo son Tideswell y Faulkner (1999, 2003) y Nicolau y Más (2005). La presente tesis desarrolla un estudio empírico de los determinantes individuales del consumo de MTTs, examinando proposiciones teóricas y resultados empíricos de estudios anteriores. El turismo receptivo internacional de Brasil es estudiado empíricamente. El conjunto de variables explicativas utilizado por estudios anteriores es ampliado para incluir determinantes adicionales, como el nivel de educación, el tipo de alojamiento y la

estación del año. La evolución de la variable dependiente en el tiempo también es examinada. Un modelo binomial negativo cero-inflado censurado es utilizado para superar algunas deficiencias econométricas de estudios anteriores. Los resultados aclaran argumentos teóricos conflictivos propuestos en la literatura, especialmente los relacionados al efecto de las restricciones presupuestarias de renta y tiempo. Se halla una diferencia cualitativa entre viajes de destinos únicos y viajes multidestino. Por tanto, se evidencia que la decisión de los turistas de consumir uno u otro tipo de viaje es diferente, y en cierto modo independiente de la decisión de cuantos destinos visitar.

En tercer lugar, se estudia el tiempo de estancia de los turistas en diferentes localidades en viajes de multidestino. Los determinantes de esa variable son estudiados por medio de un modelo de duración con heterogeneidad compartida donde el conjunto de variables independientes incluye características de los individuos, atributos del viaje, características y la identidad de la destinación y un conjunto de variables dummy para captar variaciones temporales. Cinco conclusiones principales son alcanzadas. Primero, son obtenidas diversas evidencias empíricas sobre los efectos de variables explicativas. Algunos resultados de estudios anteriores son confirmados y otros son desmentidos. Algunos ejemplos de resultados novedosos son los relacionados con los efectos de la renta, tamaño del grupo de viaje y viajes anteriores al destino. El efecto de la renta sobre la duración de la estancia de los turistas en los destinos dentro del paradigma de MTTs fue estimado como no significante para el caso brasileño. Los resultados muestran que la duración de la estancia sigue una función convexa del tamaño del grupo de viaje. Es esperado que los turistas que visitan Brasil por la primera vez tengan estancias más cortas que los que repiten la visita al país. Visitas anteriores adicionales para los turistas que ya han visitado Brasil tienen un efecto negativo sobre la duración esperada de la estancia. En segundo lugar, se comprobó que la tendencia creciente de la duración de la estancia de turistas internacionales en los destinos brasileños fue causada por variaciones en la composición del flujo receptivo internacional de turistas. En tercer lugar, se comprobó que la duración media de la estancia varía entre regiones y estados. Eso permite la identificación de destinos que pueden ser considerados benchmarks para la gestión del turismo en otros destinos. También se comprobó que la duración de la estancia de los turistas varía según las características de los destinos. En cuarto lugar, se verifico que las distribuciones de densidad de la duración de la estancia de los turistas con asimetría positiva son las más adecuadas. Finalmente, se comprobó que la heterogeneidad compartida entre observaciones mejora estadísticamente la capacidad explicativa de los modelos de duración cuando se utilizan datos de MTTs.

Table of Contents

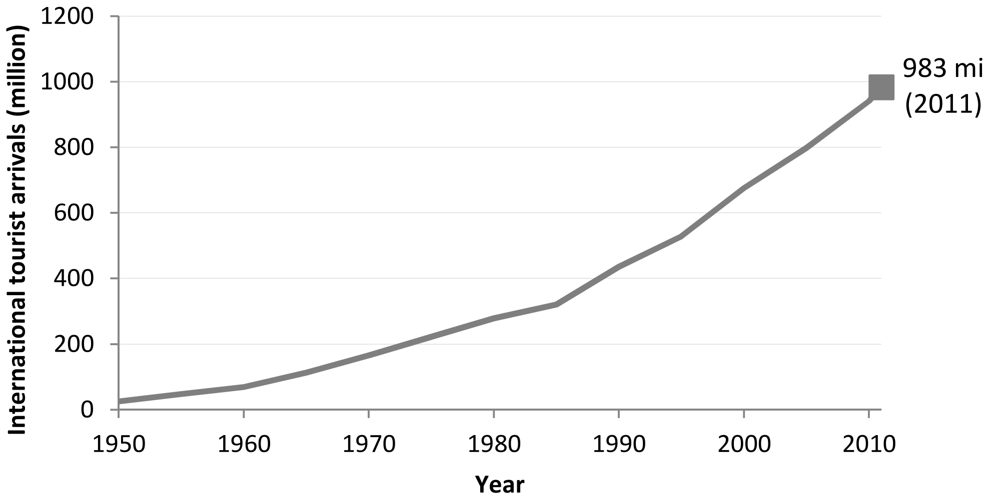
List of Tables

List of Figures

1 Introduction

Tourism has become one of the largest economic activities worldwide. According to the World Tourism Organization (2012), the overall number of international tourist arrivals increased from 25 million in 1950 to 983 million in 2011, as shown in Figure 1. The average annual growth rate during this period was 6.2%. The largest international destination region is Europe, followed by Asia and the Pacific. Americas are the third largest destination region. The overall international tourism receipts exceed US$ 1 trillion in 2011. Tourism exports account for about 30% of world's service exports and 6% of overall exports of goods and services. Tourism is the fourth largest exporting sector in the world after fuel, chemicals and automotive products.

Figure 1: International tourist arrivals from 1950 to 2011



Source: World Tourism Organization (2006, 2012)

These figures highlight the economic importance of international tourism. However, this is only one part of the whole tourism sector. Domestic tourism generates the largest share of tourism income in several countries. For example, more than 80% of total internal tourism consumption is domestic in countries such as China, India, Japan, Mexico, Netherlands, Philippines and United States (World Tourism Organization, 2010). Therefore, the economic importance of the tourism sector goes far beyond international tourism.

The prominent relevance of the tourism sector has created a strong competitive environment among tourism companies and destinations. The need for efficient governance has increased substantially, and managers are required to search for improved policies and strategies. These needs favoured tourism applied research and the increasing interest on a deeper understanding of tourism as a complex social and economic phenomenon.

The set of tourism research subjects is large and heterogeneous. One of the most studied subjects in this field is tourism demand, which have been analysed through a vast number of perspectives, including psychology, anthropology, sociology and economics. In tourism economics, demand modelling is one of the most important topics of research. The prominence of tourism demand modelling among economic studies is evidenced, for instance, by the number of submitted papers to the Tourism Economics journal. According to Wanhill (2011), 19% of all submitted papers are focused on demand modelling. Another evidence is the large number of papers published in different academic journals. According to the literature review on tourism demand conducted by Li, Song and Witt (2005) and Song and Li (2008), almost five hundred studies on this topic were published from 1960 to 2007. The main aspects remarked in these reviews are:

- The economic consumer theory is usually adopted as the theoretical framework underpinning tourism demand modelling studies.
- Tourism demand has been usually measured by arrivals or tourist expenditure.
- Some studies have differentiated tourism demand by purpose of the trip and transport mode.
- Different types of data are used in these studies, including both aggregate and individual data, cross section and time series, revealed and stated preferences, primary and secondary databases.
- Time series methodology has mainly employed yearly, quarterly and monthly data.
- Explanatory studies have employed a wide sort of independent variables. The most usual independent variables are income, relative prices, substitute prices, travel costs, exchange rates, dummies and deterministic trends.
- Most studies are geographically concentrated on Europe, United States, Asia and Pacific.
- Statistical and econometric techniques employed are also diverse and fast evolving.

Tourism demand studies have provided detailed explanations and accurate forecasts. Consistency and explanatory power of statistical models improved considerably over the last decades. The relationships between tourism demand and a large set of relevant variables were empirically tested. These achievements have provided an enormous support for the management of tourism firms and destinations. According to Song, Witt and Li (2009, p. 1)

> Tourism demand is the foundation on which all tourism-related business decisions ultimately rest. Companies such as airlines, tour operators, hotels,

> cruise ship lines, and many recreation facility providers and shop owners are interested in the demand for their products by tourists. The success of many businesses depends largely or totally on the state of tourism demand, and ultimate management failure is quite often due to the failure to meet market demand.

However, some aspects of tourism demand have been insufficiently studied in the academic literature. An important topic that has not received appropriate attention is multidestination tourism trips (MTTs), which can be defined as a single tourism trip in which more than one destination is visited. MTT is opposed to a single destination tourism trip, in which only one destination is visited. Following the UNWTO's recommendations for tourism statistics, a tourism trip may be defined as a trip to destinations outside traveller's usual environment, for less than a year, for any main purpose other than to be employed by a resident entity in the destination (United Nations, 2010). According to these recommendations, visitation is defined as the stay in a place during a tourism trip. Destination may be defined as the place visited.

Operational definitions of all these concepts are usually necessary to conduct empirical analysis. Time duration of a visit is an important aspect to be considered in these definitions. How short a stop may last for it to be considered a visit? The minimum length of the stay required by operational definitions of a visit has varied from a few moments up to an overnight. Another pertinent aspect to be considered is the geographical definition of a destination. How large is the geographical limits of a destination? In practice, the operationally defined size of a destination has varied from large areas, such as countries, to small locations, such as towns.

The UNWTO defines that "a trip is made up of visits to different places" (United Nations, 2010, p. 10). In fact, this definition does not leave space for a single destination tourism trip since a visit supposes a stop, and a tourist may stop in a single location during his or her trip. A more adequate definition would be that "a trip *may* be made up of visits to different places", indicating that single and multiple destination trips are possible alternatives. This imprecision in the UNWTO definitions is a preliminary evidence of the lack of attention given to MTTs.

Besides, it is interesting to note that the UNWTO proposes that "a tourism trip is characterized by its main destination, among other characteristics" (United Nations, 2010, p. 13). Main destination is understood as "the place visited that is central to the decision to take the trip" (United Nations, 2010, p. 14). This trip characterization perspective clearly minimizes the importance of MTTs and does not recognize the complex nature of this type of tourism trips.

Statistical information about MTTs is very scarce. Official tourism statistics usually do not record any information about this topic because destination management organizations are typically not interested on tourists' visits to other destinations (Leiper, 1989). Even when the destination management organization is responsible for a whole region or country, including more than one destination, it usually does not produce statistics concerning MTTs within its own region or country.

However, it is possible to find some data capturing the occurrence of MTT in international tourist survey reports from countries where tourism statistics are best developed. For example, MTTs account for 33% of inbound tourism in the United States, while the average number of destinations visited is two (U.S. Department of Commerce, 2011). For the outbound American tourism, the average number of countries visited is 1.3 (U.S. Department of Commerce, 2010). In Canada, the average number of provinces visited by inbound tourists is 1.14 (Canadian Tourism Commission, 2011). In New Zealand, considering only the top ten destinations and grouping all the remaining destinations in a eleventh category, the average number of destinations visited by inbound tourists is 2.1. Besides, 35% of these tourists visit other countries in the same trip to New Zealand (Tourism Strategy Group, 2011). In Australia, MTTs represent 53% of total inbound tourism and the average number of destinations visited by international tourists is 3.2 (Tourism Research Australia, 2008). Finally, according to Oppermann (1995), 41.6% of inbound tourists in Malaysia visit more than one city, while 52.6% visit other countries in the same trip.

The remaining of this chapter is organized as follows. Section 1.1 presents the research objectives of this thesis and the content of the next chapters. Then, a general literature review of the topics developed in the following chapters is presented. Section 1.2 reviews the previous literature on MTTs, while Section 1.3 focuses on studies about tourists' length of stay as a general literature review for Chapter 4. The data source used at the empirical chapters 3 and 4 is presented in detail at Section 1.4. Finally, a brief overlook of inbound tourism in Brazil is presented at Section 1.5.

1.1 Research objectives

The general objective of this thesis is to enlarge the understanding of multidestination tourism trips. The general relevance of MTTs and the lack of knowledge about this topic are the capital justifications for this work. The substantial incidence of MTTs in Brazil provides further support to the proposal of this study.

Three main lacks of the academic literature on MTTs are studied in this thesis. First, economic theoretical models of tourist choices developed hitherto are insufficient to explain MTTs consumption. On one hand, traditional microeconomic theory of consumer choices, such as Morley (1992) and Stabler, Papatheodorou and Sinclair (2010), were developed within the single destination paradigm, overlooking the possibility of tourists choosing to visit multiple destinations in a single trip. On the other hand, available economic theoretical models of consumer choices within the MTT paradigm, such as Rugg (1973) and Tussyadiah, Kono and Morisugi (2006), were based on the Lancaster (1966) characteristics theory. However, two important criticisms can be made to this approach. First, although some characteristics are clearly common among different destinations, other relevant aspects are site specific. Lancaster's theory does not allow for the consideration of singularities, overlooking the importance of uniqueness of tourist destinations. Second, it is not clear why one should consider Lancaster's theory in the first place since MTT choices can be explained by a simpler and more common approach – the traditional economics' consumer theory. Lancaster's theory seems to add unnecessary issues to the discussion. Characteristics are argued to be an unrequired concept to explain why people have interest in visiting more than one destination in a single trip. The consideration of characteristics in the MTT choice model only makes it pointlessly complex. Finally, the resulting complexity apparently prevented these researchers to account for some important aspects of consumers' choices in the MTTs paradigm. An alternative theoretical model for consumers' choices in the MTT paradigm based on the traditional economic theory is presented in Chapter 2[1].

Second, despite of the relatively large theoretical literature on the determinants of MTTs consumption, few studies had analysed this topic from an empirical perspective. The most relevant studies in this field are those conducted by Tideswell and Faulkner (1999, 2003) and Nicolau and Más (2005). The objective of the study presented in Chapter 3 is to further develop the analysis of the individual determinants of MTTs consumption, assessing theoretical propositions and empirical findings of previous studies. Inbound tourism in Brazil is empirically studied. The earlier set of explanatory variables is extended to include some additional determinants, such as education level, type of accommodation and season. The time evolution of the dependent variable is also assessed. Finally, count data models are employed, overcoming some deficiencies of previous studies and providing better estimates. The family of models used also allows for the analysis of potential qualitative differences between single and multiple destination trips.

Third, MTTs introduce an interesting additional issue on the modelling process of tourists' length of stay. When the durations of the stays at different destinations in the same trip are regarded, correlation among observations arises. Therefore, the usual modelling techniques are no longer valid. In the fourth chapter, tourists' length of stay is modelled using shared heterogeneity duration models which takes proper account of this particular characteristic of data. To our knowledge, this study is the first one to use shared heterogeneity duration models in the tourism context. Besides, differences between the length of stay at the destination and in the destination region are an important issue arising from this perspective. Different specific models are assessed in order to find the best fitting one. The empirical study conducted at Chapter 4 also expands the set of explanatory variables used by earlier studies. Besides, it analyses the recent evolution of the tourists' average length of stay in Brazil. Interestingly, this variable has presented an increasing trend in Brazil, an unexpected reality when compared to most countries. The causes of this trend are carefully analysed.

[1] The content of Chapter 2 corresponds to the complete version of the work published as a condensed article in Santos, Ramos, & Rey-Maquieira (2011).

1.2 Previous literature on MTTs

Academic research about multidestination tourism emerged in different areas and has distinct approaches. Four main approaches might be identified, as discussed in the following.

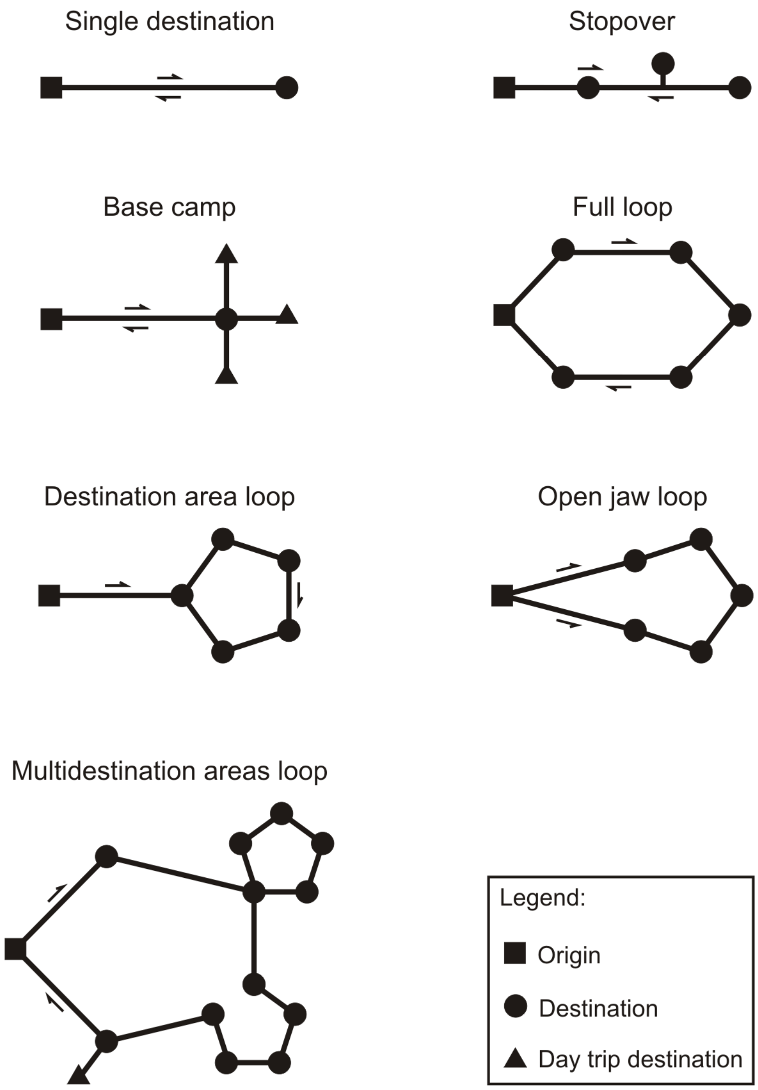
1.2.1 Descriptive studies

MTTs were described in several different ways throughout the academic literature. Initial qualitative descriptions of the MTTs phenomenon were usually done by geographical studies (for example Pearce, 1990, 1995, 1999). Exploratory quantitative analysis of MTTs employed different measures, such as the number of tourists visiting more than one destination (e.g., Hu & Morrison, 2002; Hunt & Crompton, 2008; McKercher, 2001; Oppermann, 1995; Tideswell & Faulkner, 1999; Tourism Research Australia, 2008; Wall, 1978), the distribution of tourists according to the number of destinations visited (Hwang & Fesenmaier, 2003; Lew & McKercher, 2002; Parroco, Vaccina, Cantis, & Ferrante, 2012; U.S. Department of Commerce, 2011), and the average number of destinations visited (Hwang & Fesenmaier, 2003; Hwang, Gretzel, & Fesenmaier, 2006; Oppermann, 1992; Tideswell & Faulkner, 1999, 2003, U.S. Department of Commerce). Relatively detailed descriptions of MTTs have outlined the inventory of destinations visited (e.g., Baxter & Ewing, 1981; Hunt & Crompton, 2008; Wall, 1978) and the most frequent combinations of destinations (Hwang et al., 2006; Lew & McKercher, 2002; Oppermann, 1992; Parroco et al., 2012).

Leiper (1989) proposed a specific index to describe MTTs patterns that may overcome usual lack of raw data. The Main Destination Ratio proposed is calculated by combining data collected separately at the origin and the destination. This index refers to the flow between origin and destination, and reveals the percentage of tourists from a given origin for whom a visited place is the main destination of their trips. It can be calculated by comparing departure statistics where tourists declare which is the main destination of their trips, with arrivals statistics collected at the destination.

Several authors studied spatial patterns of multidestination tourism by describing itinerary types. Each study presents a different set of patterns to describe the most frequent itineraries. Flognfeldt (1992) described three itinerary patterns of tourists in Norway. Mings and McHugh (1992) identified four patterns among tourists visiting the Yellowstone National Park in the United States. Lue, Crompton and Fesenmaier (1993), which is perhaps the most influential paper on MTTs, proposed a model of five itineraries in a theoretical study. Oppermann (1995) studied inbound tourism in Malaysia, adopting the model of Lue et al. (1993) and proposing two more itinerary patterns, as presented in Figure 2. These patterns are more clearly defined by Lew and McKercher (2002) as follows:

Figure 2: Schematic representations of Oppermann's itinerary patterns



Source: Oppermann (1995)

- Single destination: same route to and from destination with no stopovers or destination area day trips.
- Stopover: same route to and from destination with stopovers.
- Base camp: same route to and from destination with side day trips departing from it.
- Full loop: full circular route starting and ending at the origin with stopover destinations.
- Destination area loop: same route to and from destination area with a full circle route in the destination region.
- Open jaw loop: different routes to and from a destination area with a partial circular route at the destination area.
- Multidestination areas loop: full circular route with day trips or separate circular trips departing from stopover destinations.

Despite of Lew and McKercher's (2002) effort to integrate previously described itinerary patterns, further operational definitions may be required for practical use. Concepts such as destination, destination area, stopover and route may need further specifications. For example, although destinations were frequently defined as cities or towns (for example Flognfeldt, 1992; Hwang & Fesenmaier, 2003; Hwang et al., 2006; Jeng & Fesenmaier, 1998; Koo, Woo, & Dwyer, 2009), other geographical definitions were employed, such provinces (Tourism Research Australia, 2008), states (U.S. Department of Commerce, 2011) and countries (Lew & McKercher, 2002). Moreover, the set of previously described itinerary patters is incomplete, providing no proper classification for fuzzy or redundant itineraries. Thus, further classifications may be needed for specific applications.

Besides theoretical discussions, several empirical studies on tourist itineraries were published. Travel itinerary patterns in Branson (USA) were empirically analysed by Stewart and Vogt (1997). Hwang and Fesenmaier (2003) studied domestic tourism itinerary patterns in the same country. Yang, Hui-Min and Ryan (2009) studied stated preferences of Chinese tourists with respect to itineraries in South-Western USA. Taplin and McGinley (2000) analysed car travel itineraries departing from Perth (Australia) using a linear program. Zillinger (2007) provided detailed descriptions of MTTs of

German tourists in Sweden through the analysis of travel diaries where individual spatiotemporal travel patterns were registered. Wu and Carson (2008) used a Geographic Information System (GIS) to study spatiotemporal characteristics of MTTs in South Australia, providing one of the most detailed descriptions of this phenomenon. GIS was also used to describe survey collected data about travel itineraries at the Loch Lomond and Trossachs National Park in Scotland (Connell & Page, 2008).

Lew and McKercher (2002) argue that the identification of the itinerary pattern is less relevant to destination management organizations than information about the relative position of the destination within the itinerary. The authors present definitions for five types of destinations according to their relative position within the travel itinerary, proposing that each one requires different services and infrastructure. A gateway destination is the first site visited at the beginning of a multiple destination itinerary. Egress destination is the last one in the itinerary, being visited just before the tourist returns home. Touring destinations are the ones in the middle of the trip. Finally, the hub is defined as any destination visited at least twice during the trip. It is argued that transport facilities and services requirements are usually exceptional in hub destinations, good in gateways and egress destinations and just ordinary in touring and single destinations. Single destinations need a larger, more complete and more diversified mix of attractions and activities than the other ones. Finally, gateway and egress destinations have important roles in relation to psychological experiences, providing the feeling of transition between home and away, and also working as cushions for cultural-shocks. The performance of Hong Kong in different positions is empirically studied by these authors.

Some studies described spatial patterns of multidestination tourism through social network measures, describing the structure of relations between social entities through the calculation of indicators about the whole network and the position of individuals in the network structure. Oppermann (1992) presented the number of connections between destinations as an elementary network centrality measure to analyse tourism in Malaysia. Hwang et al. (2006) studied several network measures and their relationships with variables such as origin markets, previous trips to the destination, transportation system and geographical situation of destinations. Shih

(2006) applied social network measures to self-driving tourists in Nantou, Taiwan. These studies proposed some policy implications that could be derived from social network measures. However, no empirical results were used to validate these propositions.

1.2.2 Cumulative attraction

The concept of cumulative attraction was proposed by Nelson (1958, p. 58), and it states that "a given number of stores dealing in the same merchandise will do more business if they are located adjacent or in proximity to each other than if they are widely scattered". Cumulative attraction is intimately linked to the concept of cluster, particularly to the idea of spatial agglomeration benefits related to the attraction of consumers (Porter, 1990). In economic terms, cumulative attraction means that some businesses, even in the same sector, may be gross complements rather than substitutes because of spatial matters.

In the tourism context, Lue et al. (1993) argue that cumulative attraction may happen to businesses or destinations located in proximity or in a spatial sequence to each other. Destinations are said to display cumulative attraction if the sum of their demands is greater when they are closely located than when they are distant from each other (Jeng & Fesenmaier 1998). According to Papatheodorou (1999), on one hand, two close destinations compete for tourists as standard substitutes. But on the other hand, the attraction of multidestination tourists by one destination may increase the number of tourists on the other site, resulting in some degree of complementarity. The net effect is a priori ambiguous; positive net effects might happen in some cases, while in other instances negative net effects might be the actual outcome.

Several empirical studies discuss cumulative attraction in the tourism context. Crompton and Gitelson (1979) were pioneers in this field by studying cumulative attraction between commercial leisure enterprises. Lue, Crompton and Stewart (1996) used an experimental approach to search for evidence of cumulative attraction among tourist destinations. This article was strongly criticized by Beaman, Jeng and Fesenmaier (1997), who argued that the previous research confused cumulative attraction arising from multiple destinations and from destinations' attributes. Jeng and Fesenmaier (1998) also adopted an experimental approach in order to analyse the cumulative attraction among European cities in the MTT context. . Hunt and Crompton (2008) tried to measure the degree of cumulative attraction among different tourist attractions of the same city using the original scale proposed by Nelson (1958), which is measured in percentage of shared tourists. Finally, Weidenfeld, Butler and Williams (2010) deeply analysed the relationship between the nature of the tourism product, spatial clustering and cumulative attraction among attractions using in-depth interviews with tourist attraction managers and a tourist survey.

1.2.3 Theoretical approaches

Some studies proposed comprehensive theoretical backgrounds for the individual choice process regarding MTTs based on the economic theory. The application of Lancaster's (1966) characteristics approach to consumer theory provides a formalized and comprehensive approach to multidestination tourism trips choice. Several studies applied this approach to tourism consumption (e.g., Morley 1992; Papatheodorou 2001, 2006; Seddighi and Theocharous 2002), but only two of them analyse multidestination trips (Rugg 1973; Tussyadiah et al., 2006).

Rugg (1973) proposed a model of tourism demand by stating that tourists maximize their utility derived from characteristics of the destinations visited. This consumption is determined by convex preferences, as well as monetary and time budget restrictions. Accordingly, the problem faced by potential tourism consumers is to maximize

$$U = U(\boldsymbol{z})$$

subject to

$$\boldsymbol{z} = \boldsymbol{b}(\boldsymbol{d})$$

$$Y \geq \boldsymbol{p_d}.\boldsymbol{d} + \boldsymbol{p_t}.\boldsymbol{m}$$

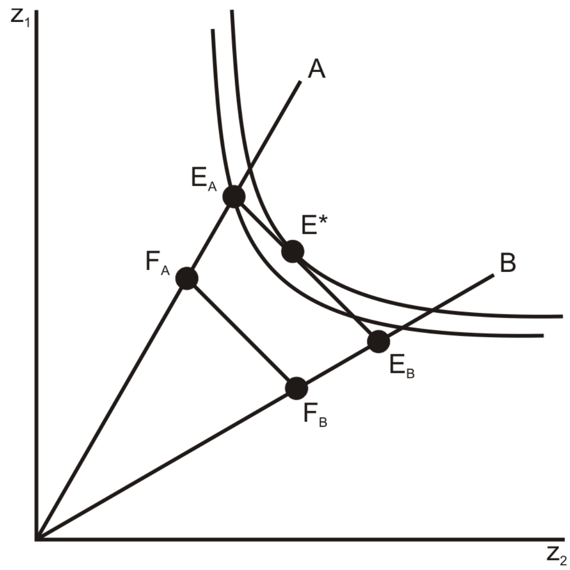
$$T \geq \boldsymbol{c}.\boldsymbol{d} + \boldsymbol{t}.\boldsymbol{m}$$

$$\boldsymbol{z}, \boldsymbol{d}, \boldsymbol{p_d}, \boldsymbol{p_t}, \boldsymbol{m}, \boldsymbol{c}, \boldsymbol{t} \geq 0$$

where **z** represents characteristics' quantities, **b** describes the production of characteristics by each destination, **d** is the time spent in each destination, $\mathbf{p_d}$ and $\mathbf{p_t}$ are destination and transport prices, respectively, **m** is a permutation vector of zeros and ones identifying trips between pairs of locations, **c** is a vector of ones, and **t** represents transport time between locations. Bold characters are used to identify vectors.

For illustration purposes, consider a case in which there are two characteristics, z_1 and z_2, such as in Figure 3. Rays A and B represent the production of characteristics by two different destinations. Given prices and available monetary resources, it is possible to calculate the maximum amount of time that the tourist may spend in each destination. If this time is actually spent in a single destination, the tourist obtains quantities of characteristics represented by points E_A or E_B, respectively related to the consumption of destination A or B. Assuming zero transport costs between both destinations, the segment joining these two points represents the maximum obtainable amount of characteristics if the tourist visits both destinations in the same trip. Assuming convex preferences such as those represented in Figure 3, the optimum combination E* is chosen by the consumer.

Figure 3: Representation of the characteristics approach



Rugg's model is also adapted for cases with positive costs of transport between destinations. In these circumstances, visiting a second destination in the same trip implies some additional expenditure. In Figure 3, the point F_A represents the quantity of characteristics obtained when the tourist goes to both destinations, spending all available time in destination A. The interpretation of point F_B is symmetric. The segment linking F_A and F_B represents the maximum obtainable amount of characteristics if the tourist visits both destinations. Note that, in this case, points E_A and E_B are still achievable if the individual travels to a single destination. In the example provided by Figure 3, the tourist chooses to travel only to destination A.

Tussyadiah et al. (2006) adds to Rugg's model the case of a stopover destination that does not require any route change. According to these authors, this case represents no extra cost of transport and it does not change the budget restriction. Their representation of the stopover case is identical to the one where zero transport cost between destinations is assumed. However, this is not correct, since the inexistence of transport costs associated with stopover itineraries happens only in one direction. If no route change is required, the visit to a stopover destination implies no extra transport cost. Nonetheless, when the stopover destination is already being visited, a visit to the furthest destination still requires additional efforts.

Tussyadiah et al. (2006) improve Rugg's model by introducing a conservation law to ensure that transport is consumed in pairs of journeys for each destination, that is, every inwards journey implies the consumption of one outwards journey. This restriction arises from the definition of tourism activity and mathematically ensures that if a tourist visits one location, he also leaves this location after some time. A restriction ignored by both studies is that the time spent at the destination is necessarily zero when no inwards journey to that particular destination is consumed.

The characteristics approach to demand presents two major limitations. First, it is difficult to objectively measure many relevant characteristics. Secondly, the characteristics approach is unable to deal properly with exclusive characteristics of products. However, tourism destinations always have some degree of uniqueness, and just a subset of all existent characteristics is present in different places. These criticisms are some of the main arguments in support of the development of the theoretical model of MTTs consumption presented in Chapter 2.

1.2.4 Empirical modelling studies

There are a few modelling studies about MTTs. Tideswell and Faulkner (1999) modelled the number of overnight destinations visited by international tourists in Queensland, Australia. They used a linear regression model with seven independent variables: party size, distance from country of origin to the destination region, number of different purposes of travel, number of different sources used to obtain information about the Queensland, and three dummy variables identifying the use of own or rented vehicles, the business and conference purpose of travel, and the visiting friends and relatives purpose of travel.

Tideswell and Faulkner (2003) analysed the consumption of MTTs in Australia. A linear regression was used to model the number of regions visited. Eleven independent variables were tested. Three non-significant variables were omitted in the final model: party size, the proportion of time away from home spent in Australia, and the use of travel package. The eight significant variables were distance from country of origin to Australia, length of stay in the destination country, number of different purposes of travel, number of different sources used to obtain information about Australia, and four dummy variables indicating use of own or rented vehicles, repeat visit to Australia, group tour, the visiting friends and relatives purpose of travel.

Both studies by Tideswell and Faulkner (1999, 2003) can be criticized by the use of an inappropriate modelling technique. The linear regression model adopted by these studies supposes that the dependent variable is normally distributed. Empirical evidence has shown that this is not the case of the number of destinations visited in a trip. Besides, linear regression models are able to provide negative estimates of the dependent variable, which would be an absurd outcome for the number of destinations visited.

Nicolau and Más (2005) used a random coefficients logit model to predict the dichotomous decision of taking a single or a multiple destination tourism trip. This prediction was made jointly with predictions about taking a vacation or not and visiting foreign or domestic destinations. Independent variables used to predict the choice of MTTs included organization through travel agents, interest in new places and cultural interest. It is relevant to stress that dichotomous dependent variable used is not as informative as the variable number of destinations visited used by Tideswell and Faulkner (1999, 2003).

More recently, two modelling studies approached MTTs by different perspectives. Koo et al (2009) adopted an experimental design to model the choice of different transport alternatives to visit the North Queensland region in Australia. Alternatives included car, rental car, bus or coach, train and four different air transport options. The multidestination nature of the trip was used as an explanatory variable in order to capture how the visitation of more than one location influences the utility of each transport alternative. It was shown that the utility of air transport is lower for MTTs than for single destination trips.

Finally, Wu, Zhang and Fujiwara (2012) developed a study about tourist destination choices within the MTTs paradigm. These authors took the interdependency across different destinations into consideration by using a universal logit model to explain tourists' destination choices. The probability of choosing one specific destination was modelled as a function of the characteristics of other destinations and their distances. The effect of other destinations over the choice of a given destination was shown to be significant. This finding evidenced the interdependency among different destinations of a MTT. They concluded that "it is inadequate to analyse destination choice separately and independently without considering the interrelationships that may exist among choices and future dependency should be incorporated into choice analysis" (Wu et al., 2012, p. 128).

