

Influence of Tourism Road Transport on the Generation of Externalities:
A Case Study for the Balearic Islands

Doctoral thesis/tesis doctoral

Author: Oscar Saenz de Miera Berglind

Supervisor: Dr. Jaume Rosselló i Nadal

Doctor in Tourism and Environmental Economics

University of the Balearic Islands

2011

Acknowledgements

Firstly I would like to thank the University of the Balearic Islands (UIB), my thesis advisor, Jaume Rosselló, and all those members of the Department of Applied Economics, whose help and support were crucial for my doctoral studies. I also want to gratefully acknowledge the financial support that I received from the University of the Balearic Islands and “Cátedra Sol Meliá de Estudios Turísticos”.

I gratefully appreciate the help provided by the *Centre de Recerca Econòmica*, Agustí Jansá (Agencia Estatal de Meteorología) and the members of the *Grupo de Meteorología* of the Physics Department of the UIB, in the estimation and/or provision of the databases. I would also like to thank the valuable and enriching comments and suggestions from Prof. Tom Brijs (Transportation Research Institute, Belgium), Daniel Eisenberg (Health Management and Policy, School of Public Health, University of Michigan, USA) and Lasse Fridström (Institute of Transport Economics, Norway), for the chapter on road accidents; as well as those from Angel Bujosa, Antoni Riera, William Nilsson and Joaquín Alegre for the chapter on traffic congestion.

I also want to thank my wife, Laura, my parents and friends and Amaya Repes for their support throughout this period.

Synthetic Index

I. Introduction.....	1
II. Empirical Chapters.....	23
1. Road accidents and tourism: The case of the Balearic Islands (Spain).....	25
2. The responsibility of tourism in traffic congestion and hyper-congestion: a case study from Mallorca, Spain.....	45
3. Influence of tourism on air pollution: the case study of tropospheric ozone in Mallorca (Spain).....	77
III. Conclusions.....	99

Table of Contents

I. Introduction.....	1
1. Motivation and General Aim of the Thesis.....	3
2. Road transport externalities and tourism.....	7
3. Daily Indicator of Tourism Pressure.....	10
3.1. Obtaining the Indicator.....	11
3.2. Indicator of population pressure with data from the Balearic Islands.....	12
3. References.....	16
II. Empirical Chapters.....	23
Chapter 1. Road accidents and tourism: The case of the Balearic Islands (Spain).....	25
1.1. Introduction.....	25
1.2. Literature Review.....	27
1.3. Methods.....	29
1.4. Data description.....	33
1.5. Empirical Analysis.....	34
1.6. Discussion and Conclusions.....	39
1.7. References.....	40

Chapter 2. The responsibility of tourism in traffic congestion and hyper-congestion: a case study from Mallorca, Spain.....	45
2.1. Introduction.....	45
2.2. Background: The determinants of traffic volume and traffic speed.....	47
2.3. Methodology and data.....	51
2.3.1. Indicators of congestion.....	51
2.3.2. Method.....	55
2.3.3. Data.....	58
2.4. Empirical Analysis.....	63
2.5. Discussion and conclusions.....	69
2.6. References.....	71
 Chapter 3. Influence of tourism on air pollution: the case study of tropospheric ozone in Mallorca (Spain).....	77
3.1. Introduction.....	77
3.2. Air pollution and tourism.....	79
3.3. Methodology.....	81
3.4. Empirical analysis.....	86
3.5. Conclusions.....	93
3.6. References.....	95
 III. Conclusions.....	99

I. Introduction

1. Motivation and General Aim of the Thesis

Throughout the last decades, awareness of the growing trends shown by the main indicators of tourism has resulted in a vast amount of studies addressing the impacts of such activity. Despite being a key task for public administrations and decision makers in many countries, the assessment of these impacts has faced problems that stem mainly from the lack of an accepted definition of a specific “tourism industry”. According to Jones and Munday (2004), the needs of tourists are not met by discrete sectors such as hotels, restaurants and leisure facilities alone, as visitors may use only a proportion of, for instance, postal, health, and sewerage services. Moreover, the authors suggest that the tourism industry does not have “natural borders” either and this hinders reconciliation with national or regional aggregates.

Faced with these problems, the specialists proposed to implement the Tourism Satellite Accounts (TSAs), which measure the size of tourism in an economy, in a manner which is consistent with the National System of Statistical Accounts (Dwyer et al., 2004; Ivanov and Webster, 2007). In spite of this, while the TSAs estimate the importance of tourism through different variables, their results are limited in terms of assessing the total impact of tourism, mainly because (i) they do not consider indirect effects caused by economic interlinkages like, for example, the purchasing links between hotels and restaurants and other firms; and (ii) tourism is usually measured from the demand-side and it is difficult to compare results with other sectors that are usually measured from the supply-side (Blake et al., 2001; Ivanov and Webster, 2007; Smeral, 2006). In order to overcome this issue, the TSAs are often “complemented” with Input-Output and multiplier based models which in turn, are gradually being replaced by General Equilibrium Models (GEMs), a technique that is recognized to outperform the former (Dwyer et al., 2004; Blake et al., 2001).

INTRODUCTION

Nevertheless, there is increasing recognition of other types of social and environmental impacts of tourism (Dwyer et al., 2004; Sheng and Tsui, 2009), that have been usually left out in the aforementioned analyses, in spite of the growing concern shown by residents of highly specialized tourism destinations (Aguiló and Rosselló, 2005). These impacts include the negative externalities created by the damage to fragile landscapes, crowding out of local populations, pollution, and erosion of socio-cultural assets (Caffyn and Lutz, 1999). In this context, one of the tourist activities that has stood out in the latest years for being an important source of externalities is the transportation of passengers within destinations. The importance of addressing issues related to tourist mobility at the destination stems not only from the increasing global trends in tourist numbers, but also from certain characteristics that nowadays define the tourist experience for a significant proportion of travelers. This refers, to the increasing emphasis on individual, non-package holidays and the preference of many tourists for a higher number of shorter breaks to short-distance, short-haul destinations which is favored by the continued growth of low-cost airlines, and leads to increased mobility (Palmer et al., 2007). In addition, an in depth analysis of the different types of tourists at the Balearics by Aguiló et al. (2010) reveals that tourists who arrive through a low cost carrier tend to hire a car instead of using the public bus; visitors classified under the high-income category are more likely to use private cars; and those visitors qualified as repeating tourists often opt for the use of private cars as well. Further, within the typology of tourists, these two last groups take a crucial role in mature destinations that seek to attract new high-income tourists while maintaining recurrent visitors.

In response to the need to value tourism associated externalities, various perspectives have emerged, including the identification of destinations' lifecycle (Hernández and León, 2007; Becken and Simmons, 2008), ecological footprint analysis (Gössling et al., 2002; Patterson et al. 2007; Rendeiro and Ramírez, 2010) or economic valuation techniques (Smith and Huang, 1995; Wardman and Bristow, 2004). Therefore, although tourism's contribution to social and environmental costs

has been acknowledged (Gössling, 2002), only recently have researchers started to analyze the direct environmental impact that certain leisure activities entail. For instance, regarding the emission of greenhouse gases, it has been argued that tourism is one of the main activities in terms of energy consumption because of the need to transport passengers and provide services at the destination (Becken and Simmons, 2002, Becken et al., 2001, 2003; Gössling, 2000; Tabatchnaia-Tamirisa et al., 1997). But it is important to point out that this sort of conclusions have been drawn mostly from a partial perspective through the analysis of some of the main economic sectors that are related to tourism (mainly transport and accommodation). Thus, the need to assign the tourism sector as a whole its responsibility in the generation of externalities becomes evident especially through most policies of development and tourist promotion; and the consensus shown by residents that tourism leads to the deterioration or destruction of certain resources (Bujosa and Rosselló, 2007).

Taking these considerations into account, the main objective of this thesis is to assess the impact of tourism on three different road transport externalities (accidents, congestion and air pollution). Such task is undertaken through the possibility to link a measure of tourist population pressure with the generation of external costs, which need to be attributed to a given economic sector. For this purpose, the Balearic Islands are taken as case a study in the three empirical chapters. This Spanish region constitutes an adequate example of the type of destinations that receive significant pressure from the massive influx of tourists, especially during the summer months. The Balearic Archipelago is located in the western Mediterranean, a region that includes some of the main destinations that keep Europe on top of the ranking of international tourist arrivals. Furthermore, not only are the Balearics among the most consolidated and representative sun-and-sand destinations that characterize the Spanish tourism product, but also they represent 1% of international tourism and 3% of international tourist arrivals to the Mediterranean.

The thesis is structured as follows: the remainder of the Introduction is divided in two additional Sections: 1) provides further detail about road transport

INTRODUCTION

externalities and their relevance and 2) presents a discussion on the elaboration of the measure of population pressure and the advantages it offers as an indicator of the influence of tourism. Part II is divided into three empirical chapters, each one modeling the influence of tourism (and the variables suggested by the literature) on a specific externality, adapting to case-specific methodological requirements. Chapter 1 evaluates the influence of tourism on daily road traffic accidents. Considering that data on accidents are discrete and non-negative, the application is carried out using the Negative Binomial estimation technique, that is suitable for modeling count data and offers the advantage of dealing with over dispersion, an issue commonly found in applications of this kind. Chapter 2 addresses the issue of traffic congestion, one of the most important road transport externalities but also one of the most complex in terms of definition and measurement. For this reason, the study devotes special attention to creating a measure that encompasses key variables included in the definition of congestion, as opposed to using proxies that can only reflect certain features of traffic activity. Once the measure of congestion is obtained, it is explained through a series of variables including tourism. In Chapter 3 the focus is on air pollution from mobile sources, focusing on the use of road vehicles by tourists. Such analysis is the product of previous applications where both the literature review and the significance of non-linear (i.e. quadratic) forms of the explanatory variables evidenced the weakness of linear estimation techniques for this kind of models. Therefore, a Generalized Additive Model is used for it is a more suitable technique to account for the fact that the relationships among variables in air pollution models are unlikely to be linear and additive. Finally, Part 3 presents the conclusions, based on the results from the empirical chapters.

2. Road transport externalities and tourism

It is widely acknowledged that transport activities impose significant costs upon society which, in contrast to the benefits, are generally not borne by transport users. Without policy intervention, the so called external costs are not taken into account by the transport users when they make a transport decision. Transport users are thus faced with incorrect incentives, leading to welfare losses (Maibach et al., 2008).

Within the different transport modes, road transport deserves special attention because few activities are the potential source of more external effects than is vehicular use of the road system (Vitalino and Held, 1991). Moreover, according to Verhoef (2000) it is important to study road transport related externalities because road transport is generally identified as the most important inland transport mode in terms of external cost generation.

External costs caused by road transport use consist not only of costs in the monetary sense, but also of time losses, pollution, noise, health effects, regional environmental effects, global warming, barrier effects, road damage, accidents, and so on (Johansson, 1997; Mayeres, 2002). While there are many different road transport externalities, the list can be narrowed down to four main categories that are usually considered in the literature (Maibach et al., 2008; Qingyu et al., 2007; Verhoef, 2000): (1) congestion, which implies time delay and fuel consumption at very slow speed; (2) noise; (3) accidents and; (4) air pollution. Qingyu et al. (2007), Verhoef (2000) and Lakshmanan et al. (2001) have further divided them into “intra-sectoral externalities”, which include costs posed upon one-another by road users, like accidents and congestion and; “inter-sectoral externalities” like pollution and noise which are posed upon society at large and could also be categorized as environmental externalities.

Even though these four externalities have been the most widely studied, it should be noted how in certain cases, air pollution, congestion and accident costs

INTRODUCTION

have each been identified as the highest in terms of external cost generation, depending on the situation (network considered, volume of traffic and vehicle type) (e.g. Mayeres, 2002) and; within the environmental externalities, most studies have found that the cost of air pollution is greater than the cost of water pollution, noise or climate change (Delucchi, 2000). With this in mind and considering that the quantification of the impacts caused by noise on health is more difficult to achieve than in the case of other externalities (De Rus et al., 2003), the present thesis does not consider noise, focusing instead on accidents, congestion and air pollution.

Regarding traffic accidents, although the discussion on external costs has been centered mostly on environmental costs, Lindberg (2001) and Lindberg et al. (1999) point out that accidents are often the main component of external costs from road transport. Additionally, De Rus et al. (2003) point out that, in absolute terms of number of victims and wounded people from transport accidents, the main problem in almost all countries stems from road transport.

With respect to congestion, several decades have been devoted to studying it, not only because it is among the greatest issues facing urbanized areas nowadays (Boarnet et al., 1998; Verhoef, 2001) but also because it has repeatedly been related to other externalities, such as road traffic emissions (Nesamani et al., 2007; Smit et al., 2008). Moreover, the costs of road traffic congestion have been regarded to be very high by authors like De Rus et al. (2003), who sustain that in developed countries, the global estimates of these costs are around 2% of the GDP per country; and Vickrey (1963), whose estimations suggested that the real economic cost of the transport infrastructure in the United States was about three times the total gasoline and vehicular taxes generated by automobile use of city streets.

As regards air pollution, the importance of studying it stems mainly from the issue that, while emissions from industrial and domestic sources has decreased in most countries, there has been a substantial increase of air pollution caused by vehicular exhaust emissions (VEEs) (Nagendra and Khare, 2002). In fact, it has been suggested that road traffic is the dominant anthropogenic source of air pollution in

urban areas, not only because of the magnitude of its emissions, but also because pollutants are emitted in close proximity to people, thus enhancing exposure levels (Smit et al., 2008). Also, it is important to point out that analyzing the emission of pollutants from road transport, is a way to consider the broader issue of global warming. According to Maibach et al. (2008), climate change or global warming impacts of transport are mainly caused by emissions of the greenhouse gases, carbon dioxide (CO₂), nitrous oxide (N₂O) and methane (CH₄).

Since road transport is a key element of tourism and destination development, during the last years there has been growing concern about tourism as a source of the aforementioned transport externalities. For instance, tourists being injured or killed in motor vehicle crashes has become a key area of concern in travel medicine (Wilks, 1999); and authors like Gössling et al. (2005) sustain that the use of fossil fuels and related emissions of greenhouse gases is, from a global point of view, the most pressing environmental problem related to tourism.

However, despite such evidence, according to some authors the negative impacts of tourist road transport remains a neglected area of tourism research (Peeters et al., 2007; Rendeiro and Ramirez, 2010). Some exceptions to this can be found throughout the literature, including Levine et al. (1995) and Eisenberg (2004), in the case of road crashes; and Cools et al. (2007), Datla and Sharma (2008), Keay and Simmonds (2005) and Liu and Sharma (2006) in the case of traffic congestion. But, it should be noted that in all of these cases, the role of tourism in determining the level of a given externality is only approximated indirectly through the use of dummy variables, but not using a real variable to compute their impact.

3. Daily Indicator of Tourism Pressure

As the literature review has shown, the nonexistence of a tourism sector in the traditional economic classifications on National Accounts has made it difficult to assess some impacts of tourism. While this issue has been overcome in the economic context thanks to certain techniques, there has been a failure to account for other kind of impacts. Hence, an effort should be made so that in applying tourism policies, public administration officials can attribute the environmental and social costs of tourism to the industry as a whole and not only to its most representative economic sectors.

Based on the definition of tourism and in order to consider it as an aggregate, a researcher may opt for tourist arrivals as an indicator of pressure and different impacts. Nevertheless, this approach, often available through monthly data, presents three main drawbacks. Firstly, when trying to reflect the impact of tourism on a certain month, the presence of a time lag between the arrival date and the day when a tourist is effectively impacting could result in a bias, particularly when including tourists arriving at the end of the month. Secondly, an indicator of such nature would not be able to capture the mean length of stay which in turn, may change over the years or seasons, and omitting this sort of information in the analysis could largely affect the estimated results. Thirdly, if it is necessary to consider the variability in tourist activity, observed over periods shorter than a month, it is difficult to find adequate indicators of tourist arrivals.

In this context, the present section shows the development of a measure of tourist pressure capable of being related to most environmental indicators and other socioeconomic indicators overcoming the limitations of the simple use of monthly aggregated tourist arrivals. In concrete, this is achieved by using data from the Balearic Islands (Spain), one of the most important Mediterranean resort destinations, that offers the advantage of having detailed records of the number of people that arrive and leave the archipelago on a daily basis. In addition, thanks to

the use of a measure based on daily data, it is possible to address the fact that the three externalities considered in this thesis need to be studied at a disaggregated level in order to accurately capture the effect of key influencing factors like the weather and to observe patterns of seasonality.

3.1. Obtaining the Indicator

From a practical point of view, one of the most important advantages that characterize islands is that all tourists arrive by plane or by boat, thus simplifying controls and statistics. Using this special feature, Riera and Mateu (2007) developed a daily indicator of human pressure (DIHP) quantifying the demographic burden that an isolated territory supports at a given day using the case study of the Balearic Islands. The methodology for computing the DIHP takes as a reference point the number of people on the first day of each year (P_0) as the sum of the resident population (RP_0) and the number of tourists (TP_0). Then, for each one of the following days, the daily balance of people between arrivals (A_t) and departures (D_t) is added, plus an estimation of the population growth (V_t), which is a consequence of the “natural” evolution of the population for any given period (t), in this case a day. Analytically:

$$DIHP_t = P_0 + \sum_{t=1}^{365} (A_t - D_t) + \sum_{t=1}^{365} V_t \quad (1)$$

Although equation 1 summarizes the basic idea of the daily population pressure evidencing, for instance, great differences between summertime and wintertime in the Balearic Islands (Riera and Mateu, 2007), an additional improvement can be considered in order to discriminate between tourist and resident population. Thus, if the amount of residents who are not present on the territory (RPO_t) is subtracted, the DIHP can be expressed as:

INTRODUCTION

$$DIHP_t = TP_0 + \sum_{t=1}^{365} (A_t - D_t) + RP_0 + \sum_{t=1}^{365} V_t - RPO_t \quad (2)$$

where the resident population present on the territory at a given day (RP_t) can be identified as:

$$RP_t = RP_0 + \sum_{t=1}^{365} V_t - RPO_t \quad (3)$$

Then, the tourist population present on the territory at a given day (TP_t) can be calculated from:

$$TP_t = DIHP_t - RP_t \quad (4)$$

It should, thus, be noted how the daily stock of people in a given day on a territory can be split into the daily stock of residents and the daily stock of tourists in a given day. Analytically:

$$DIHP_t = RP_t + TP_t \quad (5)$$

3.2. Indicator of population pressure with data from the Balearic Islands

The case study of the Balearic Islands (Spain) is analyzed because of the special importance of tourism in the archipelago (with more than 13 millions of tourists and 1 million of inhabitants in 2008), which has motivated different tourism impact analyses, and also due to the geographical characteristics of the islands' territories that facilitate the computation of the stock measures. The Balearic Archipelago is located in the West of the Mediterranean Sea and composed by four

populated islands: Mallorca, the biggest one, which absorbs more than 75% of tourist arrivals; Menorca with less than 10%, and Eivissa and Formentera (known jointly as the Pitiüses, these islands share one single airport located in Eivissa) with more than 15% of the total arrivals of the archipelago. Tourists in the Balearic Islands usually belong in the medium or medium–low income brackets and their principal motivations are related to climate and beaches, a circumstance that explains why tourism in the islands has a high degree of seasonality, with more than 60% of tourist arrivals concentrated from June to September.

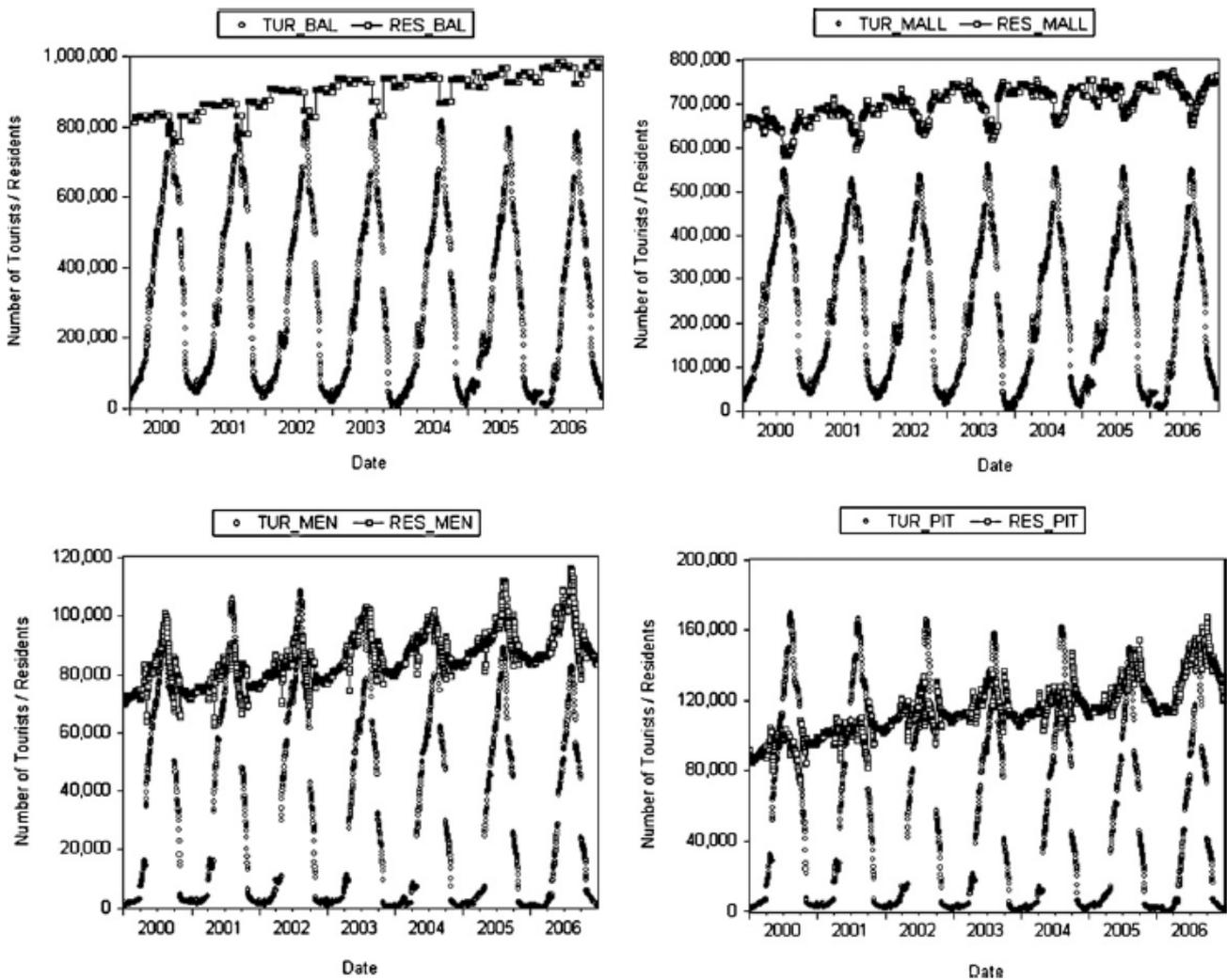
Using data collected from the 1st of January 2002 to the 31st of December 2008, the computation of the DIHP for the Balearics takes the resident population of the first day of each year (RP_0) from the municipal register that can be downloaded from the national statistical institute (INE - Instituto Nacional de Estadística, <http://www.ine.es>). The amount of tourists during the first day of the year (TP_0) is estimated using the statistics of the “Hotel Occupancy Survey” and the “Apartment Occupancy Survey” (also accessible from INE) and an estimation of the tourists lodged in other typology of accommodation (i.e. friends or relatives) from the international tourism survey for Spain (Frontur) elaborated by the Institute for Tourism Studies. Daily arrivals (A_p) and daily departures (D_p) are derived from passenger transit at ports and airports, and were provided by AENA, the Spanish national airport manager, and *Autoritat Portuaria de Balears*. V_p is calculated under the assumption that the probability of either being born or dying is the same for every day of the year; hence, it is derived from dividing the natural population growth by 365. Finally, since the domestic tourism survey of the Spanish residents (Familitur) collects monthly information on Spanish people traveling abroad (domestic and international trips), a monthly estimation of the residents outside the Balearics is accessible, with the possibility of obtaining the daily estimation of residents not present on the islands (RP_t) by using the mean length that is already available from the same survey.

INTRODUCTION

All of these variables are available for each one of the islands with the exception of RP_t . In this case, an indicator for each one of the main islands can be obtained and the amount of resident population abroad has to be estimated using the weight of resident population for each island and, then, assuming that the travel propensity is the same for all the residents.

Thus, using the data mentioned above, the decomposition of the DIHP into tourist population stock (TP) and resident population stock (RP), for the entire archipelago (BAL) and each one of the main islands, Mallorca (MALL), Menorca (MEN) and Pitiüses (PIT) (Eivissa and Formentera), is shown in Figure 1.

Figure 1: Daily tourist and resident stock of population



Through the calculation of the TP and RP variables it is shown how tourism is responsible for the strong seasonal component shown in the DIHP. Meanwhile, the residents, following their usual practices, tend to take their holidays during summertime too, reducing their presence during the peak months. Moreover, from a regional point of view, it can be seen how in the cases of Menorca and Pitiüses, the number of tourists during some weeks is even higher than that of residents, at least during the first few years of the sample.

3. References

- Aguiló, E.; Palmer, T., Rosselló , J. 2010. Road Transport for Tourism: Evaluating Policy Measures from Consumer Profiles. *Tourism Economics* (In press).
- Aguiló, E., Rosselló, J. 2005. Host Community Perceptions and Attitudes toward Tourism. A Cluster Analysis. *Annals of Tourism Research* 32(4), 925-941.
- Becken, S., Simmons, D. 2002. Understanding Energy Consumption Patterns of Tourist Attractions and Activities in New Zealand. *Tourism Management* 23(4), 343-354.
- Becken, S., Simmons, D. 2008. Using the concept of yield to assess the sustainability of different tourist types. *Ecological Economics* 67(3), 420-429.
- Becken, S., Frampton, C., Simmons, D. 2001. Energy Consumption Patterns in the Accommodation Sector – The New Zealand Case. *Ecological Economics* 39(3), 371-386.
- Becken, S., D. Simmons, Frampton, C. 2003. Energy Use Associated with Different Travel Choices. *Tourism Management* 24(3), 267-277.
- Blake, A., Durbarry, R., Sinclair, T., Sugiyarto, G. 2001. Modelling Tourism and Travel Using Tourism Satellite Accounts and Tourism Policy and Forecasting Models. *Tourism and Travel Research Institute Discussion Paper 2001/4*. Nottingham, UK: Tourism and Travel Research Institute.
- Boarnet, M., Kim, E., Parkany, E. 1998. Measuring Traffic Congestion. Institute of Transportation Studies University of California, Irvine, U.S.A. Presented at the 77th Annual Meeting of the Transportation Research Board, Washington, D.C. UCI-ITS-WP-98-3.
- Brida, J. G., Sanchez, E. J., Accinelli, E. 2008. Preservation of Environmental Quality and Tourist Investments. *European Journal of Tourism Research* 1(1), 3-15.

- Bujosa-Bestard, A., Rosselló-Nadal, J. 2007. Modelling Environmental Attitudes Toward Tourism. *Tourism Management* 28, 688–695.
- Caffyn, A., Lutz, J. 1999. Developing the Heritage Tourism Product in Multi-Ethnic Cities. *Tourism Management* 20, 213–21.
- Cools, M., Moons, E., Wets, G. 2007. Investigating Effect of Holidays on Daily Traffic Counts: Time Series Approach. *Transportation Research Record* 2019, 22-31.
- Datla, S., Sharma, S. 2008. Impact of Cold and Snow on Temporal and Spatial Variations of Highway Traffic Volumes. *Journal of Transport Geography* 16, 358–372.
- De Rus, G., Campos, J., Nombela, G. 2003. *Economía del Transporte*. Barcelona : A. Bosch. (Chapter 8).
- Delucchi, M. A. 2000. Environmental Externalities of Motor-Vehicle Use in the US. *Journal of Transport Economics and Policy* 34(2), 135-168.
- Dwyer, L., Forsyth, P., Spurr, R. 2004. Evaluating Tourism's Economic Effects: New and Old Approaches. *Tourism Management* 25, 307–317.
- Eisenberg, D. 2004. The Mixed Effects of Precipitation on Traffic Crashes. *Accident Analysis and Prevention* 36, 637-647.
- Gray, R., J. Bebbington. 1993. *Accounting for the Environment*. London: Sage.
- Gössling, S. 2000. Sustainable tourism development in developing countries: Some aspects of energy use. *Journal of Sustainable Tourism* 8(5), 410-25.
- Gössling, S. 2002. Global Environmental Consequences of Tourism. *Global Environmental Change* 12(4), 283-302.
- Gössling, S., C. Borgström Hansson, O. Hörstmeier, S. Saggel. 2002. Ecological Footprint Analysis as a Tool to Assess Tourism Sustainability. *Ecological Economics* 43(2-3), 199-211.
- Gossling, S., Peeters, P., Ceron, J., Dubois, G., Patterson, T., Richardson, R. B. 2005. The eco-efficiency of tourism. *Ecological Economics* 54, 417– 434.