1.3 Previous literature on tourists' length of stay

As explained at Section 1.1, the last chapter of this thesis deals with the analysis of tourists' length of stay within the MTT paradigm, uncovering potential misleading conclusions obtained when this variable is examined under the single destination paradigm. This section presents an overview of the study of tourists' length of stay, while the modelling particularities are discussed at Chapter 4.

It must be emphasize that the academic interest on tourists' length of stay is relatively old. For instance, Wurst (1955), a pioneer in this field, proposed ways of calculating tourists' average length of stay from aggregate data. Two decades after, Archer and Shea (1975) provided further analysis on Wurst's analysis. Different objectives and approaches are considered in more recent studies. Several descriptive and univariate analyses of tourists' length of stay and its determinants have been conducted, as in Oppermann (1994, 1995, 1997), Seaton and Palmer (1997), Sung, Morrison, Hong and O'Leary (2001) and Tierney (1993). This topic was also analysed on the context of non-market valuation of recreational assets. Some studies using the travel cost method examined the role played by the length of stay over consumer surplus estimates (for example Berman & Kim, 1999; Larson, 1993; McConnell, 1992; Shaw & Feather, 1999).

Another part of the literature is devoted to the detailed assessments of the determinants of the length of stay or the duration of a trip[2] using statistical modelling procedures were recently conducted in the academic literature. Despite of Pulina's (2010) study using aggregate data, most modelling studies about tourists' length of stay or the duration of the trip used microdata. Twenty four papers modelling tourists' length of stay from microdata have been published in the main tourism academic journals. Table 1 summarizes the main features of these studies.

Pioneer studies modelling tourists' length of stay were published in the late 1970s, but most studies were conducted since late 1990's. Studied regions include Latin America, China, United States, Portugal, Spain, Scandinavia, Turkey, Israel, Saudi Arabia and Madagascar.

[2] Length of stay denotes the amount of time that the tourist spends at a given destination, while duration of a tourism trip refers to the length of time between departure and return to home. This issue is discussed in detail at Chapter 4.

Table 1: Summary of modelling studies on tourists' length of stay

Study	Region focused (Origin / Destination)	Dependent variable	Data collection procedure	Sample size	Model	Model distributions or specification form	Special features
Mak, Moncur, and Yonamine (1977)	Hawaii, USA	Length of stay	Arrival survey	4,990	MSLS	linear	Two stages
Mak and Nishimura (1979)	Hawaii, USA	Length of stay	-	690	OLS	log-log	
Walsh and Davitt (1983)	Aspen, USA	Length of stay	During the stay	837	OLS	linear, quadratic, log-linear*, log-log	
Silberman (1985)	Virginia Beach, USA	Length of stay	During the stay	621	MSLS	linear, quadratic, log-linear, log-log*	Two stages
Uysal, McDonald, and O'Leary (1988)	USA	Length of stay	Source market	6,720	OLS, MSLS*	lin-log	Two stages
Paul and Rimmawi (1992)	Asir National Park, Saudi Arabia	Length of stay	During the stay	208	OLS	linear	
Blaine, Mohammad, and Var (1993)	Fredericksburg, USA	Length of stay	During the stay	-	OLS	log-log	
Fleischer and Pizam (2002)	Israel	Duration of the trip	Source market	400	Tobit		
Alegre and Pou (2006)	Balearic Islands, Spain	Length of stay	Departure survey	24,896	Binary	logit	

(continued)

Table 1: Summary of modelling studies on tourists' length of stay (continued)

Study	Region focused (Origin / Destination)	Dependent variable	Data collection procedure	Sample size	Model	Model distributions or specification form	Special features
Hellström (2006)	Stockholm, Gothenburg and Malmö, Sweden	Length of stay	Source market	2,000	Count data	bivariate Poisson lognormal	Inflation, Truncation
Gokovali, Bahar, and Kozak (2007)	Bodrum, Turkey	Length of stay	Departure survey	672	Duration	Cox*, exponential, Weibull*, Gompertz	
Barros, Correia, Crouch (2008)	Portugal / Latin America	Ambivalent variable	On flight survey	442	Duration	Cox, Weibull**, loglogistic	Heterogeneity
Martínez-Garcia and Raya (2008)	Catalonia, Spain	Length of stay	Departure survey	990	Duration	Cox, exponential, Weibull, Gompertz, lognormal, loglogistic*, Gamma	Heterogeneity
Menezes, Moniz, and Vieira (2008)	Azores, Portugal	Length of stay	Departure survey	400	Duration	Cox	
Fleischer and Rivlin Byk (2009)	Israel	Duration of the trip	Source market	~4,480	MSLS	log-log	Three stages
Nicolau and Más (2009)	Spain	Duration of the trip	Source market	3,781	Multinomial	logit	Random parameters
Barros, Butler and Correia (2010)	Algarve, Portugal	Length of stay	Arrival survey	593	Duration	Cox, Weibull**	Heterogeneity, sample selection

(continued)

Table 1: Summary of modelling studies on tourists' length of stay (continued)

Study	Region focused (Origin / Destination)	Dependent variable	Data collection procedure	Sample size	Model	Model distributions or specification form	Special features
Barros and Machado (2010)	Madeira, Portugal	Length of stay	Departure survey	346	Duration	Weibull	Sample selection
Machado (2010)	Madeira, Portugal	Length of stay	Departure survey	346	Duration	Weibull	SUDCD
Alegre, Mateo, and Pou (2011)	Balearic Islands, Spain	Length of stay	Departure survey	29,162	Count data	Poisson	Latent class
Raya and Martínez-Garcia (2011)	Catalonia, Spain	Length of stay	Departure survey	-	Duration	Cox, exponential, Weibull, Gompertz, lognormal, loglogistic*, Gamma	Heterogeneity
Yang et al. (2011)	Yixing, China	Length of stay	During the stay	417	Ordered discrete	logit	
Peypoch, Randriamboarison, Rasoamananjara and Solonandrasana (2012)	Madagascar	Length of stay	Departure survey	615	Duration	fractional polynomial, Weibull	
Thrane (2012)	Scandinavia	Duration of the trip	Source market	539	OLS, Duration	log-linear (OLS) Weibull*, lognormal, loglogistic (Dur.)	Heterogeneity, sample selection

** Preferred alternative; ** Preferred when together with the special feature; - Missing information*

Different theories were used to explain tourists' choices of how long to travel or to stay at the destination. Mak and Nishimura (1979) based their analysis on simple demand theory. Several authors developed utility maximization models explaining this sort of tourists' choices (Alegre et al., 2011; Alegre & Pou, 2006; Hellström, 2006; Raya & Martínez-Garcia, 2011). Other studies applied the theory of consumer behaviour arising from Lancaster's (1966) model and the concept of hedonic prices (Rosen, 1974) as their theoretical background (Barros et al., 2010; Barros et al., 2008; Barros & Machado,2010; Machado, 2010; Peypoch et al., 2012). Although these theories focus on the characteristics of available alternatives, explanatory variables used in these studies were mainly related to individual' characteristics rather than destinations' attributes. Lastly, Fleischer and Rivlin Byk (2009) used Becker and Lewis' (1973) utility maximization model allowing for quantity and quality choices. Adapting this model to tourism consumption, choices about the number of trips, quality of travel services and length of stay were jointly regarded.

At most cases, theories adopted to explain tourists' length of stay presented a high level of generality. These models were usually used only as loose frameworks for identifying relevant explanatory variables rather than actually guiding the construction of the statistical model applied. A remarkable exception was Fleischer and Rivlin Byk's (2009) theoretical framework, which besides being substantially complex, actually allowed the authors to estimate compensated elasticities for price and time costs.

Nineteen studies focused on the length of stay at tourist destinations and four modelled the duration of the trip. The study of Barros et al. (2008) focused on Portuguese tourists travelling to Latin America. This study did not report if inbound and outbound travel time was considered or not, thus providing no information to distinguish the studied variable between length of stay and duration of the trip.

Regarding the source of the empirical data, Barros et al. (2010) and Mak et al. (1977) used data of entry surveys gathered during tourists' arrival at the destination. Several studies used on-site survey data collected along tourists' stay at the destination (Blaine et al., 1993; Paul & Rimmawi, 1992; Silberman, 1985; Walsh & Davitt, 1983; Yang, Wong, & Zhang, 2011). These two types of data collection procedures conducted before the actual end of the stay might display some inaccuracy if tourists decide to change their length stay during this period. Some authors analysed data about past trips collected through household surveys (Fleischer & Pizam, 2002; Fleischer & Rivlin Byk, 2009; Hellström, 2006; Nicolau & Más, 2009; Uysal et al., 1988). Thrane (2012) analysed data about past trips gathered through a survey conducted at a university college. One major criticism to data gathered at source markets refer to inaccuracy arising from respondents' recall bias (Frechtling, 2006). Barros et al. (2008) conducted on flight surveys. Finally, the remaining studies used exit surveys data collected at gateways during tourists' return journey. This method may be considered the most adequate since it avoids all inaccuracy sources aforementioned.

Different types of tourists were analysed in previous papers due to distinct research objectives and data availability. Ten studies considered tourists with any profile (Barros & Machado, 2010; Blaine et al., 1993; Fleischer & Rivlin Byk, 2009; Machado, 2010; Menezes et al., 2008; Nicolau & Más, 2009; Paul & Rimmawi, 1992; Peypoch et al., 2012; Silberman, 1985; Walsh & Davitt, 1983). Hellström (2006) focused exclusively on leisure tourists, while Barros et al. (2010) focused on golfers. Martínez-Garcia and Raya (2008) and Raya and Martínez-Garcia (2011) studied only international tourists, while other authors focused their studies on domestic tourists (Mak et al., 1977; Mak & Nishimura, 1979; Uysal et al., 1988; Yang et al., 2011). Some studies determined their population of interest according to specific tourists' nationalities (Alegre et al., 2011; Alegre & Pou, 2006; Barros et al., 2008; Gokovali et al., 2007). Alegre et al. (2011) and Alegre and Pou (2006) also defined their population in terms of the type of accommodation used, excluding those staying at friends or relatives' homes or their own houses. Finally, two studies defined their population of interest in terms of socio-demographic characteristics. Fleischer and Pizam (2002) studied individuals above 54 years old, while Thrane (2012) studied undergraduate students attending to a medium-sized university college.

Sample sizes varied substantially. Several studies used fairly small samples, especially when these are compared to the relatively large number of parameters estimated. Small sample sizes led to a substantial number of non-significant parameters estimated at different studies. Moreover, even when significant parameters were obtained, small sample sizes caused large standard errors and loose parameter estimates. Two remarkable exceptions were the studies of Alegre et al. (2011) and Alegre and Pou (2006) which used large datasets.

1.4 Data source for empirical studies

Chapters 3 and 4 conduct empirical studies using a secondary database generated from the Brazilian Inbound Tourism Survey (BITS). The BITS is the Brazilian official source of statistics on international tourists' and travels' characteristics. This survey is conducted by the Foundation Institute of Economic Research (FIPE) and financially supported by the Tourism Ministry of Brazil since 2004. The BITS substituted a previous survey that was conducted by the National Tourist Board (Embratur) since 1980.

The survey is based on personal interviews with international tourists visiting Brazil. Interviews are conducted at international gateways a few moments before tourists' departure from the country. A total of 27 international gateways of Brazil were surveyed, including 15 international airports and 12 land borders. The set of international airports covered by the BITS accounts for more than 99% of total international air traffic in Brazil, while the set of land borders account for about 95% of total road traffic. The use of other transport modes for international passenger transport in Brazil is marginal (2,9% in 2010 and 3,4% in 2011). The list of gateways surveyed by the BITS is presented in Figure 4. Note that in Foz do Iguaçu there are two survey locations, one at the border with Paraguay (PIA) and another at the border with Argentina (PTN).

The frequency of survey periods at each gateway varies from one to six times a year. Survey periods last from one to two weeks and daily lengths of survey varies from six to 24 hours. Larger and more frequent survey periods were undertaken at the most important gateways, such as São Paulo, Rio de Janeiro and Foz do Iguaçu.

From 2004 to 2010, a total sample of 183.232 valid interviews was obtained. The sample size increased significantly from 2004 to 2005 and from 2007 to 2008. The distribution of the sample by years is presented at Table 2.

Figure 4: Gateways surveyed by the BITS

Table 2: Sample size of the BITS by year

Year	Sample size	%
2004	18,021	9.8
2005	22,692	12.4
2006	22,557	12.3
2007	23,246	12.7
2008	32,760	17.9
2009	31,739	17.3
2010	32,217	17.6
Total	*183,232*	*100.0*

Sampling quotas are defined for the main source markets in order to ensure institutional standards of accepted error. Sample weights are calculated according to source market, survey period and gateway using tourist arrivals statistics from migration control.

Statistical definitions suggested by the UNWTO and UN are adopted by the BITS. Tourist destination is specifically defined as a municipality where tourists stayed overnight. The use of municipalities is not significantly restrictive, as according to raw data from the BITS, very few international tourists stay overnight at more than one place at the same Brazilian municipality.

The questionnaire of the BITS is presented at Appendix 2. It includes about 30 questions covering:

- Tourists' characteristics: age, gender, place of residence, education level and income;
- Travel's characteristics: purpose, destinations visited, length of stay, expenditure by items, transport mode, accommodation, use of tour packages, type of travel party, sources of information, previous trips to Brazil; and
- Tourists' opinions: evaluation of different aspects, fulfilment of personal expectations and intention to return.

1.5 Inbound tourism in Brazil

International tourism in Brazil is relatively incipient despite of its relevant natural and cultural assets. In 2011 the country received almost 5.4 million tourists, 0.2 million less than Argentina, which is the leading destination country in South America.

South America represents 48.4% of total Brazilian inbound tourism, while Europe and North America respond for 29.8% and 13.4%, respectively. Argentina is the largest tourist source market, responding alone for 29.3% of total arrivals. The distribution of Brazilian inbound tourist flows according to their origin is presented at Table 3.

Table 3: Brazilian inbound tourist flows according to the origin (2011)

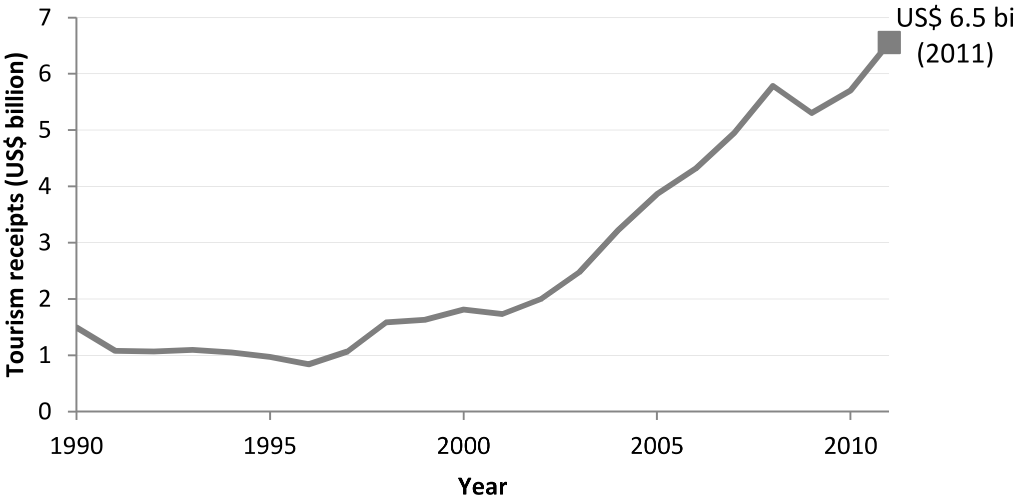
Origin	Arrivals	%
South America	**2,628,957**	**48.4**
Argentina	1,593,775	29.3
Chile	217,200	4.0
Paraguay	192,730	3.5
Uruguay	261,204	4.8
Other countries	364,048	6.7
North America	**729,756**	**13.4**
United States	594,947	10.9
Other countries	134,809	2.5
Europe	**1,621,183**	**29.8**
Germany	241,739	4.4
Italy	229,484	4.2
France	207,890	3.8
Spain	190,392	3.5
Portugal	183,728	3.4
England	149,564	2.8
Other countries	418,386	7.7
Other continents	**453,458**	**8.3**
Total	**5,433,354**	**100.0**

Source: Ministério do Turismo (2012)

International inbound tourism generated receipts of US$ 6.5 billion in 2011, what represents 17.2% of Brazil's service exports[3]. The average growth rate of Brazilian international tourism receipts from 2000 to 2011 was 12.4%, despite of the fall in 2009 due to the world economic crisis. This was a relatively fast growing period as compared to the 1990s when the average growth rate was 2.0%. The evolution of Brazil's international tourism receipts is presented in Figure 5.

[3] Data from the Balance of Payments of the Central Bank of Brazil.

Figure 5: Evolution of international tourism receipts in Brazil



Source: Banco Central do Brasil (2011)

Leisure tourism accounts for 45.5% of total international arrivals, while business represents 22.9% of this flow[4]. Other travel purposes, including visiting friends and relatives account for the remaining 31.6%. The average daily expenditure of leisure tourists is US$ 63, while business tourists spend US$ 106 and tourists with other travel purposes[5] spend US$ 42. Among leisure tourists, the most relevant specific motivations are sun and sea (61.5%), ecotourism (23.2%) and cultural tourism (9.7%). Most tourists travel by air (69.9%), while 27.1% travel by road and the remaining share uses water transport modes. About 85% of total road arrivals is composed of South American tourists. The remaining 15% share refers mainly to long haul tourists visiting multiple countries.

Most tourists accommodate at hotels and other similar commercial establishments. Friends and relatives' dwellings account for 28.2% of total accommodations, while rented dwellings represent 9.2%. Only 14% of inbound tourists travel in organized package groups. About two thirds (67.3%) of tourists visiting Brazil are repeaters, while 94.9% intend to visit the country again in the future.

---

[4] Data in the remaining part of this section comes from "Estudo da demanda turística internacional 2004-2009" and "Anuário Estatístico de Turismo 2012", both published by the Tourism Ministry of Brazil.

[5] This category is ordinarily used by the Tourism Ministry of Brazil to report the statistics generated from the BITS. It includes the sub-categories visiting friends and relatives, study, health, religion, shopping, and other. The largest share of tourists in category is actually visiting friends or relatives.

The most visited Brazilian destinations among leisure tourists are Rio de Janeiro (30.0%), Foz do Iguaçu (21.4%), Florianópolis (16.7%) and São Paulo (11.5%). Business tourists are highly concentrated in São Paulo (48.8%) and Rio de Janeiro (24.9%). The most important destinations for other travel purposes are São Paulo (27.3%), Rio de Janeiro (21.6%), Belo Horizonte (6.5%) and Salvador (5.8%).

According to the BITS, 39.5% of inbound tourists in Brazil take MTTs. More than half of multidestination tourists visit two destinations (20.8% of total tourists). A small share of total inbound tourists visits six or more destinations (1.7%). The distribution of inbound tourists according to the number of destinations visited is presented in Table 4.

Table 4: Distribution of inbound trips by number of destinations visited

Destinations visited	**Relative frequency (%)**	**Cumulative frequency (%)**
1	60,5	60,5
2	20,8	81,3
3	9,3	90,6
4	4,9	95,5
5	2,8	98,3
6 or more	1,7	100,0
Total	*100,0*	

The average number of destinations visited in Brazil by inbound tourists is 1.7. This proportion varies according to tourists' and travels' characteristics. The participation of MTTs is considerably lower among low income, South American and business motivated tourists. Road travellers and those accommodating in rented dwellings also tend to visit a smaller number of destinations. The average number of destinations visited by inbound tourists in Brazil did not varied substantially from 2004 to 2010. Complete cross tabulations of the number of destinations visited and other relevant variables are presented in Appendix 1. A detailed study of the relationships between these variables is presented in Chapter 3.

References

Alegre, J., Mateo, S., & Pou, L. (2011). A latent class approach to tourists' length of stay. *Tourism Management, 32*(3), 555-563.

Alegre, J., & Pou, L. (2006). The length of stay in the demand for tourism. *Tourism Management, 27*(6), 1343-1355.

Archer, B. H., & Shea, S. (1975). Length of stay problems in tourist research. *Journal of Travel Research, 13*(3), 8-10.

Banco Central do Brasil. (2011). *Série histórica do balanço de pagamentos*. Retrieved from http://www.bcb.gov.br/htms/infecon/Seriehist.asp

Barros, C. P., Butler, R., & Correia, A. (2010). The length of stay of golf tourism: a survival analysis. *Tourism Management, 31*(1), 13-21.

Barros, C. P., Correia, A., & Crouch, G. (2008). Determinants of the length of stay in Latin American tourism destinations. *Tourism Analysis, 13*(4), 329-340.

Barros, C. P., & Machado, L. P. (2010). The length of stay in tourism. *Annals of Tourism Research, 37*(3), 692-706.

Baxter, M. J., & Ewing, G. O. (1981). Models of recreational trip distribution. *Regional Studies, 15*(5), 327-344.

Beaman, J., Jeng, J.-M., & Fesenmaier, D. R. (1997). Clarification of cumulative attractivity as a concept and its measurement: comments on Lue, Crompton, and Stewart. *Journal of Travel Research, 36*(2), 74-78.

Becker, G. S., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy, 81*(2), S279-S288.

Berman, M. D., & Kim, H. J. (1999). Endogenous on-site time in the recreation demand model. *Land Economics, 75*(4), 603-619.

Blaine, T. W., Mohammad, G., & Var, T. (1993). Demand for rural tourism: an exploratory study. *Annals of Tourism Research, 20*(4), 770-773.

Canadian Tourism Commission. (2011). *Tourism snapshot: 2010 year-in-review*. Retrieved from http://en-corporate.canada.travel/sites/default/files/pdf/yearinreview_2010_eng_final_rev.pdf

Connell, J., & Page, S. J. (2008). Exploring the spatial patterns of car-based tourist travel in Loch Lomond and Trossachs National Park, Scotland. *Tourism Management, 29*(3), 561-580.

Crompton, J. L., & Gitelson, R. J. (1979). The theory of cumulative attraction and compatibility: a case study of two major commercial leisure enterprises. *Baylor Business Studies, 10*, 7-16.

Fleischer, A., & Pizam, A. (2002). Tourism constraints among Israeli seniors. *Annals of Tourism Research, 29*(1), 106-123.

Fleischer, A., & Rivlin Byk, J. (2009). Quality, quantity and duration decisions in household demand for vacations. *Tourism Economics, 15*(3), 513-530.

Flognfeldt, T. (1992). Area, site or route: the different movement patterns of travel in Norway. *Tourism Management, 13*(1), 145-151.

Frechtling, D. C. (2006). An assessment of visitor expenditure methods and models. *Journal of Travel Research, 45*(1), 26-35.

Gokovali, U., Bahar, O., & Kozak, M. (2007). Determinants of length of stay: a practical use of survival analysis. *Tourism Management, 28*(3), 736-746.

Hellström, J. (2006). A bivariate count data model for household tourism demand. *Journal of Applied Econometrics, 21*(2), 213-226.

Hu, B., & Morrison, A. M. (2002). Tripography: can destination use patterns enhance understanding of the VFR market? *Journal of Vacation Marketing, 8*(3), 201-220.

Hunt, M. A., & Crompton, J. L. (2008). Investigating attraction compatibility in an East Texas city. *International Journal of Tourism Research, 10*(3), 237-246.

Hwang, Y.-H., & Fesenmaier, D. R. (2003). Multidestination pleasure travel patterns: empirical evidence from the American Travel Survey. *Journal of Travel Research, 42*(2), 166-171.

Hwang, Y.-H., Gretzel, U., & Fesenmaier, D. R. (2006). Multicity trip patterns: tourists to the United States. *Annals of Tourism Research, 33*(4), 1057-1078.

Jeng, J. M., & Fesenmaier, D. R. (1998). Destination compatibility in multidestination pleasure travel. *Tourism Analysis, 3*(2), 77-87.

Koo, T. T. R., Wu, C.-L., & Dwyer, L. (2009). Transport and regional dispersal of tourists: is travel modal substitution a source of conflict between low-fare air services and regional dispersal? *Journal of Travel Research, 49*(1), 106-120.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy, 74*(2), 132-157.

Larson, D. M. (1993). Joint recreation choices and implied values of time. *Land Economics, 69*(3), 270-286.

Leiper, N. (1989). Main destination ratios: analyses of tourist flows. *Annals of Tourism Research, 16*(4), 530-541.

Lew, A. A., & McKercher, B. (2002). Trip destinations, gateways and itineraries: the example of Hong Kong. *Tourism Management, 23*(6), 609-621.

Li, G., Song, H., & Witt, S. F. (2005). Recent developments in econometric modeling and forecasting. *Journal of Travel Research, 44*(1), 82-99.

Lue, C.-C., Crompton, J. L., & Fesenmaier, D. R. (1993). Conceptualization of multi-destination pleasure trips. *Annals of Tourism Research, 20*(2), 289-301.

Lue, C.-C., Crompton, J. L., & Stewart, W. P. (1996). Evidence of cumulative attraction in multidestination recreational trip decisions. *Journal of Travel Research, 35*(1), 41-49.

Machado, L. P. (2010). Does destination image influence the length of stay in a tourism destination? *Tourism Economics, 16*(2), 443-456.

Mak, J., Moncur, J., & Yonamine, D. (1977). Determinants of visitor expenditures and visitor lengths of stay: a cross-section analysis of U.S. visitors to Hawaii. *Journal of Travel Research, 15*(3), 5-8.

Mak, J., & Nishimura, E. (1979). The economics of a hotel room tax. *Journal of Travel Research, 17*(4), 2-6.

Martínez-Garcia, E., & Raya, J. M. (2008). Length of stay for low-cost tourism. *Tourism Management, 29*(6), 1064-1075.

McConnell, K. E. (1992). On-site time in the demand for recreation. *American Journal of Agricultural Economics, 74*(4), 918-925.

McKercher, B. (2001). A comparison of main-destination visitors and through travelers at a dual-purpose destination. *Journal of Travel Research, 39*(4), 433-441.

Menezes, A. G. d., Moniz, A., & Vieira, J. C. (2008). The determinants of length of stay of tourists in the Azores. *Tourism Economics, 14*(1), 205-222.

Mings, R. C., & McHugh, K. E. (1992). The spatial configuration of travel to Yellowstone National Park. *Journal of Travel Research, 30*(4), 38-46.

Ministério do Turismo. (2012). *Anuário estatístico de turismo - 2012*.

Morley, C. L. (1992). A microeconomic theory of international tourism demand. *Annals of Tourism Research, 19*(2), 250-267.

Nelson, R. L. (1958). *The selection of retail location*. New York: FW Dodge.

Nicolau, J. L., & Más, F. J. (2005). Stochastic modeling: a three-stage tourist choice process. *Annals of Tourism Research, 32*(1), 49-69.

Nicolau, J. L., & Mas, F. J. (2009). Simultaneous analysis of whether and how long to go on holidays. *Service Industries Journal*, 29(8), 1077-1092.

Oppermann, M. (1992). Intranational tourist flows in Malaysia. *Annals of Tourism Research, 19*(3), 482-500.

Oppermann, M. (1994). Length of stay and spatial distribution. *Annals of Tourism Research, 21*(4), 834-836.

Oppermann, M. (1995). A model of travel itineraries. *Journal of Travel Research, 33*(4), 57-61.

Oppermann, M. (1997). First-time and repeat visitors to New Zealand. *Tourism Management, 18*(3), 177-181.

Papatheodorou, A. (1999). The demand for international tourism in the Mediterranean region. *Applied Economics, 31*(5), 619-630.

Papatheodorou, A. (2006). Microfoundations of tourist choice. In L. Dwyer & P. Forsyth (Eds.), *International handbook on the economics of tourism* (pp. 73-88). Cheltenham (UK): Edward Elgar.

Parroco, A. M., Vaccina, F., Cantis, S. D., & Ferrante, M. (2012). Multi-destination trips: a survey on incoming tourism in Sicily. *Economics, 2012-21*, 1-19.

Paul, B. K., & Rimmawi, H. S. (1992). Tourism in Saudi Arabia: Asir National Park. *Annals of Tourism Research, 19*(3), 501-515.

Pearce, D. G. (1990). Tourism, the regions and restructuring in New Zealand. *Journal of Tourism Studies, 1*(2), 33-42.

Pearce, D. G. (1995). *Tourism today: a geographical analysis*. Harlow (UK): Longman Scientific & Technical.

Pearce, D. G. (1999). Towards a geography of the geography of tourism: Issues and examples from New Zealand. *Tourism Geographies, 1*(4), 406-424.

Peypoch, N., Randriamboarison, R., Rasoamananjara, F., & Solonandrasana, B. (2012). The length of stay of tourists in Madagascar. *Tourism Management, 33*(5), 1230-1235.

Porter, M. E. (1990). *The competitive advantage of nations*. New York: The Free Press.

Pulina, M. (2010). Modelling and forecasting length of stay. *Anatolia, 21*(2), 305-321.

Raya-Vilchez, J. M., & Martínez-Garcia, E. (2011). Nationality and low-cost trip duration. a microeconometric analysis. *Journal of Air Transport Management, 17*(3), 168-174.

Rosen, S. (1974). Hedonic prices and implicit markets: products differentiation in pure competition. *Journal of Political Economy, 82*(1), 34-55.

Rugg, D. (1973). The choice of journey destination: a theoretical and empirical analysis. *Review of Economics and Statistics, 55*(1), 64-72.

Santos, G. E. d. O., Ramos, V., & Rey-Maquieira, J. (2011). A microeconomic model of multidestination tourism trips. *Tourism Economics, 17*(3), 509-529.

Seaton, A. V., & Palmer, C. (1997). Understanding VFR tourism behaviour: the first five years of the United Kingdom Tourism Survey. *Tourism Management, 18*(6), 345-355.

Seddighi, H. R., & Theocharous, A. L. (2002). A model of tourism destination choice: a theoretical and empirical analysis. *Tourism Management, 23*(5), 475-487.

Shaw, W. D., & Feather, P. (1999). Possibilities for including the opportunity cost of time in recreation demand systems. *Land Economics, 75*(4), 592-602.

Shih, H.-Y. (2006). Network characteristics of drive tourism destinations: an application of network analysis in tourism. *Tourism Management, 27*(5), 1029-1039.

Silberman, J. (1985). A demand function for length of stay: the evidence for Virginia Beach. *Journal of Travel Research, 23*(4), 16-23.

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting: a review of recent research. *Tourism Management, 29*(2), 203-220.

Song, H., Witt, S. F., & Li, G. (2009). *The advanced econometrics of tourism demand*. New York: Routledge.

Stewart, S. I., & Vogt, C. A. (1997). Multi-destination trip patterns. *Annals of Tourism Research, 24*(2), 458-461.

Sung, H. H., Morrison, A. M., Hong, G.-S., & O'Leary, J. T. (2001). The effects of household and trip characteristics on trip types: a consumer behavioral approach for segmenting the U.S. domestic leisure travel market. *Journal of Hospitality & Tourism Research, 25*(1), 46-68.

Taplin, J. H. E., & McGinley, C. (2000). A linear program to model daily car touring choices. *Annals of Tourism Research, 27*(2), 451-467.

Thrane, C., & Farstad, E. (2011). Domestic tourism expenditures: the non-linear effects of length of stay and travel party size. *Tourism Management, 32*(1), 46-52.

Tideswell, C., & Faulkner, B. (1999). Multidestination travel patterns of international visitors to Queensland. *Journal of Travel Research, 37*(4), 364-374.

Tideswell, C., & Faulkner, B. (2003). Identifying antecedent factors to the traveler's pursuit of a multidestination travel itinerary. *Tourism Analysis, 7*(3/4), 177-190.

Tierney, P. T. (1993). The influence of state traveler information centers on tourist length of stay and expenditures. *Journal of Travel Research, 31*(3), 28-32.

Tourism Research Australia. (2008). *Travel journeys in Australia: itineraries from the top 7 international markets 2004 & 2005*. Retrieved from http://www.ret.gov.au/tourism/Documents/tra/International%20Analysis/Travel%20Journeys%20Introduction%20Final.pdf

Tourism Strategy Group. (2011). *International visitor survey*. Retrieved from http://www.tourismresearch.govt.nz/Data--Analysis/International-tourism/International-Visitors/IVS-Online-Databases/

Tussyadiah, I. P., Kono, T., & Morisugi, H. (2006). A model of multidestination travel: implications for marketing strategies. *Journal of Travel Research, 44*(4), 407-417.

U.S. Department of Commerce. (2010). *Profile of U.S. resident travelers visiting overseas destinations: 2009 outbound*. Retrieved from http://tinet.ita.doc.gov/outreachpages/download_data_table/2009_Outbound_Profile.pdf

U.S. Department of Commerce. (2011). *Profile of overseas travelers to the United States: 2010 inbound*. Retrieved from http://tinet.ita.doc.gov/outreachpages/download_data_table/2010_Overseas_Visitor_Profile.pdf

United Nations. (2010). *International recommendations for tourism statistics 2008*. Retrieved from http://unstats.un.org/unsd/publication/Seriesm/SeriesM_83rev1e.pdf

Uysal, M., McDonald, C. D., & O'Leary, J. T. (1988). Length of stay: a macro analysis for cross-country skiing trips. *Journal of Travel Research, 26*(3), 29-31.

Wall, G. (1978). Competition and complementarity: a study in park visitation. *International Journal of Environmental Studies, 13*(1), 35-41.

Walsh, R. G., & Davitt, G. J. (1983). A demand function for length of stay on ski trips to Aspen. *Journal of Travel Research, 21*(4), 23-29.

Wanhill, S. (2011). What tourism economists do. their contribution to understanding tourism. *Estudios de Economía Aplicada, 29*(3), 679-692.

Weidenfeld, A., Butler, R. W., & Williams, A. M. (2010). Clustering and compatibility between tourism attractions. *International Journal of Tourism Research, 12*(1), 1-16.

World Tourism Organization. (2006). *Tourism market trends: world overview & tourism topics*. Madrid: World Tourism Organization.