INTRODUCTION

- Hau, T. D. 1998. Congestion Pricing and Road Investment, in *Road Pricing, Traffic Congestion and the Environment*, by Kenneth J. Button and Erik T. Verhoef, eds., Edward Elgar Publishing Limited, Cheltenham, England, Chapter 3, pp. 39-78.
- Hernández, J.M., León, C. 2007. The Interactions between Natural and Physical Capitals in the Tourist Lifecycle Model. *Ecological Economics* 62(1), 184-193.
- Ivanov, S., Webster, C. 2007. Measuring the Impact of Tourism on Economic Growth. *Tourism Economics* 13(3), 379–388.
- Johansson, O. 1997. Optimal Road-Pricing: Simultaneous Treatment of Time Losses, Increased Fuel Consumption, and Emissions. *Transportation Research D* 2(2), 77-87.
- Jones, C., Munday, M. 2004. Evaluating the Economic Benefits from Tourism Spending through Input-Output Frameworks: Issues and Cases. *Local Economy* 19(2), 117–133.
- Keay, K., Simmonds, I. 2005. The Association of Rainfall and Other Weather Variables with Road Traffic Volume in Melbourne, Australia. *Accident Analysis and Prevention* 37, 109–124.
- Lakshmanan, T.R., Nijkamp, P., Rietveld, P., Verhoef, E.T. 2001. Benefits and Costs of Transport Classification, Methodologies and Policies. *Papers in Regional Science* 80, 139–164.
- Levine, N., Kim, K. E., Nitz, L. H. 1995. Daily Fluctuations in Honolulu Motor Vehicle Accidents. *Accident Analysis and Prevention*. 27, 785-796.
- Lindberg, G. 2001. Traffic Insurance and Accident Externality Charges. *Journal of Transportation Economics and Policy* 35(3), 399-416.
- Lindberg, G., Ajo, P., Brito da Silva, A., Krupp, R., Osório Nunes, M. R., Peirson, J., Schneglberger, G., Smolders, W. 1999. Calculating Transport Accident Costs. Final Report of the Expert Advisors to the High Level Group on Infrastructure Charging (Working Group 3). Sweden.

- Liu, Z., Sharma, S. 2006. Statistical Investigations of Statutory Holiday Effects on Traffic Volumes. *Transportation Research Record* 1945, 40-48.
- Maibach, M., Schreyer, C., Sutter, D., Van Essen, H.P., Boon, B.H., Smokers, R., Schroten, A., Doll, C., Pawlowska, B., Bak, M. 2008. Handbook on Estimation of External Costs in the Transport Sector. Internalisation Measures and Policies for all External Cost of Transport (IMPACT) Delft: CE.
- Martin, I. 2009. Economic Development versus Environmental Sustainability: The Case of Tourist Marinas in Andalusia. *European Journal of Tourism Research* 2(2), 162-177.
- Mayeres, I. 2002. Taxes and Transport Externalities. Center for Economic Studies Energy, Transport & Environment Working Paper Series n°2002-11.
- Nagendra, S., Khare, M. 2002. Line Source Emission Modelling. *Atmospheric Environment* 36, 2083–2098.
- Nesamani, K. S., Chu, L., McNally, M. G., Jayakrishnan, R. 2007. Estimation of Vehicular Emissions by Capturing Traffic Variations. *Atmospheric Environment* 41, 2996–3008.
- Palmer, T., Riera, A., Rosselló, J. 2007. Taxing Tourism: The Case of Rental Cars in Mallorca. *Tourism Management* 28, 271–279.
- Patterson, T. M., Niccolucci, V., Bastianoni, S. 2007. Beyond ‘More Is Better’: Ecological Footprint Accounting for Tourism and Consumption in Val di Merse, Italy. *Ecological Economics* 62(3-4), 747-756.
- Peeters, P., Szimba, E., Duijnsveld, M. 2007. Major Environmental Impacts of European Tourist Transport. *Journal of Transport Geography* 15, 83–93.
- Qingyu, L., Zhicai, J., Baofeng, S., Hongfei, J. 2007. Method Research on Measuring the External Costs of Urban Traffic Congestion. *Journal of Transportation Systems Engineering and Information Technology* 7(5), 9–12.
- Rendeiro R., Ramírez, P. P. 2010. Ecological Footprint Analysis of Road Transport Related to Tourism Activity: The Case for Lanzarote Island. *Tourism Management* 31, 98–103.

INTRODUCTION

- Riera, A., Mateu, J. 2007. Aproximación al Volumen de Turismo Residencial en la Comunidad Autónoma de las Illes Balears a Partir del Cómputo de la Carga Demográfica Real. *Estudios Turísticos* 174, 59-72.
- Sheng, L., Tsui, Y. 2009. A General Equilibrium Approach to Tourism and Welfare: The Case of Macao. *Habitat International* 33, 419–424.
- Smeral, E. 2006. Tourism Satellite Accounts: A Critical Assessment. *Journal of Travel Research* 45, 92-98.
- Smit, R., Brown, A. L., Chan, Y. C. 2008. Do Air Pollution Emissions and Fuel Consumption Models for Roadways Include the Effects of Congestion in the Roadway Traffic Flow? *Environmental Modelling & Software* 23, 1262–1270.
- Smith, V. K., Huang, J. C. 1995. Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models. *The Journal of Political Economy* 103(1), 209-227.
- Tabatchnaia-Tamirisa, N., Loke, M. K., Leung, P., Tucker, K.A. 1997. Energy and Tourism in Hawaii. *Annals of Tourism Research* 24(2), 390-401.
- Verhoef, E. T. 2000. The Implementation of Marginal External Cost Pricing in Road Transport. Long Run vs Short Run and First-best vs Second-best. *Papers in Regional Science* 79, 307–332.
- Verhoef, E. T. 2001. An Integrated Dynamic Model of Road Traffic Congestion Based on Simple Car-Following Theory: Exploring Hypercongestion. *Journal of Urban Economics* 49, 505-542.
- Vickrey, W. S. 1963. Pricing in Urban and Suburban Transport. *American Economic Review: Papers and Proceedings* 53(2), 452-465.
- Vitalino, D. F., Held, J. 1991. Road Accident External Effects: an Empirical Assessment. *Applied Economics* 23, 373-378.
- Wardman, M., Bristow, A. L. 2004. Traffic Related Noise and Air Quality Valuations: Evidence from Stated Preference Residential Choice Models. *Transportation Research D* 9(1), 1-27.

Wilks, J. 1999. International Tourists, Motor Vehicles and Road Safety: A Review of the Literature Leading Up to the Sydney 2000 Olympics. *Journal of Travel Medicine* 6, 115-121.

II. Empirical Chapters

Chapter 1. Road accidents and tourism: The case of the Balearic Islands (Spain)¹

Abstract:

The increase in the number of tourists for many destinations and their increased mobility within host countries or regions has implied a rise in tourism-associated externalities, with vehicle crashes as the most common cause of injury for tourists. Within the transport literature, the number and variation in the amount of accidents has been related to a large set of determining variables, including weather conditions, socio-economic characteristics, exposure, physical characteristics of the road and a variety of dummies that try to capture effects such as safety laws and seasonal variations. However, the presence of tourism has been neglected. Using the case study of the Balearic Islands, the present study estimates the role of tourism in determining the number of accidents in a daily context, using the set of variables suggested by the literature and incorporating a daily measure for the stock of tourists at a host destination. Results show how tourism can be associated with a significant amount of the accidents that take place in the Balearics.

1.1. Introduction

Within the increasing body of literature in which travel has been recognized as a contributor to various problems (e.g., Dickinson and Robbins, 2008; Dickinson et al., 2009; Gursoy et al., 2002; Hall, 1999), road traffic accidents have stood out because they are often the main component of external costs from road transport activities (Lindberg et al., 1999; Lindberg, 2001).

In the tourism literature, the issue of road traffic accidents has been addressed by various studies, like those of Bentley et al. (2001), Howard (2009), Page and Meyer (1996) and Page (2009). Tourists injured or killed in motor vehicle crashes have also become a key area of concern in travel medicine (Carey and Aitken, 1996; Leggat and Leggat, 2003; Wilks et al., 1999). This follows from the

¹ This chapter benefited from the valuable and enriching comments from Prof. Tom Brijs (Transportation Research Institute, Belgium), Daniel Eisenberg (Health Management and Policy, School of Public Health, University of Michigan, USA) and Lasse Fridström (Institute of Transport Economics, Norway), none of whom share any responsibility for any flaws in the article.

EMPIRICAL CHAPTERS

consideration that motor vehicle crashes have been identified, but not quantified, as the most common cause of injury or death for tourists (Wilks, 1999). Moreover, there are studies such as those of Page and Meyer (1996), Page (2009) and Wilks et al. (1999) that raise awareness of the problems that accidents may pose for the tourism industry, which has the image of a business selling positive holiday experiences (Clift and Page, 1996) and of a catalyst for change in both individuals and communities (Ryan, 1997). This concern is relevant in the global context because of the increasing numbers in international tourist arrivals, which was 806 million in 2005 and is expected to reach an estimated 1.6 billion in 2020 (UNWTO, 2008, 2009).

In spite of this, the nonexistence of a single economic sector that can be associated with tourism has made it difficult to account for the external costs of tourism activity. While the literature has been able to develop a set of techniques aimed at valuing the economic impacts of tourism (Tourism Satellite Accounts and General Equilibrium Models), there still exists a void when trying to assess the non-economic impacts of tourism.

In this context, while a broad range of studies (Brijs et al., 2008; Eisenberg, 2004; Fridstrom, 1999; García-Ferrer et al., 2006; Hermans et al., 2006; Hutchings et al., 2003; Keay and Simmonds, 2006; Levine et al., 1995; Lord and Persaud, 2000; Quddus, 2008; Van den Bossche et al., 2005a, b; Wang and Abdel-Aty, 2006) have addressed the underlying causes of road traffic accidents, only few have linked them to tourism. Studies that do consider the influence of tourism on the externality that road accidents represent do so through the use of dummy variables for labor/school holidays and vacation periods. The only exception found in the literature is Levine et al. (1995), who consider monthly tourist arrivals but do not include them in the definitive results due to the non-significance of the variable.

Thus, the present study tries to fill this gap by explaining the number of accidents in a daily context, using most of the variables suggested by the literature and including a daily measure for the amount of tourists at a host destination. The

empirical application is carried out with data from the Balearic Islands which are especially suited for this study because of the relative importance of the tourism activity in the region and the availability of the necessary variables, including a tourism pressure indicator.

The remainder of this chapter is structured as follows. Section 1.2 contains a literature review on the variables commonly included in studies of this nature. Section 1.3 reviews the various methods used to approach the modeling of road accidents. Section 1.4 provides a description of the data. Section 1.5 contains the estimation results. Finally, concluding remarks and some limitations can be found in Section 1.6.

1.2. Literature Review

Previous research suggests a number of well-identified variables that have proved to be significant determinants of the number of crashes. They can be classified into broad categories, and a literature review reveals that the weather conditions constitute some of the most commonly used variables in empirical studies (Eisenberg, 2004; Andreescu and Frost, 1998; Brijs et al., 2008; Fridstrom et al., 1995; Hermans et al., 2006). Common examples of these variables are precipitation, hours of sun, temperature, air pressure, wind speed, relative humidity or snow for places with more extreme conditions.

Besides the weather, most studies (Van den Bossche et al., 2005b; Fridstrom, 1999; Fridstrom et al., 1995; Brijs et al., 2008; Quddus, 2008) have agreed that the volume of exposure (the amount of units exposed to accident risk) is a key variable to be included in any road accident model. The most commonly used measure to account for exposure is the traffic volume in terms of the number of vehicle-kilometers driven on the road network. However, there are various cases in which no valid measure of exposure is available (Hermans et al., 2006; Van den Bossche et al., 2005a). This issue has been discussed by Brijs et al. (2008) and Van den Bossche

EMPIRICAL CHAPTERS

et al. (2005a, b), who wanted to test how well the safety model could perform without any measure of exposure and concluded that even without a variable such as exposure, valid models can be constructed.

Dummy variables can be considered as a valid alternative to address the consideration of exposure. Thus, for instance, if information on daily traffic exposure is not available for some reason, day-of-the-week dummies may account quite well for the day-of-the-week variability in exposure and still produce consistent results for the weather effects (Brijs et al., 2008).

Additional special situations have been frequently approached through the use of dummy variables. Hermans et al. (2006), Quddus (2008) and García-Ferrer et al. (2006) employed dummies in order to capture the effect of certain policies and countermeasures for the occurrence of traffic accidents, like the compulsory use of seat belts. Keay and Simmonds (2006) used dummies to capture the effect of school and public holidays, and Fridstrom (1999) presents a brief discussion on the use of the calendar effects and their relationship with fuel sales, a variable commonly used to proxy traffic volume. Of particular note for this study, Eisenberg (2004) included spatial dummies to account for tourism in states in which this activity is relatively higher than in others.

Nevertheless, in spite of the importance of evaluating tourism activity in understanding crash counts for an increasing set of regions, it seems clear that the need to take such influence into account through a quantitative measure has been neglected. The exception to this finding is Levine et al. (1995), who incorporate tourist arrivals within the set of explanatory variables. However, the variable turns out to be non-significant in the main model. It is important to stress that in that study, arrivals were introduced on a monthly basis, even though the model was set out in terms of daily data and thus unable to accurately capture daily variations in tourism.

This example demonstrates the relevance of using disaggregated temporal data. Often considered the most common context when exploring crash counts

(Quddus, 2008; Lord and Persaud, 2000; Brijs et al., 2008), time series data can be classified into aggregated and disaggregated data (either at a temporal or a spatial level). This distinction is important because, while changes in crash counts on a highly aggregated level can be explained by structural changes, they cannot easily pick up patterns of seasonality or weather effects. In contrast, the lower the level of aggregation, the more feasible it is to study the effects of weather conditions, traffic volume, and holidays on changes in crash counts (Brijs et al., 2008). Several authors have therefore issued warnings regarding biases that are introduced by modeling crash counts at high levels of aggregation (Golob et al., 1990; Jovanis and Chang, 1989).

Evidence on how the use of daily (disaggregated) versus monthly (aggregated) data improves the results can be found in Eisenberg (2004), where both monthly and daily data are used to study the effects of precipitation on the number of accidents. With monthly data, a rather surprising negative relationship between such variables is found, while further disaggregation (daily) helps to find the expected positive relationship. The source of the contrasting results, as the author suggests, appears to be a substantial negative lagged effect of precipitation across days within a state-month.

1.3. Methods

Within the regression framework, the choice of an adequate estimation technique for road traffic accidents depends on specific characteristics of the data. If the distribution followed by the number of accidents tends towards the normal, Ordinary Least Squares (OLS) can be an appropriate estimation technique (Levine et al., 1995; Keay and Simmonds, 2006). However, while the normality assumption is reasonable for (roughly) continuous dependent variables that can take on a large range of values, a count variable such as accidents cannot have a normal distribution (since the normal distribution is for continuous variables that can take on all values).

EMPIRICAL CHAPTERS

If it takes on very few values, the distribution can be very different from the normal (Wooldridge, 2000). Furthermore, the assumptions for OLS methods are violated by the discrete, non-negative nature of accident count data and the reality that the variance in the number of accidents increases as the traffic flow increases (Lord and Persaud, 2000). For these reasons, there are alternative methods that are better suited to this particular type of data, whose starting point is the Poisson, a probability distribution that is specifically used for count data (Cameron and Trivedi, 2001; Gujarati, 2004; Wooldridge, 2000).

The Poisson for the number of occurrences of the event has a probability density function given by

$$\Pr(Y = y) = \frac{e^{-\mu} \mu^y}{y!}, \quad y = 0, 1, 2, \dots, \quad (1.1)$$

where μ is the intensity or rate parameter (Cameron and Trivedi, 2001, Gujarati, 2004). The expression can also be referred to as $P[\mu]$, and its two first moments are $E[Y] = \mu$ and $V[Y] = \mu$. The Poisson regression model can then be derived from the Poisson distribution by parametrizing the relation between the mean parameter μ and covariates (regressors) x (Wooldridge, 2000). The standard assumption is to use the exponential mean parametrization

$$\mu = \exp(x'_i \beta), \quad i = 1, \dots, n, \quad (1.2)$$

where by assumption there are k linearly independent covariates, usually including a constant.