World Tourism Organization. (2010). *TSA data around the world: worldwide summary*. Retrieved from http://www.unwto.org/statistics/bali/tsa_data.pdf

World Tourism Organization. (2012). *UNWTO tourism highlights: 2012 edition*. Retrieved from http://mkt.unwto.org/en/content/tourism-highlights

Wu, C.-L., & Carson, D. (2008). Spatial and temporal tourist dispersal analysis in multiple destination travel. *Journal of Travel Research, 46*(3), 311-317.

Wu, L., Zhang, J., & Fujiwara, A. (2012). A tourist's multi-destination choice model with future dependency. *Asia Pacific Journal of Tourism Research, 17*(2), 121-132.

Wurst, C. (1955). The length-of-stay problem in tourist studies. *The Journal of Marketing, 19*(4), 357-359.

Yang, O. L. W., Hui-Min, G., & Ryan, C. (2009). Itinerary planning and structured travel - preferences by outbound Chinese holidaymakers. *Anatolia, 20*(1), 119-133.

Yang, Y., Wong, K. K. F., & Zhang, J. (2011). Determinants of length of stay for domestic tourists: case study of Yixing. *Asia Pacific Journal of Tourism Research, 16*(6), 619-633.

Zillinger, M. (2007). Tourist routes: a time-geographical approach on German car-tourists in Sweden. *Tourism Geographies, 9*(1), 64-83.

2 A microeconomic model of multidestination tourism trips

Autores: Santos, G.; Ramos, V.; Rey-Maquieira, J.A

Título: Microeconomic model of multidestination tourism trips
Año: 2011
Revista: Tourism Economics (The Business and Finance of Tourism and Recreation). ISSN 1354-8166. Nº17,pp. 509 - 529.
(United Kingdom)

3 Determinants of multidestination tourism trips in Brazil

Destination choice is a major issue in tourism research. However, although multidestination tourism trips (MTTs) frequently account for a significant share of total trips (Hwang, Gretzel, & Fesenmaier, 2006; Mings & McHugh, 1992; Oppermann, 1995; Shih; 2006; Stewart & Vogt, 1997; Tideswell & Faulkner, 1999, 2003; Wall, 1978), most tourism demand studies have assumed that all trips visit a single destination. This assumption may be misleading since destination choices may be interdependent (Wu, Zhang, & Fujiwara, 2012), especially due to interactions on preferences and on transport costs (Santos, Ramos, & Rey-Maquieira, 2011[6]).

Understanding MTTs is relevant for several tourism organizations. Transport companies may benefit from this knowledge by designing efficient transport routes and special products, such as air or train passes. Countries and tour operators may benefit from marketing optimal sets of destinations. Besides, countries can take advantage of the spatial distribution and enhanced tourist expenditure of MTTs. This type of trips is also relevant for car rental agencies since tourists frequently rent a car at one

[6] This reference corresponds to the article published from the content of Chapter 2 of this thesis.

destination and deliver it at a different site. Even firms that operate within the boundaries of the destination, such as attractions, hotels and ground operators, may benefit from knowledge about MTTs through cooperation and product bundling across destinations. Interdependent discounts for hotels at different destinations are a good example of that. Finally, single destinations may take advantage of the complementarity between destinations, developing better single and cooperative marketing strategies.

Studies focusing on MTTs are numerous and their approaches are varied. Some studies attempted to describe itinerary patterns of MTTs, proposing some theoretical categories of spatial distribution of trips (e.g., Flognfeldt, 1992; Lew & McKercher, 2002; Lue, Crompton, & Fesenmaier, 1993; Mings & McHugh, 1992; Oppermann, 1995; Stewart & Vogt, 1997). Lew and McKercher (2002) focused on the relative position of destinations within trip itineraries, stressing the importance of different types of destinations. Some other studies applied social network measures to describe MTTs' outcomes (Hwang et al., 2006; Oppermann, 1992; Shih, 2006).

Comprehensive theoretical backgrounds for the individual choice process regarding MTTs were proposed by some studies based on the economic theory. Rugg (1973) and Tussyadiah, Kono and Morisugi (2006) built theoretical choice models based on Lancaster's (1966) characteristics theory, while Santos et al. (2011) developed a theoretical model based on the traditional economic theory of consumers' choices. Cumulative attraction among destinations in a MTT itinerary has also been a research topic (Beaman, Jeng, & Fesenmaier, 1997; Crompton & Gitelson, 1979; Hunt & Crompton, 2008; Jeng & Fesenmaier, 1998; Lue, Crompton, & Stewart, 1996). These studies explain how the spatial distribution of destinations influences MTTs consumption.

A last group of studies attempted to explain MTTs consumption from empirical observations of individual choices. Nicolau and Más (2005) used a random coefficients logit model to predict the dichotomous decision of taking a single or a multiple destination tourism trip. Tideswell and Faulkner (1999, 2003) modelled the number of destinations visited on a single trip using multiple linear regressions. Some of the studies mentioned on the previous paragraphs also provided empirical assessments of MTTs' determinants through simple statistics comparisons or hypothesis tests for sample means (Hwang et al., 2006; Mings & McHugh, 1992; Oppermann, 1992, 1995; Stewart & Vogt, 1997).

The objective of this chapter is to further develop the empirical analysis of the determinants of MTTs consumption, testing theoretical propositions and providing evidences for comparison with previous studies. The set of explanatory variables is extended beyond previous analysis to include some additional determinants, such as education level, type of accommodation and season. Variations of the expected number of destinations visited across different years are also assessed. Count data models are employed, overcoming some deficiencies of previous studies and providing more reliable estimates. The family of models used also allows for the analysis of potential qualitative differences between single and multiple destination trips.

3.1 Determinants of MTTs consumption

According to the theoretical microeconomic model proposed by Santos et al. (2011), there are three sets of variables influencing individual choices regarding MTTs: transport costs, budget restrictions and preferences. These aspects are discussed in the following subsections.

3.1.1 Transport costs

The cost of visiting a given destination depends on the geographical distribution of the origin and destinations of the trip. For the sake of explanation's simplicity, let's assume that the spatial distances are a good proxy for transport costs. When destination A is in the route from origin to destination B, the transport cost of visiting A is zero if B is already visited. This contingent transport cost reduction is not limited to destinations in the same route. In fact, almost any spatial relationship between A and B will entail some reduction of the transport cost to A when B is visited. The transport cost of visiting A is only independent from the visitation of B when both destinations are at completely opposite sides from the tourist origin. Let x_A be the transport cost of visiting A when B is not visited. When B is chosen, the cost of visiting A varies from zero to x_A according to the spatial (transport cost) relationship between A and B. This contingent transport cost reduction is one the leading forces of MTTs consumption.

The transport cost reduction of visiting a given destination tends to be larger as the number of destinations included in the trip increases. Besides, this transport cost reduction associated with MTTs tends to be larger for destination located at further regions, what supports the theoretical proposition that MTTs are positively associated with distance between origin and destinations (Lue et al., 1993; Tideswell & Faulkner, 1999, 2003). Previous empirical evidence supports this positive relationship (Hwang et al., 2006; Mings & McHugh, 1992; Stewart & Vogt, 1997; Tideswell & Faulkner, 2003).

Transport costs are largely influenced by the transport mode used. Tideswell and Faulkner (1999, 2003) argue that more flexible transport modes impose lower costs for including an additional destination in the itinerary, favouring the visitation of a larger number of destinations. The flexibility argument was empirically supported by findings of association of MTTs with rented cars (Koo, Woo, & Dwyer, 2010; Tideswell & Faulkner, 2003) and owned cars (Koo et al., 2010). However, road transport might not be more flexible than air transport when distances among destinations are significantly large. In cases like Brazil, visiting the most prominent destinations by car or bus may require several days on the road. This may turn air transport into a more flexible alternative.

3.1.2 Budget restrictions

Tourism consumption requires the employment of two main resources: money and time. The effect of the availability of these resources on the choices of whether to take a MTT or a single destination trip, and how many destinations to visit is a point of discussion in the academic literature. There are theoretical arguments in favour of both positive and negative effects.

On the positive side, visiting an additional destination in a single trip may provide some extra utility for the tourist, but it also demands some additional monetary and time expenditures. Since less constrained consumers face lower opportunity costs for their resources, one could expect the average number of destinations visited in a single trip to be higher among tourists with larger monetary (Tideswell & Faulkner, 1999) and time budgets (Tideswell & Faulkner, 2003).

On the other hand, the argument of a negative impact of resources availability on MTTs consumption can be based on comparisons of the current travel with future travel opportunities. According to the usual decreasing marginal utility assumption, the utility obtained from visiting a destination is lower on larger itineraries. Apart from that, the transport expenditure required for visiting an additional destination also tends to be lower on larger itineraries due to contingent transport cost reduction. Therefore, visiting an additional destination on the current trip provides lower utility at a lower cost as compared to visiting the same destination on a future trip. In this case, less restricted consumers could prefer to visit additional destinations on future trips, taking trips more frequently but with shorter itineraries. Conversely, more constrained individuals could be expected to travel less but with larger itineraries on each trip. This proposition is compatible with the arguments presented by several studies (Ben-Akiva & Lerman, 1985; Lue et al., 1993; Tideswell & Faulkner, 1999, 2003). Tideswell and Faulkner (1999, p. 365), for instance, argued that consumers with smaller monetary budgets tend to "fit as much into the travel itinerary as possible to ensure that the economic costs incurred in making the trip are justified".

Therefore, positive and negative effects of resources availability over MTTs consumption are justifiable in theoretical terms. In fact, according to the traditional microeconomic theory, both positive and negative effects may be observed at the individual level depending on consumer's preferences (Santos et al., 2011). Nevertheless, the negative effect can only be sustained up to a point where resources scarcity starts binding overall consumption. In practice, there is an upper limit to the number of destinations that can be visited with given monetary and time budgets. In other words, it is impossible to visit a large number of destinations with very small time or monetary budgets. Thus, even if the net effect of resources availability on MTTs consumption is negative for the less constrained individuals, it still has to be positive for a set of the most constrained ones.

The academic literature has considered monetary and time constraints as completely different determinants of MTTs consumption. However, both resources constitute budgets with similar properties in terms of their use for tourism and other types of consumption (see DeSerpa, 1971). Thus, it is possible that the effects of both variables on MTTs consumption are analogous. Although the positive effect of both resources availability on MTTs consumption has been recognized in the academic literature, the possibility of a negative effect has been recognized only with respect to the monetary constraint. These arguments seem to be incomplete since, for instance, the desire to "fit as much into the travel itinerary as possible" of tourists with smaller budgets could also be applied to time constraint at certain level.

Empirical evidence has preponderantly indicated a positive effect of resources availability on the number of destinations visited. Mings and McHugh (1992) presented empirical evidence regarding the effect of monetary constraint by analysing tourists at the Yellowstone National Park in the USA. These authors found a positive relationship between monetary budgets and MTTs. Using a multinomial variable for income, Nicolau and Más (2005) estimated that the propensity to take a MTT is positively associated with income up to the category between 1,200 and 2,400 euros of monthly income. For higher categories the relationship between both variables was negative.

No direct measures of time constraint were employed in the academic literature as explanatory variables of MTTs consumption. However, length of stay was employed several times as a surrogate measure for that. That implies assuming that time availability is actually binding tourists' length of stay. Under this assumption, several studies found that tourists staying longer at the destination region tend to visit a larger number of destinations (Mings & McHugh, 1992; Oppermann, 1992, 1994; Pearce, 1990; Tideswell & Faulkner, 1999, 2003; Tourism Research Australia, 2008).

3.1.3 Preferences

Tourists' individual preferences are likely to be key determinant for MTTs consumption. According to Santos et al. (2011), besides transportation issues, preferences are the only other rational reason for choosing MTTs. Convex indifference curves lead to the preference for splitting travel time across different destinations, what allows MTTs to provide higher utility than single destination trips in some conditions.

On the other hand, concave preferences are a realistic case within the tourism destinations choice process regarding a single trip (Santos et al., 2011). This type of preferences implies that single destination trips provide higher utility than MTTs.

Several variables are related to tourists' preferences, including individuals' and travels' characteristics. The impact of personal characteristics on preferences regarding MTTs has not been theoretically discussed in the academic literature. Oppermann's (1992) empirical analysis found that younger tourists tend to visit more destinations on a single trip, while no differences between genders were found.

Tourists' preferences regarding MTTs might be influenced by their place of residence due to cultural and social aspects, as well as due to the supply of goods and services for outbound tourism. Since these variables are numerous and some of them are not easily measurable, some authors have considered the origin of tourists itself as an explanatory variable of MTTs choices (Oppermann, 1995; Stewart & Vogt, 1997).

MTTs consumption is intrinsically related to the travel purposes. Empirical evidence has shown that leisure motivations favour the visitation of a larger number of destinations in a single trip when compared to business purposes and visiting friends and relatives (Oppermann, 1992, 1995; Tideswell & Faulkner, 1999, 2003; U.S. Department of Commerce, 2011). Although visits to friends and relatives have been suggested as a potential cause of MTTs (Lue et al., 1993), empirical studies have shown that this motivation is even less associated with MTTs than business purposes (Oppermann, 1995, Tideswell & Faulkner, 1999).

Different types of leisure motivations may also affect MTTs propensity. Motivations associated with "need to know" could be argued to affect positively the number of destinations visited in a single trip (Nicolau & Más, 2005). This would be supportive of the argument that MTTs can be encouraged by the boredom of staying too long at a single site (Lue et al., 1993) since this feeling seems more likely to come up among inquisitive tourists. However, findings of Nicolau and Más (2005) were partially contradictory to this proposition. Surprisingly, these authors found that both the interest on discovering new places and broadening cultural knowledge influence the consumption of MTTs negatively. According to them "it seems that the intellectual 'need to know' is satisfied by remaining in one destination during their vacation in order to 'learn' in detail the characteristics of the place" (Nicolau & Más, 2005, p. 64).

The number of destinations visited on a single trip is not only related to tourists' main purpose of travelling, but also to their total range of purposes. Trips aimed to fulfil a broader set of purposes are more likely to visit a larger number of destinations (Lue et al., 1993). According to Tideswell and Faulkner (1999, p. 365) "as the benefits sought expand from one to many, and the capacity of individual destinations to provide the full range of benefits diminishes, the propensity to seek variety by visiting many destinations increases". This proposition was supported by empirical findings of Tideswell and Faulkner (1999, 2003).

Multiple travel purposes may be related not only to a single individual with several interests, but also to the composition of the travel party. Larger travel parties are expected to present broader sets of purposes, thus being more likely to take MTTs (Lue et al., 1993, Tideswell & Faulkner, 1999). On the other hand, small travel groups may be more inclined to "explore" the destination region due to their enhanced agility. According to this argument, smaller travel parties are expected to be more associated with MTTs (Oppermann 1992). Empirical findings are not conclusive in this point. The relationship between the travel party size and the number of destinations visited on a single trip was found to be negative but statistically insignificant by Oppermann (1992), while Tideswell and Faulkner (1999) found it to be significantly positive.

The type of travel arrangement can also be associated with the number of destinations visited. Some studies found that individuals on package tours are likely to visit fewer destinations (Oppermann, 1992; Tideswell & Faulkner, 2003, Tourism Research Australia, 2008). Nevertheless, findings of Nicolau and Más (2005) are partially conflicting with these evidences. These authors found that the use of intermediaries is positively associated with MTTs consumption.

Destination familiarity and previous visits might constitute another relevant explanatory variable of MTTs consumption. Several studies found that first-time visitors are more likely to visit a larger number of sites (Mings & McHugh, 1992; Oppermann, 1992; Tideswell & Faulkner, 2003; Tourism Research Australia, 2008). Nonetheless, Hwang et al. (2006) found this tendency only for first-time tourists from certain source markets, while Debbage (1991) found that the effect of previous travel experience on the spatial dispersion of tourists within a destination region was insignificant.

Some authors have also suggested that risk aversion can be a determinant of MTTs consumption (Lue et al., 1993; Tideswell & Faulkner, 1999). As Tideswell and Faulkner (1999, p. 366) argued, “the aggregation of a set of destinations into an itinerary reduces the risk of being disappointed or dissatisfied with the tourist experience”. However, the existent academic literature has not discussed properly which are the determinants of risk aversion for tourists within the MTT paradigm. One may expect that more risk averse tourists would be those investing a larger share of their resources on the trip. Those investing a smaller share of their resources would tend to be less worried about the risk of spending their time and money on an unsatisfactory travel. Following that, individuals with scarce monetary and time resources would tend to visit more destinations since the investment on the trip is relatively larger for them, what would add to the argument of the negative impact of resources availability on MTTs consumption. Moreover, from the same argument it is possible to predict that more distant and longer trips would imply a tendency to visit more destinations, what reinforces previous expectations regarding these variables (Tideswell & Faulkner, 1999).

It is necessary to stress that some explanatory variables discussed here may also be choice outcomes. Transport mode, type of travel arrangements and party size are some of the most evident examples of variables that are chosen by the tourist. However, most explanatory variables may present the same restriction at some level. As an example, even personal income may be considered a choice outcome when labour-leisure trade-off is regarded (Stabler et al., 2010). In fact, tourists’ choices are often analysed as a multistage structure where some variables are chosen before others (Dellaert, Ettema, & Lindh, 1998; Eugenio-Martin, 2003; Morley, 1992; Nicolau & Más, 2005, 2008; Vassallo & Oliveira, 2009). This structure is probably not the same for all tourists and it may change over time and according to the occasion. Moreover, it is reasonable to assume that this multistage process is fluid instead of rigid, meaning that the choice process can flow back and take different paths, crossing the same choice stage more than once.

This complex structure of tourists' choices gives support to the use of choice outcomes as explanatory variables of MTTs consumption. Causality cannot be taken as a major concern in this context under the penalty of having almost no completely independent explanatory variables to work with. In fact, causality has not yet been studied or tested for most variables within the MTT paradigm. A remarkable exception is the study of Ye, Pendyala and Gottardi (2007) who tested the causality direction between transport mode and multidestination daily work and non-work trips. These authors found evidences in favour of the hypothesis that transport mode is influenced by the number of sites to be visited on a single trip.

3.2 Modelling consumers' choices of MTTs

MTTs consumption among international tourists visiting Brazil was modelled in order to assess theoretical propositions and empirical findings of previous studies. A large data set of 183,000 observations was obtained from the official Brazilian International Tourist Survey (BITS). The BITS was conducted by the Foundation Institute of Economic Research and financially supported by the Brazilian Tourism Ministry. Its main objective was to provide official tourism statistics for the country. Data was collected through personal interview from 2004 to 2010 at the 27 main gateways of the country, including 15 airports and 12 land borders[7].

According to the BITS, roughly 40% of all tourists visiting Brazil took MTTs. The average number of destinations visited was 1.74. The distribution of tourists according to the number of destinations visited is presented at Table 5.

In the modelling process, the multidestination nature of trips was described by the number of destinations visited supplementary to the first one. Compared with the dichotomous variable identifying MTTs versus single destination trips used by Nicolau and Más (2005), this variable provides a more detailed representation of MTTs consumption. The consideration of destinations supplementary to the first one is justified by the fact that all individuals in the sample visited at least one destination in Brazil. The actual variable considered was, therefore, the total number of destinations

[7] A detailed description of this data source is presented at Section 1.4 of this thesis.

visited minus one. This definition allowed for testing the existence of qualitative differences between MTTs and single destination trips, as discussed afterwards.

Table 5: Distribution of tourists according to the number of destinations visited

Number of destinations visited	Frequency (%)
1	60.5
2	20.8
3	9.3
4	4.9
5	2.8
6	1.7
Total	100

Other international destinations were ignored due to information inexistence in the BITS. The number of Brazilian destinations visited was obtained from the inventory of destinations where tourists stayed overnight. A destination was geographically defined as a municipality. The BITS reported a maximum of six destinations visited by each tourist. When more than six destinations were visited, the ones with longer length of stay were selected for reporting. Therefore, the number of destinations visited by each tourist was censored at six. Let y_i^* be the actual number of destinations visited, the modelled variable y_i is

$$y_i = y_i^* - 1 \quad \text{if} \quad y_i^* < 6$$

$$y_i = 5 \quad \text{if} \quad y_i^* \geq 6$$

Based on previously discussed theoretical arguments and empirical findings, the most comprehensive available set of explanatory variables was selected from the BITS database. Four groups of explanatory variables were used. The first group regards tourists' profiles, including gender, age, education level, origin and income. Origin was built as a multinomial variable specifying the most relevant countries as an independent value, while less relevant countries were aggregated by continent. Relevant selected countries were those eleven from where more than 100,000 tourists go to Brazil yearly. Income was measured as monthly household income in thousand constant US Dollars of

2010. This variable was used as a proxy for tourists' monetary constraint. Squared age and income were also introduced in the model as explanatory variables in order to allow for non-monotonic effects over MTTs consumption.

The second group of explanatory variables includes the most relevant trip characteristics according to previously discussed theoretical propositions. This group includes overnights, transport mode and travel purpose. Similarly to the income variable, the total number of overnights in Brazil was employed as a proxy for tourists' time constraint. The squared number of overnights was also used to capture eventual non-monotonic effects. Transport mode regarded the use of road or air transport to return from Brazil to the country of origin. Road transport also includes scheduled bus services. The main purpose of the trip to Brazil was described as a multinomial variable, including leisure and non-leisure alternatives.

The third group of independent variables includes some additional trip characteristics, such as type of trip organization, accommodation, party size, party type, first visit to Brazil, number of previous visits to the country and season. Type of trip organization distinguished independent from package organized trips. Party size described the number of people in the travel party. The squared value of this variable was also used. Party type indicated the relationship among tourists in the same travel group. First visit to Brazil is a dichotomous variable, while the number of previous visits and the square of this value were also used as explanatory variables. Seasons were defined as high season (December to February), low season (July and August) and off-season (remaining months).

Finally, the fourth group was composed of a single element: year. This variable was numeric and where 2004 = 0. This variable aims to capture an eventual time trend.

Explanatory variables of groups two and three present some degree of simultaneity since their values are the outcomes of tourists' choices. Using these variables as regressors requires some careful interpretation. By introducing these variables at the econometric model, the coefficients of the variables regarding tourists' profiles should be interpreted as the effect of a given regressor on the expected number of destinations visited for trips with the same characteristics. For example, the coefficients of origin indicate the expected effect of residing on a particular country while keeping constant other trip's characteristics.

Three model's specifications, ranging from the simplest to the most comprehensive, were estimated in order to present the effect of including additional variables in the estimation. Specification 1 includes only tourists' profiles and year as explanatory variables, while Specification 2 additionally included the most relevant trips' characteristics and Specification 3 included all available regressors.

Furthermore, a study of correlations among all explanatory variables was conducted in order to assess the magnitude of possible multicollinearity. The estimated average absolute correlation among all explanatory variables was only 0.048, while the 90% percentile is 0.11. The largest correlation was found between the travel purpose of visiting friends and relatives and accommodation on friends' and relatives' dwellings (0.65). In sum, correlation among regressors was estimated not to be substantial. Moreover, the huge number of observations ensures enough variability to reduce any risk arising from multicollinearity.

3.2.1 Estimation method

The number of destinations visited is a count data variable taking only nonnegative integer values. Modelling this type of variable requires special statistical procedures. Nonnegativity and the integer nature of alternative values are not properly treated by the multiple linear regression approach used by Tideswell and Faulkner (1999, 2003). Instead, count data models recognize these characteristics, providing more reliable estimates.

The simplest usual count data model is the Poisson regression. In this model, the dependent variable is drawn from a Poisson distribution, such that

$$P(y_i = j) = \frac{e^{-\lambda_i}\lambda_i^j}{j!}, \;\; j = 0, 1, 2 \ldots$$

The parameter λ_i is related to the regressors' vector $\mathbf{x_i}$. The most common formulation for λ_i is loglinear, such that $\ln(\lambda_i) = \boldsymbol{\beta}\mathbf{x_i}$. In this case, the expected value of y_i equals $\exp(\boldsymbol{\beta}\mathbf{x_i})$ and the marginal effect of $\mathbf{x_i}$ is

$$\frac{\partial E(y_i|\mathbf{x_i})}{\partial \mathbf{x_i}} = \lambda_i \boldsymbol{\beta}$$

In the Poisson regression, the variance and the mean of y_i are assumed to be equal. This is a binding restriction since overdispersion (i.e., $var(y_i)>E(y_i)$) is frequently observed in the real world. A less restrictive model is the negative binomial, which is a generalization of the Poisson regression that allows for overdispersion. In the negative binomial model y_i follows a Poisson distribution with parameter μ_i such that

$$\ln(\mu_i) = \boldsymbol{\beta}\mathbf{x_i} + \varepsilon_i = \ln(\lambda_i) + \ln(u_i)$$

where ε_i is the error term. For mathematical convenience, u_i is assumed to follow a Gamma distribution with mean one and variance θ (i.e., $e^{\varepsilon_i} \sim \Gamma(1, \theta)$). Larger values of θ indicate greater overdispersion. The Poisson regression corresponds to a special case of the negative binomial model where θ=0. The adequacy of the negative binomial model can be tested through the significance of θ.

The nature of the decisions leading to the number of destinations visited on a single trip is another issue that influences the econometric design adopted. Several questions may be addressed to clarify this nature: Is the choice of taking a MTT generated by the same process that leads to the visitation of an additional destination? Or is the decision of taking a MTT qualitatively different from simply adding one destination to the itinerary? In other words, is there a qualitative difference between single and multiple destination trips or are they only quantitatively differentiated?

If a qualitative difference exists, single and multiple destination trips require different consideration in the modelling process. Distinct data generating processes may lead to an excess of observed zeros, that is, an unexpected high frequency of single destination trips. If this is the case, the negative binomial model can be altered in order to allow for zero inflation, accounting not only for the quantitative difference between zero and one, but also to the potential qualitative difference between these two values.

Zero inflated negative binomial models account for excess of zeros by introducing an extra probability mass at zero with the probability π_i and reducing the probability of other non-zero frequencies by $1 - \pi_i$. The probability parameter π_i is given by

$$\pi_i = \Omega(\tau\boldsymbol{\beta}\mathbf{x_i})$$

where Ω(.) is the logistic cumulative distribution. The probability density function of y_i is

$$P(y_i = j) = (1 - \pi_i)f(y_i) + 1(y = 0)\pi_i$$

where f(.) is the negative binomial probability density function. This model is known as zero-inflated-tau negative binomial model. The significance of the zero-inflation can be assessed through the Vuong (1989) statistic.

Finally, taking also account for the censored nature of the dependent variable, the actual econometric instrument used in this chapter was a censored zero-inflated-tau negative binomial model.

3.2.2 Results

Estimated coefficients and standard errors of the three estimated specifications are presented at Table 6. The significance of each parameter is indicated by a superscripted letter. The first line of each multinomial variable presents Wald statistic for the hypothesis that all coefficients equal zero simultaneously, that is, for the hypothesis that the multinomial variable as a whole is not significant. For example, the Wald statistic indicates that the multinomial variable origin is significant at the 0.1% level at Specification 1, despite the non-significance of the parameters associated with some specific origin countries. The same is done to quantitative variables used in their squared and non-squared forms. The first category of multinomial variables indicates their reference groups.

Table 6: Coefficients estimates

Variable	Specification 1 Coef.	Std. error	Specification 2 Coef.	Std. error	Specification 3 Coef.	Std. error
Gender						
Female†	0.00		0.00		0.00	
Male	-0.038[a]	0.0046	0.025[a]	0.0044	0.029[a]	0.0045
Age		1109.1[a]		447.4[a]		223.1[a]
Exponent 1	-0.029[a]	0.0010	-0.014[a]	0.00093	-0.011[a]	0.00094
Exponent 2	0.00030[a]	1.1E-05	0.00013[a]	1.0E-05	0.00010[a]	1.1E-05
Education level		227.2[a]		827.3[a]		774.1[a]
Under high school†	0.00		0.00		0.00	
High school	0.089[a]	0.011	0.10[a]	0.010	0.10[a]	0.010
Graduation	0.14[a]	0.011	0.20[a]	0.010	0.19[a]	0.010
Post-graduation	0.14[a]	0.011	0.24[a]	0.011	0.24[a]	0.011
Origin		5832.1[a]		5281.3[a]		4486.7[a]
Africa						
All countries†	0.00		0.00		0.00	
Asia and Oceania						
All countries	0.29[a]	0.020	0.17[a]	0.019	0.15[a]	0.019
Central America						
All countries	-0.010	0.030	0.024	0.029	0.013	0.028
Europe						
England	0.27[a]	0.020	0.18[a]	0.019	0.17[a]	0.019
France	0.37[a]	0.020	0.28[a]	0.019	0.27[a]	0.018
Germany	0.31[a]	0.020	0.22[a]	0.019	0.22[a]	0.018
Italy	0.17[a]	0.019	0.059[b]	0.018	0.066[a]	0.018
Portugal	0.039[c]	0.019	-0.033[d]	0.018	-0.022	0.018
Spain	0.19[a]	0.020	0.10[a]	0.019	0.10[a]	0.019
Other Europe	0.31[a]	0.019	0.21[a]	0.018	0.20[a]	0.017
North America						
USA	0.062[a]	0.018	0.045[c]	0.017	0.047[b]	0.017
Other North America	0.11[a]	0.020	0.093[a]	0.019	0.082[a]	0.019

(continued)

Table 7: Coefficients estimates (continued)

	Specification 1		Specification 2		Specification 3	
Variable	**Coef.**	**Std. error**	**Coef.**	**Std. error**	**Coef.**	**Std. error**
South America						
Argentina	-0.40[a]	0.019	-0.33[a]	0.018	-0.30[a]	0.018
Chile	-0.27[a]	0.021	-0.21[a]	0.020	-0.21[a]	0.020
Paraguay	-0.78[a]	0.024	-0.54[a]	0.022	-0.49[a]	0.022
Uruguay	-0.39[a]	0.024	-0.25[a]	0.023	-0.23[a]	0.023
Other South America	-0.16[a]	0.020	-0.10[a]	0.019	-0.10[a]	0.019
Income		14.6[a]		93.1[a]		100.3[a]
Exponent 1	0.0026[b]	0.00085	0.0077[a]	0.00081	0.0080[a]	0.00081
Exponent 2	-7.9E-05[a]	2.1E-05	-0.00016[a]	2.0E-05	-0.00016[a]	2.0E-05
Overnights				9942.0[a]		11143.7[a]
Exponent 1			0.014[a]	0.00014	0.015[a]	0.00014
Exponent 2			-4.4E-05[a]	4.6E-07	-4.6E-05[a]	4.6E-07
Transport mode						
Road†			0.00		0.00	
Air			0.14[a]	0.0077	0.13[a]	0.0081
Purpose				4858.6[a]		3757.7[a]
Leisure purposes						
Sun and sea†			0.00		0.00	
Ecotourism			0.28[a]	0.0080	0.25[a]	0.0081
Cultural tourism			0.21[a]	0.0085	0.19[a]	0.0085
Other			-0.062[a]	0.012	-0.072[a]	0.012
Non-leisure purposes						
Business			-0.23[a]	0.0065	-0.22[a]	0.0069
Visit friends and relatives			-0.11[a]	0.0060	-0.062[a]	0.0069
Other			-0.14[a]	0.012	-0.13[a]	0.012
Trip organization						
Independent†					0.00	
Package					-0.11[a]	0.0067

(continued)

Table 7: Coefficients estimates (continued)

Variable	Specification 1 Coef.	Specification 1 Std. error	Specification 2 Coef.	Specification 2 Std. error	Specification 3 Coef.	Specification 3 Std. error
Accommodation						751.6[a]
Hotels and counterparts†					0.00	
Friends or relatives' dwellings					-0.058[a]	0.0060
Rented dwelling					-0.14[a]	0.0084
Owned dwelling					-0.24[a]	0.010
Other					0.023[d]	0.014
Party size						38.2[a]
Exponent 1					0.024[a]	0.0046
Exponent 2					-0.0016[b]	0.00053
Party type						676.7[a]
Alone†					0.00	
Couple					0.15[a]	0.0072
Family					0.044[a]	0.0080
Friends					0.13[a]	0.0069
Other					0.18[a]	0.019
First visit to Brazil						
No†					0.00	
Yes					0.072[b]	0.023
Number of previous visits						5.3[d]
Exponent 1					7.6E-07	2.3E-05
Exponent 2					-2.9E-07[c]	1.3E-07
Season						24.7[a]
Off-season†					0.00	
Summer season					0.0075	0.0049
Winter season					0.025[a]	0.0051
Year	0.0093[a]	0.0016	0.0067[b]	0.0022	0.0082[a]	0.0022
Constants						
β_0	0.68[a]	0.030	-0.026	0.030	-0.23[a]	0.032
θ	0.12[a]	0.0075	0.016[a]	0.0046	0.0028	0.0043
τ	-1.9[a]	0.046	-2.0[a]	0.033	-2.0[a]	0.032

† indicates the reference group for each multinomial variable
a = p<0.001; b = p<0.01; c = p<0.05; d = p<0.1

Starting from Specification 1, most additional variables included in specifications 2 and 3 were significant. This indicates that the hazard of multicollinearity was overcome by the large number of observations. Moreover, it points out that simpler models might present biases related to the omission of relevant variables. This sort of bias might also be present in some previous studies that employed relatively narrow sets of explanatory variables. In this sense, the inclusion of a larger set of regressors may also be considered an important contribution of this study.

On the other hand, all three specifications provided similar estimates regarding common variables. The great majority of coefficients and significance levels indicate the same qualitative interpretation. In fact, the only significant contradictory sign of an estimated parameter is the one regarding gender at Specification 1. Considering the significance of additional variables and the similarity of qualitative interpretation of the parameters regarding common regressors, only estimates of Specification 3 are going to be analysed in the following.

As indicated by the superscript "a", most explanatory variables were significant at the 0.001 level. Besides, all multinomial variables were significant as whole at this same level. In other words, the hypothesis that all coefficients regarding the same multinomial variable are simultaneously equal to zero was not confirmed in any case. The same is true for quantitative variables introduced with exponents one and two, expect for the number of previous visits at Specification 3 ($p=0.07$). In sum, the set of explanatory variables used was highly significant to explain the consumption of MTTs.

Individuals' characteristics display relevant effects on MTTs consumption. Male tourists tend to visit a larger number of destinations according to specifications 2 and 3. The coefficients of age and age^2 were significant ($p<0.001$), indicating a non-linear relation. MTTs consumption is decreasing on tourists' age up to 54 years old. For individuals older than this, the relationship between these two variables is positive. Hence, expected MTTs consumption is lower for middle aged tourists and higher for younger and older individuals.

MTTs consumption is associated with higher levels of education. One possible explanation for this relationship is that the "need to know" is stronger among more educated individuals. As argued by Nicolau and Más (2005), this need might be a cause of MTTs consumption. Besides, higher education levels may also contribute to the visitation of several destinations through the enhanced ability to organize the trip and overcome travel situations.

Tourists' origins play a significant role on the explanation of MTTs. Europeans, Asians and Oceanians present the highest propensity to visit multiple destinations. France is the leading country in this ranking. North Americans present a medium tendency to choose this type of trip. On the other hand, South American tourists usually visit a smaller number of destinations, Paraguayans been the ones with the lowest tendency to visit multiple destinations. These results support the positive effect of distance between origin and destination region over MTTs consumption. Moreover, they show that differences on cultural, social, economic and tourism market environments across different countries have an important effect over MTTs consumption.

Income was found to have a positive effect over the number of destinations visited up to a monthly income of US$ 25,000. Only about 1.4% of all tourists visiting Brazil had an income larger than that. Therefore, the positive effect of income is dominant. However, the magnitude of this effect is decreasing on income. In sum, the marginal effect of monetary resources availability on MTTs consumption is positive and decreasing for the large majority and negative for the very less constrained tourists.

Similar results were found regarding time constraint. In this case, the maximum of the MTTs function with respect to time availability happens at a very high duration (162 overnights). Only 0.8% of all tourists visiting Brazil stayed longer than that. Therefore, the marginal effect of time availability is also positive and decreasing for most tourists. Nevertheless, this effect is negative for the very less time constrained individuals.

The positive and decreasing marginal effects of monetary and time constraints over MTTs consumption for the large majority are a novelty in the academic literature. The non-monotonic shape of the MTTs consumption function with respect to these two constraints enlightens the conflicting theoretical arguments presented by Tideswell and Faulkner (1999, 2003). Moreover, the similarity between the effects of monetary and time constraints supports the proposition discussed earlier in this chapter that both resource constraints are analogous and should be treated as such.

Air travellers tend to visit a larger number of destinations than road travellers. This finding is conflicting with theoretical propositions of Tideswell and Faulkner (1999, 2003) and empirical findings of Tideswell and Faulkner (2003) and Koo et al. (2010). The explanation for this outcome may be the large distances among most important Brazilian destinations. In this context, air transport might be the most flexible alternative.

Regarding travel purposes, ecotourism trips present the highest association with MTTs consumption, followed by cultural tourism. These findings further supports theoretical expectations that MTTs are associated with the "need to know" on tourism trips. The coefficients of these two purposes and the coefficient of sun and sea oriented trips show that, consistently with previous studies, leisure tourists are more likely to take MTTs than individuals taking non-leisure trips. Trips aimed at visiting friends and relatives are slightly less associated with MTTs than sun and sea trips, while business trips are the ones less associated with MTTs.

Individuals on package tours tend to visit a smaller number of destinations, confirming the findings of Oppermann (1992) and Tideswell and Faulkner (2003), though contradicting those from Nicolau and Más (2005). Again, this finding supports the argument that more flexible types of trips are more associated with MTTs. However, it is necessary to note that this result might be substantially associated with the characteristics of the specific destination regions under scrutiny. Besides, individuals' choices regarding the characteristics of their package tours are relevantly restricted by the supply side.

MTTs are more associated with hotels than any other specific type of accommodation. Conversely, owned and rented dwellings are less associated with this type of trip than friends' and relatives' dwellings. The explanation for these findings may be related to the flexibility of each accommodation type. Arguably, more flexible types of accommodation are more associated with MTTs.

The effect of party size over MTTs consumption is positive for the large majority of groups. The positive effect happens for parties up to 7.5 people. Besides party size, the type of travel party is also relevant, being couples and friends the most associated with larger sets of destinations visited. On the other extreme are single person and family travel parties.

First time tourists are expected to visit a larger number of destinations than repeaters. Besides, although the coefficient related to the number of previous visits is not significant, the parameter regarding the squared variable is significant at 5%. The negative sign in the latter case shows that additional previous trips to repeaters also lead to a lower number of destinations visited. Therefore, evidence shows that the more frequently the tourist visits the country, the weaker the MTT behaviour.

MTTs are more associated with the winter season. The explanation for this may be related to variety seeking since the winter season in Brazil is more appropriate to satisfy diverse travel purposes. No significant difference between off-season and the summer season was found.

The positive and significant coefficient for the variable year shows that the expected number of destinations visited by any given tourist increased over the last years. This may be associated with structural changes in the Brazilian tourism supply, such as transport prices decrease, and the development of secondary destinations. Changes in the image of the country arising from economic prosperity and the hosting of large events possibly exert some additional influence. Moreover, structural worldwide changes in tourism consumption related to economic and social changes may also have had some effect over MTTs consumption.

Finally, the negative binomial overdispersion parameter was not significant ($p>0.05$). Thus, a Poisson regression may be as adequate as a negative binomial to model the number of destinations visited in a single trip. On the other hand, the Vuong statistic was significantly positive ($p<0.001$), what leads to the rejection of the non-zero inflated model. Thus, there is evidence in favour of the introduction of zero-inflation on the model. This outcome indicates that deciding to take a MTT is not as simple as adding a second destination to the itinerary. In fact, it implies a change in the type of trip. Adding the second destination to the itinerary is significantly less likely than adding the following ones. This finding has relevant policy implications. When companies and destinations intend to sell their products to tourists by including additional destinations to their itineraries, they should target multiple destinations tourists first.

3.3 Conclusion

The determinants of MTTs consumption were analysed from the theoretical and empirical perspective. The number of destinations visited by international tourists in Brazil was modelled using a large microdata set with 183,000 observations. The censored zero-inflated-tau negative binomial model employed presents an enhanced level of sophistication when compared to most previous studies. Although overdispersion was not confirmed, zero inflation was estimated to be significant. This is an important finding of this study, indicating that there is a qualitative difference between single and multiple destination trips. Therefore, taking a MTT is not as simple as adding a destination to the itinerary. In fact, stepping from a single to a double destination trip also implies taking a different type of trip. Tourists' decision to take single or multiple destination trips is different and somehow detached from the decision of how many destinations to visit. Moreover, adding a second destination to the travel itinerary is relatively more difficult than expected. This finding should be taken into account by companies and destinations when marketing additional destinations for existing itineraries. Tourists should be considered first as single or multiple destination travellers. If a multidestination trip is to be taken, then the number of destinations visited is relevant supplementary information.

A comprehensive set of regressors was used in this study, avoiding biases related to the omission of relevant variables. This important contribution was possible due to the large dataset. Besides, it provided some new evidences on the determinants of MTTs consumption. Most findings were supportive of previous theoretical argumentations and empirical findings. Regarding personal characteristics, MTTs are associated with male, younger or older tourists with higher level of education. Income and time budgets have positive but decreasing marginal effects over MTTs consumption for the majority. This result is particularly interesting to enlighten previous discussions on the effect of these variables. It also shows that the effects of both resources are analogous.

The expected number of destinations visited is higher for long haul tourists. Moreover, MTTs are more associated with leisure purposes rather than business or visiting friends and relatives. Regarding travels' characteristics, MTTs are associated with independent trips, air transport, hotels, large travel parties, first-timers and winter season. The association between air transport and MTTs constitute another original finding of this chapter. This relationship might be related to the large distances between the most important Brazilian tourist destinations, what may turn air transport into the most flexible alternative.

Estimates provided by this study may be useful for the development of multidestination products, such as transport passes, package tours and combined products. These products should target tourists and trips with characteristics found to be more associated with MTTs. Tourists with stronger propensity to MTTs should also be targeted by discount initiatives related to multidestination product bundling. Multidestination products promotion and marketing should also be oriented towards these tourists. The knowledge on tourists' and trips' characteristics associated with MTTs may be also used to target destination cooperative marketing campaigns. Finally, this information may help to forecast inbound tourism patterns arising from changes in the value of the explanatory variables.

References

Beaman, J., Jeng, J.-M., & Fesenmaier, D. R. (1997). Clarification of cumulative attractivity as a concept and its measurement: comments on Lue, Crompton, and Stewart. *Journal of Travel Research, 36*(2), 74-78.

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand*. Cambridge, MA: MIT Press.

Crompton, J. L., & Gitelson, R. J. (1979). The theory of cumulative attraction and compatibility: a case study of two major commercial leisure enterprises. *Baylor Business Studies, 10*, 7-16.

Debbage, K. G. (1991). Spatial behavior in a Bahamian resort. *Annals of Tourism Research, 18*(2), 251-268.

Dellaert, B. G. C., Ettema, D. F., & Lindh, C. (1998). Multi-faceted tourist decisions: a constraint-based conceptual framework to describe tourists' sequential choices of travel components. *Tourism Management, 19*(4), 313-320.

DeSerpa, A. C. (1971). A theory of the economics of time *The Economic Journal, 81*(324), 828-846.

Eugenio-Martin, J. L. (2003). Modelling determinants of tourism demand as a five-stage process: a discrete choice methodological approach. *Tourism and Hospitality Research, 4*(4), 341-354.

Flognfeldt, T. (1992). Area, site or route: the different movement patterns of travel in Norway. *Tourism Management, 13*(1), 145-151.

Hunt, M. A., & Crompton, J. L. (2008). Investigating attraction compatibility in an East Texas city. *International Journal of Tourism Research, 10*(3), 237-246.

Hwang, Y.-H., Gretzel, U., & Fesenmaier, D. R. (2006). Multicity trip patterns: tourists to the United States. *Annals of Tourism Research, 33*(4), 1057-1078.

Jeng, J. M., & Fesenmaier, D. R. (1998). Destination compatibility in multidestination pleasure travel. *Tourism Analysis, 3*(2), 77-87.

Koo, T. T. R., Wu, C.-L., & Dwyer, L. (2009). Transport and regional dispersal of tourists: is travel modal substitution a source of conflict between low-fare air services and regional dispersal? *Journal of Travel Research, 49*(1), 106-120.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy, 74*(2), 132-157.

Lew, A. A., & McKercher, B. (2002). Trip destinations, gateways and itineraries: the example of Hong Kong. *Tourism Management, 23*(6), 609-621.

Lue, C.-C., Crompton, J. L., & Fesenmaier, D. R. (1993). Conceptualization of multi-destination pleasure trips. *Annals of Tourism Research, 20*(2), 289-301.

Lue, C.-C., Crompton, J. L., & Stewart, W. P. (1996). Evidence of cumulative attraction in multidestination recreational trip decisions. *Journal of Travel Research, 35*(1), 41-49.

Mings, R. C., & McHugh, K. E. (1992). The spatial configuration of travel to Yellowstone National Park. *Journal of Travel Research, 30*(4), 38-46.

Morley, C. L. (1992). A microeconomic theory of international tourism demand. *Annals of Tourism Research, 19*(2), 250-267.

Nicolau, J. L., & Más, F. J. (2005). Stochastic modeling: a three-stage tourist choice process. *Annals of Tourism Research, 32*(1), 49-69.

Oppermann, M. (1992). Intranational tourist flows in Malaysia. *Annals of Tourism Research, 19*(3), 482-500.

Oppermann, M. (1994). Length of stay and spatial distribution. *Annals of Tourism Research, 21*(4), 834-836.

Oppermann, M. (1995). A model of travel itineraries. *Journal of Travel Research, 33*(4), 57-61.

Pearce, D. G. (1990). Tourism, the regions and restructuring in New Zealand. *Journal of Tourism Studies, 1*(2), 33-42.

Rugg, D. (1973). The choice of journey destination: a theoretical and empirical analysis. *Review of Economics and Statistics, 55*(1), 64-72.

Santos, G. E. d. O., Ramos, V., & Rey-Maquieira, J. (2011). A microeconomic model of multidestination tourism trips. *Tourism Economics, 17*(3), 509-529.

Shih, H.-Y. (2006). Network characteristics of drive tourism destinations: an application of network analysis in tourism. *Tourism Management, 27*(5), 1029-1039.

Stabler, M., Papatheodorou, A., & Sinclair, M. T. (2010). *The economics of tourism* (2 ed.). London: Routledge.

Stewart, S. I., & Vogt, C. A. (1997). Multi-destination trip patterns. *Annals of Tourism Research, 24*(2), 458-461.

Tideswell, C., & Faulkner, B. (1999). Multidestination travel patterns of international visitors to Queensland. *Journal of Travel Research, 37*(4), 364-374.

Tideswell, C., & Faulkner, B. (2003). Identifying antecedent factors to the traveler's pursuit of a multidestination travel itinerary. *Tourism Analysis, 7*(3/4), 177-190.

Tourism Research Australia. (2008). *Travel journeys in Australia: itineraries from the top 7 international markets 2004 & 2005*. Retrieved from http://www.ret.gov.au/tourism/Documents/tra/International%20Analysis/Travel%20Journeys%20Introduction%20Final.pdf

Tussyadiah, I. P., Kono, T., & Morisugi, H. (2006). A model of multidestination travel: implications for marketing strategies. *Journal of Travel Research, 44*(4), 407-417.

U.S. Department of Commerce. (2011). *Profile of overseas travelers to the United States: 2010 inbound*. Retrieved from http://tinet.ita.doc.gov/outreachpages/download_data_table/2010_Overseas_Visitor_Profile.pdf

Vassallo, M. D., & Oliveira, A. V. M. d. (2009). Modelagem dos determinantes da escolha por tipo de viagem e por destinos de viagens turísticas no Brasil. *Revista de Literatura dos Transportes, 3*(1), 40-67.

Vuong, Q. (1989). Likelihood ratio tests for model selection of non-nested hypothesis. *Econometrica, 57*(2), 307-333.

Wall, G. (1978). Competition and complementarity: a study in park visitation. *International Journal of Environmental Studies, 13*(1), 35-41.

Wu, L., Zhang, J., & Fujiwara, A. (2012). A tourist's multi-destination choice model with future dependency. *Asia Pacific Journal of Tourism Research, 17*(2), 121-132.

Ye, X., Pendyala, R. M., & Gottardi, G. (2007). An exploration of the relationship between mode choice and complexity of trip chaining patterns. *Transportation Research: Part B, 41*(1), 96-113.

4 Length of stay at multiple destinations of a tourism trip

Aggregate tourism is measured in several ways, such as tourists' arrivals, receipts and overnights. These aggregate measures have individual counterparts, such as tourism participation, destination choice, expenditure and length of stay. Therefore, length of stay is an alternative measure of individual tourism demand (Lim, 1997).

Tourists' length of stay and trips' durations have been sometimes mistakenly used as synonyms. Length of stay denotes the amount of time that the tourist spends at a given destination, while duration of a tourism trip refers to the length of time between departure and return to home. Two main differences between these terms must be stressed. First, length of stay does not include the time spent on transport, while duration of the trip does (United Nations, 2010). Second, length of stay refers to a single destination, while duration of the trip may include stays at several different locations when a multidestination tourism trip (MTT) is regarded.

The issue of the number of destinations is particularly interesting. Duration of the trip and length of stay at each destination may move in different directions when the number of destinations visited in the trip increases. As compared to a single destination trip, an overall longer MTT might be constituted of several shorter stays at different destinations. Moreover, a given factor, such as income or tastes, might have different impacts on tourists' length of stay and overall duration of the trip since this factor may also influence the number of destinations visited, as has been evidenced at Chapter 3. Therefore, the length of stay at tourism destinations in the multidestination trip context requires particular consideration. Nevertheless, no previous study has analysed tourists' length of stay within the MTT context. While some previous studies analysed the duration of trips, all existent studies about tourists' length of stay focused on single destination trips. Thus, researchers may have overlooked particularities of the length of stay when more than one destination is visited in a single trip.

Understanding the determinants of tourists' length of stay[8] is useful in three main ways. First, knowledge in this area can be used to forecast variations of tourists' length of stay caused by changes in its determinants. For instance, which is the effect of the aging process of world population over the average tourists' length of stay? Which is the impact of an economic crisis over this variable? How the average length of stay will be affected by the expansion of a given source market? Tourism destinations can benefit from precise knowledge about the determinants of MTT by developing appropriate infrastructure, offering convenient services and assessing tourism costs. For instance, tourists staying shorter demand proportionally more transport and information services than those staying longer. Besides, these tourists require larger administration costs for some companies (Martínez-Garcia & Raya, 2008). On the other hand, longer stays may imply a larger stress of local resources (Alegre & Pou, 2006).

Second, some destinations might be interested in attracting tourists with longer stays due to their propensity for larger total expenditures, while others might be interested on tourists with the opposite profile because they usually present larger daily expenditures (Thrane & Farstad, 2011). Both types of destinations can benefit from understanding tourists' behaviour by focusing on specific market segments, and consequently pushing the average length of stay towards the desired direction.

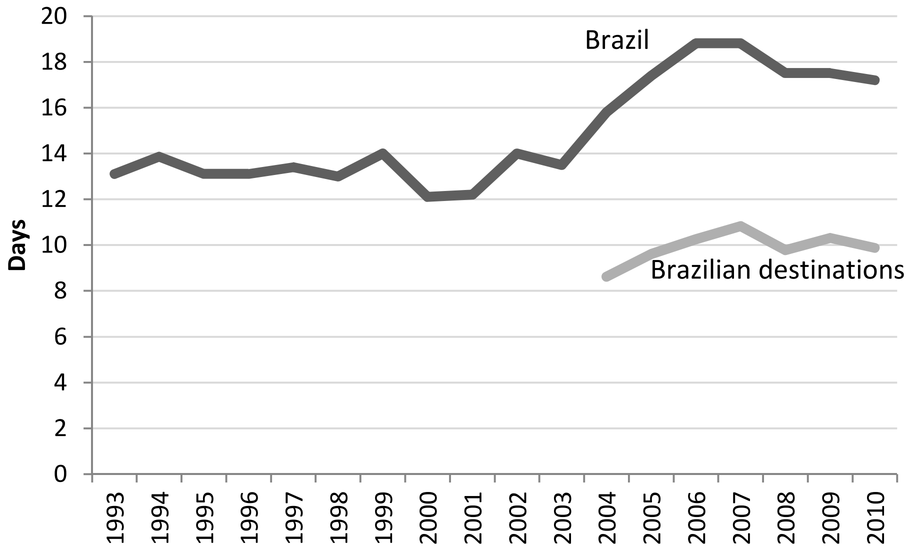
[8] From this point onwards 'length of stay' will be used as synonym of 'duration of the trip', unless stated otherwise.

Third, tourism managers can benefit from research on tourists' length of stay by developing efficient 'on the fly' strategies of pricing or service provision. For instance, discounts or free additional services might be offered during a tourist's stay according to his or her expected additional length of stay. It would not be optimal to offer a 'on the fly' discount for a tourist if his or her probability of going away is very low. On the other hand, using this sort of strategy at the right moment might be very useful for converting existent potential demand for longer stays into actual revenue. Empirical support for these strategies might be developed through duration dependence analysis.

Tourism has faced a worldwide trend of decreasing length of stay as reported by the World Tourism Organization (2006, 2007) and several other authors (for example Alegre & Pou, 2006; Barros, Correia, & Crouch, 2008; Barros & Machado, 2010; Fleischer & Rivlin Byk, 2009; Martínez-Garcia & Raya, 2008). Curiously, this worldwide decreasing trend of tourists' length of stay is not observed in Brazil. According to official statistics of the Brazilian Tourism Ministry, the average length of stay of international tourists in this country was relatively constant from 1993 to 2003. In 1993 the average length of stay was 13.1 days, while in 2003 it was 13.5. From 2003 to 2007 the average length of stay increased almost 40%, reaching a maximum of 18.8 days in 2006 and 2007. Over the following three years the average length of stay fell roughly 10%. In 2010 the average duration of international tourists' stays in Brazil was 17.2 days. A yearly average growth rate of 1.6% was observed from 1993 to 2010.

This increasing trend is also present when tourists' length of stay at Brazilian destinations is assessed. According to the Brazilian International Tourist Survey (BITS), between 2004 to 2010 the average length of stay at Brazilian destinations increased 14.7% from 8.6 days up to 9.9 days. The highest value was achieved in 2007, when the length of stay at destinations reached 10.8 days. These values are represented at Figure 6.

Figure 6: Average length of stay evolution of inbound tourists in Brazil and at Brazilian destinations



Source: Tourism Ministry of Brazil and BITS

The increasing trend of international tourists' length of stay in Brazil and at Brazilian destinations is unusual. One potential explanation for the unexpected reality of international tourism in Brazil is that tourists' behaviour is facing an uncommon trend in this particular case. The confirmation of this hypothesis would pose an interesting question for further research: why tourists' behaviour in Brazil is evolving in the opposite direction of most countries? A second hypothesis is that there have been relevant changes in the composition of the inbound tourism flow in Brazil. This hypothesis is supported by some univariate analyses of time series. For instance, in 1993 roughly 65% of inbound tourists in Brazil were South Americans. In 2010 this share had dropped to 45%. Since South Americans tend to stay shorter in Brazil, the overall average length of stay might have increased due to the decrease in the participation of South Americans in the Brazilian inbound tourism. However, univariate analysis is not enough to support the second hypothesis since tourists' behaviour is affected by multiple variables. Therefore, distinguishing between these two hypotheses requires multivariate analysis.

This study models international tourists' length of stay at different Brazilian destinations according to individuals' characteristics, travels' attributes and visited destinations. Contributions to the understanding of tourists' length of stay are made both in the empirical and methodological domains. There are five main empirical contributions.

(1) Determinants of tourists' length of stay are analysed in the MTTs' context instead of been analysed with respect to a single destination trip. In particular, the difference between the determinants of the length of stay at the destination and in the destination region is an important issue arising from this perspective. This approach was never adopted in the academic literature before. Thus it has potential to provide particular and innovative evidence about tourists' behaviour.

(2) The effects of several variables tested in previous studies about tourists' length of stay are analysed for the Brazilian case with a large dataset (309,413 observations) that include rich information about tourists and the characteristics of their trips.

(3) The effects of some variables, such as income and party size, are analysed in a deeper level of detail than before.

(4) Variations of the average length of stay across a large number of destinations is analysed for the first time in the academic literature. This analysis provides information about which regions and states can be used as benchmarks for tourism management regarding length of stay. Besides, it also provides some information about how length of stay varies according to destinations' characteristics.

(5) The evolution of tourists' length of stay in Brazilian destinations is also analysed in order to identify the actual cause of the unexpected increasing trend. This is implemented by the use of dummy variables built to capture variations of the average length of stay over the years. All these contributions provide further support for the development of an empirical theory of tourists' length of stay since they provide factual evidence on behaviour. Besides, these empirical results may support the development of specific tourism management strategies for Brazilian destinations.

Some contributions are also made on the methodological side. Previous studies were revised and several econometric models were considered. Duration models were finally chosen as the most appropriate alternative. Besides, inbound tourists' length of stay at Brazilian destinations is studied within the MTT paradigm (Santos, Ramos, & Rey-Maquieira, 2011). Lengths of stay at all destinations visited by each tourist constitute the dependent variable of the econometric model. For example, three observations of the dependent variable would be obtained from a tourist that visits three destinations within Brazil in the same trip. Differences between single and multiple destination trips might be relevant since some variables might jointly influence length of stay and the number of destinations. Thus, the analysis of tourists' length of stay in the MTTs context might present particularities as compared to the single destination paradigm.

Modelling this type of data requires special attention since observations from the same tourist are correlated. In order to account for this feature, shared heterogeneity is introduced in the econometric model. To our knowledge, employing shared heterogeneity duration models to explain tourists' length of stay is an innovation in the tourism academic literature. This approach may improve the econometric model applied to observations of multidestination tourism trips, providing better fit and more accurate estimates of covariates' effects.

The rest of the chapter is organized as follows. The next section presents a general review of modelling studies about tourists' length of stay. Since duration models are argued to be the most appropriate statistical technique for modelling tourists' length of stay, these models are presented in detail at Section 4.2. That section also includes a specific literature review of studies applying duration models to tourists' length of stay. Section 4.3 presents a review of empirical findings regarding the determinants of tourists' length of stay. Once these revisions are covered, Section 4.4 presents the empirical modelling study of inbound tourists' lengths of stay at Brazilian destinations and it is followed by the conclusion.

4.1 Previous modelling studies about tourists' length of stay

Length of stay has some particular characteristics that influence its modelling process. First, it is a strictly positive variable. No negative length of stay can be measured, neither should it be estimated. Second, the length of stay is ultimately a continuous variable, despite of its usual measurement in discrete terms, such as days or overnights. Discreteness of duration data is a collection issue rather than a fundamental characteristic of this variable. Although measuring length of stay in a continuous scale is usually unpractical, continuous estimates of the length of stay are perfectly sensible.

In this regard, it is necessary to point out that some studies have focused on days or overnights as variables of interest instead of using these variables as indirect measures of the length of stay. In these cases the variable of interest is discrete by nature. Although most interpretations of the analysis of overnights and length of stay are very similar, their actual meanings are not exactly the same. Overnights might provide more useful information for those sectors whose quantities are commonly measured on a discrete daily basis, such as accommodation. On the other hand, information about length of stay may be more useful for other sectors whose attachment to the daily cycle is not that strong, such as food, sightseeing and entertainment.

It is worthy to stress that all previous studies about tourists' length of stay are restricted to the single destination trip paradigm, while the present study focuses on the MTTs' paradigm. Studies on duration of trips did not differentiate destinations visited, while studies on tourists' length of stay focused a single destination. In the latter case, destinations could be part of a larger trip itinerary, but no information about other destinations included in the same trip was analysed. Therefore, although most previous empirical evidence might help to understand tourists' behaviour within the MTT context, some differences between both paradigms might exist. Explanations for the consumption of longer stays in the single destination paradigm might be distorted if applied to the MTTs context. For instance, tourists may trade off longer stays for a larger number of destinations. Therefore, analysing tourists' length of stay within the MTTs paradigm may provide particular and useful insight.

Twenty four papers modelling tourists' length of stay from microdata have been published in the main tourism academic journals. A general presentation of these studies was conducted at Section 1.3 and a summary of the main characteristics of these studies were presented at Table 1.

Different econometric methods have been applied to model tourists' length of stay depending on the assumed properties of the dependent variable. Used methods include Ordinary Least Squares (OLS), multiple stage least squares (MSLS), Tobit, binary and multinomial logit, ordered logit, latent class truncated Poisson model, bivariate

Poisson lognormal model and several duration models. As this chapter applies duration models, a detailed discussed is presented in Section 4.2, while the econometric particularities of the remaining models are briefly discussed in the following.

The simplest econometric approach was adopted by Paul and Rimmawi (1992), who used an OLS regression. No transformation of the dependent variable was conducted, what means that the model could potentially predict negative lengths of stay.

OLS regressions were also used by Blaine, Mohammad and Var (1993), Mak and Nishimura (1979), and Walsh and Davitt (1983) to model tourists' length of stay. These studies used different specification forms. Blaine et al. (1993) and Mak and Nishimura (1979) used a log-log functional form which is able to predict only positive durations. Walsh and Davitt (1983) tested several specification forms, such as linear, quadratic, log-linear and log-log. The log-linear form was finally chosen because of its better fit.

Silberman (1985) and Uysal, McDonald and O'Leary (1988) studied tourists' length of stay and daily expenditure. Separate estimations were conducted using OLS regressions, while jointly estimates were provided by a MSLS model with two stages. Mak, Moncur and Yonamine (1977) also used a MSLS model with two stages to jointly predict length of stay and daily expenditure. Unfortunately, no information about the use of instrumental variables was provided in any of these three studies, leaving the reader to guess how this issue was handled.

Different functional forms were employed by studies using MSLS. Mak et al. (1977) employed a linear equation form. Silberman (1985) used quadratic, log-linear and log-log specifications, additionally to the linear form. Uysal et al. (1988) used independent variables in the logarithmic form while the dependent variable was considered in its linear form.

Fleischer and Rivlin Byk (2009) broke down tourists' demand into three variables: number of trips, quality of consumed tourism services and length of stay. Since household survey data was used, no objective information on services' quality and no direct measurements of prices were available. In order to indirectly obtain these variables, the authors estimated a secondary model regressing daily expenditure on accommodation type. The predicted expenditure of this model was taken as a quality proxy, while the error term was assumed to indicate quality-adjusted prices. The main model was a MSLS model with three stages and it had a log-log functional form. Income, prices, party size and season of the year were used as explanatory variables. Two instrumental variables were used: educational level and tourist season, both assumed to explain price and to be uncorrelated with the other endogenous variables. For instance, regarding the level of education, the authors argued that a person with higher level of education has better access to information than other, what allows them to find lower prices. Although this proposition might be true, assuming that this is the only role played by education in tourism consumption seems too restrictive. It is likely that level of education play a relevant direct influence on tourists' preferences regarding number of trips, quality of consumed tourism services and length of stay. Therefore, using this as an instrumental variable looks inappropriate. The same excessively restrictive assumption can be made with respect to the variable season, also potentially leading to inadequate model specification.

Fleischer and Pizam (2002) considered the duration of trips taken by individuals. Duration zero was imputed for those who had not travelled at all. In this case, the meaning of zero was qualitatively different from what is implied by the subjacent continuous scale of time. The decisions of taking a trip and how long to travel are related, but different. Therefore, the decision process and the effects of independent variables may be distinct to each dependent variable. This characteristic of the dependent variables was properly considered by the use of a Tobit model.

Alegre and Pou (2006) obtained information about the number of days that tourists stayed in the destination. However, a 'strong bimodal nature' of the dependent variable was observed, what was supposed to be related to the fact that most trips were booked for periods of one or two weeks. This characteristic distinguished the length of stay's distribution from usual continuous distributions. Due to this bimodal nature, the authors transformed the original variable into a dichotomous one, differentiating stays shorter and longer than seven days. The dichotomization of the length of stay led to considerable loss of information. As Menezes, Moniz and Vieira (2008, p. 208) stressed, by adopting this procedure "the ensuing policy implications are less far-reaching in the sense that all lengths of stay shorter than, say, one week are treated alike, be they one-day stays or six-day stays". Finally, the dichotomous variable was modelled though a binary choice model where a logistic distribution of the error term was assumed (i.e.; a binary logit model).

Nicolau and Más (2009) analysed the choices of whether to take a trip and how long to stay away from home. These two choices were considered to be part of a single decision process. The duration of the trip was considered as a three level multinomial variable. Hence, together with the no-trip alternative, the dependent variable had four different values. In order to account for this particular design, a random parameters multinomial logit model was used. The model allowed for different correlation patterns among all four alternatives and the two-step structure of tourists' choices was confirmed by empirical findings.

Yang et al. (2011) also applied a discrete model to the length of stay at a tourist destination. In this case, the model used was an ordered logit. The discrete nature of the model was allegedly imposed by the type of measurement adopted by the survey from where data was obtained. Length of stay was coded in four categories: 1 day, 2 days, 3 days and 4 days or longer. In fact, the ordered logit model is able to handle this type data. However, it might be overparameterized since this model does not recognize the underlying linearity between these categories of duration. On the other hand, a censored linear regression or a censored duration model would have been able to consider this linearity while still accounting for the censored nature of the fourth category.

Alegre, Mateo and Pou (2011) used the same data source of the previous study carried out by Alegre and Pou (2006), and, therefore, they faced the same high frequency of stays around seven and fourteen days. Recognizing the arbitrariness of establishing categories from the original variable as done by the previous study, the authors applied a latent class truncated Poisson model. They assumed the hypothesis that tourists are divided in two segments characterized by a preference for either short or long stays. The model used is able to assign individuals endogenously to a segment adopting a binary logit model in its first stage. Subsequently, the model predicts the length of stay according to segment preferences and individuals' profiles using a count data model with a Poisson distribution.

A count data model was also used by Hellström's (2006) study of tourists' length of stay. Households' joint choice of the number of leisure trips and the number of total nights spent on these trips were studied. The excess amount of zeros was handled by the introduction of a bivariate Poisson lognormal model. Truncation was also considered since the number of nights possibly observed is conditional on the number of trips taken. For example, as far as length of stay was measured by the number of nights, observing a total of two trips and one night would be impossible. Finally, the model took into account the fact that leisure time is usually concentrated, by institutional regulations, on weekends. The excess of particular count observations arising from this fact was handled by introducing inflation at specific points of the distribution.

Duration models have also been frequently applied to tourists' length of stay. These models are usually considered the most appropriate econometric instruments available for the study of tourists' length of stay. A detailed presentation and discussion of this family of models is conducted in the next section.

4.2 Duration models and tourists' length of stay

The term 'duration model' refers to a family of statistical models used to explain and predict the time length of spells. These models have also been termed 'survival models' due to their frequent application in biomedical sciences. Duration data is otherwise called 'transition data' since spells are periods terminating when there is transition from one state to another. When tourists' length of stay is regarded, the term 'duration model' seems to be more appropriate than 'survival model'. For consistency, 'duration data' will also be used instead of 'transition data'. The nature of these models is presented in the following. A more detailed discussion of this family of models is provided because it is finally going to be selected for the empirical analysis of tourists' length of stay in Brazil. Reasons of this selection are discussed in Section 4.2.2.