If we are interested in the effects of the x_j on the mean response, there is little reason to go beyond the Poisson regression (Wooldridge, 2000). In fact, the purpose of using a statistical model in the present study lies basically in the effects of the regressors on the dependent variable. However, the Poisson's first two moments

evidence the well-known equality of the mean and variance property of the distribution (Cameron and Trivedi, 2001; Gujarati, 2004), a feature better known as equi-dispersion. Since count data are often characterized by exhibiting over-dispersion (i.e., the variance is greater than the mean), the Negative Binomial has gained further popularity (Dionne et al., 1995; Fridstrom et al., 1995; Eisenberg, 2004; Lord, 2006; Martin, 2002; Noland and Oh, 2004) and has been considered more suitable, due to its ability to deal with this issue (Lord, 2006; Eisenberg, 2004; Fridstrom et al., 1995; Dionne et al., 1995; Quddus, 2008).

Analytically, it is hypothesized that the distribution of a random count y is Poisson, conditional on the parameter λ , so that $f(y|\lambda) = \exp(-\lambda)\lambda^y/y!$ and $\lambda = \mu v$, where μ is a deterministic function of x , a matrix of $N \times K$ explanatory variables; for example, $\exp(x'\beta)$, and $v > 0$ is iid distributed with density $g(v|\alpha)$. Different observations may have different λ (heterogeneity), but part of this difference is due to a random (unobserved) component v . If $f(y|\lambda)$ is the Poisson density and $g(v)$, $v > 0$, is the gamma density with $E[v] = 1$ and $V[v] = \alpha$, we obtain the negative binomial density,

$$h(y|\mu, \alpha) = \frac{\Gamma(\alpha^{-1} + y)}{\Gamma(\alpha^{-1})\Gamma(y + 1)} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \mu} \right)^{\alpha^{-1}} \left(\frac{\mu}{\mu + \alpha^{-1}} \right)^y, \quad \alpha > 0, \quad (1.3)$$

where $\Gamma(\bullet)$ denotes the gamma integral that becomes a factorial for an integer argument (Cameron and Trivedi, 2001).

Since $\exp(\bullet)$ is a nonlinear function, we can neither use linear regression methods nor nonlinear squares because all standard count data distributions exhibit heteroskedasticity. Instead, we will use maximum likelihood estimation methods (Wooldridge, 2000).

While it is common to estimate the parameters of the model using the maximum likelihood (ML) of a negative binomial specification, the ML estimation can also be performed under a number of alternative distributional assumptions.

Moreover, when doubt exists about the form of the variance function, the use of the Quasi-ML (QML), also referred to as Pseudo-ML (PML), is recommended. Computationally, it is the same as Poisson ML, with the qualification that the variance matrix must be recomputed (Cameron and Trivedi, 2001).

The quasi-maximum likelihood estimators are robust in the sense that they produce consistent estimates of the parameters of a correctly specified conditional mean, even if the distribution is incorrectly specified. More specifically, this is done through a quasi-generalized pseudo-maximum likelihood estimator (QGPMLE) (Gourieroux et al., 1984a, b). If the variance of the Negative Binomial is $\omega_i = \mu_i + \alpha\mu_i^2$ and

$$\hat{\alpha} = \frac{1}{n-k} \sum_{i=1}^n \frac{\{(y_i - \hat{\mu}_i)^2 - \hat{\mu}_i\}}{\hat{\mu}_i^2} \quad (1.4)$$

is a consistent estimator of α , then the QGPMLE $\hat{\beta}_{QGPMLE}$ maximizes

$$\ln L_{LEFN} = \sum_{i=1}^n \left\{ -\alpha^{-1} \ln(1 + \alpha\mu_i) + y_i \ln\left(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\right) + b(y_i, \alpha) \right\}, \quad (1.5)$$

where LEFN is a “linear exponential family with nuisance parameter” density (Cameron and Trivedi, 2005).

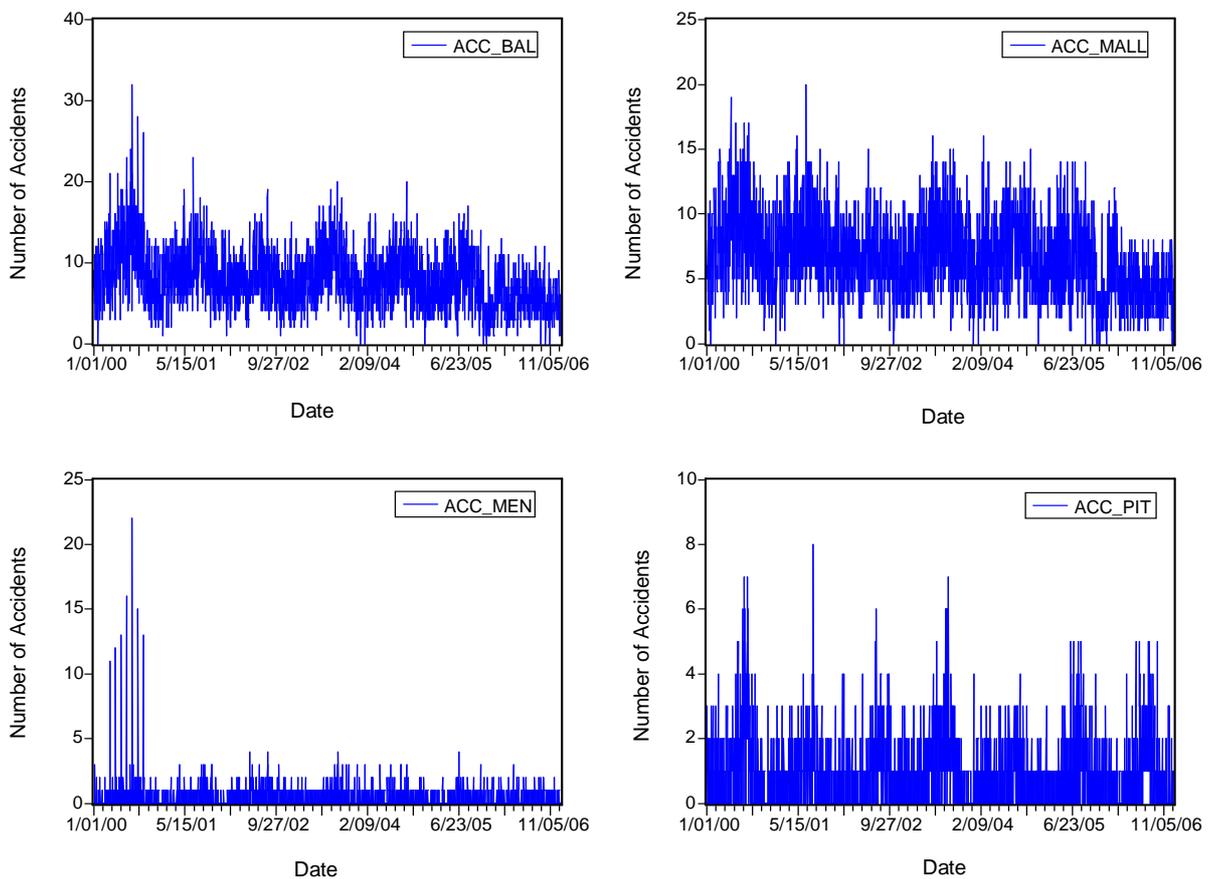
If the count data used for an application constitute a time series, an additional issue has to be considered. In particular, the independence of event occurrences in successive time intervals can be a reasonable assumption when the underlying stochastic process for such events, conditional on covariates, has no memory. But if the series presents serial correlation, there are various models to choose from. In the present study, we will focus on an autoregressive, or Markov model, which is a simple adjustment to the cross-section model detailed above that enters lagged

values of y into the formula for the conditional mean of current y (Wooldridge, 2000).

1.4. Data description

This study benefits from different databases collected and/or elaborated by the *Centre de Recerca Econòmica*² (Center for Economic Research) that come from different sources.

Figure 1.1 Road accidents in the Balearic Islands



²

EMPIRICAL CHAPTERS

Thus, daily counts of road traffic accidents (where accidents involving pedestrians are not included) were provided by the *Dirección General de Tráfico* (Traffic Department) and consist of daily vehicle crashes that take place in each one of the Balearic Islands: Mallorca, Menorca and the Pitiüses (Eivissa and Formentera). The time series data range from January 1, 2000 to December 31, 2006 and can be seen in Figure 1.1.

Meteorological variables are based on data provided by the *Instituto Nacional de Meteorología* (National Institute of Meteorology) of the Balearic Islands and are taken from the airport stations. The variable used to account for the daily stock of tourist and resident population is the DIHP (Part I).

1.5. Empirical Analysis

Following the methodological considerations explained in Section 3, Poisson and Negative Binomial models were initially estimated. However, given the presence of over-dispersion (the regression based tests³ in Table 1.1 are lead to reject the Poisson restriction, $v(x_i, \beta) = m(x_i, \beta)$), the results for the Poisson are not reported here. Estimation results from Negative Binomial models, using both ML and QML,⁴ are shown in Table 1.2. The presence of first order autocorrelation in the time series suggested the estimation of the dynamic models, including different lags of the dependent variable.

Based on the data collected and described above, the equations for each island and for the entire archipelago were estimated. The variables included in the final estimations are: dummies to control for different levels of exposure observed over the week and for specific public/school holidays; hours of sunshine; precipitation in liters per square meter; mean daily temperature in Celsius; a dummy

³ Cameron and Trivedi's (1990) test is based on an auxiliary regression of the squared difference between the dependent variable and the forecasted dependent variable, minus the dependent variable $((y-y_f)^2-y)$, on the square of the forecasted dependent variable $(y_f)^2$; Wooldridge's (1990) test is based on the estimation of the square of the standardized residuals minus one $(sresid^2-1)$ on the forecasted dependent variable (y_f) .

⁴ The variance is computed using the estimate obtained from Wooldridge's (1990) over-dispersion test.

to control for the implementation of a safety law through which drivers are entitled to a given amount of “points” that are subtracted when committing traffic violations (CPP);⁵ a linear trend, usually included in similar exercises in order to control for improvements in the infrastructure; and the variables that account for the stock of residents (DIHP residents) and tourists (DIHP tourists). Interaction and quadratic forms of the DIHP (both residents and tourists) were also considered but turned out to be non-significant and therefore were not included in the final estimations presented here.

Table 1.1 Regression based tests for overdispersion

	Test proposed by:	
	Cameron and Trivedi (1990)	Wooldridge (1990)
	Coeff. (<i>p-value</i>)	Coeff. (<i>p-value</i>)
Balearic Islands	0.025604 (0.0000)	0.023755 (0.0000)
Mallorca	0.01405 (0.0006)	0.017871 (0.0002)
Menorca	3.675888 (0.0000)	0.033414 (0.7853)
Pitiuses	0.114426 (0.0000)	0.10987 (0.0122)

Bearing in mind that the reference observation is the number of crashes on Sunday, in all cases except for Menorca, the signs of the estimated parameters for the day-of-the-week dummies follow the expected relationship, which indicates that there is a lower level of exposure on Sundays. Regarding the dummy variables for holidays, up to nineteen were initially considered, but only five turned out to be significant, indicating a lower level of accidents on important religious, national and local holidays. The CPP dummy was also significant in the cases of Mallorca and the entire archipelago, with a negative sign that shows the effectiveness of the law approved by the Spanish Government. The linear trend, as in similar applications, was significant and negative for all cases under study.

⁵ This variable takes the value 1 for all days since February 1, 2006 and 0 otherwise.

Table 1.2 Static estimations for crash counts

Method	Balearics		Mallorca		Menorca		Pitiuses	
	ML	QML	ML	QML	ML	QML	ML	QML
Intercept	Coeff. (p-value) 0.9084 (0.00)	Coeff. (p-value) 0.9084 (0.00)	Coeff. (p-value) 0.1606 (0.54)	Coeff. (p-value) 0.1648 (0.54)	Coeff. (p-value) -1.4286 (0.00)	Coeff. (p-value) -1.3885 (0.01)	Coeff. (p-value) -2.1255 (0.00)	Coeff. (p-value) -1.6299 (0.00)
Monday	0.0626 (0.03)	0.0626 (0.03)	0.1086 (0.00)	0.1029 (0.00)	-0.3214 (0.01)	-0.2843 (0.05)	0.0659 (0.48)	0.1172 (0.17)
Tuesday	0.0576 (0.05)	0.0576 (0.05)	0.0993 (0.00)	0.0929 (0.00)	-0.5638 (0.00)	-0.5557 (0.00)	0.0952 (0.31)	0.0858 (0.31)
Wednesday	0.0363 (0.22)	0.0363 (0.22)	0.0682 (0.03)	0.0681 (0.03)	-0.3788 (0.00)	-0.3571 (0.02)	0.0356 (0.71)	0.0475 (0.58)
Thursday	0.1029 (0.00)	0.1029 (0.00)	0.1643 (0.00)	0.1643 (0.00)	-0.1584 (0.21)	-0.1333 (0.34)	-0.0665 (0.49)	-0.0003 (0.99)
Friday	0.1398 (0.00)	0.1398 (0.00)	0.1791 (0.00)	0.1791 (0.00)	-0.0969 (0.43)	-0.0521 (0.70)	0.1073 (0.25)	0.1372 (0.10)
Saturday	0.0954 (0.00)	0.0954 (0.00)	0.1119 (0.00)	0.1119 (0.00)	-0.2129 (0.09)	-0.2195 (0.12)	0.1150 (0.21)	0.1685 (0.04)
Hours of sunshine	0.0059 (0.02)	0.0059 (0.02)	-	-	-	-	0.0119 (0.14)	-
Precipitation	0.0070 (0.00)	0.0070 (0.00)	0.0049 (0.00)	0.0046 (0.01)	0.0212 (0.00)	0.0209 (0.00)	0.0048 (0.13)	-
Average temperature	0.0050 (0.08)	0.0050 (0.08)	-	-	-	-	-	-
CPP	-0.1621 (0.00)	-0.1621 (0.00)	-0.1886 (0.00)	-0.1891 (0.00)	-	-	-0.0379 (0.67)	-
Trend	-0.0002 (0.00)	-0.0002 (0.00)	-0.0002 (0.00)	-0.0002 (0.00)	-0.0003 (0.00)	-0.0003 (0.00)	-0.0001 (0.14)	-0.0002 (0.00)
DIHP tourists	5x10 ⁻⁷ (0.00)	5x10 ⁻⁷ (0.00)	9x10 ⁻⁷ (0.00)	9x10 ⁻⁷ (0.00)	9x10 ⁻⁶ (0.00)	9x10 ⁻⁶ (0.00)	8x10 ⁻⁶ (0.00)	8x10 ⁻⁶ (0.00)
DIHP residents	1x10 ⁻⁶ (0.00)	1x10 ⁻⁶ (0.00)	3x10 ⁻⁶ (0.00)	3x10 ⁻⁶ (0.00)	1.1x10 ⁻⁵ (0.09)	1.1x10 ⁻⁵ (0.14)	1.3x10 ⁻⁵ (0.00)	1.1x10 ⁻⁵ (0.00)
Holidays								
August 15	-0.3906 (0.02)	-0.3906 (0.02)	-0.4407 (0.05)	-	-	-	-1.2103 (0.04)	-
October 12	-0.4333 (0.02)	-0.4333 (0.02)	-0.3291 (0.08)	-0.3274 (0.09)	-	-	-	-
December 8	-0.5706 (0.01)	-0.5706 (0.01)	-	-0.4400 (0.05)	-	-	-	-
December 26	-0.6114 (0.01)	-0.6114 (0.01)	-0.9919 (0.00)	-0.9911 (0.00)	-	-	-	-
December 31	-0.8107 (0.01)	-0.8107 (0.01)	-0.6964 (0.02)	-0.6953 (0.02)	-	-	-	-
Equation Statistics								
Adj. R-squared	0.2592	0.2595	0.1943	0.1941	0.0427	0.0444	0.1846	0.1862
LR statistic	1102.289	587773	738.0664	857376.8	433.0241	363606.3	464.344	55777.25
Probability(LR stat)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mean dep. Var.	7.9339	7.9339	6.6326	6.6326	0.4216	0.4216	0.8647	0.8911
S.D. dep. var	3.5579	3.5579	3.0337	3.0338	0.9968	0.9968	1.0719	1.1029
LR index (Pseudo-R2)	0.0794	0.9787	0.0569	0.9859	0.0924	0.988	0.089	0.9013
AIC	5.0156	5.0148	4.8036	4.8033	1.6732	1.7339	2.3626	2.4019