Duration models can be described as follows. Let $T \geq 0$ be the variable denoting the duration of the spell and t denote a particular value of T. For instance, T may represent the length of stay at a destination or the duration of the trip. F(t) is the cumulative distribution function of T, which refers to the probability that the duration is less than or equal to t. For example, F(t) might indicate the probability that a given tourist stays at the destination for t days or less. If f(t) denotes the density function of T, F(t) can also be defined as the integral of f(t) from zero to t, such that:

$$F(t) = \int_0^t f(t) = P(T \leq t)$$

The survival function of T indicates the probability that the spell is larger than t. When applied to tourists' length of stay, the survival function refers to the probability that the tourist stays at the destination longer than t. The survival function is given by:

$$S(t) = 1 - F(t) = P(T > t)$$

The survival function refers to the total duration of a spell. An alternative way of characterizing duration is by focusing on the probability that the spell ends in the next period given that it has already lasted until actual time. This instantaneous rate of ending the spell is called hazard rate (λ_t). When tourists' length of stay is regarded, the hazard rate indicates the marginal probability that a tourist leaves the destination in the next period. The hazard function can be defined as:

$$\lambda(t) = \lim_{\Delta \to 0} \frac{P(t \leq T \leq t + \Delta | T \geq t)}{\Delta} = \frac{f(t)}{S(t)}$$

Once the hazard function is known, obtaining the density and the cumulative distributions of T is straightforward, as presented in the following.

$$f(t) = \lambda(t) e^{-\int_0^t \lambda(s) ds}$$

$$F(t) = 1 - e^{-\int_0^t \lambda(s) ds}$$

The duration of a spell may be conditional on a set of explanatory variables or covariates **x**. For example, the length of stay at a tourist destination may depend on tourists' profiles, such as available time, income and tastes, and on destination's characteristics, such as attractions, services and prices. There are two ways of introducing **x** in the duration model: the proportional hazards approach and the accelerated failure time models. In the former case, the set of covariates are included in the hazard function such that:

$$\lambda(\mathrm{t}, \mathbf{x}) = \mathrm{k}(\mathbf{x})\lambda_0(\mathrm{t})$$

where $k(\cdot)$ is a nonnegative function of $\mathbf{x}$ and $\lambda_0(t)$ is called the baseline hazard. The function $k(\cdot)$ is usually parameterized as $k(\mathbf{x})=e^{\boldsymbol{\beta}\mathbf{x}}$, where $\boldsymbol{\beta}$ are parameters. In this case, β_k measures the semielasticity of the hazard with respect to x_k, that is:

$$\beta_k = \frac{\partial \ln[\lambda(\mathrm{t}, \mathbf{x})]}{\partial \mathbf{x_k}}$$

The shape of the hazard function is typically unknown *a priori*. Theoretical propositions regarding the distribution of T over time are usually scarce. Leaving $\lambda_0(t)$ unspecified yields the semi-parametric Cox Proportional Hazards model (Cox, 1972) which is estimated through a partial likelihood method. This is an advantageous approach when the distribution of T is unknown, since different distributional assumptions might lead to significantly different estimates at certain cases. On the other hand, Cox's approach may be seen as disadvantageous if one is interested in the shape of the baseline hazard function for its own sake. Implications of different hazard functions are discussed at Section 4.2.3. Besides, the underlying hypothesis of the Cox Proportional Hazards model can be tested through Schoenfeld's residuals (Grambsch, Therneau, 1994; Schoenfeld, 1982).

Assuming some particular functional form for the hazard rate yields parametric duration models. If the hazard rate is assumed to be constant, the generating process of T is said to be memoryless (Wooldridge, 2002), that is, the probability of ending the spell is independent from the amount of time elapsed since its beginning. In this case, the survival function follows an exponential form such as

$$\mathrm{S}(\mathrm{t}) = e^{-\lambda t}$$

However, a constant hazard rate is no more than a special case within a large variety of alternatives. In practice, the hazard rate is usually dependent on the amount of time elapsed since the beginning of the spell. Particularly, in the case of tourism it seems logical to consider that the probability of leaving the destination the next day would not be the same for a tourist that has just arrived than for a tourist that arrived four days ago.

Processes with non-constant hazard functions are said to be duration dependent. Positive duration dependence happens at time t when $\frac{d\lambda(t)}{dt} > 0$, that is, when the probability of ending the spell increases over time. In this case, the probability of ending the stay at a destination on the following day would be higher for those who have already stayed longer. The opposite situation is referred as negative duration dependence.

Besides the exponential distribution of T given by a constant hazard function, several other distributions have been proposed in the literature. Some of the most usual distributions used are Weibull, Gompertz, lognormal, loglogistic and Gamma. However, several other complex distributions might be used. Among usual alternatives, proportional hazards models allow for the adoption of the exponential, Weibull and Gompertz. The hazard functions of the Weibull duration model is given by

$$\lambda(t, \mathbf{x}) = \alpha t^{\alpha-1} e^{\boldsymbol{\beta}\mathbf{x}}$$

where α>0 is a parameter. The hazard rate increases monotonically with time when α>1. If α<1 the hazard rate decreases monotonically. If α=1 the model shrinks to an exponential model since the hazard function becomes time-invariant.

The hazard function of the Gompertz model is

$$\lambda(t, \mathbf{x}) = e^{\boldsymbol{\beta}\mathbf{x}+\gamma t}$$

where γ is a shape parameter. The hazard rate is monotonically increasing if γ>0 and monotonically decreasing if γ<0.

The second way of introducing covariates in duration models is known as accelerated failure time models. In this case, it is assumed that the set of covariates have a linear relationship with the logarithm of the latent survival time T such that

$$\ln(T) = \boldsymbol{\beta}\mathbf{x} + \sigma u$$

where σ is a scale parameter and u is an error term with density function f(u). Accordingly, accelerated failure time models are also properly classified as parametric. The parameters of these models reveal the proportionate change in the survival time given by a unit change in the covariate, that is

$$\beta_k = \frac{\partial \ln[T]}{\partial \mathbf{x_k}}$$

Accelerated failure time models are called this way because of the time scaling role played by the term **βx**. When **βx** is positive the spell tends to last longer, while when **βx** is negative the time scale is contracted.

Distributional assumptions about u lead to different distributions of T. If u is assumed to follow an Extreme Value distribution with two parameters, T follows a Weibull distribution. In this case the survival function is given by

$$S(t, \mathbf{x}) = e^{-e^{-\boldsymbol{\beta}\mathbf{x}} t^{\alpha}}$$

If α=1 this survival function equals the exponential model. If f(u) is a Normal distribution, T follows a lognormal distribution such that

$$S(t, \mathbf{x}) = 1 - \Phi\left[\frac{\ln(t) - \boldsymbol{\beta}\mathbf{x}}{\sigma}\right]$$

In this case the accelerated failure time model equals a usual loglinear model estimated by OLS (Wooldridge, 2002). Therefore, the loglinear OLS model may be interpreted as a particular type of duration model.

If u is assumed to follow a logistic distribution, T is distributed loglogistic and the survival function is given by

$$S(t, \mathbf{x}) = \left[1 + \left(e^{-\boldsymbol{\beta}\mathbf{x}} t\right)^{\frac{1}{\gamma}}\right]^{-1}$$

where γ is a scale parameter.

Finally, if u follows a log-Gamma distribution with three parameters, T follows a Generalized Gamma distribution. The survival function in this case is

$$S(t, \mathbf{x}) = \begin{cases} 1 - I(\gamma, u), & \text{if } k > 0 \\ 1 - \Phi(z), & \text{if } k = 0 \\ I(\gamma, u), & \text{if } k < 0 \end{cases}$$

where Φ is the standard Normal cumulative distribution, k is a parameter, $\gamma=|k|^{-2}$, $u=\gamma e^{|k|z}$, $I(\cdot,\cdot)$ is the incomplete gamma function, z=sign(k)ln(t)μ/σ, and μ=**βx**. This model incorporates some of the previous models as special cases. The generalized gamma model shrinks to a Weibull if k=1 and to an exponential model if additionally σ=1. When k=0 the model becomes a lognormal. Thus, this model is frequently used to select the distribution of T.

The hazard rate given by the Weibull or the Gompertz distribution can be monotonically increasing, monotonically decreasing or constant according to the estimated parameters. Note that the exponential distribution is a special case of these distributions. On the other hand, the lognormal and loglogistic hazard rate distributions are positively skewed. In this case, the hazard rate first raises and then falls monotonically displaying a long right tail. Finally, the gamma distribution can be either monotonically decreasing or a positively skewed distribution depending on its parameters.

Duration dependence and the shape of the hazard function of tourists' length of stay have to be carefully regarded. The usual meaning of duration dependence is that the spell's past influences the future probability of ending the spell. However, tourists' length of stay is frequently defined before the beginning of the trip. Therefore, when the length of stay is defined previously, the decision of leaving the destination at a given moment is independent from the stay time elapsed up to this point. In this case, the past does not influence the future because the ulterior is defined beforehand.

It seems likely that the density function of tourists' length of stay follows a positively skewed distribution due to the opposition of two effects. First, there are some relevant initial costs of visiting a destination (e.g., McKercher, Chan, Lam, 2008; McKercher, Lew, 2003; Nicolau, 2008). Transport to the destination is an important example of this sort of cost for the tourist. Accommodating and other practical arrangements, such as getting information about the destination, also require some initial time and money expenses. According to Nicolau (2010, p. 261)

> a tourist will be prepared to make a long journey if he or she stays at the destination for at least the minimum number of days, which will compensate for the effort made on the journey and allows individuals to spread the fixed costs associated with the long journey over a sufficiently long period.

Therefore, the initial costs of the visitation implies on a low propensity to short stays. Of course, the value of 'short' here depends both on tourists' preferences and destinations' characteristics.

On the other hand, tourists' stays usually display decreasing marginal utility. As the length of stay becomes larger, opportunities for new experiences decrease. At the same time, there is an increase in the additional costs of being away from home. The decreasing marginal utility of stays implies on a low propensity to stay for too long.

Together, initial costs of the visitation and decreasing marginal utility of the stay lead to a positively skewed density distribution of tourists' length of stay. This sort of density function can be derived both from monotonically increasing and positively skewed hazard functions. However, it cannot be derived from constant or monotonically decreasing hazard functions. These arguments support the criticism made by Thrane (2012) about the usage of exponential duration models to represent tourists' length of stay. The hypothesis of a positively skewed density distribution of tourists' length of stay is tested in the empirical study presented at Section 4.4.

Different duration models are tested in the empirical study presented at Section 4.4 in order to choose the best fitting distribution. A more detailed explanation of duration models is provided by Hougaard (1999), Kiefer (1988) and Lancaster (1990).

4.2.1 Heterogeneity in duration models

Unobserved heterogeneity is a major concern in duration models arising from the omission of relevant explanatory variables or measurement errors (Lancaster, 1979; Vaupel, Manton, & Stallard, 1979). According to Heckman and Singer (1984), heterogeneity in duration data is never rejected when tested in microeconomic studies.

Duration models allowing for unobserved heterogeneity are also called 'frailty models'. Besides their use in biomedical sciences, these models have been used in Economics and other social sciences (e.g., Boehmke, 2006; Jones, 2011; Kau, Keenan, & Li, 2011). Unobserved heterogeneity may be introduced in a usual duration model through a random variable in the hazard function such that

$$\lambda(t, \mathbf{x}|v) = vk(\mathbf{x})\lambda_0(t)$$

where v is an unobservable individual effect. The random variable v is assumed to have the following properties:

- $v > 0$
- $E(v) = 1$
- finite variance $\sigma^2>0$
- distributed independently from t and x

The random variable v scales the hazard rate. *Ceteris paribus*, if v>1 the hazard rate is higher than average and the spell is shorter. Conversely, if v<1 the hazard rate is lower than average and the spell lasts longer.

The estimation of heterogeneity duration models requires the assumption of a specific distribution for v. Let g(v) be the probability density function of v, the survival function is obtained by integrating S(t,**x**) on v as given by

$$S(t, \mathbf{x}) = \int_0^{\infty} [S(t, \mathbf{x})]^v g(v)dv$$

The distributions most commonly adopted to describe v are the Gamma and the Inverse Gaussian (Hougaard, 1984). These distributions are usually employed mainly due to computational facility.

Unobserved heterogeneity may vary across each observation or across groups of observations. The latter case yields 'shared heterogeneity' duration models, also known as 'shared frailty' models (Collier, 2005; Hougaard, 1986; Jones, 2011; Whitmore, 1991). This type of model might be understood as the duration model counterpart of the random effects model for OLS. Shared heterogeneity may happen when there is potentially more than one observation from the same individual or group. This is the case when lengths of stay on MTTs are regarded. If observations refer to stays of the same tourist at different destinations, the error term is no longer uncorrelated among observations, what may be interpreted as shared heterogeneity. In this case, the length of stay at one destination is correlated with other observations coming from the same tourist. Therefore, modelling tourists' length of stay in the MTT duration models context requires shared heterogeneity to be taken into account.

According to Box-Steffensmeier and Boef (2006), shared heterogeneity on repeated events may arise from two different causes. First, it may be caused by usual heterogeneity across individuals since people are different, and only a small share of their characteristics is observed. Second, shared heterogeneity may arise from event dependence, that is, the influence of an event over the following ones. For example, on a multidestination trip, the visitation of a given destination might decrease (or increase) the expected length of stay at a similar destination visited later.

Ignoring unobserved heterogeneity has two main consequences (Lancaster, 1979). First, the hazard function in the omitted heterogeneity model increases slower or falls faster than in the correctly specified model. Thus, failing to allow for heterogeneity prevent unbiased estimation of duration dependence. Second, in the model with no heterogeneity the proportionate variation of the hazard rate caused by changes in $\mathbf{x}$ is not constant. In fact, in this model the effect of $\mathbf{x}$ tends to zero for large t. In this case, the effect of $\mathbf{x}$ on the hazard rate is correctly estimated only for $t=0$. The absolute value of this effect is underestimated for any $t>0$. The cause of both problems is known as the 'weeding out' effect, that is, the result of the fact that the unobserved characteristics of the surviving population are not constant over time.

While non-shared heterogeneity is only compatible with parametric models, shared heterogeneity is also compatible with the Cox Proportional Hazards model. However, the semi-parametric case requires the estimation of an independent parameter for each group. When the number of groups is large, the computational cost may become excessively large. This will usually be the case when a dataset of multiple observations from same individuals is used. This is also the case of the empirical study presented at Section 4.4. Thus, as will be explained latter, some tests were conducted to choose between the Cox Proportional Hazards model without heterogeneity and parametric models with shared-heterogeneity.

4.2.2 Advantages and disadvantages of duration models

Provided that tourists' length of stay are a continuous variable, the adequacy of duration models to explain and predict it could be compared to OLS and derivations of this technique. The best argument in favour of duration models as a statistical tool to explain and predict tourists' length of stay is their flexibility in terms of distributions allowed for the dependent variable. OLS models require that the error term is normally distributed, whereas duration models offer a variety of distributions. Besides the lognormal distribution, the length of stay may be assumed to follow an exponential, Weibull, loglogistic, Gamma or many other distributions. This variety of distributions may provide better parameter estimates, besides allowing for an appropriate analysis of the duration dependence and the baseline hazard function.

Several weaker arguments in favour of duration models have been proposed in the academic literature. A very common argument is that duration models take full account of data positiveness. In fact, the construction of duration models departs from this premise. This has been sometimes argued as one advantage of duration models as compared to usual models such as OLS. However, this is a weak argument since OLS models can be easily adapted to satisfy the non-negativity requirement by adopting strictly positive functions at the right hand side. A usual example of this is the log-linear equation such as $y = e^{\boldsymbol{\beta x}}$.

Duration models are quite convenient to deal with censored data, such as when the exact moment of the beginning or the end of the spell is not observed. This feature has favoured the application of these models in areas such as biomedical sciences where censored duration data is frequently used. However, censored data about tourists' length of stay is relatively unusual. Most tourist surveys are conducted at the end of the stay or after tourists have returned home. Therefore, the actual length of stay is usually known without censoring. Even when surveys are conducted before the end of the trip, censoring usually does not happen since planned length of stay is taken as the dependent variable. Besides, duration models are not the only alternative for dealing with censored data. Several linear models adjusted for censored data have been developed, such as Tobit, censored normal regression and interval regression.

A major advantage of duration models for some applications is the appropriate consideration of time-varying covariates. If values of the covariates change along the duration of the spell, usual statistical models are not able to provide correct estimates. Once again, these cases are frequent in areas such as engineering and biomedical sciences, though not in tourism. Therefore, despite of the clear superiority of duration models when time-varying covariates are studied, this quality is usually not relevant for the study of tourists' length of stay.

Studies on tourists' length of stay have usually advocated in favour of duration models by alleging their qualities regarding strictly positivity, appropriate consideration of censoring and time-varying covariates. Thrane (2012) strongly criticized these arguments since the first one is easily overcome by OLS and the two latter qualities have not been put into actual use in the studies about tourists' length of stay. Thrane proposed that a log-linear OLS model could be a satisfactory alternative to duration models. Unfortunately, Thrane did not recognize that a log-linear OLS is equivalent to the lognormal duration model (Wooldridge, 2002). That omission led the author to conduct an inappropriate comparison of two models with the same structure. Estimates obtained by Thrane differed only because of the introduction of heterogeneity in the duration model, while no heterogeneity was allowed in the OLS model. If no heterogeneity was allowed in the lognormal duration model the author would have found identical estimates for both models. It is worthy to note that heterogeneity could also have been included in the log-linear model through random effects.

4.2.3 Tourists' length of stay with duration models

Duration models have recently been applied to the study of tourists' length of stay. The first published study was conducted by Gokovali, Bahar and Kozak (2007). These authors analysed tourists' behaviour in a Turkish sun and sea destination. This study was followed by eight studies from different authors and focusing on distinct regions. Five of these studies were related to Portugal as tourism origin or destination.

Most studies on tourists' length of stay using duration models focused their attention both on methodological issues and on the effects of explanatory variables. On the methodological side, the debate included the fitting of different statistical distributions, the consideration of heterogeneity, sample selection issues and endogeneity. The methodological presentation provided by some studies was excessively short, offering incomplete descriptions of the models used (for example Machado, 2010; Martínez-Garcia & Raya, 2008; Peypoch, Randriamboarison, Rasoamananjara, & Solonandrasana, 2012). Occasionally the letters used to identify parameters at the tables did not correspond to those used at the formulas, also inhibiting readers' own analysis (Barros, Butler, & Correia, 2010; Barros et al., 2008; Barros & Machado, 2010).

Different statistical distributions were analysed by the set of previous studies. The semi-parametric Cox Proportional Hazards model was employed by five studies (Barros et al., 2010; Barros et al., 2008; Gokovali et al., 2007; Martínez-Garcia & Raya, 2008; Menezes et al., 2008; Raya & Martínez-Garcia, 2011). In particular, Menezes et al. (2008) used Cox's model alone, following an agnostic approach in order to avoid potential biases arising from a mistaken selection of a parametric functional form. However, no test of the proportional hazards hypothesis assumed by this model was reported.

Parametric models were more frequently used by previous studies than the semi-parametric approach. Starting from the most simple parametric model, the exponential distribution was used by three studies (Gokovali et al., 2007; Martínez-Garcia & Raya, 2008; Raya and Martínez-Garcia, 2011). Nevertheless, Thrane (2012) criticized this practice by arguing that the assumptions arising from the exponential distribution are unrealistic when tourists' length of stay is regarded. The most frequently used parametric distribution was the Weibull. Studies focusing on other methodological issues, rather than the best fitting distribution, adopted the Weibull by default (Barros & Machado, 2010; Machado, 2010).

The Gompertz was the only distribution exclusive for the proportional hazards approach used by previous studies (Gokovali et al., 2007; Martínez-Garcia & Raya, 2008; Raya & Martínez-Garcia, 2011). It is interesting to note that this distribution was used only by studies that tested a relatively large number of distributions, what indicates a resistance of more selective scholars to use it. Distributions exclusive for accelerated failure time models used by previous studies include lognormal (Martínez-Garcia & Raya, 2008; Raya & Martínez-Garcia, 2011; Thrane, 2012), loglogistic (Barros et al., 2008; Martínez-Garcia & Raya, 2008; Raya & Martínez-Garcia, 2011; Thrane, 2012) and Gamma (Martínez-Garcia & Raya, 2008; Raya & Martínez-Garcia, 2011). Finally, Peypoch et al. (2012) used a proportional hazards distribution based on fractional polynomials in order to allow for complex nonlinear effects of regressors on the hazard function.

Six studies compared the fitting of different models. Gokovali et al. (2007) preferred the Weibull distributed model among other parametric models, although Cox's model was not rejected. Martínez-Garcia and Raya (2008) and Raya and Martínez-Garcia (2011) tested seven types of parametric models and they found evidence in favour of the loglogistic distribution. Two studies allowed for heterogeneity on models with a Weibull distributed length of stay (Barros et al., 2010; Barros et al., 2008). Both studies preferred the Weibull with heterogeneity as an alternative to other models. All duration models tested by Thrane (2012) allowed for heterogeneity and the Weibull distribution was preferred to the lognormal and the loglogistic models. Though providing no empirical justification for that, Peypoch et al. (2012) preferred the fractional polynomial distribution against the Weibull distribution. Finally, it is necessary to stress that in most studies the superiority of one or another model was minor.

No previous study examined duration dependence in detail and any possible policy implication of this aspect was overlooked. Martínez-Garcia & Raya (2008) did not even report the structural parameters of their estimated models. Still, duration dependence might have relevant implications for tourism management. For instance, hotels usually have different prices according to the total length of stay. If efficient discounts are those formulated according to the probability of leaving the hotel, non-monotonic hazard functions would require non-monotonic discount functions with respect to the total length of stay. In that case, increasing (decreasing) discounts should be offered when the length of stay presents positive (negative) duration dependence. Duration dependence might also be used to develop 'on the fly' marketing strategies oriented to influence tourists' decisions during their stays. This type of analysis is conducted in the empirical study presented at Section 4.4.

Unobserved heterogeneity of tourists' length of stay was analysed by five studies using duration models (Barros et al., 2010; Barros et al., 2008; Martínez-Garcia & Raya, 2008; Raya & Martínez-Garcia, 2011; Thrane, 2012). All five studies allowed the unobserved heterogeneity term to vary across every observation since data registries referred to different tourists. Three studies (Barros et al., 2010; Barros et al., 2008; Thrane, 2012) found evidences in favour of the heterogeneity model. Barros et al. (2008) textually declared their preference for the heterogeneity model based on the comparison of the log likelihood statistic across different estimated models. However, the log likelihood statistics reported in that study indicates a smaller value (larger in absolute value) for the heterogeneity model, pointing out a worse result. Exactly the same model selection problem happened in Barros and Machado (2010) and Machado (2010). The analysis of the findings of those studies in this chapter assumes that the original textual statements done by those scholars are correct, although they contradict the log likelihood statistics presented. However, results from those papers should be considered carefully.

Sample selection biases were considered by two studies. Barros et al. (2010) and Barros and Machado (2010) assumed the hypothesis that tourists tend to elect to answer a questionnaire if they are satisfied with the destination, while dissatisfied tourists tend to decline to answer it. They further assumed that respondents on a repeat visit to the destination tend to be enthusiastic towards this destination, thus tending to have a higher response rate in the survey. In both studies a sample selection model as proposed by Boehmke, Morey and Shannon (2006) was compared to a model without sample selection. Both models provided similar parameter estimates except with regard to the variable considered to be sample-selected.

The sample-selection approach in Barros et al. (2010) and Barros and Machado (2010) pose some questions with respect to the specific treatment of the sample selection variable 'repeat'. The authors were not clear about which independent variables were used to explain the selection equation and which were used in the outcome equation (Thrane, 2012). This leaves a question about how the models were actually estimated and what is the meaning of their results. This point is particularly problematic since there is no obvious categorization of available variables, and defining the specific model structure would be delicate. Therefore, the sample-selection approach as adopted by Barros et al. (2010) and Barros and Machado (2010) might be considered not absolutely revealed and explained.

Another methodological issue was considered by Machado (2010), who analysed the relationship between image of a destination and length of stay through duration models by considering that both variables were endogenous and inextricably linked. This perspective imposes a very particular situation for the econometric modelling. A seemingly unrelated discrete choice duration (SUDCD) model based in Boehmke (2006) was applied to this problem. The main advantage of seemingly unrelated models is that they take advantage from the relationship between both endogenous variables in order to estimate parameters more efficiently. However, when two endogenous variables are related in a way that seemingly unrelated models become useful, separate models do not provide unbiased estimates. Therefore, Machado's approach may be slightly better when two endogenous variables are included in the model, but even in these cases this approach is not essential.

Finally, Thrane (2012) criticized the application of duration models to tourists' length of stay with respect to the very nature of this variable. The author argues that tourists' length of stay is usually decided in advance and that this characteristic is inappropriate for the use of duration models. This criticism can be contested in two different ways. First, tourists' stays might not last the exact planned time. Unplanned changes during the trip may either lead to shorter or longer stays. Particularly, in the multidestination context, changes in the length of stay at one particular destination seem to be more likely than changes at the total duration of the trip. For example, it seems more likely that a tourist reschedules his stays at different Brazilian destinations in the same trip than that he or she changes the total planned length of stay at Brazil.

Second, Thrane's criticism can be contested with regard to the adequacy of duration models to explain previously defined lengths of stay. In fact, most general applications of duration models regard spells with no advanced defined duration, such as duration of life or unemployment. However, there is nothing incompatible between duration models without time-varying covariates and spells with previously defined durations. The family of duration models simply states that the dependent variable should be the length of a spell. The idea that a duration variable should not have its own value defined in advance might come from usual applications and interpretations of duration models, but certainly it does not come from the statistical construction of the model. The only situation where advanced planned duration would constitute a problem is the one where time-varying covariates are used. However, this is usually not the case in studies about tourists' length of stay, as it was previously discussed.

4.3 Determinants of tourists' length of stay

Tourism demand have been explained and predicted by a large set of variables. At the aggregate level, the most usual explanatory variables include tourists' income, relative prices, substitute prices, travel costs and exchange rates (Li, Song, & Witt, 2005; Lim, 1999; Song & Li, 2008). At the individual level, the usual explanatory variables include age, gender, household size and composition, marital status, education level, occupation and place of residence (Wang & Davidson, 2010). All these variables probably also influence tourists' length of stay. However, the effect of each relevant variable may differ according to the specific tourism demand variable analysed. Thus, tourists' length of stay might present a particular explanatory structure that should be analysed separately from other measures of tourism demand.

Empirical findings of previous studies about the effects of different variables over tourists' length of stay are discussed in the following. Variables are divided into four groups: individuals' characteristics, travel characteristics, destination characteristics and assessments, and price and expenditure. It must be noted that the effect of the explanatory variables discussed is influenced by the characteristics of each study, such as source market, destination, type of tourists considered and statistical methodology.

Besides, two criticisms regarding previous uses of explanatory variables must be discussed. First, there is an inconsistency in Barros et al. (2008) and Barros et al. (2010) regarding interpretations of estimated parameters. Presented formulas follow the proportional hazards metrics. However, the interpretation textually proposed is consistent with the accelerated failure time metrics, contradicting the first information.

The authors first textually state that positive parameters indicate positive effects of explanatory variables over duration, what is consistent with the accelerated failure time metrics. However, at some points the metrics confusion is present also in the text, such as when authors use both opposing metrics in the same sentence: "the variable budget has a positive effect on the hazard, which means that tourists with relatively high budgets tend to stay longer" (Barros et al., 2008, p. 337). Another example of the same confusion is "expenditure has a negative impact on hazards, which means that high-spending tourists stay shorter lengths of time" (Barros et al, 2010, p. 702). Results of those studies are discussed in the following by assuming that the coefficients are presented in the accelerated failure time metrics, what enables us to agree with the conclusions discussed by the authors. However, those results should be considered carefully.

Second, Barros et al.'s (2010) use of some explanatory variables raises some relevant issues. Travel motivation, for instance, was a nominal variable with eight categories, such as holiday, business and visiting friends and relatives. Each category received an index number (e.g.; holiday=1; business=2). Finally, these numbers seem to have been mistakenly considered to constitute a continuous scale. Therefore, a multinomial variable was included in the model as a continuous explanatory variable. The same problem occurred with other variables, such as type of accommodation, mean of transport, reservation of tourism services and source of information. Findings regarding these variables were omitted in the following analysis.

4.3.1 Individuals' characteristics

Gender was not found to be a significant explanatory variable of tourists' length of stay by most studies (Barros et al., 2010; Fleischer & Pizam, 2002; Machado, 2010; Martínez-Garcia & Raya, 2008; Menezes et al., 2008; Raya & Martínez-Garcia, 2011). The only exceptions were the studies of Barros and Machado (2010) and Peypoch et al. (2012) which found a significantly higher length of stay for male tourists. This finding in Barros and Machado (2010) was obtained only in the sample-selected model, while their non-sample-selected model provided non-significant parameter estimates for gender.

Empirical estimates indicate that age is usually a significant covariate of tourists' length of stay. Only three studies found non-significant effects (Gokovali et al., 2007; Menezes et al., 2008; Raya & Martínez-Garcia, 2011). Besides, by analysing the effect of age as represented by four categories, Alegre et al. (2011) found that this multinomial explanatory variable was significant, but no clear pattern among categories was found. When the effect of age is considered to be monotonic, older tourists usually tend to stay longer. One study found evidence in favour of a negative effect of age on the length of stay (Barros et al., 2008), while nine studies found evidences in favour of a positive influence (Alegre & Pou, 2006; Barros et al., 2010; Barros & Machado, 2010; Hellström, 2006; Machado, 2010; Mak et al., 1977; Martínez-Garcia & Raya, 2008; Nicolau & Más, 2009; Peypoch et al., 2012). When considered to be non-monotonic, the effect of this variable on the length of stay is given by a concave function according to Fleischer and Pizam (2002) and a convex function according to Yang et al. (2011). Besides, estimated vertexes vary considerably. In Fleischer and Pizam (2002), who studied only individuals above 54 years old, the maximum expected length of stay corresponded to age 65. In Yang et al. (2011) the minimum point took place between 34 and 38 years old according to different estimated models.

Empirical evidences on the effect of education level over tourists' length of stay do not provide a clear picture. Three studies found that higher education leads to shorter stays (Gokovali et al., 2007; Martínez-Garcia & Raya, 2008; Menezes et al., 2008), while other three found evidences of the opposite (Barros et al., 2010; Barros & Machado, 2010; Peypoch et al., 2012). However, it is relevant to note that Barros and Machado (2010) found a positive effect of education only at their sample-selected model, while the non-sample-selected instrument provided no significant parameter for this variable. Finally, three studies found non-significant parameters for education (Fleischer & Pizam, 2002; Machado, 2010; Raya & Martínez-Garcia, 2011).

The effect of labour status on tourists' length of stay seems to be related to time availability since pensioners and students tend to choose longer stays (Alegre & Pou, 2006; Alegre et al., 2011; Blaine et al., 1993; Martínez-Garcia & Raya, 2008; Raya & Martínez-Garcia, 2011). The same applies for unemployed individuals (Martínez-Garcia & Raya, 2008). On the other hand, high level professionals (Alegre & Pou, 2006; Alegre et al., 2011) and self-employed individuals (Martínez-Garcia & Raya, 2008) tend to stay shorter. Non-significant estimates or results with no clear pattern regarding labour status were found by Gokovali et al. (2007) and Menezes et al. (2008).

The influence of marital status was tested by seven studies, but six of them found no significant results (Barros et al., 2008; Fleischer & Pizam, 2002; Gokovali et al., 2007; Martínez-Garcia & Raya, 2008; Menezes et al., 2008; Raya & Martínez-Garcia, 2011). Mak et al. (1977) found that married tourists in Hawaii tend to stay shorter.

Despite of its evident relevance, only two studies analysed the effect of tourists' time availability over their length of stay. This lack might be related to the difficulty of defining and measuring individuals' available time for travelling. Hellström's (2006) estimated that the relationship between this variable and tourists' length of stay is non-significant, while Nicolau and Más (2009) found a positive relationship.

Length of stay was most frequently found to be a normal good, that is, higher income leads to longer stays (Barros et al., 2008; Fleischer & Pizam, 2002; Fleischer & Rivlin Byk, 2009; Gokovali et al., 2007; Mak et al., 1977; Peypoch et al., 2012; Silberman, 1985; Walsh & Davitt, 1983). Nevertheless, Hellström (2006) and Barros et al. (2010) found no evidence of a significant relationship between these two variables. Blaine et al. (1993) and Mak and Nishimura (1979) found evidence of a negative relationship between income and length of stay, indicating that this is an inferior good. Finally, Nicolau and Más (2009) found that income display a non-linear effect over tourists' length of stay. According to these authors, the effect of income is positive for the most constrained individuals, while for the less constrained ones its effect is negative.

It is necessary to stress that the theoretical relationship between income and tourists' length of stay is more complicated than for a usual good since the consumption of tourism requires the joint expenditure of monetary and time resources. For instance, although higher income individuals might face a smaller opportunity cost for the monetary resources required for consuming tourism, they usually face a higher opportunity cost for the time necessary for this activity.

Fleischer and Rivlin Byk (2009) estimated that income elasticities regarding length of stay was 0.104, while elasticities regarding the number of trips and tourism services quality were 0.154 and 0.044, respectively. Therefore, according to these results, when income increases, individuals tend to intensify their tourism consumption first by travelling more frequently and second by staying longer, while quality is a less relevant variable to adjust for the new income level.

The effect of tourists' nationality was analysed by ten different studies (Alegre et al., 2011; Alegre & Pou, 2006; Barros et al., 2010; Barros & Machado, 2010; Gokovali et al., 2007; Machado, 2010; Martínez-Garcia & Raya, 2008; Menezes et al., 2008; Peypoch et al., 2012; Raya & Martínez-Garcia, 2011). Estimated parameters by most studies were significant, indicating that expected tourists' length of stay varies across different source markets. However, no clear pattern was recognized regarding countries' characteristics.

Empirical evidences show contradictory results regarding the effect of travel behaviour on tourists' length of stay. Some studies found that individuals travelling more frequently tend to stay longer (Barros et al., 2008; Fleischer & Pizam, 2002; Gokovali et al., 2007). However, contrary evidences were found at three studies (Alegre et al., 2011; Alegre & Pou, 2006; Uysal et al., 1988). Alegre and Pou's (2006) and Alegre et al. (2011) found a non-monotonic relationship between these two variables, the shorter expected stay being attributed to those individuals who travelled between two and four times in the previous year, while those who travelled once or more than four times are expected to stay longer than the first group.

The effect of some additional individual characteristics were analysed by single studies. Hellström (2006) estimated that the length of stay is negatively affected by the number of adults and positively affected by the number of children in the household. Regarding senior tourists in Israel, Fleischer and Pizam (2002) found that healthy individuals tend to travel for longer than unhealthy ones. These authors also analysed the influence of belonging to retirement associations and the level of religious orthodoxy. However, both variables were found to be non-significant. Nicolau and Más (2009) estimated that tourists living in larger cities tend to travel for longer periods. The authors argued that this relationship is due to a higher level of stress and a more intense need for relaxation of individuals living in big cities.

Analysing tourists' behaviour in Azores, Menezes et al. (2008) gave attention to the effect of individuals' attitudes towards sustainability initiatives. Tourists were asked to rank the importance of different sustainability practices or environmental initiatives in the tourism industry as an integral part of a high-quality holiday experience. The authors found that awareness towards waste and environmental management are associated with longer stays, while positive attitudes towards water management display no significant effect on tourists' length of stay.

4.3.2 Travel characteristics

Travel characteristics may be considered simultaneous variables since they are outcomes of tourists' choice processes. When planning their trips, tourists have to decide several aspects, such as destination, length of stay, type of organization, mean of transport and accommodation type. Using travel characteristics as explanatory variables of tourists' length of stay might violate the exogeneity assumption of the modelling process. Yet, most studies have adopted this procedure. The study of Alegre and Pou (2006) was the only one to explicitly assume exogeneity of travel characteristics as a working hypothesis.

The result of underlying simultaneity is that estimated parameters are biased. Almost no empirical assessments of the relevance of this bias are provided in the academic literature. Maybe the only case is the one reported by Silberman (1985), who found no significant differences between models with and without controlling for endogeneity. On the other hand, the introduction of travel characteristics in the model of tourists' length of stay might help segmenting the demand. In this way, managers become able to better identify tourists and predict their length of stay. Being aware of these restrictions and advantages, empirical findings about travel characteristics as explanatory variables of tourists' length of stay are discussed in the following.

Travel purpose is a fundamental explanatory variable of tourists' length of stay. Most studies considered the effect of this variable and non-significant results were found only by Martínez-Garcia and Raya (2008) and Raya and Martínez-Garcia (2011). Empirical evidences show that the effect of specific purposes depends on the destination. In Azores, Portugal, Menezes et al. (2008) found that business tourists are expected to stay shorter than tourists visiting friends and relatives, while leisure tourists are expected to stay longer than both groups. Leisure tourists were also found to have longer expected stays than other tourists by Mak et al. (1977), and specifically longer than those visiting friends and relatives by Hellström (2006). Contrarily, Yang et al. (2011) found that tourists on vacation at Yixing, China, have a shorter average stay than business tourists, while those visiting friends and relatives display the longest average length of stay.

Tourists' lengths of stay also vary according to activities carried out by individuals. Tourists participating in sports, fishing, golf or tennis tend to stay longer at the destination according to Silberman (1985). In a skiing resort such as Aspen, USA, tourists with skiing abilities also tend to stay longer (Walsh & Davitt, 1983). Buying wine (Barros & Machado, 2010; Machado, 2010) and visiting natural attractions or casinos (Barros & Machado, 2010) in Madeira Island, Portugal, also increases expected tourists' length of stay.

Mak et al. (1977) and Hellström (2006) estimated that price and duration of transport has a positive effect over the length of stay. Conversely, Mak and Nishimura (1979) found a negative effect of airfares. More flexible means of transport seems to be associated with shorter stays. Menezes et al. (2008) estimated that tourists travelling on regular flights tend to stay shorter than those using charter flights. Yang et al. (2011) found that expected length of stay increases according to the mean of transport following the sequence self-driving, coach and bus, airplane and train.

Although longer stays are associated with organized trips according to Walsh and Davitt (1983), most studies found that they are associated with independent tourists (Alegre & Pou, 2006; Gokovali et al., 2007; Mak et al., 1977; Yang et al., 2011). No significant relationships between these two variables were found by Martínez-Garcia and Raya (2008) and Raya and Martínez-Garcia (2011). Gokovali et al. (2007) also did not find found relevant differences in the expected length of stay between tourists with full and partial package tours.

The effect of the type of accommodation on tourists' length of stay is not definite. Higher quality hotels were found to be more associated with longer stays than lower quality hotels by Alegre and Pou (2006). Conversely, higher quality hotels were estimated to be associated with shorter stays by Martínez-Garcia and Raya (2008). Hotels were found to have a negative relationship with length of stay by Mak et al. (1979). Alternative types of accommodation, such as campgrounds, rented and owned dwellings, are associated with longer stays (Alegre et al., 2011; Alegre & Pou, 2006; Martínez-Garcia & Raya, 2008; Silberman, 1985; Raya & Martínez-Garcia, 2011). Yachts as an accommodation type were also found to be associated with longer stays (Gokovali et al., 2007).

The effect of party size on tourists' length of stay is ambiguous. Four studies found that larger travel parties tend to stay shorter at the destination (Alegre et al., 2011; Alegre & Pou, 2006; Fleischer & Rivlin Byk, 2009; Walsh & Davitt, 1983). Fleischer and Rivlin Byk (2009) estimated that as travel parties increase, tourists' expenditures tend to be reduced more due to shorter stays than to a smaller number of trips or lower quality of services consumed. A positive effect of travel party size on tourists' length of stay was found by three studies (Barros et al., 2008; Mak & Nishimura, 1979; Uysal et al., 1988;). Gokovali et al. (2007) also found a positive relationship when using a Cox Proportional Hazards model, nevertheless no significant relationship was found when the Weibull model was applied. Other cases of non-significant relationships were reported by Barros et al. (2010), Martínez-Garcia and Raya (2008) and Raya and Martínez-Garcia (2011).

Two studies analysed the relationship between tourists' length of stay and the type of board contracted. Both empirical evidences were contradictory. Alegre and Pou (2006) estimated that more complete boards, such as full board, are associated with longer stays, while Gokovali et al. (2007) estimated the contrary.

Tourists' length of stay tends to be longer during high season (Fleischer & Rivlin Byk, 2009; Martínez-Garcia & Raya, 2008; Raya & Martínez-Garcia, 2011; 2006). Advanced planning or booking is also associated with longer stays (Alegre et al., 2011; Gokovali et al., 2007; Silberman, 1985). Regarding sources of information for the trip, longer stays were found to be associated with brochure (Barros et al., 2008) and advertising (Silberman, 1985).

Repeaters are expected to stay longer according to Menezes et al. (2008), while the opposite was estimated by Paul and Rimmawi (1992) and Silberman (1985). When the number of previous visits to the destination is regarded, all evidences indicate a positive relationship with tourists' length of stay (Alegre et al., 2011; Alegre & Pou, 2006; Barros & Machado, 2010; Gokovali et al., 2007; Mak et al., 1977; Yang et al., 2011). Walsh and Davitt (1983) found that tourists in Aspen tend to stay longer if they have spent larger shares of their previous skiing trips in this same destination. Non-significant relationships between previous visits and tourists' length of stay were found by Machado (2010) and Barros et al. (2008). It is interesting to note that previous visits to the destination was the sample selection variable in Barros and Machado (2010).

4.3.3 Destination characteristics and tourists perceptions

Martínez-Garcia and Raya (2008) and Raya and Martínez-Garcia (2011) estimated that tourists on a sun and sea resort of Catalonia, Spain, are expected to stay longer than those on medium and large cities such as Gerona and Barcelona. These were the only previous studies where variations of tourists' length of stay were analysed across specific destinations. All other studies focused on the influence of destinations' characteristics rather than the effect of destinations themselves. However, it is worthy to note that the number of destinations analysed in these cases was small since only Gerona, Barcelona and a third category consisting of 'sun and sand destinations'.

Uysal et al. (1988), while studying tourists' length of stay at skiing resorts, found that shorter stays are associated with man-made structures and with the offer of non-recreational activities. These authors also found that longer stays are associated with crowded areas.

The most studied objective characteristic of tourist destinations affecting the length of stay is distance from tourists' origin. Most studies found that tourists travelling further are expected to stay longer at the destination (Blaine et al., 1993; Nicolau & Más, 2009; Paul & Rimmawi, 1992; Silberman, 1985; Yang et al., 2011; Walsh & Davitt, 1983), while only one study found the opposite (Uysal et al., 1988).

A few subjectively measured variables regarding destination characteristics were used to explain length of stay. Silberman (1985) estimated that expected stay is longer for tourists who perceive the destination as 'classy' and shorter for those who perceive the destination as 'rundown'.

The importance of different destination aspects over tourists' decisions were analysed by several studies. Although some authors have interpreted these variables as destination's attributes (Barros et al., 2008; Barros et al.; 2010; Barros & Machado, 2010; Machado, 2010; Peypoch et al., 2012), in reality their meaning is strongly contaminated by individuals' preferences. The importance of climate, for instance, depends both on destination's actual climate and on tourist's preference and perception of this aspect. Estimated effects vary according to the destination and other travel characteristics. A typical example is that Alegre and Pou (2006) estimated that tourists who attribute large importance to beaches tend to stay longer in Mallorca, Spain, while Barros et al. (2010) estimated that golfers with the same opinion about the relevance of beaches tend to stay shorter in Algarve. This difference is reasonable since both Mallorca and Algarve are a sun and sea destinations and only golfers were considered in the latter study.

The importance of climate over tourists' decisions was found to have a positive effect over the length of stay by several studies (Barros et al., 2008; Barros et al., 2010; Machado, 2010; Menezes et al., 2008; Nicolau & Más, 2009; Peypoch et al., 2012), while no significant effect was found by Alegre and Pou (2006). A positive influence of the attributed importance to nature was found by Barros et al. (2008), Menezes et al. (2008) and Peypoch et al. (2012). Attributed importance to cultural heritage leads to shorter stays according to Menezes et al. (2008), while Barros et al. (2008) estimated that tourists who find culture an important aspect tend to stay longer. While Barros et al. (2008) found a positive influence of gastronomy, Peypoch et al. (2012) found the opposite. Opposite effects were also found with respect to security (Barros et al., 2008; Peypoch et al., 2012). The importance credited to price as a destination choice determinant was found to have a negative effect over tourists' length of stay by Alegre and Pou (2006) and Nicolau and Más (2009), whereas no significant effect by Menezes et al. (2008). While Barros et al. (2008) found a negative effect for the importance of distance, Menezes et al. (2008) found a positive effect for destination's remoteness. Other aspects whose importance have positive effects over tourists' length of stay are events and hospitality (Barros et al., 2010), physical appearance of the population (Peypoch et al., 2012), quality of hotel (Alegre & Pou, 2006) sun and sea (Peypoch et al., 2012), and wine (Machado, 2010). Other aspects with negative effects are life style (Peypoch et al., 2012), ethnicity, and exoticism (Barros et al., 2008). Finally, other aspects with no significant effects are quality of surroundings (Alegre & Pou, 2006), golf court attributes (Barros et al., 2010), availability of packages and flights, safety and hospitality (Menezes et al., 2008). It is worthy to stress that Gokovali et al. (2007) used variables somehow related to travel characteristics as explanatory variables. However, parameter estimates for these variables were not discussed here since their actual meaning was not reported in that study.

Positive tourists' assessments of the destination were found to be associated with longer stays by Machado (2010), while no significant relationship was found by Menezes et al. (2010). Positive assessments of accommodation were also found to be associated with longer expected lengths of stay by Yang et al. (2011). Tourists' intention to return to the destination was estimated to be associated with longer stays by Machado (2010), while Silberman (1985) estimated the opposite. Finally, Menezes et al. (2008) found no relevant association between these two variables. It is worthy to note that the use of tourists' assessments of the destination as explanatory variables is controversial. If most tourists decide their length of stay before the trip, then there is no reason to expect any causal effect of satisfaction on the dependent variable.

4.3.4 Prices and expenditure

Some studies used tourism prices as an explanatory variable of tourists' length of stay (Alegre et al., 2011; Alegre & Pou, 2006; Fleischer & Rivlin Byk, 2009; Hellström, 2006; Mak et al., 1977; Mak & Nishimura, 1979; Silberman, 1985; Walsh & Davitt, 1983). All these studies obtained price estimates from data about tourists' expenditure gathered through demand surveys. Most authors considered tourists' expenditure as a straight proxy of prices. This approach implies assuming severe inelasticity of demand. More careful price estimates were analysed in Fleischer and Rivlin Byk (2009) and

Silberman (1985), who estimated prices from tourists' expenditures by controlling for the type of accommodation. The latter used a MSLS model with two stages, though no information about the instrumental variables was reported.

Estimating prices from expenditure can be seriously criticized since the tourism product is qualified by a vast number of variables. Different expenditures may arise from different qualities, rather than prices. The difference between expenditure and prices was early noted by Mak and Nishimura (1979). For example, two tourists may spend different amounts per day because they consume different types of accommodation or different quality levels of food services. Besides, tourists may consume different amounts of additional services, such as transportation, entertainment and shopping. In our view, expenditure variation seems more strongly related to quality and quantity of additional services than to price. In any case, 'price' variations among individuals estimated from actual expenditure is at least partially correlated with individuals' preferences and constraints, what leads to biased estimates of 'price' effects on tourists' length of stay.

Even when expenditure is controlled by relevant variables, such as type of accommodation, major quality variations remain. Chen and Rothschild (2010), for instance, found evidences that the star rating of hotels might play an insignificant role in price determination, while using a set of fourteen variables altogether explain only 70% of total price variations. Other studies provide further support to the argument that hotel type is far from explaining a substantial share of accommodation prices (Juaneda, Raya, & Sastre, 2011; Espinet, Saez, Coenders, & Fluvià, 2003). Therefore, the price measurement procedure adopted by Silberman (1985) and Fleischer and Rivlin Byk (2009) may still provide substantially biased estimates of price effects on tourists' length of stay.

When total travel expenditure is regarded, most studies found positive association with tourists' length of stay (Alegre et al., 2011; Alegre & Pou, 2006; Machado, 2010; Peypoch et al., 2012). The study of Machado (2010) actually found a significant positive effect only from the separate equations model used, which was not the preferred one. From the preferred model (SUDCD) no significant relationship between total expenditure and length of stay was obtained. Barros and Machado (2010) obtained unexpected results. These authors defined total expenditure as an ordinal discrete variable with six categories, the higher value expressing the higher category of expenditure. Using this ordinal variable as a continuous covariate, the estimated parameters were negative. Therefore, according to this unusual result, larger total expenditures would be associated with shorter lengths of stay.

Studies that analysed the effect of daily expenditure on the length of stay are conclusive. Larger daily expenditures are associated with shorter stays (Alegre et al., 2011; Alegre & Pou, 2006; Fleischer & Rivlin Byk, 2009; Hellström, 2006; Mak et al., 1977; Mak & Nishimura, 1979; Silberman, 1985; Thrane & Farstad, 2011; Uysal et al., 1988; Walsh & Davitt, 1983). If this variable was a true proxy for tourism prices, this finding would be no more than expected according to the demand theory. In that case, Silberman's (1985) parameter estimate -0.126 could be interpreted as the price elasticity. The same is true for Fleischer and Rivlin Byk's (2009) parameter estimate of -0.124. In any case, this later estimate can be compared with parameter's estimates obtained for other models in the same study where the dependent variables were the number of trips and the quality of tourism services consumed. The result of this comparison is that the parameter in the length of stay model is the larger (in absolute value). Therefore, as daily expenditure increases, tourists tend to adjust their budgets first by reducing the length of stay, second by downgrading the quality of services consumed, and lastly by travelling less.

4.4 Empirical analysis of tourists' length of stay in Brazil

The analysis of the determinants of inbound tourists' length of stay at different destinations in Brazil uses data from the BITS, a survey conducted by the Foundation Institute of Economic Research and financially supported by the Tourism Ministry of Brazil. The main objective of the BITS is to provide official tourism statistics for the country. Data was collected through personal interviews from 2004 to 2010 at the 27 main gateways of the country, including 15 airports and 12 land borders[9].

[9] A detailed description of this data source is presented at Section 1.4 of this thesis.

The BITS gathered information about tourists' visits to multiple Brazilian destinations, which were geographically defined as a municipality. A maximum of six different destinations were registered for each tourist. When more than six destinations were visited by the tourist, the survey registered information on the ones with longer stays.

The dependent variable of this study is the length of stay at different Brazilian destinations. Note that this variable may differ considerably from total length of stay in Brazil. Tourists that stay longer at each location may visit a smaller number of destinations in the same trip. Therefore, it might be possible that longer stays at specific destinations were associated with shorter total stays in the country. A total of 181 thousand tourist interviews were obtained from the BITS. Since each tourist visited an average of 1.7 destinations, 309 thousand observations of lengths of stay were available.

The survival distribution of tourists' length of stay in Brazilian destinations according to the BITS is presented in Figure 7. The hazard distribution for this data is presented in Figure 8.

Figure 7: Survival distribution of tourists in Brazilian destinations

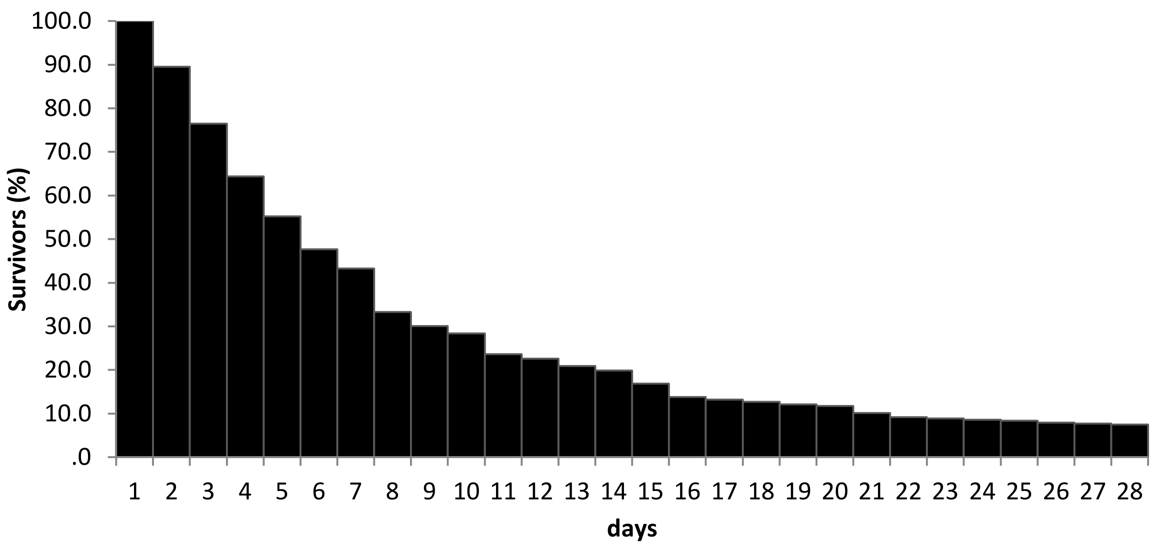
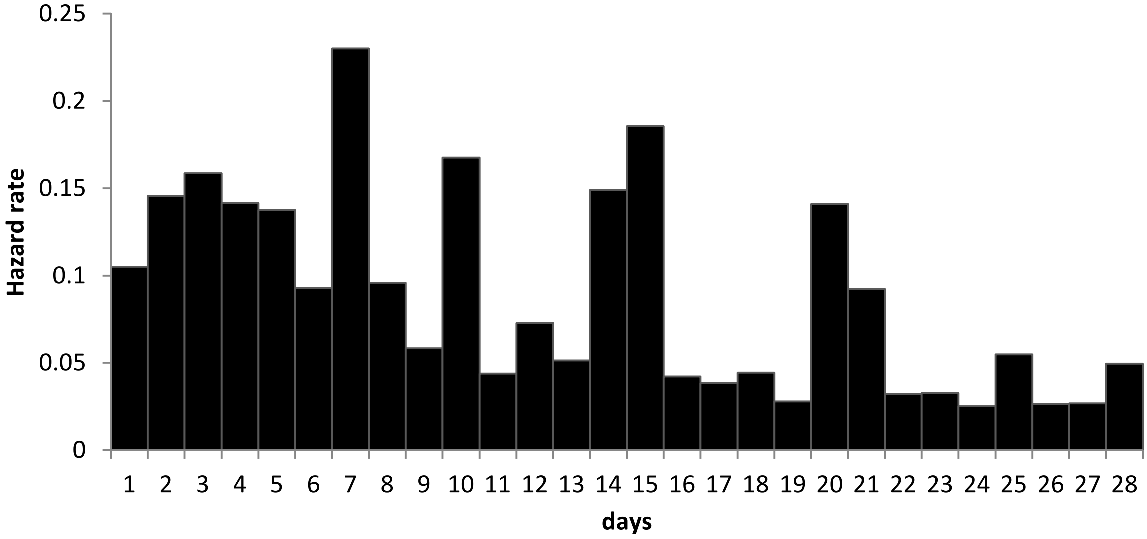


Figure 8: Hazard distribution of tourists in Brazilian destinations



The analysis of Figure 8 show the existence of discontinuous mass points at 7, 10, 14, 15, 20 and 21 days. These points may have two different explanations. First, some concentrations might be related to vacation time constraints, as well as to supply side restrictions. Many people take one or two weeks vacations. Besides, several organized trips are one or two weeks long. Second, the existence of some mass points might be related to a measurement error arising from imprecision on tourists' responses. This source of error is more likely to happen for larger lengths of stay (Biemer, Groves, Lyberg, Mathiowetz, & Sudman, 2004). For instance, instead of answering 19 days, some tourists might tend to make an approximate mental account and finally answer 20 days.

Covariates used in the explanation of the length of stay were categorized into five groups. The first group includes different tourists' characteristics, such as gender, age, level of education, household income and country of residence. Age is introduced in its linear and squared forms, allowing for non-monotonic relationships of this variable with the length of stay. Level of education is measured in four ordinal categories: under high school, high school, graduation, and post-graduation. Household income is measured in thousands of constant dollars of 2010. Squared income was also used, allowing for a non-monotonic effect. Eleven countries from where more than 100 thousand tourists go to Brazil every year are identified by specific dummy variables. The remaining countries are identified by continent specific dummy variables.

The second set of explanatory variables includes several travel characteristics, such as the number of destinations visited in Brazil, travel purpose, mean of international transport used to leave Brazil, type of trip organization, type of accommodation, party size, number of previous visits to Brazil, season and tourists' *per capita* daily expenditure. The number of destinations visited and its squared value are used additionally to a dummy variable indicating single destination trips as a qualitatively different case. Travel purpose is a multinomial variable including four leisure and three non-leisure purposes. Squared and non-squared values are also used for party size and per capita daily expenditure. Seasons were defined as high season (December to February), low season (July and August) and off-season (remaining months). Expenditure is measured in constant dollars of 2010.

These two first groups of explanatory variables follow the tradition of most previous studies on modelling tourists' length of stay aforementioned. The estimates of this study might be particularly useful for tourism managers in Brazil. Besides, the study of these effects also allows for comparisons with previous studies, providing further empirical evidences for a broader understanding of tourists' length of stay. However, it is necessary to stress that the effects of some variables on tourists' length of stay within the single destination paradigm might be different from their effects within the MTTs context. This may happen mainly because tourists' and travels' characteristics may influence both the number of destinations visited and the length of stay at each location. Thus, comparisons between the present study and previous evidence should be conducted carefully.

It is relevant to stress that the use of travel characteristics as explanatory variables might incur in the problem of simultaneity, as discussed at Section 4.3.2. By introducing these covariates at econometric models, the coefficients of variables regarding tourists' profiles should be interpreted as the effect of a given covariate on the expected length of stay for trips with the same characteristics. The coefficients of income, for instance, indicate the expected effect of a unitary increase of income on the length of stay while keeping constant other travel characteristics. The set of travel characteristics used as explanatory variables includes the number of destinations visited, what may reduce differences between effects of other variables in the single and multiple destination paradigms. Travel characteristics were used in order to allow for a more detailed understanding of tourists' behaviour regarding the length of stay and facilitating demand segmentation analysis.

The third category of explanatory variables regards the set of different destinations visited. A total of 190 destinations with more than a hundred observations in the database are identified by a specific dummy variable. All destinations are also categorized according to its state and region[10]. This approach provides evidences about the heterogeneity of destinations regarding tourists' length of stay. Besides, it provides information on which state or region has the longest expected length of stay. The analysis of this information may help to identify best tourism management practices. This is the first time in the academic literature where the effects a large number of specific destinations over tourists' length of stay are estimated. All destinations are also characterized by their population and by a binary variable indicating coastal localization.

The forth category of variables includes a single information, the year of the trip as identified by specific dummy variables. The objective of introducing this information in the model is to test the hypothesis that the increasing trend of tourists' length of stay in Brazilian destinations has been caused by changes on tourists' behaviour. This hypothesis is tested through the set of year specific dummies since they capture time variations of the expected length of stay for tourists with constant profiles. If this set of dummy variables displays an increasing trend, then the hypothesis would be confirmed. Otherwise, the estimates would lead to the acceptance of the null hypothesis that the increasing trend has been caused by changes in the composition of the inbound tourism flow. The confirmation of the null hypothesis would be consistent with the observed worldwide decreasing trend of tourists' length of stay.

Finally, the fifth category of covariates consists of instrumental dummy variables used to identify lacking data. This procedure aims to avoid unnecessary data loss, as well as a potential sample selection problem. For instance, 21.5% of the total dataset has missing values at the income variable. It is reasonable to believe that these lacking observations are not random with regard to the level of income. Therefore,

[10] Brazil is officially divided into 27 states, which are grouped into five regions.

omitting observations with lacking data could incur on a relevant sample selection problem.

It is possible that the effect of each variable on the expected length of stay varied from 2004 to 2010. However, taking account of this characteristic would require specific coefficients for each variable in each year. The number of parameters to be estimated would exceed two thousand, implying on a great estimation cost and interpretation complexity. As an alternative solution for this issue, two models were estimated. Model 1 used all 309 thousand observations from 2004 to 2010. This model included the set of dummy variables identifying each year, but the effects of all remaining covariates were considered to be constant over time. A total of 285 explanatory variables were used in this model. Alternatively, Model 2 used only the set of 54 thousand observations available for 2010, avoiding the parameter time-variation issue by analysing a single year. This model used 279 explanatory variables.

Duration models were employed in the analysis. The dependent variable was inbound tourists' length of stay at different Brazilian destinations. This variable was measured in number of overnights. Observations are not independent from each other since the MTTs paradigm implies that some data come from the same tourists. Tourists' usual heterogeneity in this case leads to shared heterogeneity on data. This is a special feature of the data on tourists' length of stay in the multidestination trips context. In order to account for this characteristic, shared heterogeneity was introduced in the duration models applied.

The Cox Proportional Hazards model with shared heterogeneity is not appropriate to this case since most explanatory variables do not vary across different observations of the same tourist. For instance, the income of a given tourist is constant with respect to all different destinations visited in the same trip. The only exception is the set of dummy variables used to identify the specific destination visited. Therefore, the Cox Proportional Hazards model would prevent the estimation of most parameters of interest since its shared heterogeneity scheme requires an exclusive parameter for each tourist. Therefore, almost all variation would be attributed to the heterogeneity parameters.

Anyway, the proportional hazards hypothesis of the Cox model was tested using Schoenfeld's residuals by taking two different approaches. First, the Schoenfeld test was conducted by using all observations while shared heterogeneity was omitted. Second, the test was applied to a sub-sample of the dataset where a single observation from each tourist was randomly selected. The proportional hazards hypothesis was rejected by both approaches with respect to almost all variables. Therefore, the Cox Proportional Hazards model was discarded.

The estimation process of parametric duration models was constrained by availability of statistical packages and computational capacity. According to Kelly (2004) there are five statistical packages able to handle shared heterogeneity duration models without resorting to advanced programing of mathematical optimization packages. WinBugs (Lunn, Thomas, Best, & Spiegelhalter, 2007) was discarded due to its small set of available distributions for the duration variable. S-Plus (Insightful Corporation, 2007), R (R Development Core Team, 2011) and MLinN 2.24 (Rasbash, Browne, Healy, Cameron, & Charlton, 2011) were discarded since they were not able to handle all covariates using a sample of more than 20 thousand tourists. Finally, Stata (StataCorp, 2010) was able to estimate models using a much larger number of observations. Stata offers four distributions for the duration variable (Weibull, Gompertz, lognormal and loglogistic) that can be used together with two distributions for the heterogeneity term (gamma and inverse Gaussian). All eight possible combinations were considered. Unfortunately, even Stata presented some estimation restriction. The statistical package was not able to estimate models with Gompertz and loglogistic distributions using the whole dataset from 2004 to 2010. Therefore, only Weibull and lognormal distributions for the duration variable were tested at Model 1. Table 7 presents the log likelihood and the Akaike Information Criterion (AIC) of all twelve estimated models.

Table 7: Summary statistics of estimated models

Duration distribution	Heterogeneity distribution	Model 1 (2004-2010)		Model 2 (2010)	
		Log likelihood	AIC	Log likelihood	AIC
Weibull	Gamma	-398,890	798,356	-69,869	140,299
	Inverse Gaussian	-404,858	810,292	-70,886	142,333
Gompertz	Gamma			-73,692	147,944
	Inverse Gaussian			-73,694	147,948
Lognormal	Gamma	-377,694	755,963	-66,130	132,821
	Inverse Gaussian	-378,081	756,737	-66,191	132,942
Loglogistic	Gamma			-65,590	131,739
	Inverse Gaussian			-65,599	131,757

The gamma heterogeneity distribution provided slightly better results with any duration distribution for both models. At Model 1, the lognormal distribution for the duration variable provided the best results, while at Model 2 the best distribution was the loglogistic. These two best fitting models were selected for analysis. It is relevant to note that both performance measures indicate that the superiority of the loglogistic distribution to the lognormal at Model 2 was relatively small (0.8%), while the difference to other distributions was considerably larger (7% to Weibull and 12% to Gompertz). This result supports the theoretical proposition that skewed density distributions of tourists' length of stay are appropriate to represent tourists' length of stay.

Estimated coefficients, standard errors and p-values of both selected models are presented in three different tables. Table 8 displays estimates referring to individuals' characteristics, while Table 9 refers to travel characteristics and Table 10 presents estimates regarding different destinations, years and models' constants. The first line of each multinomial variable presents the p-values for the Wald test of the hypothesis that all coefficients are simultaneously equal to zero, that is, for the hypothesis that the multinomial variable as a whole is not significant. The same is done to quantitative variables used in their squared and non-squared forms. The first category of multinomial variables indicates their reference groups.

Table 8: Model estimates – Individuals' characteristics

Variable	Model 1 (2004-2010)			Model 2 (2010)		
	Coefficient	Std. error	p-value	Coefficient	Std. error	p-value
Gender						
Female*	0.00			0.00		
Male	-0.049	0.0035	0.00	-0.048	0.0072	0.00
Age			0.00			0.00
Exponent 1	-0.012	0.00074	0.00	-0.0094	0.0016	0.00
Exponent 2	0.00013	8.4E-06	0.00	0.00011	1.8E-05	0.00
Education level			0.00			0.00
Under high school*	0.00			0.00		
High school	0.023	0.0080	0.00	0.033	0.020	0.09
Graduation	-0.038	0.0080	0.00	-0.029	0.019	0.14
Post-graduation	-0.064	0.0086	0.00	-0.057	0.020	0.01
Income			0.87			0.01
Exponent 1	-0.00034	0.00065	0.60	0.0012	0.0015	0.40
Exponent 2	8.1E-06	1.6E-05	0.62	2.1E-05	3.5E-05	0.55
Origin			0.00			0.00
Africa*						
*All countries**	0.00			0.00		
Asia and Oceania						
All countries	0.071	0.015	0.00	-0.14	0.035	0.00
Central America						
All countries	-0.17	0.022	0.00	-0.35	0.047	0.00
Europe						
England	-0.014	0.015	0.36	-0.14	0.035	0.00
France	-0.045	0.015	0.00	-0.20	0.034	0.00
Germany	-0.00059	0.015	0.97	-0.17	0.034	0.00
Italy	0.038	0.015	0.01	-0.071	0.035	0.04
Portugal	-0.048	0.015	0.00	-0.14	0.035	0.00
Spain	-0.028	0.015	0.06	-0.15	0.035	0.00
Other Europe	0.0066	0.014	0.64	-0.14	0.032	0.00
North America						
USA	-0.17	0.014	0.00	-0.31	0.032	0.00
Other North America	-0.060	0.015	0.00	-0.18	0.034	0.00
South America						
Argentina	-0.37	0.014	0.00	-0.43	0.033	0.00
Chile	-0.37	0.015	0.00	-0.50	0.034	0.00
Paraguay	-0.83	0.017	0.00	-0.89	0.037	0.00
Uruguay	-0.56	0.018	0.00	-0.68	0.038	0.00
Other South America	-0.23	0.015	0.00	-0.39	0.033	0.00

* reference group

Estimates of models 1 and 2 presented only slight differences. Magnitudes diverged mostly with respect to the less frequent categories of multinomial variables. Qualitative interpretations of estimated coefficients in both models are the same. No relevant differences on signs were obtained. In other words, both models indicate the same qualitative explanatory structure of tourists' length of stay. Therefore, the cost of assuming time invariant effects was considered smaller than the benefit of using a larger number of observations and consequently obtaining smaller standard deviations on parameters' estimates. Thus, the following discussion majorly focuses on estimates of Model 1.

Table 8 shows that men tend to stay roughly 4.7%[11] less time at destinations than women. This is the first study to find gender differences in this direction. Previous studies have mostly found no significant differences, while Barros and Machado (2010) and Peypoch et al. (2012) found that men tend to stay longer.

The expected length of stay follows a significant convex function of age[12]. This result is consistent with Yang et al. (2011) and conflicting with Fleischer and Pizam (2002). The effect of age is negative for young tourists, while for older tourists this influence is positive. The minimum length of stay is expected for tourists age 46. Tourists age 56, for example, are expected to stay 1.3%[13] longer than those ten years younger.