Table 1.3 Dynamic estimations for crash counts

Method	Balearics		Mallorca		Menorca		Pitiuses	
	ML	QML	ML	QML	ML	QML	ML	QML
Intercept	Coeff. (p-value)							
Day 1	0.764461 (0.002)	0.764461 (0.0019)	0.346556 (0.1842)	0.346556 (0.1921)	-1.275087 (0.006)	-1.264015 (0.0184)	-1.579641 (0.0000)	-1.579795 (0.0000)
Day 2	0.069312 (0.018)	0.069312 (0.0173)	0.114684 (0.0002)	0.114684 (0.0003)	-0.333571 (0.0104)	-0.359488 (0.0143)	0.126735 (0.1276)	0.126419 (0.1359)
Day 3	0.060275 (0.041)	0.060275 (0.0399)	0.105373 (0.0007)	0.105373 (0.0009)	-0.577062 (0.0000)	-0.581423 (0.0001)	0.105987 (0.2047)	0.105517 (0.2154)
Day 4	0.038964 (0.1874)	0.038964 (0.1848)	0.072641 (0.0206)	0.072641 (0.0228)	-0.381701 (0.0038)	-0.396217 (0.0071)	0.063676 (0.4495)	0.063654 (0.4582)
Day 5	0.114113 (0.0001)	0.114113 (0.0001)	0.171462 (0.0000)	0.171462 (0.0000)	-0.17502 (0.1639)	-0.202643 (0.1558)	0.014187 (0.8676)	0.013728 (0.8742)
Day 6	0.146103 (0.0000)	0.146103 (0.0000)	0.189312 (0.0000)	0.189312 (0.0000)	-0.10278 (0.4078)	-0.136301 (0.3367)	0.150906 (0.0675)	0.151341 (0.0721)
Hours of sunshine	0.094433 (0.0011)	0.094433 (0.0011)	0.107828 (0.0005)	0.107828 (0.0006)	-0.204535 (0.1084)	-0.194616 (0.1764)	0.192315 (0.0188)	0.192497 (0.021)
Precipitation	0.005265 (0.0389)	0.005266 (0.0377)	-	-	-	-	-	-
CPP	0.007021 (0.0000)	0.007021 (0.0000)	0.005212 (0.0021)	0.005212 (0.0026)	0.021041 (0.0000)	0.021054 (0.0008)	-	-
Trend	-0.100521 (0.0014)	-0.100521 (0.0013)	-0.102324 (0.0034)	-0.102324 (0.0039)	-	-	-	-
DIHP tourists	-0.000155 (0.0000)	-0.000155 (0.0000)	-0.000148 (0.0000)	-0.000148 (0.0000)	-0.000311 (0.0000)	-0.00028 (0.0006)	-0.000185 (0.0022)	-0.000185 (0.0027)
DIHP residents	0.000004 (0.000)	0.000004 (0.000)	0.0000051 (0.000)	0.0000051 (0.000)	0.0000988 (0.000)	0.0000983 (0.000)	0.0000685 (0.000)	0.0000685 (0.000)
Accidents (-1)	0.000011 (0.0005)	0.000011 (0.0005)	0.000016 (0.000)	0.000016 (0.0001)	0.000011 (0.1053)	0.0000102 (0.1914)	0.0000101 (0.0002)	0.0000101 (0.0003)
Accidents (-2)	0.012076 (0.0000)	0.012077 (0.0000)	0.019111 (0.0000)	0.019112 (0.0000)	-0.051905 (0.1624)	-0.045001 (0.2743)	0.036081 (0.0619)	0.036115 (0.0675)
Accidents (-3)	0.009467 (0.0001)	0.009467 (0.0001)	0.010213 (0.0005)	0.010214 (0.0006)	-0.026142 (0.4498)	-0.028924 (0.4816)	0.047609 (0.0116)	0.047574 (0.0135)
Accidents (-4)	0.003195 (0.1924)	0.003196 (0.1893)	0.007119 (0.015)	0.007119 (0.0169)	-0.048657 (0.1764)	-0.040258 (0.3)	0.030394 (0.1124)	0.030111 (0.1236)
Accidents (-5)	0.010294 (0.0000)	0.010294 (0.0000)	0.007947 (0.0067)	0.007948 (0.0077)	0.026574 (0.3409)	0.032024 (0.3674)	0.019551 (0.314)	0.01952 (0.3254)
Accidents (-6)	0.00541 (0.0272)	0.00541 (0.0261)	0.009377 (0.0014)	0.009378 (0.0016)	-0.039044 (0.2698)	-0.035025 (0.3557)	-0.005399 (0.7825)	-0.005218 (0.7939)
Accidents (-7)	0.004413 (0.0711)	0.004413 (0.0692)	0.007799 (0.0076)	0.007799 (0.0087)	-0.053052 (0.156)	-0.045684 (0.2675)	0.034606 (0.0705)	0.034648 (0.0764)
Accidents (-7)	0.003319 (0.172)	0.003318 (0.1692)	0.006359 (0.0291)	0.00636 (0.0319)	-0.053161 (0.1617)	-0.047537 (0.2652)	-0.023122 (0.2417)	-0.023201 (0.2499)
Holidays	-	-	-	-	-	-	-	-
August 15	-0.481115 (0.004)	-0.481115 (0.0038)	-0.304333 (0.0813)	-0.304333 (0.0862)	-	-	-1.166328 (0.0238)	-1.169085 (0.0261)
October 12	-0.413658 (0.0254)	-0.413658 (0.0247)	-	-	-	-	-	-
December 8	-0.6136 (0.0063)	-0.6136 (0.0062)	-0.522655 (0.0191)	-0.522655 (0.0206)	-	-	-	-
December 26	-0.593584 (0.01)	-0.593584 (0.0098)	-0.960147 (0.0009)	-0.960147 (0.001)	-	-	-	-
December 31	-0.735074 (0.0149)	-0.735074 (0.0148)	-0.602012 (0.044)	-0.602012 (0.0462)	-	-	-	-
Equation Statistics								
Adj. R-squared	0.273673	0.273962	0.204179	0.204488	0.048549	0.046611	0.203923	0.204195
LR statistic	1173.666	831802	789.5542	1181.438	442.7436	3222.762	624.6829	90613.91
Probability(LR stat)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Mean dep. Var.	7.936863	7.936863	6.666005	6.666005	0.422745	0.422745	0.889804	0.889804
S.D. dep. var	3.561753	3.561753	3.035226	3.035226	0.997895	0.997895	1.102804	1.102804
LR index (Pseudo-R2)	0.084728	0.084987	0.061745	0.061745	0.094613	0.094613	0.093438	0.0937305
AIC	4.993942	4.992939	4.786286	4.785398	1.677166	1.720646	2.393281	2.392572

EMPIRICAL CHAPTERS

In reference to the weather, both sunshine and precipitation were significant in most equations and followed the expected relationships, based on similar exercises (Eisenberg, 2004; Andreescu and Frost, 1998; Brijs et al., 2008; Fridstrom et al., 1995; Hermans et al., 2006). Thus, more hours of sun and more precipitation could be associated with higher numbers of accidents. Meanwhile, the average temperature was significant only in one equation, showing a positive relation with accident counts.

The population stock was highly significant in all cases, indicating a positive relationship between the population and crashes. The distinction between tourist and resident population shows, in a direct way, the impact of tourists in determining the level of accidents.

When comparing the ML and QML models in terms of the log-likelihood and the Akaike Information Criteria (AIC), the statistics do not differ much from one model to another. Thus, the better statistics and the most suitable estimation will depend on the island. Meanwhile, the comparison between static and dynamic alternatives shows a slight improvement for the latter; hence the dynamic equations are preferred in terms of the AIC.

Using the estimation results, since the population was split into residents and non-residents, simulations were carried out to evaluate the impact of tourism on road traffic accidents. From the data, we can see that the average number of daily traffic accidents within the observed period is 7.9 for the Balearics, 6.6 for Mallorca, 0.4 for Menorca and 0.9 for the Pitiuses. We then reestimated the equation with the substituted coefficients obtained from the estimation results under a hypothetical scenario in which the non-resident population was zero (Table 1.4).

Table 1.4 Simulation of the impact of tourism on accidents

Method	ML		QML	
	Static	Dynamic	Static	Dynamic
Balearic Islands	-15.80%	-14,1%	-15.80%	-14,1%
Mallorca	-14.00%	-13,4%	-14.00%	-13,5%
Menorca	-27.80%	-28,5%	-28.10%	-28,9%
Pitiuses	-40.40%	-41,8%	-40.40%	-40,2%

It is worth noting that the percent decrease is higher in the cases of Menorca and the Pitiüses, which can be explained by the fact that during the high season, on both islands, tourists outnumber the resident population. Furthermore, differences between Menorca and the Pitiüses could be explained by the different typology of tourists that often visit the islands, with more family-oriented tourism in the case of Menorca. In any case, we note that all the estimations carried out yielded similar results in the estimation of the impact of tourism on crash counts.

1.6. Discussion and Conclusions

Using a case study of the Balearic Islands, we relate the number and variation in the amount of accidents to the most common set of determining variables used in the reference literature and also include a new explanatory variable that incorporates tourism activity. Our results show how weather-related variables play a key role in the determination of road traffic accidents and also how tourism can be associated with a significant amount of the accidents that take place in the Balearics.

From the estimated results, different simulations were formulated to show the impact of tourism on road accidents. The simulation considered showed how, *ceteris paribus*, the level of accidents would decrease in 15.8% for the Balearics. These results should be connected to the fact that population in the Balearics grew from 875,000 in 2000 to 1 million in 2006, while the number of tourist arrivals grew from 11 to 12.5 million over the same period. This gives a constant rate of about 12.5 tourists per inhabitant in the sample.

EMPIRICAL CHAPTERS

Tourism is often regarded as an important source of externalities, and this study provided empirical evidence to prove that, indeed, the presence of tourists leads to an increase in the number of road traffic accidents. The present study provides a quantification of this relationship for the case of the Balearic Islands. We obtain the expected positive relationship according to which the increase in population due to the presence of tourists leads to an increase in road accidents. The application of the model should be extended to other tourist areas to assign a tourism responsibility to crash counts.

1.7. References

- Andreescu, M. P., Frost, D. 1998. Weather and Traffic Accidents in Montreal, Canada. *Climate Research* 9, 225-230.
- Bentley, T., Meyer, D., Page, S., Chalmers, D. 2001. Recreational Tourism Injuries among Visitors to New Zealand: an Exploratory Analysis Using Hospital Discharge Data. *Tourism Management* 22, 373-381.
- Brijs, T., Karlis, D., Wets, G. 2008. Studying the Effect of Weather Conditions on Daily Crash Counts Using a Discrete Time-Series Model. *Accident Analysis and Prevention* 40, 1180-1190.
- Cameron, A. C., Trivedi, P. K. 1990. Regression-Based Tests for Over Dispersion in the Poisson Model. *Journal of Econometrics* 46 (3), 347-364.
- Cameron, C., Trivedi, P. K. 2001. Essentials of Count Data Regression, in Badi H. Baltagi ed., *A Companion to Theoretical Econometrics*, 331-348. Blackwell, Oxford (U.K.).
- Cameron, C., Trivedi, P. K. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.
- Carey, M., Aitken, M. 1996. Motorbike Injuries in Bermuda: A Risk for Tourists. *Annals of Emergency Medicine* 28 (4), 424-429.

- Clift, S., Page, S. J. 1996. *Health and the International Tourist*. London: Routledge.
- Dionne, G., Desjardins, D., Laberge-Nadeau, C., Maag, U. 1995. Medical Conditions, Risk Exposure, and Truck Drivers' Accidents: An Analysis with Count Data Regression Models. *Accident Analysis and Prevention* 27 (3), 295-305.
- Dickinson, J. E., Robbins, D. 2008. Representations of Tourism Transport Problems in a Rural Destination. *Tourism Management* 29, 1110–1121.
- Dickinson, J. E., Robbins, D., Fletcher, J. 2009. Representation of Transport a Rural Destination Analysis. *Annals of Tourism Research* 30 (1), 103-123.
- Eisenberg, D. 2004. The Mixed Effects of Precipitation on Traffic Crashes. *Accident Analysis and Prevention* 36, 637-647.
- Fridstrom, L., Ifver, J., Ingebrigsten, S., Kulmala, R., Krogsgard-Thomsen, L. 1995. Measuring the Contribution of Randomness, Exposure, Weather, and Daylight to The Variation in Road Accident Counts. *Accident Analysis and Prevention* 27, 1-20.
- Fridstrom, L. 1999. *Econometric Models of Road Use, Accidents, and Road Investment Decisions, Volume II: An Econometric Model of Car Ownership, Road Use, Accidents, and Their Severity (Essay 3)*. Institute of Transport Economics, Oslo, Norway.
- García-Ferrer A., de Juan, A., Poncela, P. 2006. Forecasting Traffic Accidents Using Disaggregated Data. *International Journal of Forecasting* 22, 203-222.
- Golob, T.F., Recker, W.W., Levine, D.W. 1990. Safety of Freeway Median High Occupancy Vehicle Lanes: A Comparison Of Aggregate and Disaggregate Analyses. *Accident Analysis and Prevention* 22 (1), 19–34.
- Gourieroux, C., Monfort, A., Trognon, C. 1984a. Pseudo-Maximum Likelihood Methods: Theory. *Econometrica*, 52, 681–700.
- Gourieroux, C., Monfort, A., Trognon, C. 1984b. Pseudo-Maximum Likelihood Methods: Applications to Poisson Models. *Econometrica*, 52, 701–720.

EMPIRICAL CHAPTERS

- Gujarati, D. N. 2004. *Gujarati: Basic Econometrics, Fourth Edition*. The McGraw–Hill Companies.
- Gursoy, D., Jurowski, C., Uysal, M. 2002. Resident Attitudes a Structural Modeling Approach. *Annals of Tourism Research* 29 (1), 79–105.
- Hall, D. R. 1999. Conceptualizing Tourism Transport: Inequality and Externality Issues. *Journal of Transport Geography* 7, 181-188.
- Hermans, E., Wets, G., Van den Bossche, F. 2006. Frequency and Severity of Belgian Road Traffic Accidents Studied by State Space Methods. *Journal of Transportation and Statistics* 9 (1), 63-76.
- Howard, R. W. 2009. Risky Business? Asking Tourists What Hazards They Actually Encountered in Thailand. *Tourism Management* 30, 359–365.
- Hutchings, C. B., Knight, S., Reading, J. C. 2003. The Use of Generalized Estimating Equations in the Analysis Of Motor Vehicle Crash Data. *Accident Analysis and Prevention* 35, 3–8.
- Jovanis, P.P., Chang, H.-L. 1989. Disaggregated Model of Highway Accident Occurrence Using Survival Theory. *Accident Analysis and Prevention* 21 (5), 445–458.
- Keay, K., Simmonds, I. 2006. Road Accidents and Rainfall in a Large Australian City. *Accident Analysis and Prevention* 38, 445-454.
- Leggat, P. A., Leggat, F. W. 2003. Reported Fatal and Non-Fatal Incidents Involving Tourists in Thailand, July 1997–June 1999. *Travel Medicine and Infectious Disease* 1, 107–113.
- Levine, N., Kim, K. E., Nitz, L. H. 1995. Daily Fluctuations in Honolulu Motor Vehicle Accidents. *Accident Analysis and Prevention* 27, 785-796.
- Lindberg, G. 2001. Traffic Insurance and Accident Externality Charges. *Journal of Transport Economics and Policy* 35 (3), 399-416.
- Lindberg, G., Ajo, P., Brito da Silva, A., Krupp, R., Osório Nunes, M. R., Peirson, J., Schneglberger, G., Smolders, W. 1999. Calculating Transport Accident

- Costs. Final Report of the Expert Advisors to the High Level Group on Infrastructure Charging (Working Group 3). Sweden.
- Lord, D., Persaud, B. N. 2000. Accident Prediction Models with and Without Trend: Application of the Generalized Estimating Equations Procedure. *Transportation Research Record* 1717, 102–108.
- Lord, D. 2006. Modeling Motor Vehicle Crashes Using Poisson-Gamma Models: Examining The Effects of Low Sample Mean Values And Small Sample Size on the Estimation of the Fixed Dispersion Parameter. *Accident Analysis and Prevention* 38, 751-766.
- Martin, J. L. 2002. Relationship between Crash Rate and Hourly Traffic Flow on Interurban Motorways. *Accident Analysis and Prevention* 34, 619-629.
- Riera, A., Mateu, J. 2007. Aproximación al Volumen de Turismo Residencial en la Comunidad Autónoma de las Illes Balears a Partir del Cómputo de la Carga Demográfica Real. *Estudios Turísticos* 174, 59-71.
- Noland, R. B., Oh, L. 2004. The Effect of Infrastructure and Demographic Change on Traffic-Related Fatalities and Crashes: A Case Study of Illinois County-Level Data. *Accident Analysis and Prevention* 36, 525–532.
- Page, S. J. 2009. Current Issue in Tourism: The Evolution of Travel Medicine Research: A New Research Agenda for Tourism? *Tourism Management* 30 (2), 149–157.
- Page, S. J., Meyer, D. 1996. Tourist Accidents an Exploratory Analysis. *Annals of Tourism Research* 23 (3), 666-690.
- Quddus, M. A. 2008. Time Series Count Data Models: An Empirical Application to Traffic Accidents. *Accident Analysis and Prevention* 40, 1732-1741.
- Ryan, C. 1997. *The Tourist Experience* (1st ed.). London: Cassell.
- UNWTO. 2008. *Tourism Highlights 2008 Edition*.
- UNWTO. 2009. *Facts and Figures Information, Analysis and Know-How. Historical Perspective of World Tourism*. Retrieved April 02, 2009 from <http://unwto.org/facts/eng/historical.htm>.

EMPIRICAL CHAPTERS

- Van den Bossche, F., Wets, G., Brijs, T. 2005a. Predicting Road Crashes Using Calendar Data. Transportation Research Institute, Belgium.
- Van den Bossche, F., Wets, G., Brijs, T. 2005b. Role of Exposure in Analysis of Road Accidents: A Belgian Case Study. Transportation Research Record 1908, 96–103.
- Wang, X., Abdel-Aty, M. 2006. Temporal and Spatial Analyses of Rear-End Crashes at Signalized Intersections. Accident Analysis and Prevention 38, 1137–1150.
- Wilks, J. 1999. International Tourists, Motor Vehicles and Road Safety: A Review of the Literature Leading Up To the Sydney 2000 Olympics. Journal of Travel Medicine 6, 115-121.
- Wilks, J., Watson, B., Faulks, I. J. 1999. International Tourists and Road Safety in Australia: Developing a National Research and Management Programme. Tourism Management 20, 645-654.
- Wooldridge, J. M. 1990. Quasi-Likelihood Methods for Count Data, in Handbook of Applied Econometrics, Volume 2. Edited by M.H. Pesaran and P. Schmidt. Malden, MA: Blackwell, 352-406.
- Wooldridge, J. M. 2000. Introductory Econometrics: A Modern Approach. 2nd Ed. South Western College Publishing, Thomson Learning.

Chapter 2. The responsibility of tourism in traffic congestion and hyper-congestion: a case study from Mallorca, Spain.

Abstract:

This paper explores the relationship between tourism and traffic congestion and hyper-congestion using the case study of Mallorca (Spain), one of the most important resort destinations in the Mediterranean. After discussing different proxies to capture the associated problems to road traffic congestion, different time series models are estimated considering the days of the week, holidays and meteorological determinants jointly with a daily indicator of tourist population pressure. Results show how the tourist pressure variable is an important determinant in explaining the different alternative indicators of traffic congestion and hyper-congestion, for different roads. Hence it is possible to classify the roads in terms of usage by tourists in order to anticipate the levels of traffic intensity, especially during peak periods.

2.1. Introduction

Given the need to understand tourism from a holistic perspective, which implies considering both its “benefits” and its external costs (Dwyer and Forsyth, 1993), a growing number of studies have begun to assign the responsibility of leisure activities and holidays in generating externalities of road transport, such as accidents (Keay and Simmonds, 2006; Levine et al., 1995; Rosselló and Saenz-de-Miera, 2010) and air pollution (Dickinson and Robbins, 2008; Rendeiro and Ramírez, 2010).

Although initially congestion issues were not addressed within the main tourist road transport externalities, recent trends toward a higher use of private or hired cars in the destination (Palmer et al. 2007) and the popularization of the city-break holidays have led to a growing concern and interest about the contribution of tourism to road traffic congestion, fueled by the interest of authorities in applying

EMPIRICAL CHAPTERS

economic instruments for its regulation (Aguiló et al., 2009) conscious of how the presence of congestion can damage the tourist image and how congestion have been recently pointed out as one of the main negative impacts of tourism (Cui and Ryan, 2010). This is of special relevance because each country has its own image which is part of its tourist product, but is also susceptible to the effects of transportation problems (Teye, 1992). One of these problems is traffic congestion, which can reduce the time available for participation in tourism activities and could be perceived as an unsatisfactory experience by visitors, having a negative effect on a possible future visit (Alegre and Cladera, 2006) or even lead visitors to seek out alternative destinations (Dickinson and Robbins, 2008).

Since the seasonal fluctuations in traffic demand are usually affected by the social and economic activities of the area being served by a highway (TRB, 2000), the issue of congestion is especially relevant in tourist destinations that have witnessed the effects of exponential increase in tourist numbers (UNWTO, 2009a; WTTC, 2007). However, while studies of traffic congestion have focused on pricing (Hau, 1998; Li, 2000), alternative ways to measure congestion (Taylor et al., 2000; Wang et al., 2009) and explaining the variables that compose the fundamental diagram of traffic flow (Del Castillo and Benítez, 1995a, 1995b; Hall et al., 1986; Koetse and Rietveld, 2009), through a set of determinants that includes dummies to account for specific holidays and/or vacation periods along the year (Cools et al., 2007; Keay and Simmonds, 2005; Liu and Sharma, 2006), none has included a quantitative measure of the stock of tourists.

This study aims to fill this gap by using a measure of the daily stock of non-resident population as an explanatory variable, which is able to reflect the influence of tourism on traffic congestion, thanks to its seasonal component. Regression analysis based on traffic flow and speed data and congestion and hyper-congestion indicators from different strategically located road stations are modeled using time series analysis and including the tourism pressure indicator. Thus, the influence of tourism is considered both on a temporal and a spatial level. On the one hand, the

temporal dimension is captured by the use of time series data on a daily basis, thus reflecting the seasonal fluctuation explained by tourism. On the other hand, the spatial dimension is taken into account by considering different road traffic counters, four of which are located on roadway segments that represent the influence of tourism on vehicular activity and contrast with a fifth one located on an intersection that carries significant traffic volumes that are less affected by leisure activities. For these purposes, Mallorca is taken as case study.

The chapter is structured as follows: Section 2.2 reviews the variables commonly employed in studies of this nature and makes a brief review of the various methods used to approach the modeling of traffic activity. Section 2.3 presents the model and provides a description of the data used in this study. Section 2.4 presents the results. Finally, concluding remarks and some limitations can be found in Section 2.5.

2.2. Background: The determinants of traffic volume and traffic speed

Within the empirical literature, a set of variables have been identified as predictors of traffic activity, among which, the weather stands out for the significant attention it has deserved. According to Smith et al. (2003), it is widely accepted that the weather plays a significant role in the performance of the surface transportation system and, while extreme weather events, such as snow or thick fog, can bring traffic to a standstill, more common weather events, such as rainfall, have also been shown to impact traffic conditions. It is important to take the weather into consideration because many traffic engineering guidance and methods used to analyze traffic conditions assume standard or ideal conditions (i.e. clear weather) but there are many regions where inclement weather conditions occur during a significant portion of the year (Argawal et al., 2005; Kyte et al., 2001).

The existence of a relationship between traffic fluctuations and the weather is evident but so is the fact that they are interrelated in complex ways, therefore it is

EMPIRICAL CHAPTERS

important to bear in mind that the effects of weather conditions on the traffic are mixed (Hassan and Barker, 1999). Weather conditions can affect traffic in different ways, impairing drivers' visibility, decreasing vehicle stability, reducing vehicle controllability and influencing drivers' behavior and traffic patterns, thus affecting safety (Oh et al., 2002). Furthermore, adverse weather conditions may cause travel disruptions, shifts to other modes and trip adjustments like delays or even cancellation which could result in significant variations in highway traffic volumes (Datla and Sharma, 2008; Hassan and Barker, 1999; Keay and Simmonds, 2005). In addition, adverse weather can degrade the capacities and operating speeds on roadways, resulting in congestion and productivity loss (Argawal et al., 2005).

To account for such impacts of the weather on traffic activity, the variables that most studies have usually considered include precipitation, temperature, sunshine hours, wind, fog and snow. About precipitation, the literature indicates that it has a negative influence on traffic speeds and although some studies consider that this effect appears to be small, the response is larger for heavy rain (Cools et al., 2008; Keay and Simmonds, 2005; Koetse and Rietveld, 2009). The rain can also be associated with reduced traffic flows. In this case, it has been suggested that precipitation does not impact on mean traffic volumes but the reduction of speeds, headway and capacity of roads may manifest itself in the form of fewer vehicles passing the same point on a roadway during inclement rainfall conditions (Butler et al., 2007). Also, wet weather and severe storms can deter motorists from venturing onto the road, leading to a reduction in traffic volume (Keay and Simmonds, 2005).

The temperature is often considered in traffic studies as well and, along with the hours of sunshine, it has been associated to slight increases in traffic activity (Cools et al., 2008). In fact, Hassan and Barker (1999) point out that there exists the possibility of increased road traffic activity on unseasonably warm days or reduced traffic activity under extremely inclement conditions perhaps indicating mode shifts. Likewise, mild winters and warm summers could influence mode shifts, for instance by having a stimulating effect on bicycle use (Cools et al., 2008). Additionally, the

temperature can play an important role in places where the winter is very severe due to their geographical location. In these cases, even though driving conditions may not be affected by severe cold, a reduction in traffic volumes (mainly discretionary trips) might occur due to less desire of travelers to travel during severe cold temperatures because of increased risk associated with extreme cold in general and increased necessity for precautionary measures for safe journeys (Datla and Sharma, 2008). With regard to the effect of temperature on traffic speed, Koetse and Rietveld (2009) sustain that it appears to be small or not existent.

Wind speed and fog are other factors that can have an influence on traffic activity and, according to Cools et al. (2008), are worthwhile investigating. In the case of wind, Keay and Simmonds (2005) find that higher average wind speed leads to a volume reduction, in their daily models. Also, Kyte et al. (2001) obtain an estimated effect of a 9.0 km/h reduction in free-flow speeds for wind speeds above 48 km/h and Liang et al. (1998) find that wind speed reduces driver speed by 1.1 km/h for every km/h of wind speed exceeding 40 km/h. Nevertheless, from their literature review, Koetse and Rietveld (2009) conclude that the effects of wind on traffic speed appear to be small or not existent. In the case of fog, the same study by Koetse and Rietveld (2009) states that reduced visibility causes reductions in traffic speed, and Liang et al. (1998), also find that driver speed is reduced during fog.

The snow is another major determinant in models relating the weather to traffic activity and there seems to be consensus in the literature in that snowy weather is associated with reduced speeds and traffic volume (Keay and Simmonds 2005; Koetse and Rietveld 2009). Studies supporting such assertion include Datla and Sharma (2008), which provides evidence of reductions in traffic volume due to snow in different types of highways; and Ibrahim and Hall (1994) which shows reductions of 3 km/h in free-flow speed, caused by light snow and of 38 to 50 km/h for heavy snow.

EMPIRICAL CHAPTERS

The acknowledgement of day to day variations and differences between weekdays and weekends is also an essential component in traffic models. Dummy variables are the most simple and used formula to consider this issue (Datla and Sharma, 2008; Hassan and Barker, 1999; Keay and Simmonds, 2005; Thomas et al., 2008). In a similar way, the consideration of holidays and/or vacation periods has also been undertaken through the use of dummy variables. While various works have incorporated these variables in order to add explanatory power (Datla and Sharma, 2008; Keay and Simmonds, 2005) regardless of the objective of their study, Liu and Sharma (2006) and Cools et al. (2007) focused their objective on the relevance of including them showing how throughout the year there are specific days or periods when the traffic activity differs in an outstanding manner. Furthermore, it should be highlighted that not only do traffic volumes vary significantly during those specific days or periods but also during the previous and/or the following days (e.g. Liu and Sharma 2006).