The relationship between level of education and length of stay is significant, though not monotonic. Tourists who have completed high school tend to stay 2.4%

---

[11] All analytical calculations of the effects of non-continuous covariates considered discrete differences instead of derivatives. In the case of dichotomous covariates, for instance, the discrete effect is calculated as $1-\frac{T(x=1)}{T(x=0)}=1-\frac{e^{\beta\times 1}}{e^{\beta\times 0}}=1-e^{\beta}$. For example, the effect of male gender was calculated as $1-\frac{T(gender=male)}{T(gender=female)}=1-e^{-0.049}=4.7\%$. Further examples are provided in the next footnotes for slightly different forms of calculation. It is worthy to stress that the difference between the estimated parameter and the discrete effect is larger for greater βs.

[12] It is relevant to discuss the interpretation of the effects of quantitative variables used in their non-squared and squared forms. If the coefficient of the squared variable is positive, the expected length of stay follows a convex function of the explanatory variable. On the other hand, if the coefficient of the squared variable is negative, the expected length of stay is not a concave function. Rather, it presents a bell shape, that is, it is concave in the center and convex at both sides, tending to zero at both extremes. If the total effect of x is given by $\beta_1 x^2+\beta_2 x$, then the maximum point is at $-\beta_2/2\beta_1$, and the limits between the concave and the convex parts are at $\frac{-\beta_2 \mp \sqrt{-2\beta_1}}{2\beta_1}$.

[13] $\frac{T(age=56)}{T(age=46)}-1=\frac{e^{-0.012\times 56+0.00013\times 56^2}}{e^{-0.012\times 46+0.00013\times 46^2}}-1=1.3\%$.

longer than those who did not. On the other hand, graduated and post-graduated tourists tend to stay shorter than high school educated tourists (5.9% and 8.4%, respectively). It is interesting to recall that previous studies were inconclusive about the relationship between these two variables. Thus, despite of its significance, the influence of level of education seems to present a complex and variable pattern.

Estimates show that tourists' income does not have a significant effect on expected length of stay within the MTTs paradigm. Both squared and non-squared income variables were non-significant at models 1 and 2. Besides, the p-value of the overall significance of income at Model 1 is 0.87. The overall significance of income at Model 2 (p=0.01) does not provide any relevant interpretation since both variables considered are not significant when assessed alone. Previous studies within the single destination paradigm were not conclusive regarding the relationship between income and length of stay. Although most studies found a positive relationship between both variables, some studies found the opposite (Blaine et al., 1993; Mak & Nishimura, 1979), while case found a non-monotonic relationship (Nicolau & Más, 2009) and two other cases found that income was not a significant explanatory variable (Hellström, 2006; Barros et al., 2010). The present study supports the idea that the effect of income on the length of stay is case specific.

Place of origin is a significant determinant of tourists' length of stay (p<0.01). Asians and Oceanians are the ones staying longer. Other tourists with relatively large expected stays are those from Africa, Germany, England, Spain and Other European Countries. On the other hand, South Americans tend to stay for the shortest periods. Paraguayans present the overall shortest expected length of stay (59.4%[14] less than Asians and Oceanians). These results seem to point out distance as an underpinning variable providing sense to differences across countries. Tourists from farther countries seem to tend to stay longer at each destination visited. This outcome is consistent with most previous studies.

[14] $1-\frac{\mathrm{T(origin=Paraguay)}}{\mathrm{T(origin=Asia\ and\ Oceania)}}=1-\frac{e^{-0.83}}{e^{0.071}}=59.4\%$.

Table 9: Model estimates – Trips' attributes

Variable	**Model 1 (2004-2010)**			**Model 2 (2010)**		
	Coefficient	**Std. error**	**p-value**	**Coefficient**	**Std. error**	**p-value**
Multidestination trip						
No	0.00			0.00		
Yes	-0.19	0.0066	0.00	-0.15	0.015	0.00
Number of destinations						
Logarithm	-0.55	0.0055	0.00	-0.60	0.013	0.00
Purpose			0.00			0.00
Leisure purposes						
*Sun and sea**	0.00			0.00		
Ecotourism	-0.052	0.0064	0.00	-0.023	0.014	0.11
Cultural tourism	-0.058	0.0069	0.00	-0.031	0.017	0.06
Other	-0.11	0.010	0.00	-0.058	0.024	0.02
Non-leisure purposes						
Business	-0.29	0.0058	0.00	-0.26	0.014	0.00
VFR	0.022	0.0057	0.00	0.039	0.013	0.00
Other	0.038	0.010	0.00	0.0017	0.022	0.94
Transport mode						
Road*	0.00			0.00		
Air	0.11	0.0067	0.00	0.13	0.013	0.00
Trip organization						
Independent*	0.00			0.00		
Package	-0.11	0.0052	0.00	-0.22	0.012	0.00
Accommodation			0.00			0.00
Hotels and counterparts*	0.00			0.00		
Friends or relatives' dwellings	0.26	0.0050	0.00	0.25	0.012	0.00
Rented dwelling	0.35	0.0067	0.00	0.25	0.015	0.00
Owned dwelling	0.51	0.0088	0.00	0.53	0.022	0.00
Other	0.086	0.012	0.00	0.12	0.026	0.00
Party size			0.00			0.00
Exponent 1	-0.019	0.0034	0.00	-0.040	0.0076	0.00
Exponent 2	0.0016	0.00041	0.00	0.0030	0.00094	0.00
First visit to Brazil						
No	0.00			0.00		
Yes	-0.10	0.0038	0.00	-0.11	0.0083	0.00
Number of previous visits			0.00			0.00
Exponent 1	-0.0020	0.00010	0.00	-0.0025	0.00022	0.00
Exponent 2	1.5E-06	1.2E-07	0.00	1.9E-06	2.9E-07	0.00
Season			0.00			0.00
Summer season	0.00			0.00		
Winter season	-0.021	0.0042	0.00	-0.044	0.010	0.00
Off-season	-0.11	0.0039	0.00	-0.13	0.0085	0.00
Per capita daily expenditure			0.00			0.00
Exponent 1	-0.0026	3.7E-05	0.00	-0.0029	8.6E-05	0.00
Exponent 2	2.4E-06	6.2E-08	0.00	2.6E-06	1.5E-07	0.00

Shorter lengths of stay at destinations are associated with multidestination tourism trips. Tourists visiting two destinations are expected to stay 43.7%[15] shorter than those visiting a single destination. As the number of destinations visited increase, the duration of the stay decreases even more. The addition of a third destination in the itinerary implies a 20.0%[16] extra decrease on the expected length of stay at each destination, while for the fourth destination the expected decrease is 14.6%. It is interesting to note that stepping from a single to a multiple destination trip implies a qualitatively different change as compared to the addition of a destination to an originally multidestination trip. In other words, the number of destinations is not enough to explain the length of stay, and the difference between single and multiple destination trips is relevant. This qualitative difference is evidenced by the significance of the multidestination trip variable.

Travel purpose is significantly associated with tourists' length of stay ($p<0.01$). Sun and sea tourists are expected to stay for a relatively long period, although tourists visiting friends and relatives (VFR) tend to stay roughly 2.3% longer. Ecotourists and cultural tourists present medium expected lengths of stay. Business tourists tend to stay for the shortest period among the identified purposes; around 26.8% less than VFR tourists. As discussed earlier, the relationship between travel purpose and length of stay may vary from one to another destination. Thus, these results cannot be properly compared with previous empirical findings.

Tourists taking international trips by air are expected to stay 11.9% longer than tourists travelling by roads. This finding is consistent with previous studies. Longer stays are also expected for tourists travelling independently. Those travelling on organized package tours tend to stay about 10.8% shorter than independent tourists. This finding is also consistent with most previous studies.

Type of accommodation is a significant covariate of tourists' length of stay ($p<0.01$). Those staying at hotels are expected to stay for the shortest period. As

[15] This value is calculated from the estimated parameters of both variables multidestination trip and number of destinations, that is $1-\frac{\mathrm{T(multidestination\ trip=yes;\ number\ of\ destinations=2)}}{\mathrm{T(multidestination\ trip=no;\ number\ of\ destinations=1)}}=1-\frac{e^{-0.19\times1-0.55\ln(2)}}{e^{-0.19\times0-0.55\ln(1)}}=$ 43.7%.

[16] $1-\frac{\mathrm{T(multidestination\ trip=yes;\ number\ of\ destinations=3)}}{\mathrm{T(multidestination\ trip=no;\ number\ of\ destinations=2)}}=1-\frac{e^{-0.19\times1-0.55\,ln(3)}}{e^{-0.19\times1-0.55\,ln(2)}}=20.0\%$.

compared to this group, tourists accommodating at friends' and relatives' dwellings are expected to stay 29.9% longer, those at rented dwellings 41.5% longer, and tourists at their own dwellings 66.9% longer. These results may help to understand the relationship between type of accommodation and tourists' length of stay since previous studies were inconclusive in this issue.

The effect of party size is negative in most cases. The expected length of stay follows a convex function of party size with a minimum value at 6.0 people. About 0.5% of the sampled parties are larger than six people. This result may enlighten the ambiguity found by earlier researchers. No previous study allowed a non-monotonic effect of party size on tourists' length of stay. The adoption of monotonic functions might have been the cause of their contradictory findings.

Tourists who visit Brazil for the first time are expected to stay 9.4%[17] shorter at each destination than those who are visiting the country for the second time. This outcome is consistent with Menezes et al. (2008), though contradictory with Paul and Rimmawi (1992) and Silberman (1985). For tourists that have already visited Brazil, the expected length of stay decreases about 0.2% for each additional previous visit. Although this rate of decrease is very small, both parameters involved are significant. This result is contradictory to all previous studies (Alegre et al., 2011; Alegre & Pou, 2006; Barros & Machado, 2010; Gokovali et al., 2007; Mak et al., 1977; Yang et al., 2011). Note that the effect of the second visit is substantially positive, while the addition of further visits is only marginally negative. Moreover, earlier studies did not consider the difference between both variables regarding previous visits (i.e., first visit and number of previous visits). In those cases, the strong positive effect of being a repeater might have prevailed over the negative effect of additional previous visits. Misleading conclusions might have been obtained under these conditions. Therefore, the conflicting finding of the present study may have been obtained due to the more complete consideration of the issue of previous visits to the destination.

The summer season is associated with the longest stays, while the off-season is associated with the shortest stays. The difference of the length of stay between both is

[17] This value is calculated from the estimated parameters of both variables first visit to Brazil and number of previous visits.

10.0%. The length of stay at the winter season is closer to the one at the summer season (2.1% difference). These findings are consistent with all previous empirical evidence.

Tourists' length of stay follows a convex function of *per capita* daily expenditure. The minimum point of this function is at US$ 547. Only 1.4% of all international tourists in Brazil spend a higher daily average amount. Thus, the relationship between both variables is negative for most cases. This finding is strongly consistent with earlier studies. Besides, the marginal effect of expenditure is also decreasing (in absolute values). An additional dollar of expenditure is associated with a 0.9% shorter stay for tourists spending around US$ 100 daily, while the same figure for tourists spending US$ 200 is 0.4%.

Table 10 displays estimates relative to destinations and constants of the model. States' and destinations' specific parameters are shown in appendixes 3 and 4 in order to focus the text on the main results. Dummy variables for lacking data are not displayed because of their irrelevance for analysis.

The coastal nature of the destination was found to be a significant determinant of tourists' length of stay ($p=0.01$). Likewise, more populated destinations are positively associated with longer stays. These two outcomes indicate that the average length of stay varies according to some destinations' attributes. Further investigation in this topic should be made on future studies.

Tourists visiting the Southeast and the South are expected to stay significantly longer. Therefore, destinations in these regions should be taken as benchmarks for destinations in other regions with regard to tourists' length of stay. Future detailed analysis of the characteristics of destinations in these regions may provide useful insight about how to push up tourists' length of stay. All groups of variables identifying specific destinations were statistically significant ($p<0.01$), what indicates that the expected length of stay varies significantly across destinations.

Table 10: Model estimates – Destinations, year and constants

Variable	Model 1 (2004-2010)			Model 2 (2010)		
	Coefficient	Std. error	p-value	Coefficient	Std. error	p-value
Coast	0.042	0.017	0.01	-0.016	0.041	0.70
Population	6.6E-07	5.8E-08	0.00	6.0E-07	1.3E-07	0.00
Region			0.00			0.00
South*	0.00			0.00		
Southeast	0.16	0.036	0.00	0.25	0.088	0.00
Central-West	-0.87	0.75	0.24	-2.0	0.39	0.00
Northeast	-0.060	0.064	0.35	0.028	0.18	0.88
North	-0.49	0.14	0.00	0.19	0.51	0.70
State			0.00			0.00
...						
Destination			0.00			0.00
...						
Year			0.00			
2004*	0.00					
2005	0.0046	0.032	0.89			
2006	0.030	0.032	0.34			
2007	-0.0030	0.032	0.93			
2008	-0.018	0.032	0.56			
2009	-0.017	0.032	0.58			
2010	0.0039	0.032	0.90			
Constants						
β_0	3.2	0.086	0.00	3.4	0.20	0.00
γ				0.43	0.0019	0.00
σ	0.76	0.0013	0.00			
v	0.12	0.0018	0.00	0.054	0.0035	0.00

The set of dummy variables identifying the year of the observation is significant ($p<0.01$). Although all parameters of individual years were non-significant, the correlation between year and the value of estimated parameters for year dummies is -0.38. However, this correlation was not significant ($p>0.10$). Therefore, there is no concrete evidence of a decreasing trend in the expected length of stay for tourists with the same profiles and travel characteristics over time. However, even if the decreasing trend is not confirmed, testing the opposite trend would lead to hypothesis rejection. Thus, this result is sufficient to affirm that the increasing trend of tourists' length of stay in Brazilian destinations was not caused by changes of tourists' behaviour. The null hypothesis that this trend was caused by changes in the composition of the inbound tourism flow is accepted. In other words, since the expected length of stay for a tourist with constant characteristics is not increasing over time, the increasing trend of the average length of stay at Brazilian destinations can only be attributed to the attraction of different sorts of tourists over the years.

Finally, it is necessary to highlight the statistical significance of the v parameter. This means that shared heterogeneity is a relevant characteristic of data regarding different stays from the same tourists. Model 1 with shared heterogeneity provided an absolute percentage error 15.1% smaller than the same model without shared heterogeneity. For Model 2 the same figure was 7.5%. Besides, the non-heterogeneity Model 1 incorrectly indicates 17 variables as non-significant ($p<0.05$) when, in fact, they were significant. On the other hand, the non-heterogeneity Model 1 mistakenly indicates one non-significant variable as if it was significant. The same figures for Model 2 were six and one, respectively. Therefore, shared heterogeneity must be taken into account in this case in order to obtain improved estimates.

4.5 Conclusion

This chapter modelled international tourists' length of stay in Brazilian destinations within the MTTs paradigm. Results obtained help to understand tourists' behaviours, and consequently to forecast variations of the average length of stay according to changes in its determinants, to develop efficient marketing strategies aiming to push the average length of stay to the desired direction, and to develop individual 'on the fly' marketing strategies. The empirical study used a large dataset with more than 309 thousand observations of tourists' lengths of stay at multiple destinations. Five main conclusions were derived.

(1) Positively skewed distributions are evidenced to be appropriate for representing tourists' length of stay. In other words, too short or too long stays are less likely, while an optimum length of stay may be observed. This finding supports theoretical expectations related to the joint effects of initial costs of the visitation and decreasing

marginal utility of the stay. Moreover, the superiority of positively skewed hazard distributions indicates that the probability of ending the stay is lower when the elapsed time since the beginning of the stay is small or large. This probability is higher when the spell already lasts for a medium length. This outcome is particularly useful for the development of individual 'on the fly' marketing strategies. For instance, discounts aiming to enlarge the stay of an actual guest should be offered at medium duration of the stay, that is, when the hazard of ending the stay is higher.

(2) Shared heterogeneity across observations proved to be a necessary characteristic of the econometric model when multidestination tourism data is analysed. Introducing a gamma distributed shared heterogeneity term in the duration model improved estimates substantially. Moreover, this finding shows that tourist surveys conducted at single destinations might have an improved predictive capacity by gathering information about tourists' lengths of stay at other destinations visited in the same trip.

(3) Specific effects of different explanatory variables were analysed. Most findings were consistent with previous studies, such as those regarding the effects of origin, mean of transport, type of trip organization, season and expenditure. Further empirical evidence was provided about effects unresolved by earlier studies, such as age, level of education, type of accommodation and travel purpose. Remarkably, income was not found to be a significant determinant of tourists' length of stay within the MTTs paradigm. Further interesting evidences were found on the relationship between length of stay and party size. Different previous studies pointed out opposite effects by using monotonic functions. This study used a non-monotonic function and found that a negative relationship exists for small parties, while for relatively large parties the relationship is positive. It was also found that tourists who visit Brazil for the first time are expected to stay shorter than those who are repeating their visits to the country. However, additional previous trips for repeaters have a negative effect on the expected length of stay. This finding is conflicting with earlier studies and may have been obtained from the more complete consideration of the effect of previous visits in the present study. These additional empirical evidences help to understand how the effects of some explanatory variables change across different destination regions and specific situations. The effect of the number of destinations included in

the itinerary was analysed for the first time in the academic literature. Estimates showed that multidestination trips are associated with shorter stays. Tourists trade off stay time for a larger number of destinations in their trips. This might create some conflict between destinations and destination regions. A tourist profile expected to stay shorter at the destination might be the one expected to stay longer at the destination region since he or she is expected to visit a larger number of destinations. The difference between length of stay at the destination and in the destination region is an important issue that had not been analysed before in the academic literature.

(4) Another relevant conclusion of this study regards differences on tourists' length of stay across destinations and types of destinations. Dummy variables regarding different destinations were statistically significant, what indicates that the expected length of stay varies across destinations. Tourists' length of stay was also found to vary across regions and states, pointing out some cases that might be used as benchmarks for policymakers. Further research should be conducted to investigate the causes of these regional variations. Moreover, variations were shown to be at least partially caused by destinations' characteristics. More populated and coastal destinations were found to be associated with longer stays. The detailed analysis of the effects of other destinations' attributes is a relevant topic for further research.

(5) The increasing trend of international tourists' length of stay at Brazilian destinations was proved to be caused by changes in the composition of the inbound tourism flow. This finding shows that the Brazilian reality is not contradictory to the worldwide international tourism trend regarding length of stay. In fact, the trip of a tourist with constant characteristics does not show an increasing length of stay in Brazilian destinations.

References

Alegre, J., Mateo, S., & Pou, L. (2011). A latent class approach to tourists' length of stay. *Tourism Management, 32*(3), 555-563.

Alegre, J., & Pou, L. (2006). The length of stay in the demand for tourism. *Tourism Management, 27*(6), 1343-1355.

Barros, C. P., Butler, R., & Correia, A. (2010). The length of stay of golf tourism: a survival analysis. *Tourism Management, 31*(1), 13-21.

Barros, C. P., Correia, A., & Crouch, G. (2008). Determinants of the length of stay in Latin American tourism destinations. *Tourism Analysis, 13*(4), 329-340.

Barros, C. P., & Machado, L. P. (2010). The length of stay in tourism. *Annals of Tourism Research, 37*(3), 692-706.

Biemer, P. P., Groves, R. M., Lyberg, L. E., Mathiowetz, N. A., & Sudman, S. (2004). *Measurement errors in surveys*. Hoboken, USA: John Wiley & Sons.

Blaine, T. W., Mohammad, G., & Var, T. (1993). Demand for rural tourism: an exploratory study. *Annals of Tourism Research, 20*(4), 770-773.

Boehmke, F. J. (2006). The influence of unobserved factors on position timing and content in the NAFTA vote. *Political Analysis, 14*(4), 421-438.

Boehmke, F. J., Morey, D. S., & Shannon, M. (2006). Selection bias and continuous-time duration models: consequences and a proposed solution. *American Journal of Political Science, 50*(1), 192-207.

Box-Steffensmeier, J. M., & Boef, S. D. (2006). Repeated events survival models: the conditional frailty model. *Statistics in Medicine, 25*(20), 3518-3533.

Chen, C. F., & Rothschild, R. (2010). An application of hedonic pricing analysis to the case of hotel rooms in Taipei. *Tourism Economics, 16*(3), 685-694.

Collier, W. (2005). Unemployment duration and individual heterogeneity: a regional study. *Applied Economics, 37*(2), 133-153.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society (Series B), 43*(2), 187-220.

Espinet, J. M., Saez, M., Coenders, G., & Fluvià, M. (2003). Effect on prices of the attributes of holiday hotels: a hedonic prices approach. *Tourism Economics, 9*(2), 165-177.

Fleischer, A., & Pizam, A. (2002). Tourism constraints among Israeli seniors. *Annals of Tourism Research, 29*(1), 106-123.

Fleischer, A., & Rivlin Byk, J. (2009). Quality, quantity and duration decisions in household demand for vacations. *Tourism Economics, 15*(3), 513-530.

Gokovali, U., Bahar, O., & Kozak, M. (2007). Determinants of length of stay: a practical use of survival analysis. *Tourism Management, 28*(3), 736-746.

Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika, 81*, 515-526.

Heckman, J., & Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica, 52*(2), 271-320.

Hellström, J. (2006). A bivariate count data model for household tourism demand. *Journal of Applied Econometrics, 21*(2), 213-226.

Hougaard, P. (1984). Life table methods for heterogeneous populations: distributions describing the heterogeneity. *Biometrika, 71*(1), 75-83.

Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika, 73*(3), 671-678.

Hougaard, P. (1999). Fundamentals of survival data. *Biometrics, 55*(1), 13-22.

Insightful Corporation. (2007). S-Plus (Version 8.0) [Computer software]. Seattle (USA): Insightful Corporation.

Jones, M. K. (2011). Disability, employment and earnings: an examination of heterogeneity. *Applied Economics, 43*(7-9), 1001-1017.

Juaneda, C., Raya, J. M., & Sastre, F. (2011). Pricing the time and location of a stay at a hotel or apartment. *Tourism Economics, 17*(2), 321-338.

Kau, J. B., Keenan, D. C., & Li, X. (2011). An analysis of mortgage termination risks: a shared frailty approach with MSA-level random effects. *Journal of Real Estate Finance and Economics, 42*(1), 51-67.

Kelly, P. J. (2004). A review of software packages for analyzing correlated survival data. *The American Statistician, 58*(4), 337-342.

Kiefer, N. M. (1988). Economic duration data and hazard functions. *Journal of Economic Literature, 26*(2), 646-679.

Lancaster, T. (1979). Econometric methods for the duration of unemployment. *Econometrica, 47*(4), 939-956.

Lancaster, T. (1990). *The analysis of transition data*. New York: Cambridge University Press.

Li, G., Song, H., & Witt, S. F. (2005). Recent developments in econometric modeling and forecasting. *Journal of Travel Research, 44*(1), 82-99.

Lim, C. (1997). Review of international tourism demand models. *Annals of Tourism Research, 24*(4), 835-849.

Lim, C. (1999). A meta-analytic review of international tourism demand. *Journal of Travel Research, 37*(3), 273-284.

Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2007). WinBUGS (Version 1.4.3) [Computer software].

Machado, L. P. (2010). Does destination image influence the length of stay in a tourism destination? *Tourism Economics, 16*(2), 443-456.

Mak, J., Moncur, J., & Yonamine, D. (1977). Determinants of visitor expenditures and visitor lengths of stay: a cross-section analysis of U.S. visitors to Hawaii. *Journal of Travel Research, 15*(3), 5-8.

Mak, J., & Nishimura, E. (1979). The economics of a hotel room tax. *Journal of Travel Research, 17*(4), 2-6.

Martínez-Garcia, E., & Raya, J. M. (2008). Length of stay for low-cost tourism. *Tourism Management, 29*(6), 1064-1075.

McKercher, B., Chan, A., & Lam, C. (2008). The impact of distance on international tourist movements. *Journal of Travel Research, 47*(2), 208-224.

McKercher, B., & Lew, A. A. (2003). Distance decay and the impact of effective tourism exclusion zones on international travel flows. *Journal of Travel Research, 42*(2), 159-165.

Menezes, A. G. d., Moniz, A., & Vieira, J. C. (2008). The determinants of length of stay of tourists in the Azores. *Tourism Economics, 14*(1), 205-222.

Nicolau, J. L. (2008). Characterizing tourist sensitivity to distance. *Journal of Travel Research, 47*(1), 43-52.

Nicolau, J. L. (2010). Variety-seeking and inertial behaviour: the disutility of distance. *Tourism Economics*, 16(1), 251-264.

Nicolau, J. L., & Mas, F. J. (2009). Simultaneous analysis of whether and how long to go on holidays. *Service Industries Journal*, 29(8), 1077-1092.

Paul, B. K., & Rimmawi, H. S. (1992). Tourism in Saudi Arabia: Asir National Park. *Annals of Tourism Research, 19*(3), 501-515.

Peypoch, N., Randriamboarison, R., Rasoamananjara, F., & Solonandrasana, B. (2012). The length of stay of tourists in Madagascar. *Tourism Management, 33*(5), 1230-1235.

R Development Core Team. (2011). R: A language and environment for statistical computing, reference index (Version 2.14.0) [Computer software]. Vienna: R Foundation for Statistical Computing.

Rasbash, J., Browne, W., Healy, M., Cameron, B., & Charlton, C. (2011). MLwiN (Version 2.24) [Computer software]. Bristol (UK): Centre for Multilevel Modelling, University of Bristol.

Raya-Vilchez, J. M., & Martínez-Garcia, E. (2011). Nationality and low-cost trip duration. a microeconometric analysis. *Journal of Air Transport Management, 17*(3), 168-174.

Santos, G. E. d. O., Ramos, V., & Rey-Maquieira, J. (2011). A microeconomic model of multidestination tourism trips. *Tourism Economics, 17*(3), 509-529.

Schoenfeld, D. A. (1982). Partial residuals for the proportional hazards regression. *Biometrika, 69*(1), 239-241.

Silberman, J. (1985). A demand function for length of stay: the evidence for Virginia Beach. *Journal of Travel Research, 23*(4), 16-23.

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting: a review of recent research. *Tourism Management, 29*(2), 203-220.

StataCorp LP. (2010). Stata/SE (Version 11.1 for Windows) [Computer software]. College Station, USA: StataCorp LP.

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., & Farstad, E. (2011). Domestic tourism expenditures: the non-linear effects of length of stay and travel party size. *Tourism Management, 32*(1), 46-52.

United Nations. (2010). *International recommendations for tourism statistics 2008*. Retrieved from http://unstats.un.org/unsd/publication/Seriesm/SeriesM_83rev1e.pdf

Uysal, M., McDonald, C. D., & O'Leary, J. T. (1988). Length of stay: a macro analysis for cross-country skiing trips. *Journal of Travel Research, 26*(3), 29-31.

Vaupel, J. W., Manton, K. G., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography, 16*(3), 439-454.

Walsh, R. G., & Davitt, G. J. (1983). A demand function for length of stay on ski trips to Aspen. *Journal of Travel Research, 21*(4), 23-29.

Wang, Y., & Davidson, M. C. G. (2010). A review of micro-analyses of tourist expenditure. *Current Issues in Tourism, 13*(6), 507-524.

Wedel, M., Desarbo, W. S., Bult, J. R., & Ramaswamy, V. (1993). A latent class poisson regression model for heterogeneous count data. *Journal of Applied Econometrics, 8*(4), 397-411.

Whitmore, G. A., & Lee, M.-L. T. (1991). A multivariate survival distribution generated by an inverse Gaussian mixture of exponentials. *Technometrics, 33*(1), 39-50.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

World Tourism Organization. (2006). *Tourism market trends: world overview & tourism topics*. Madrid: World Tourism Organization.

World Tourism Organization. (2007). *Tourism: 2020 vision: global forecast*. Madrid: UNWTO.

Yang, Y., Wong, K. K. F., & Zhang, J. (2011). Determinants of length of stay for domestic tourists: case study of Yixing. *Asia Pacific Journal of Tourism Research, 16*(6), 619-633.

5 Conclusion

This thesis focused on the demand for multidestination tourism trips. The relevance of this work is justified by the large incidence of MTTs and the scarcity of research about this topic.

Through developing the academic knowledge about MTTs consumption, this thesis attempts to improve the theoretical consistency and accuracy of tourism demand models, going beyond the single destination paradigm usually adopted. The development of research in this area may help tourism firms and destinations to develop more efficient strategies to achieve their objectives. Among organizations that may find applicable insights in this research it is possible to highlight country and destination management organizations, transport companies, tour operators and car rental agencies. It may also benefit firms that operate within the boundaries of the destination, such as attractions, hotels and ground operators.

Empirical studies of this thesis focused on inbound tourism in Brazil. Therefore, the set of estimations can be especially useful for Brazilian tourism organizations. Results might aid the development of tourism in this country, which still has a secondary position in the world tourism scenario, despite of its significant potential as a destination country. In fact, the lack of reliable research about tourism in Brazil can be pointed out as one of the main barriers to the professional development of this sector in the country. Thus, this thesis also aims to contribute to the expansion and improvement of the tourism sector in Brazil.

Four chapters were presented before this conclusion. Chapter 1 introduced the object of study and provided a general literature review to establish the research context. It also delivered some basic information about tourism in Brazil and the source of data used at chapters 3 and 4. It is worthy to note that chapters 2, 3 and 4 were developed as independent articles. Thus, those chapters presented a complete structure, including conclusions. Therefore, the present chapter mainly reports previously debated conclusions.

The usual theoretical paradigm of single destination trips is not sufficient for understanding and explaining MTTs. Tourist choices within the MTTs paradigm are much more complex than in the single destination context. All previous theoretical explanations of MTTs consumption were based in Lancaster's characteristics theory. However, this approach was also insufficient to provide a complete explanation, besides being unnecessarily complex. In order to elucidate the choice structure of MTTs, a theoretical model of the demand based on the traditional economic consumer theory was presented at Chapter 2. The model disclosed the high level of complexity subjacent to consumers' choice in this context, which arises both from the considerable number of variables involved and the discontinuous nature of some of these variables. In particular, indifference curves, monetary and time budget restrictions are discontinuous in the MTTs context. Deductions showed that small generalized costs of transport favour MTTs instead of single destination trips. The effect of the cost of stay over demand functions is quite complex. Large costs of stay may lead to either single or multiple destination trips. The effect of the budget allocated to tourism consumption over the demand functions is highly indefinite. Interestingly, negative income effects may be experienced by destinations considered as normal goods, and positive price effects may be experienced by non-Giffen destinations.

Chapter 3 modelled the number of destinations visited during a trip. Previous studies in this area adopted either limited statistical techniques or minimally informative dependent variables, both cases providing restricted information. In order to overcome these limitations and provide a more informative study, a censored zero-inflated negative binomial model was applied to data about the number of destinations visited by inbound tourists in Brazil. Zero inflation was found to be significant, indicating that there is a qualitative difference between single and multiple destination trips. Thus, the choice of adding a second destination to the itinerary also represents the choice of taking a qualitatively different sort of trip. As compared to previous empirical studies, the set of explanatory variables was substantially extended to include additional determinants, such as education level, type of accommodation and season. MTTs were found to be associated with male, younger or older tourists with higher level of education. The effects of income and time budgets over MTTs consumption are positive but marginally decreasing for the majority. These results enlighten some conflicting arguments provided by previous studies about the effect of income and time availability, besides showing that the effects of both resources are analogous. The study also showed that MTTs are associated with long haul trips, leisure purposes, independent trips, air transport, hotels, large travel parties, first-timers and winter season.

Chapter 4 considered the problem of modelling tourists' length of stay at multiple destinations of a trip. Duration models were selected as the most appropriate method for modelling inbound tourists' length of stay at Brazilian destinations. The contributions of the study are both empirical and methodological. Differently from previous studies, the determinants of tourists' length of stay are analysed in the MTTs' context. This new perspective stresses existent differences between the length of stay at the destination and in the destination region. In fact, both measures may vary differently due to changes correlated with the number of destinations visited. Moreover, the effects of several explanatory variables were examined, including the effects of some variables that had not been studied in the academic literature before. Particularly interesting results in this sense were obtained with regard to income, party size and the number of previous visits. Income was not found to be a significant determinant of tourists' length of stay within the MTTs paradigm. The length of stay was found to follow a convex function of party size. Moreover, the effect of a previous visit to the country was shown to have a positive effect on the expected length of stay, while further previous visits have a negative impact on it. The empirical study also showed that the increasing trend of international tourists' lengths of stay at Brazilian destinations was caused by changes in the composition of the inbound tourism flow. The analysis also pointed out destinations that might be regarded as benchmarks for tourism management due to their ability to host tourists for longer periods. In the theoretical side, it was found that the tourists' length of stay density distribution follows a positively skewed function, what can be explained by the opposing effects of initial costs of the visitation and decreasing marginal utility of the stay. Finally, it was shown that shared heterogeneity must be included in the duration model when the lengths of stay at multiple destinations of a trip are studied.

There are some important cross relations among chapters 2, 3 and 4. The selection of the set of explanatory variables used in the empirical study presented at Chapter 3 was partially sustained by the theoretical propositions of Chapter 2. Tourists' length of stay is the main variable of the theoretical model presented at Chapter 2 as well as of the study presented at Chapter 4. The empirical studies conducted at chapters 3 and 4 are also closely related. Tourists' total length of stay in Brazil was used as an explanatory variable at the model of the number of destinations visited presented at Chapter 3. On the other hand, the number of destinations visited was used as an explanatory variable at the model of tourists' length of stay presented at Chapter 4. If information about exclusive determinants of these two dependent variables were available, it would be quite interesting to develop a simultaneous equations model. Unfortunately, data available in the BITS is not sufficient to estimate such a model.

There are many opportunities for future research on MTTs. The set of determinants of the number of destinations visited and the length of stay at multiple destinations of a trip can be considerably extended. In particular, the analysis of some deeper consumer behavioural variables, such as attitudes and psychographic profiles, are likely to provide interesting insights. The study of other dependent variables related to MTTs may also be an interesting opportunity. Notably relevant analysis may come from destination choice studies in the context of MTTs. This perspective may also provide relevant contributions by explaining correlation across destination choices. Extensive data about destinations' attributes are required for such a study. An extent of that would be the analysis of substitution and complementarity patterns across multiple destinations of trip.