In this context, the relevance of incorporating an explicit measure of the tourism activity as an explanatory variable can be understood when considering the importance and magnitude of the tourism sector, particularly in certain regions which have witnessed an exponential growth in the number of tourists in relatively short periods of time, and where traffic congestion has been pointed out as one of the most important externalities from tourism (Aguiló and Rosselló, 2005; Lee et al., 2010; Lorde et al., 2011). In fact, the number of international arrivals have shown an evolution from 25 million international arrivals in 1950 to an estimated 806 million in 2005, corresponding to an average annual growth rate of 6.5% (UNWTO, 2009b); and are expected to keep growing, to reach an estimated 1.6 billion by the year 2020 (UNWTO, 2008). These trends are especially relevant when considering an increase in the amount of tourists using road vehicles. Thus, the present study tries to relate different traffic congestion measures and effects using most of the variables suggested by the literature and including a measure for the amount of tourists.

2.3. Methodology and data

2.3.1. Indicators of congestion

Road traffic congestion is a complex phenomenon which results from the dynamic behavior of and interactions between many road users (Verhoef and Rouwendal, 2004). Described as one of the major liabilities of modern life, congestion is a price that people pay for the various benefits derived from agglomeration of population and economic activity (Lindsay and Verhoef, 2000). Traffic congestion is said to occur when the cost of travel is increased by the presence of other vehicles, either because speeds fall or because greater attention is required to drive safely (Lindsay and Verhoef, 2000).

Congestion manifests in long queues of vehicles saturating the roads on specific times of the day (peak hours) or specific days along the year (e.g. holiday transport to and from big cities) and its main reason is that the decision of each user to enter the highway is taken according to her/his own private benefits, without considering the external costs imposed on the rest of the users of the same infrastructure (De Rus et al., 2003). Further, although many studies have measured congestion externalities borne directly by drivers, more recently researchers have started devoting attention to congestion's broader economic impact on urban areas as a whole (Hymel, 2009).

In view of these considerations regarding the complexity of road traffic congestion and in order to determine the responsibility of tourism in this issue, this chapter considers three different indicators: traffic volume, speed and an artificial variable combining both.

Firstly, as traffic volumes (expressed in vehicles per hour) increase they can reach or exceed the capacity of a given road, which leads to congestion. This idea is usually captured through the use of volume to capacity (V/C) ratios, where traffic volumes are divided by capacity. Although it has been suggested that V/C ratios

EMPIRICAL CHAPTERS

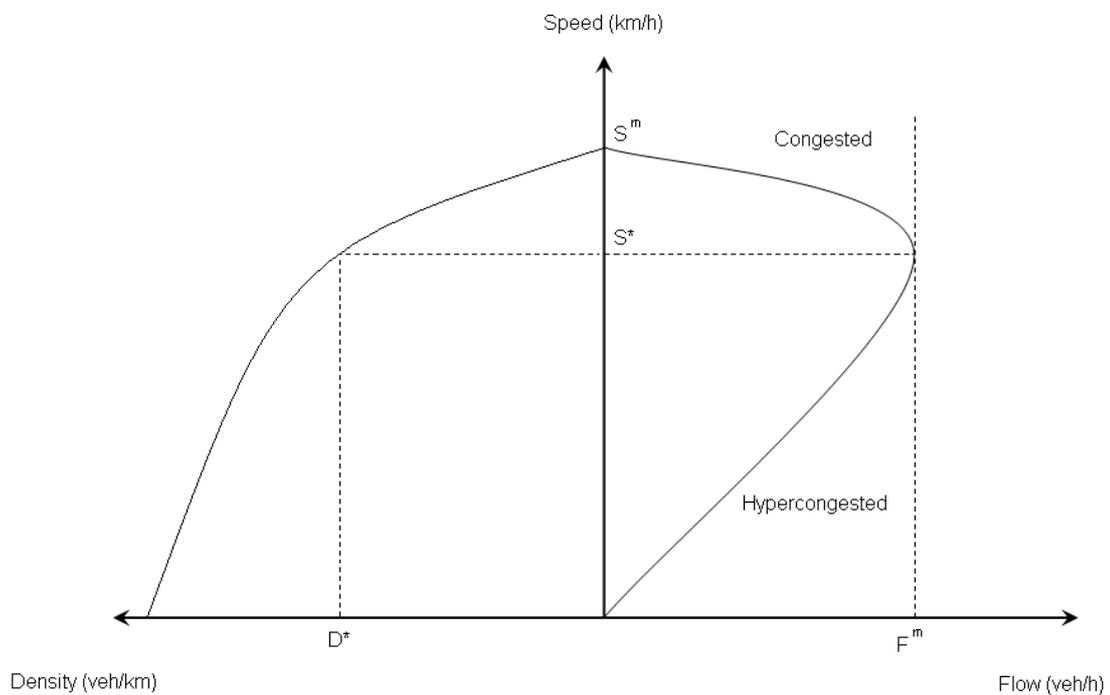
might not be the most suitable way to account for congestion (Wang et al., 2009), they constitute a popular way to proxy this issue, especially due to the virtue of allowing for comparisons across locations or times and the advantage that they are built upon widely available data (Boarnet et al., 1998). Using time series analysis, since the value of capacity is often constant and based on default values from the Highway Capacity Manual (TRB, 2000) that might not entirely reflect the characteristics of the road segments under study, volumes can be used directly because of the analogous behavior exhibited by both series.

Secondly, one of the most visible external effects of congestion stems from the time lost during journeys but the availability of this variable depends on the use of questionnaires (Cantos and Alvarez, 2009) or specific technology such as GPS-equipped probe vehicles (Taylor et al., 2000). Therefore, vehicle speed (in kilometers per hour) can be used as dependent variable because the external effect that drivers impose upon each other is that the generalized travel costs of each driver increase by the presence of other drivers and an important component of these costs is the travel time, which is inversely related to the speed during the trip (Verhoef et al., 1999).

Thirdly, since neither high volume nor low speed can always be related with the concept of congestion (for instance, a road can be able to carry high levels of traffic without congestion; or lower speeds could be originated by different causes) two additional indicators are developed based on the speed-flow diagram (Figure 2.1) to measure congestion and hyper-congestion. In concrete, the indicators are built upon scatter plots of our data and the intuition behind the observed shape of the curves is based on the standard economic model for analyzing traffic congestion, which describes traffic streams by three variables, density (number of vehicles per unit of distance), speed and flow (number of vehicles per unit of time), through the *fundamental diagram of traffic flow* (Haight, 1963). From the diagram it can be shown how as the density increases, speeds decrease until a maximum combination of density and speed is reached, corresponding with the point where the speed-flow

curve bends. This denotes a negative relationship between flow and speed on the upper branch part of the curve where more congestion is observed as flows increase and speed decreases. On the other hand, when a capacity limit is exceeded somewhere in the system, local queuing begins, which becomes more severe the more cars are added to the input flows (Small and Chu, 2003). Empirical measurements indicate this situation on the lower branch of the curve, known as hyper-congestion, in which speed increases with flow (Small and Chu, 2003).

Figure 2.1: Speed-flow-density relationship



Source: Own elaboration, based on Haight (1963, in Lindsay and Verhoef, 2000).

The fundamental diagram, has received much criticism mainly because the standard model of traffic congestion is static and provides a description of traffic conditions that evolve only slowly. Hence it is not the adequate framework for studies of congestion and hyper-congestion aimed at understanding travel costs and determining highway tolls in more realistic, dynamic situations (Lindsay and Verhoef, 2000; Verhoef, 2003). However, there are several examples, including Hall

EMPIRICAL CHAPTERS

et al. (1986), Li (2002), Nair et al. (2001) and Smith et al. (1996), in which the empirical speed-density and speed-flow relationships are used to characterize congested and hyper-congested conditions.

Although the upper branch of a speed-flow curve is often referred to as uncongested, unrestricted or free flow, it should be acknowledged that data pairs on this area can reflect congestion, in particular those below the corresponding free flow speed (Lindsay and Verhoef, 2000). For this study, in the cases where scatter plots only exhibit “normal congestion”, the indicator is defined by:

$$C = (S_m - S_i) \times F_i \quad (2.1)$$

Where, C stands for the congestion indicator; S_m is the highest observed speed; and S_i and F_i are each pair of observed speeds and flows (volumes), respectively. According to this measure, more congestion is observed on the speed-flow curve as the distance between the highest and the observed speeds (multiplied by the corresponding flow) increases.

Under hyper-congestion, the measure of expression (2.1) is not valid and has to be adapted to account for this special situation. In order to address this problem it is necessary to determine the point where the speed-flow curve bends, to be able to clearly differentiate the congested region from the hyper-congested one. For this reason a fourth degree polynomial can be estimated using the following formulation:

$$F = a_0 + a_1 S_i + a_2 S_i^2 + a_3 S_i^3 + a_4 S_i^4 \quad (2.2)$$

Where a_0, a_1, \dots, a_4 are parameters to be determined and the speed that maximizes the estimated speed-flow curve (S^*) can be easily derived. Then, in order to obtain a measure that considers both the normally congested and the hyper-

congested observations, separated by the speed that maximizes the speed-flow curve (S^*), the following formula is used:

$$y = \begin{cases} C = (S_m - S_i) \times F_i & \text{if } S_i > S^* \\ HC = [F_m \times (S_m - S^*)] + (S^* - S_i) \times (F_m - F_i) & \text{if } S_i \leq S^* \end{cases} \quad (2.3)$$

Where, HC stands for hypercongestion; and F_m and S_m are the highest observed flow and speed, respectively. What the indicator does for hyper-congested observations is considering the product of the highest flow (F_m) by the difference between the highest speed (S_m) and the maximizing speed (S^*). This is the reference point, which accounts for the congestion present at S^* , and the indicator increases with the right-hand addend: lower flows and speeds, observed as we move to the left, would denote more congestion. Observations in this region correspond with the segment of a speed-density curve where the amount of vehicles per kilometer increase so much that speed starts dropping toward zero and less vehicles cross the same point per unit of time.

Finally, for the applied exercise, in order to account only for those observations that lie below S^* , a special variable was considered including only hyper congestion values and ascribing a zero value for the rest, thus excluding “normally congested” observations. Since the speed-flow curves for the stations of Lluçmajor are the only ones that exhibit a bending shape (as can be appreciated in Figure 2.4 of the next section), with both a congested and a hyper-congested branch, the indicator of hyper-congestion will be used only in these two cases.

2.3.2. Method

Following the previous considerations, the model analyzed under a daily time-dimensional perspective, can be analytically expressed as:

$$CON_t^i = f(x_t^z, D_t^k, POP_t^j) \quad (2.4)$$

Where CON_t^i represents the i traffic congestion indicators analyzed above, indexed in time t ; x_t^z stands for z weather variables; D_t^k are k day-of-the week and special holiday dummies; and POP_t^j represents the population pressure with special emphasis on the distinction between residents and tourists.

For estimation purposes, volume, speed and the congestion plus hyper-congestion (HC) cases can be approached using standard linear regression analysis following other similar studies (Datla and Sharma, 2008; Keay and Simmonds, 2005). In this case the quantification of the relationship between the explanatory and the congestion measures is expressed as follows:

$$CON_t^i = \alpha + \beta_z x_t^z + \beta_k D_t^k + \beta_j POP_t^j + u_t \quad (2.5)$$

Where the vectors α and β contain the parameters to be determined and u_t is an error term, distributed normally and independently.

In the cases where only hyper-congestion is modeled, as a function of the aforementioned explanatory variables, there is a key methodological issue that has to be considered: if there are days when the combination of low traffic volume and low speed does not occur, we are talking about days when there is no hyper-congestion, hence the dependent variable would take on zero values with positive probability. According to Wooldridge (2000), nothing prevents us from using a linear model for hyper-congestion (HC). In fact, a linear model might be a good approximation to $E(HC|x_1, x_2, \dots, x_k)$, especially for x_j near the mean values. But we would possibly obtain negative fitted values, which leads to negative predictions for HC .

Alternatively, let's say that those days when hyper-congestion is zero, are taken as days for which there are no data (n_1) and we have such data only on days when hyper-congestion occurs (n_2). In a situation like this, we could not estimate an

Ordinary Least Squares regression without considering the n_1 observations. If we did, the estimates of the parameters would be biased and inconsistent; if a regression line based only on the n_2 observations were estimated, the resulting intercept and slope coefficients would be different than if all the (n_1+n_2) observations were taken into account.

A sample with these characteristics, in which information on the regressand is available only for some observations, is known as a censored sample (Gujarati, 2004). A censored regression to fit such data, better known as a Tobit model, can be expressed as

$$HC_i = \begin{cases} 0 & \text{if } HC_i^* \leq 0 \\ HC_i^* & \text{if } HC_i^* > 0 \end{cases} \quad (2.6)$$

for the model $HC_i^* = x_i'\beta + \sigma\varepsilon$,

where σ is a scale parameter. Note that in contrast with binary dependent variable models, the scale parameter σ is identified, and will be estimated along with the β .

All negative values of HC_i^* are coded as a single value 0. This situation differs from a truncated regression model where all negative values of HC_i^* are simply dropped from the sample. We say that these data are left censored. In this study, both left and right censoring at arbitrary limit points is considered, so that

$$HC_i = \begin{cases} \underline{c}_i & \text{if } HC_i^* \leq \underline{c}_i \\ HC_i^* & \text{if } \underline{c}_i < HC_i^* < \bar{c}_i \\ \bar{c}_i & \text{if } \bar{c}_i \leq HC_i^* \end{cases} \quad (2.7)$$

EMPIRICAL CHAPTERS

where \underline{c}_i , \bar{c}_i are fixed numbers representing the censoring points. If there is no left censoring, then we can set $\underline{c}_i = -\infty$. If there is no right censoring, then we set $\bar{c}_i = \infty$. The canonical Tobit model is a special case with $\underline{c}_i = 0$ and $\bar{c}_i = \infty$. The parameters β , σ are estimated by maximizing the log likelihood function

$$\begin{aligned} \ell(\beta, \sigma) = & \sum_{i \ni HC_i = \underline{c}_i} \log F((\underline{c}_i - x_i' \beta) / \sigma) + \\ & \sum_{i \ni \underline{c}_i < HC_i < \bar{c}_i} \log f((HC_i - x_i' \beta) / \sigma) + \sum_{i \ni HC_i < \bar{c}_i} \log(1 - F((\bar{c}_i - x_i' \beta) / \sigma)) \end{aligned} \quad (2.8)$$

where f , F are the density and cumulative distribution functions of ε .