In general, MTTs was shown to be a highly complex issue. Tourists' choices in this context are non-linearly influenced by many variables. The study of MTTs requires detailed theoretical explanations and sophisticated statistical techniques. The complexity of the theoretical deductions of Chapter 2 and of the empirical findings of chapters 3 and 4 showed that common sense in the MTTs context may not always lead to the correct analysis. Moreover, elementary statistical techniques may not provide precise estimates of causal relationships. As a result, the hazard for tourism organizations of setting inefficient policies from imprecise information provided by inconsistent analysis is permanent in this field. Therefore, the most important recommendation of this thesis regards the necessity of recognizing the complexity of MTTs, what takes to the need of using proper methods for analysing tourism demand in this context.

Appendix

Appendix 1: Number of destinations visited by other relevant variables

Variable	Number of destinations visited (%)						Average number of destinations visited
	1	2	3	4	5	6	
Gender							
Female	59,1	21,6	9,4	5,1	2,9	1,9	1,8
Male	61,4	20,4	9,2	4,9	2,7	1,5	1,7
Age							
18 to 32 years	54,8	21,4	10,9	6,3	4,0	2,6	1,9
33 to 45 years	63,7	20,6	8,5	4,0	2,1	1,1	1,6
More than 45 years	63,0	20,5	8,5	4,3	2,4	1,3	1,7
Education level							
Under high school	68,7	16,8	7,6	3,8	1,8	1,2	1,6
High school	61,9	19,7	9,2	4,7	2,8	1,7	1,7
Graduation	60,1	21,3	9,4	4,7	2,8	1,8	1,7
Post-graduation	55,7	23,5	10,2	5,4	2,9	2,3	1,8
Household yearly income							
Up to US$ 1800	69,9	17,1	6,7	3,3	2,0	1,0	1,5
US$ 1801 to US$ 4500	58,0	21,9	9,8	5,5	3,1	1,8	1,8
US$ 4501 or more	54,5	23,6	11,0	5,8	3,1	1,9	1,9
Overnights							
Up to 6 days	81,9	14,6	2,6	0,6	0,2	0,1	1,2
7 to 14 days	59,6	24,5	9,8	4,0	1,7	0,5	1,6
More than 14 days	42,2	22,5	14,8	9,6	6,4	4,5	2,3
Origin							
Africa							
All countries	60,0	24,1	8,9	4,1	2,0	1,0	1,7
Asia and Oceania							
All countries	44,5	24,6	14,8	8,1	5,1	2,8	2,1
Central America							
All countries	61,8	21,9	6,9	6,1	2,8	0,4	1,7
Europe							

Variable	Number of destinations visited (%)						Average number of destinations visited
	1	2	3	4	5	6	
England	43,6	24,6	14,8	8,5	5,6	2,9	2,2
France	42,8	22,1	14,6	9,4	6,4	4,6	2,3
Germany	46,2	22,9	13,0	8,3	5,6	4,1	2,2
Italy	51,7	23,0	12,2	6,7	3,9	2,4	2,0
Portugal	55,9	23,6	10,8	5,5	2,8	1,4	1,8
Spain	52,0	22,6	12,3	6,6	4,1	2,4	2,0
Other Europe	45,8	22,3	13,7	8,6	5,6	4,1	2,2
North America							
USA	54,4	25,3	11,6	5,2	2,2	1,4	1,8
Other North America	52,4	25,3	11,6	5,4	3,2	2,1	1,9
South America							
Argentina	78,2	15,6	3,9	1,4	0,6	0,2	1,3
Chile	66,6	22,1	5,9	2,7	1,9	0,8	1,5
Paraguay	84,8	12,4	2,1	0,6	0,1	0,1	1,2
Uruguay	82,7	11,9	2,9	1,6	0,8	0,2	1,3
Other South America	66,8	20,7	7,0	3,6	1,5	0,5	1,5
Purpose							
Leisure purposes							
Sun and sea	65,3	20,4	7,5	3,6	2,1	1,2	1,6
Ecotourism	59,3	15,0	9,2	7,8	5,2	3,5	2,0
Cultural tourism	41,9	23,1	14,0	8,9	6,9	5,1	2,3
Other	60,7	20,7	8,3	4,8	3,7	1,8	1,8
Non-leisure purposes							
Business	68,2	20,1	6,9	2,9	1,4	0,6	1,5
VFR	52,5	23,8	12,8	6,3	3,0	1,7	1,9
Other	57,6	21,7	9,2	4,9	3,6	3,0	1,8
Transport mode							
Road	75,6	13,5	5,1	3,0	1,8	1,0	1,4
Air	55,9	23,0	10,6	5,4	3,1	1,9	1,8
Trip organization							
Independent	59,8	20,9	9,7	5,1	2,8	1,7	1,8
Package	66,8	19,8	5,9	3,6	2,5	1,3	1,6
Accommodation							
Hotels and counterparts*	62,9	20,1	8,1	4,4	2,7	1,7	1,7

Variable	Number of destinations visited (%)						Average number of destinations visited
	1	2	3	4	5	6	
Friends or relatives' dwellings	52,3	23,6	12,6	6,5	3,2	1,8	1,9
Rented dwelling	71,0	17,4	6,4	2,6	1,6	1,0	1,5
Owned dwelling	58,9	21,8	10,2	5,5	2,4	1,2	1,7
Other	54,6	19,4	11,3	6,0	4,9	3,9	2,0
Party size							
1 person	59,7	21,0	9,5	5,0	2,9	1,8	1,8
2 people	56,6	22,3	10,2	5,7	3,2	2,0	1,8
3 to 4 people	68,9	17,7	7,4	3,5	1,8	0,8	1,5
5 people or more	70,7	18,0	6,4	3,3	1,2	0,5	1,5
Party type							
Alone	60,4	21,3	9,6	4,6	2,2	1,9	1,7
Couple	54,9	22,5	10,6	5,7	3,6	2,8	1,9
Family	66,6	19,4	7,6	3,7	1,7	1,0	1,6
Friends	54,6	20,8	10,6	6,3	4,3	3,4	2,0
Other	66,5	21,0	7,4	2,9	1,3	0,8	1,5
Previous visits							
None	53,6	21,6	11,0	6,5	4,4	2,9	2,0
1 to 5	62,7	20,6	8,7	4,4	2,3	1,3	1,7
More than 5	65,6	20,1	8,1	3,8	1,6	0,7	1,6
Season							
Off-season*	59,3	22,0	9,3	4,8	2,9	1,7	1,8
Summer season	63,6	19,8	8,6	4,3	2,3	1,4	1,7
Winter season	61,6	20,9	8,9	4,5	2,6	1,6	1,7
Year							
2004	59,1	21,3	9,8	4,4	3,3	2,0	1,8
2005	62,1	20,4	8,6	5,2	2,6	1,1	1,7
2006	60,5	19,7	9,4	6,0	3,9	0,4	1,7
2007	59,3	21,0	9,4	5,3	3,3	1,8	1,8
2008	59,7	20,9	9,8	4,8	2,4	2,4	1,8
2009	60,4	21,5	9,2	4,5	2,4	2,0	1,7
2010	61,8	20,8	8,9	4,4	2,3	1,8	1,7

Appendix 2: BITS' questionnaire

INFRAERO Aeroportos Brasileiros | fipe | Ministério do Turismo | BRASIL UM PAÍS DE TODOS GOVERNO FEDERAL

RECEPTIVO AÉREO (Inglês)

NÃO É NECESSÁRIO IDENTIFICAR O ENTREVISTADO

CIDADE: ______________________ CÓDIGO DE QUESTIONÁRIO: ______________

ENTREVISTADOR: ______________________ DATA: ____/____/____

ENTREVISTAR APENAS MAIORES DE 18 ANOS E RESIDENTES NO EXTERIOR

1 What's the number of your flight to leave Brazil?

Airline: ______________ Flight: ______________

2 Where is your permanent residence?

City: ______________ Country: ______________

3 How old are you?

Age: __________

4 What was the MAIN reason of this trip to Brazil?

1. Leisure
2. Visit friends and relatives
3. Business or work
4. Congresses, fairs or conventions
5. Study purposes or courses
6. Health reasons
7. Religion or pilgrimage
8. Shopping
9. Others: ______________

5 (If LEISURE), what was the main kind of attraction?

1. Sun and beach
2. Nature or ecotourism
3. Culture
4. Sports
5. Incentive travel
6. Nightlife
7. Others: ______________

6 How many nights have you spent in Brazil for this trip?

Overnights: __________

7 Which cities have you visited while in Brazil for this trip and how many nights have you spent in each one of them?

	City	UF	Overnights
1.	________	____	____
2.	________	____	____
3.	________	____	____
4.	________	____	____
5.	________	____	____
6.	________	____	____
7.	**EM TRÂNSITO**		____

8 What main kind of accommodation was used in the city you spent most time? *(Caso haja empate entre cidades, considerar aquela na qual houve maior gasto e identificá-la com um asterisco na questão anterior)*

1. Resort
2. Posada
3. Hotel or flat
4. Hostel or bed & breakfast
5. Camping Site
6. House of friends and relatives
7. Own house / apartment
8. Rented house
9. Others: ______________

9 Was any service bought on travel agency OUTSIDE BRAZIL?

1. Yes, package
2. Yes, separate services
3. No *(ir para a questão 11)*

10 (If you used travel agency), which services were bought at the travel agency OUTSIDE BRAZIL? *(Resposta múltipla)*

1. International transport
2. Inner transport while in Brazil
3. Accommodation
4. Food
5. Attractions and sightseeing

11 NOT INCLUDING YOUR INTERNATIONAL TRANSPORT TICKET, what was the total amount spent in this trip?

Não representa a totalidade dos gastos → **Inclui Passagem Intl? (S/N)**

1. Total ________ Currency ____ ☐ ____
2. In your country ________ Currency ____ ☐ ____
3. In Brasil ________ Currency ____ ☐ ____

12 INCLUDING YOURSELF, how many people account for those expenses IN BRAZIL? (Over 6 years old and residents outside Brazil)

People: __________

13 Specify the amount spent on each of the following items IN BRAZIL: *(A soma dos valores desta questão deve corresponder ao valor descrito no item 3 da questão 11)*

	Gasto Zero	Amount	Currency
1. Accommodation	☐	______	____
2. Food	☐	______	____
3. Inner transport	☐	______	____
4. Shopping	☐	______	____
5. Attractions and sightseeing	☐	______	____
6. Others: ______	☐	______	____

14 Have you bought any property in Brazil for LEISURE or VACATION during this trip?

1. No
2. Yes. What: ______________

City: ______________ UF: ______

Amount: ______________ Currency: ______

15 Who came with you in this trip?

1. No one
2. Couple with no children
3. Couple with children
4. Family group
5. Friends
6. Workmates
7. Others: ______________

What was the main source of information used by you to organise this trip? 16

1. Tourist guidebooks
2. Internet
3. Friends and relatives
4. Place where you work
5. Travel agency
6. Fairs, events and congresses
7. Already knew the destination
8. Folders and brochures
9. Articles in newspapers and magazines
10. TV shows and Radio
11. Ads, advertising campaign
12. Embassies and Consulates
13. Brazilian Tourism Bureaux
14. Others:___________________________

Did you seek information, reserve or buy any service by the Internet for this travel? *(Resposta múltipla)* 17

	Seek Information	Reserve	Buy	No
1. Tour package	☐	☐	☐	☐
2. Internacional transport	☐	☐	☐	☐
3. Accommodation	☐	☐	☐	☐
4. Car rental	☐	☐	☐	☐
5. Attractions and sightseeing	☐	☐	☐	☐
6. Others: ______________	☐	☐	☐	☐

Regarding the CITY you spent most time in Brazil, evaluate the following: 18

Cidade: _____________________ UF: _____ Pernoites: _____

Very Good (VG), Good (G), Bad (B), Very Bad (VB), Not Applicable (NA)

General Infrastructure	VG	G	B	VB	NA
1. Public Cleanliness	☐	☐	☐	☐	☐
2. Public Safety	☐	☐	☐	☐	☐
3. Taxi Service	☐	☐	☐	☐	☐
4. Public Transportation	☐	☐	☐	☐	☐
5. Telecommunication / Internet	☐	☐	☐	☐	☐
6. Tourist Signalization	☐	☐	☐	☐	☐
Tourist Infrastructure	VG	G	B	VB	NA
1. Restaurants	☐	☐	☐	☐	☐
2. Accommodation	☐	☐	☐	☐	☐
3. Nightlife	☐	☐	☐	☐	☐
Tourist Services	VG	G	B	VB	NA
1. Tourist guiding services	☐	☐	☐	☐	☐
2. Tourist information	☐	☐	☐	☐	☐
3. Cuisine	☐	☐	☐	☐	☐
4. Prices	☐	☐	☐	☐	☐

Regarding BRAZIL, evaluate the following: 19

	VG	G	B	VB	NA
1. Highways	☐	☐	☐	☐	☐
2. Hospitality of the people	☐	☐	☐	☐	☐
3. THIS airport	☐	☐	☐	☐	☐

During this trip, YOUR EXPECTATIONS regarding Brazil were: 20

1. Exceeded
2. Totally fulfilled
3. Partialy fulfilled
4. Not fulfilled or disappointed

Gênero: *(não perguntar)* 21

1. Masculino 2. Feminino

What is your level of formal education? 22

1. No formal education / Uncomplete Primary School
2. Elementary / Primary School
3. Middle School / Secondary School
4. High School
5. College *(perguntar se tem pós)*
6. Postgraduate

What is your monthly income? 23

Não tem

1. Personal:__________________ Currency:________ ☐

2. Family:____________________ Currency:________ ☐

INCLUDING YOURSELF, how many people depend on this family income? 24

People: _________________

Is this your first trip to Brazil? 25

1. Yes *(comparar com o item 7 da questão 16)*
2. No.
BEFORE THIS TRIP, how many times have you visited?_______

Do you intend to visit Brazil again? 26

1. Yes
2. No

Did you consider going to another country instead of Brazil for this trip? 27

1. No
2. Yes. Which one: 1.______________________________
2. ______________________________
3. ______________________________

If LEISURE, what was the main reason that made you choose Brazil as your destination for this trip? 28

Reason:_______________________________________

____

Have you ever seen this logo? 29
(Apresentar a Marca Brasil - resposta múltipla)

1. No
2. Yes. Where:
   1. Newspaper, magazine
   2. Brochure, folders
   3. Internet
   4. TV
   5. Billboard
   6. Other: ____________________

Observações:

Appendix 3: Estimated parameters of tourists' length of stay for states

State	Model 1 (2004-2010)			Model 2 (2010)		
	Coefficient	Std. error	p-value	Coefficient	Std. error	p-value
Alagoas*	0.00			0.00		
Amapá	-0.024	0.22	0.91	-1.4	0.69	0.05
Amazonas	0.51	0.15	0.00	-0.21	0.51	0.68
Bahia	0.18	0.064	0.01	0.20	0.18	0.26
Ceará	-0.070	0.067	0.30	0.10	0.19	0.59
Goiás	1.2	0.75	0.12	2.4	0.38	0.00
Maranhão	0.13	0.076	0.08	0.46	0.22	0.04
Mato Grosso	1.1	0.75	0.13	2.3	0.39	0.00
Mato Grosso do Sul	1.1	0.75	0.16	2.3	0.40	0.00
Minas Gerais	-0.074	0.034	0.03	-0.074	0.082	0.37
Pará	0.61	0.15	0.00	0.26	0.51	0.62
Paraíba	-0.090	0.083	0.28	0.17	0.27	0.53
Pernambuco	0.047	0.069	0.50	-0.23	0.20	0.24
Piauí	0.22	0.10	0.03	0.036	0.25	0.89
Rio de Janeiro	-0.20	0.042	0.00	-0.11	0.10	0.25
Rio Grande do Norte	-0.15	0.072	0.03	0.21	0.20	0.29
Rio Grande do Sul	-0.34	0.027	0.00	-0.44	0.064	0.00
Rondônia	1.0	0.15	0.00	0.16	0.53	0.76
Roraima	0.78	0.23	0.00	-0.28	0.58	0.63
Santa Catarina	0.059	0.028	0.04	0.21	0.068	0.00
São Paulo	-0.10	0.033	0.00	-0.13	0.079	0.09
Sergipe	0.19	0.17	0.26	0.69	0.76	0.37
Tocantins	0.87	0.16	0.00	0.059	0.54	0.91

Appendix 4: Estimated parameters of tourists' length of stay for states

State	Model 1 (2004-2010)			Model 2 (2010)		
	Coefficient	Std. error	p-value	Coefficient	Std. error	p-value
Abadiânia - GO*	0.00					
Americana - SP	-0.42	0.10	-4.22	-0.11	0.10	0.25
Anápolis - GO	-0.29	0.086	-3.34	0.21	0.20	0.29
Angra dos Reis - RJ	-0.57	0.082	-6.92	0.16	0.53	0.76
Aquiraz - CE	-0.21	0.11	-1.99	-0.28	0.58	0.63
Aracaju - SE	-0.80	0.18	-4.52	-0.44	0.064	0.00
Aracati - CE	-0.56	0.085	-6.52	0.21	0.068	0.00
Araçatuba - SP	-0.47	0.10	-4.48	0.69	0.76	0.37
Araraquara - SP	-0.47	0.10	-4.54	-0.13	0.079	0.09
Armação dos Búzios - RJ	-0.36	0.082	-4.46	0.059	0.54	0.91
Arraial do Cabo - RJ	-0.62	0.093	-6.70	-0.45	0.24	0.06
Atibaia - SP	-0.61	0.10	-6.01	-0.38	0.21	0.07
Balneário Camboriú - SC	-0.33	0.079	-4.13	-0.70	0.20	0.00
Barreirinhas - MA	-0.79	0.10	-8.08	-0.47	0.38	0.22
Barueri - SP	-0.49	0.10	-5.02	-1.4	0.77	0.07
Bauru - SP	-0.64	0.092	-6.92	-0.81	0.23	0.00
Beberibe - CE	-0.47	0.094	-4.93	-0.045	0.29	0.88
Belém - PA	-1.3	0.11	-11.69	-0.64	0.25	0.01
Belo Horizonte - MG	-1.9	0.16	-12.34	-0.54	0.20	0.01
Bento Gonçalves - RS	-0.27	0.10	-2.81	-0.80	0.22	0.00
Bertioga - SP	-0.88	0.10	-8.93	-0.91	0.26	0.00
Blumenau - SC	-0.92	0.086	-10.65	-0.53	0.20	0.01
Boa Vista - RR	-1.3	0.21	-6.01	-1.1	0.26	0.00
Bombinhas - SC	-0.12	0.080	-1.52	-0.37	0.24	0.12
Bonito - MS	-0.77	0.092	-8.46	-0.94	0.23	0.00
Botucatu - SP	-0.43	0.11	-4.07	-0.65	0.29	0.02
Brasília - DF	-1.3	0.77	-1.71	-1.5	0.28	0.00
Cabo de Santo Agostinho - PE	-0.46	0.11	-4.20	-1.9	0.36	0.00
Cabo Frio - RJ	-0.79	0.086	-9.19	-0.26	0.24	0.27
Cairu - BA	-0.46	0.080	-5.78	-0.75	0.25	0.00
Caldas Novas - GO	-1.1	0.079	-14.16	-1.2	0.21	0.00
Camaçari - BA	-0.68	0.094	-7.30	-0.39	0.40	0.33
Campina Grande - PB	-0.50	0.12	-4.31	-0.26	0.20	0.18

State	Model 1 (2004-2010)			Model 2 (2010)		
	Coefficient	Std. error	p-value	Coefficient	Std. error	p-value
Campinas - SP	-1.2	0.10	-12.53	-0.90	0.23	0.00
Campo Grande - MS	-1.2	0.10	-11.95	-0.31	0.26	0.23
Campos do Jordão - SP	-1.1	0.090	-11.72	#VALOR!	0.00	0.00
Campos dos Goytacazes - RJ	-0.60	0.11	-5.39	-0.34	0.33	0.30
Canela - RS	-0.63	0.10	-6.19	-0.91	0.21	0.00
Canoas - RS	-0.46	0.11	-4.00	-0.46	0.20	0.02
Capão da Canoa - RS	-0.14	0.087	-1.63	-1.3	0.20	0.00
Caraguatatuba - SP	-0.68	0.10	-7.04	-0.69	0.23	0.00
Caruaru - PE	-0.83	0.11	-7.42	-1.3	0.33	0.00
Cascavel - PR	-1.1	0.086	-13.35	-1.3	0.24	0.00
Caucaia - CE	-0.36	0.090	-4.04	-1.2	0.25	0.00
Caxias do Sul - RS	-0.51	0.088	-5.77	-1.4	0.23	0.00
Chapecó - SC	-0.88	0.11	-8.42	-1.0	0.26	0.00
Chuí - RS	-1.2	0.10	-12.36	-0.45	0.25	0.07
Corumbá - MS	-1.2	0.10	-11.50	-0.93	0.28	0.00
Criciúma - SC	-0.14	0.091	-1.55	-0.022	0.21	0.92
Cuiabá - MT	-0.89	0.093	-9.54	-0.67	0.24	0.01
Curitiba - PR	-1.6	0.13	-12.79	-0.90	0.30	0.00
Diamantina - MG	-0.92	0.11	-8.67	-1.3	0.21	0.00
Dourados - MS	-0.74	0.11	-6.48	-0.44	0.23	0.05
Entre Rios - BA	-0.30	0.10	-2.91	-0.71	0.22	0.00
Feira de Santana - BA	-0.89	0.11	-8.16	-1.0	0.27	0.00
Fernando de Noronha - PE	-0.24	0.090	-2.69	-0.76	0.33	0.02
Florianópolis - SC	-0.49	0.082	-5.95	-1.5	0.26	0.00
Fortaleza - CE	-1.8	0.16	-10.86	-0.32	0.23	0.16
Foz do Iguaçu - PR	-1.2	0.079	-15.60	-0.92	0.22	0.00
Franca - SP	-0.46	0.11	-4.40	-1.6	0.30	0.00
Garopaba - SC	-0.20	0.081	-2.51	-1.0	0.26	0.00
Goiânia - GO	-1.1	0.10	-11.01	-0.65	0.32	0.04
Governador Valadares - MG	-0.30	0.084	-3.53	0.81	0.64	0.21
Gramado - RS	-0.61	0.084	-7.34	-0.44	0.28	0.12
Guarapari - ES	-0.69	0.093	-7.35	-0.076	0.24	0.75
Guarapuava - PR	-1.2	0.11	-10.70	-0.63	0.20	0.00
Guaratuba - PR	-0.28	0.090	-3.15	-1.8	0.38	0.00
Guarujá - SP	-0.93	0.083	-11.32	-1.2	0.20	0.00

State	Model 1 (2004-2010)			Model 2 (2010)		
	Coefficient	Std. error	p-value	Coefficient	Std. error	p-value
Guarulhos - SP	-1.7	0.11	-15.77	-0.30	0.25	0.23
Ijuí - RS	-1.4	0.10	-14.27	-0.41	0.20	0.04
Ilha Grande - PI	-0.62	0.12	-5.31	-1.2	0.25	0.00
Ilhabela - SP	-0.74	0.084	-8.85	-0.28	0.21	0.19
Ilhéus - BA	-0.62	0.085	-7.30	-0.58	0.21	0.00
Imbituba - SC	-0.22	0.088	-2.49	-0.69	0.23	0.00
Indaiatuba - SP	-0.43	0.10	-4.27	-1.2	0.24	0.00
Ipatinga - MG	-0.32	0.088	-3.64	-0.27	0.23	0.25
Ipojuca - PE	-0.41	0.084	-4.84	-1.0	0.21	0.00
Itabuna - BA	-0.52	0.11	-4.80	-1.7	0.26	0.00
Itacaré - BA	-0.37	0.084	-4.44	-1.7	0.21	0.00
Itajaí - SC	-0.60	0.10	-6.22	-0.43	0.48	0.36
Itaparica - BA	-0.58	0.090	-6.40	-0.75	0.21	0.00
Itapema - SC	-0.16	0.081	-1.92	-0.66	0.22	0.00
Itatiaia - RJ	-0.73	0.10	-6.98	-0.29	0.21	0.17
Itu - SP	-0.66	0.10	-6.55	-0.44	0.23	0.06
Jaraguá do Sul - SC	-0.37	0.10	-3.59	-0.43	0.22	0.05
Jijoca de Jericoacoara - CE	-0.25	0.083	-3.00	-0.17	0.22	0.45
João Pessoa - PB	-0.67	0.11	-6.36	-0.46	0.28	0.11
Joinville - SC	-0.81	0.089	-9.11	-0.46	0.21	0.03
Juiz de Fora - MG	-0.73	0.090	-8.10	-0.84	0.25	0.00
Jundiaí - SP	-0.67	0.088	-7.59	-0.54	0.24	0.03
Lages - SC	-1.6	0.10	-14.98	-0.36	0.20	0.07
Laguna - SC	-0.38	0.090	-4.28	-1.0	0.31	0.00
Lençóis - BA	-0.57	0.084	-6.73	-0.77	0.23	0.00
Limeira - SP	-0.65	0.10	-6.25	-0.59	0.24	0.01
Londrina - PR	-0.61	0.087	-7.00	-0.29	0.21	0.17
Macaé - RJ	-0.10	0.085	-1.23	-1.0	0.31	0.00
Macapá - AP	-0.32	0.19	-1.68	-0.83	0.22	0.00
Maceió - AL	-0.78	0.11	-6.99	-0.64	0.22	0.00
Manaus - AM	-1.6	0.14	-11.53	-0.89	0.21	0.00
Mangaratiba - RJ	-0.35	0.11	-3.26	-2.2	0.22	0.00
Maragogi - AL	-0.47	0.11	-4.33	-0.65	0.22	0.00
Maraú - BA	-0.39	0.11	-3.71	-0.73	0.21	0.00
Marechal Cândido Rondon - PR	-1.2	0.10	-12.83	-0.71	0.25	0.01

State	Model 1 (2004-2010)			Model 2 (2010)		
	Coefficient	Std. error	p-value	Coefficient	Std. error	p-value
Marília - SP	-0.50	0.10	-4.95	-0.71	0.22	0.00
Maringá - PR	-0.47	0.087	-5.47	-0.064	0.21	0.76
Mata de São João - BA	-0.47	0.080	-5.81	0.49	0.53	0.36
Matinhos - PR	-0.20	0.088	-2.29	-0.80	0.29	0.01
Medianeira - PR	-1.4	0.094	-15.43	-1.4	0.32	0.00
Missal - PR	-1.3	0.10	-13.08	-0.083	0.26	0.75
Mogi das Cruzes - SP	-0.44	0.10	-4.43	-0.32	0.28	0.26
Montes Claros - MG	-0.29	0.10	-2.74	-0.60	0.27	0.02
Mossoró - RN	-0.35	0.12	-2.98	-1.3	0.23	0.00
Natal - RN	-0.62	0.10	-6.46	-0.64	0.24	0.01
Niterói - RJ	-0.57	0.088	-6.47	-0.36	0.21	0.10
Nova Friburgo - RJ	-0.66	0.10	-6.76	-0.61	0.20	0.00
Novo Hamburgo - RS	-0.28	0.088	-3.16	-0.060	0.22	0.79
Olinda - PE	-0.89	0.089	-9.92	-1.6	0.24	0.00
Ouro Preto - MG	-1.0	0.079	-12.31	-1.6	0.25	0.00
Palhoça - SC	-0.41	0.10	-4.13	-0.31	0.27	0.25
Palmas - TO	-0.87	0.12	-7.38	-0.68	0.27	0.01
Paracuru - CE	-0.13	0.10	-1.24	-1.0	0.24	0.00
Paranaguá - PR	-0.66	0.086	-7.67	-0.74	0.21	0.00
Paraty - RJ	-0.75	0.082	-9.14	-0.58	0.24	0.02
Parnaíba - PI	-1.0	0.13	-7.95	-0.087	0.21	0.68
Passo Fundo - RS	-1.0	0.094	-11.04	-0.87	0.24	0.00
Paulista - PE	-0.43	0.11	-4.02	-1.1	0.20	0.00
Pelotas - RS	-0.79	0.10	-8.24	-0.63	0.19	0.00
Petrolina - PE	-0.32	0.11	-2.81	-0.63	0.23	0.01
Petrópolis - RJ	-1.1	0.087	-12.43	-0.69	0.30	0.02
Piracicaba - SP	-0.66	0.089	-7.42	-0.31	0.26	0.24
Pirenópolis - GO	-1.3	0.10	-12.31	-0.63	0.21	0.00
Poconé - MT	-0.50	0.092	-5.46	-0.92	0.20	0.00
Poços de Caldas - MG	-0.54	0.095	-5.69	-0.66	0.32	0.04
Ponta Grossa - PR	-0.71	0.10	-7.17	-1.5	0.21	0.00
Porto Alegre - RS	-1.2	0.11	-10.63	-0.29	0.34	0.39
Porto Belo - SC	-0.27	0.090	-2.99	-0.83	0.23	0.00
Porto Seguro - BA	-0.49	0.080	-6.20	-0.078	0.30	0.80
Porto Velho - RO	-1.0	0.12	-8.80	-1.3	0.21	0.00

State	Model 1 (2004-2010)			Model 2 (2010)		
	Coefficient	Std. error	p-value	Coefficient	Std. error	p-value
Praia Grande - SP	-0.87	0.10	-8.74	-0.84	0.22	0.00
Presidente Prudente - SP	-0.43	0.10	-4.29	-1.3	0.25	0.00
Recife - PE	-1.4	0.12	-11.81	-0.48	0.22	0.03
Resende - RJ	-0.63	0.10	-6.35	-0.64	0.24	0.01
Ribeirão Preto - SP	-0.84	0.087	-9.58	-0.19	0.25	0.44
Rio Branco - AC	-0.42	0.17	-2.40	-1.1	0.27	0.00
Rio Claro - SP	-0.49	0.11	-4.55	-0.44	0.26	0.09
Rio das Ostras - RJ	-0.43	0.11	-3.97	-0.54	0.20	0.01
Rio de Janeiro - RJ	-4.6	0.37	-12.46	-0.81	0.30	0.01
Rio Grande - RS	-0.45	0.094	-4.77	-0.70	0.24	0.00
Salvador - BA	-2.2	0.17	-12.91	-0.13	0.26	0.62
Sant'Ana do Livramento - RS	-1.3	0.11	-11.45	-1.0	0.29	0.00
Santa Cruz do Sul - RS	-0.046	0.11	-0.42	-0.61	0.24	0.01
Santa Helena - PR	-1.1	0.10	-11.31	-0.84	0.21	0.00
Santa Maria - RS	-0.30	0.095	-3.17	-0.89	0.56	0.11
Santa Terezinha de Itaipu - PR	-1.5	0.085	-17.79	-0.45	0.25	0.08
Santa Vitória do Palmar - RS	-0.67	0.091	-7.31	-0.78	0.25	0.00
Santarém - PA	-0.76	0.10	-7.84	-4.4	0.83	0.00
Santo André - SP	-0.73	0.10	-7.38	-0.52	0.23	0.02
Santos - SP	-1.0	0.083	-11.61	-2.2	0.40	0.00
São Bernardo do Campo - SP	-0.88	0.11	-8.29	-1.0	0.35	0.01
São Carlos - SP	-0.50	0.094	-5.36	-0.19	0.27	0.48
São Francisco do Sul - SC	-0.35	0.10	-3.52	-1.1	0.22	0.00
São Gabriel - RS	-1.4	0.087	-15.62	-0.25	0.22	0.24
São João del Rei - MG	-1.0	0.10	-9.56	-1.6	0.21	0.00
São José do Rio Preto - SP	-0.74	0.091	-8.12	-0.66	0.26	0.01
São José dos Campos - SP	-0.86	0.087	-9.94	-0.77	0.24	0.00
São Leopoldo - RS	-0.15	0.11	-1.42	-0.86	0.24	0.00
São Luís - MA	-1.0	0.11	-9.64	-1.1	0.21	0.00
São Miguel do Iguaçu - PR	-1.5	0.093	-15.84	-0.85	0.26	0.00
São Paulo - SP	-8.1	0.66	-12.28	-0.66	0.23	0.00
São Sebastião - SP	-0.81	0.083	-9.71	-0.37	0.24	0.12
São Vicente - SP	-0.65	0.11	-5.70	-1.5	0.21	0.00
Saquarema - RJ	-0.50	0.10	-4.88	-1.0	0.30	0.00
Sorocaba - SP	-0.79	0.089	-8.96	-0.77	0.22	0.00

State	Model 1 (2004-2010)			Model 2 (2010)		
	Coefficient	Std. error	p-value	Coefficient	Std. error	p-value
Tamandaré - PE	-0.47	0.12	-4.06	-1.0	0.22	0.00
Taubaté - SP	-0.57	0.10	-5.86	-0.31	0.24	0.19
Teófilo Otoni - MG	-0.34	0.11	-3.17	-1.4	0.27	0.00
Teresina - PI	-0.79	0.13	-6.14	-1.5	0.22	0.00
Teresópolis - RJ	-0.83	0.094	-8.85	-7.5	1.5	0.00
Tibau do Sul - RN	-0.26	0.086	-2.98	-0.81	0.21	0.00
Tiradentes - MG	-1.2	0.087	-14.24	-0.66	0.27	0.02
Toledo - PR	-1.1	0.093	-12.35	-0.79	0.24	0.00
Torres - RS	-0.20	0.081	-2.50	-0.85	0.22	0.00
Tramandaí - RS	-0.24	0.11	-2.23	-0.29	0.32	0.37
Ubatuba - SP	-0.70	0.082	-8.49	-0.47	0.24	0.05
Uberaba - MG	-0.64	0.10	-6.39	-0.35	0.29	0.24
Uberlândia - MG	-0.64	0.090	-7.05	-0.65	0.32	0.04
Uruguaiana - RS	-1.0	0.11	-9.10	-1.2	0.24	0.00
Vila Velha - ES	-0.49	0.10	-5.05	-0.59	0.22	0.01
Vitória - ES	-0.70	0.084	-8.31	-1.5	0.22	0.00
Volta Redonda - RJ	-0.25	0.10	-2.37	-1.3	0.24	0.00
Other destinations	-0.51	0.075	-6.80	-0.54	0.24	0.03