Taking these considerations into account, the final expression for quantifying the relationship between the explanatory and the hyper-congestion measures is expressed as follows:

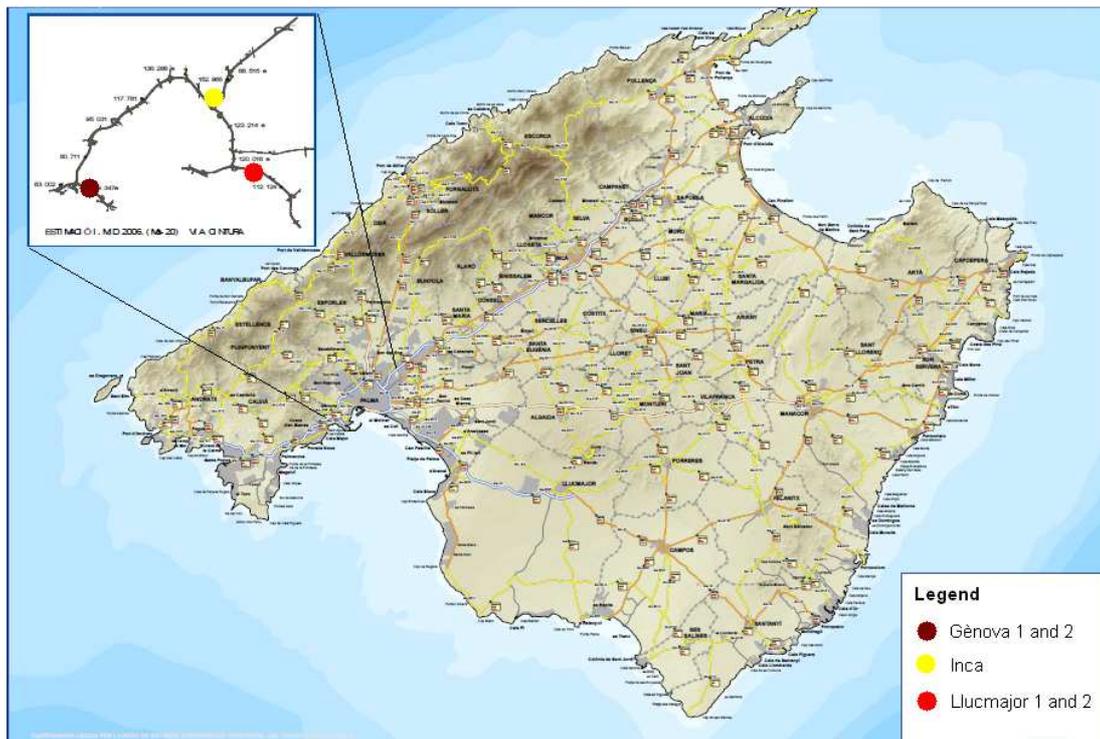
$$HC_i^* = \alpha + \beta_z x_i^z + \beta_k D_i^k + \beta_j POP_i^j + \sigma \varepsilon_i \quad (2.9)$$

2.3.3. Data

For the special purpose of this chapter, data on traffic volume and speed were collected at hourly and sub-hourly (every 15 minutes) levels, ranging from January 1, 2006 to December 31, 2006; and at daily level for the period January 1, 2004 to December 31, 2006. Sub-hourly data were used for the elaboration of the congestion and hyper-congestion indicators which were averaged in order to obtain a daily indicator that could be comparable with the rest of the variables. The volume and speed databases were obtained from five stations installed by the *Consell Insular de Mallorca* and the *Ministerio de Fomento*, located in the outskirts of Palma (the capital city of Mallorca), on roads that capture the traffic to and from the main tourist sites

and one located on at an intersection that captures traffic from the main populations of central-eastern Mallorca (Figure 2.2).

Figure 2.2 Location of traffic counters



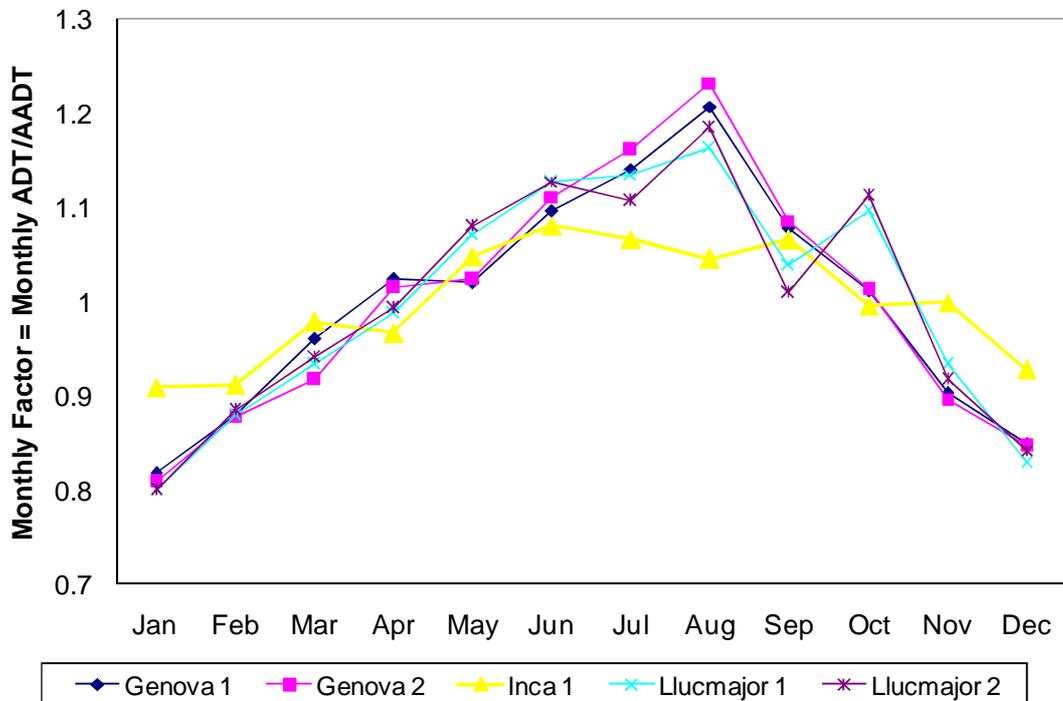
Source: Consell Insular de Mallorca

Gènova (gen_1 and gen_2) stations, are located in the area where highway M-20 intersects with M-1, a point that serves as entrance to the main populations of western Mallorca, namely Andratx and Calvià, which have a Mean Daily Intensity (MDI) –i.e. daily traffic volume– of 35,270 vehicles per day which is above the average and further increased during summer (MDI: 40,367) because it is surrounded by important tourist locations: Cala Major, Palma Nova, Magaluf, Peguera, Santa Ponça and Port d’Andratx. The Lluçmajor (lluc_1 and lluc_2) stations are located near the *Levante* loop which, through highway M-19, serves as entrance to incoming vehicles from Coll d’en Rabassa, Can Pastilla (both belong to the capital), the International Airport of Mallorca, Arenal, Lluçmajor, Campos and

EMPIRICAL CHAPTERS

Santanyí. This area presents the most substantial traffic fluctuations, oscillating between 49,991 vehicles per day during the 1st quarter of the year and 64,310 during the 3rd quarter. The Inca (inca_1) station, located on the north of the capital, is on the point that serves as entrance (through highway MA-13) for a significant proportion of incoming traffic from Mallorca’s flatlands and presents the lowest seasonal variations (24,892 vehicles per day on summer versus 22,787 on winter) which clearly indicates the lesser extent to which the population centers that use this road are oriented towards tourism activities (Riera and Aguiló, 2009). The comparison of monthly evolution in traffic volumes, amongst different road segments, reflects that Gènova and Lluçmajor indeed carry more recreational traffic than Inca and this can be seen on Figure 2.3 (traffic volumes are measured through a monthly factor obtained by dividing a month’s average daily traffic by the annual average daily traffic).

Figure 2.3 Monthly variation pattern of traffic at the selected road segment

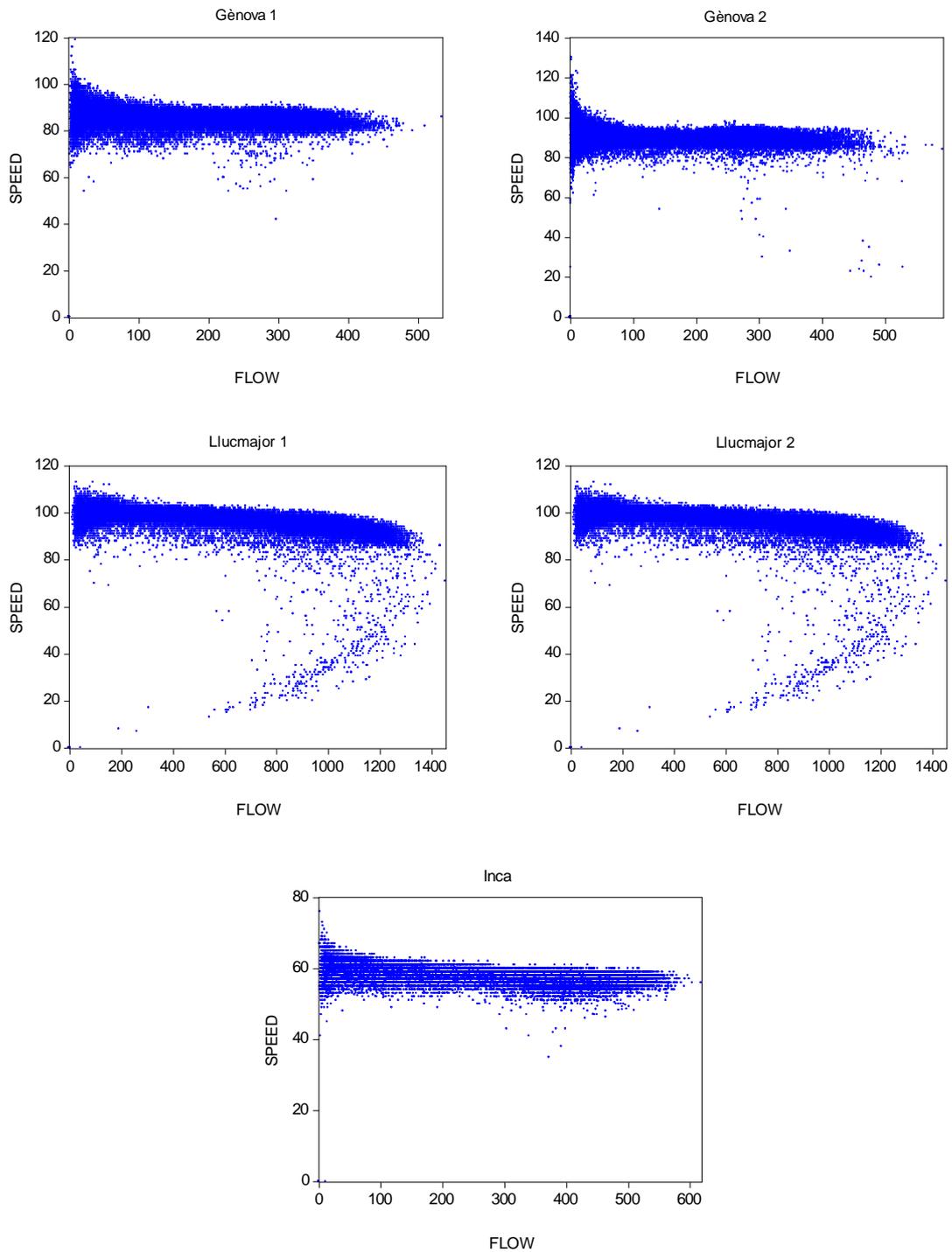


The distinction between these types of roads should be made because, as the Highway Capacity Manual (TRB, 2000) suggests, seasonal peaks in traffic demand are of importance, particularly for recreational facilities. Highways serving beach resort areas may be virtually unused during much of the year, only to be subject to oversaturated conditions during peak summer periods. Thus, while recreational traffic creates the greatest variation in volume patterns, commuter and business-oriented travel occurs in more uniform patterns (TRB, 2000). Keeping this distinction in mind, it should be expected that the roads of Gènova and Lluçmajor are more affected by the presence of tourists than that of residents, while the opposite should be expected in Inca.

Traffic volumes and speed are shown on scatter plots in Figure 2.4. Data from Lluçmajor exhibit the backward bending relationship usually observed on speed-flow diagrams and both the congestion and hyper-congestion segments can be clearly appreciated. Meanwhile, the data from the stations of Gènova 1 and 2 and Inca are almost entirely concentrated in the area known as congested, or uncongested in the engineering literature (Lindsay and Verhoef, 2000). It is important to point out that the speeds and traffic volumes vary across the different stations, a fact that is explained by the fact that the stations of Gènova and Lluçmajor are located on multilane highways, while the station of Inca is located on a ramp. According to the Highway Capacity Manual (TRB, 2000), multilane highways typically are located in suburban communities, leading into central cities, or along high-volume rural corridors connecting two cities or two significant activities that generate a substantial number of daily trips. Multilane highways usually have free-flow speeds (FFS) (i.e. the mean speed of passenger cars measured during low to moderate flows) higher than 100 km/h. On the other hand, ramps are roadways of limited length and width, where passing is often not possible and free-flow speed is frequently lower than on other types of roads.

EMPIRICAL CHAPTERS

Figure 2.4 Sub-hourly speed-flow scatter-plots



Meteorological variables were provided by the *Instituto Nacional de Meteorología*, the Spanish official meteorological bureau, and are referred to the airport station, located near Palma. For the case of the tourism pressure variable, the Daily Indicator of Human Pressure, split into residents and tourists, is used (see Part I).

2.4. Empirical Analysis

Based on the methodology and data collected and described above, equations for each road station were estimated. Explanatory variables included in the final estimations are:

- Precipitation in liters per square meter, wind speed and maximum temperature in Celsius as weather determining variables.
- Dummies from Monday (D1) to Sunday (D7), and Friday is chosen as the reference day to control for different levels of exposure observed over the week; and selected specific public/school holidays.
- Stock of resident population (DIHP Residents) and Stock of tourists (DIHP Tourists) in the Island of Mallorca.

2.4.1. Results for traffic volume

Results for the traffic volume indicators are presented in Table 2.1. Since results from the static equation showed a high level of first order autocorrelation, a lag of the independent variable was included as an explanatory variable resulting in a dynamic estimation. A more complex underlying lag structure of residuals was rejected in view of the Breusch-Godfrey Serial Correlation LM test, and the Breusch-Pagan-Godfrey and Harvey tests for heteroskedasticity. Additionally, the Ramsey RESET test was applied in order to validate the stability of the estimated parameters across the sample.

EMPIRICAL CHAPTERS

Table 2.1 Estimations for traffic volume

Counter	Gènova 1		Gènova 2		Inca 1		Llucmajor 1		Llucmajor 2	
	Coefficient	P-value								
Intercept	1196.017	0.4999	-1751.029	0.3216	11386.85	0.000	-5878.293	0.3079	-14779.42	0.0007
Day 1	2741.315	0.000	2489.284	0.000	595.944	0.018	1367.925	0.0319	1896.042	0.0001
Day 2	-665.7977	0.000	-563.8694	0.0002	-964.6574	0.000	-3140.584	0.000	-2522.61	0.000
Day 3	-767.8196	0.000	-692.7847	0.000	-1082.073	0.000	-3856.668	0.000	-2653.062	0.000
Day 4	-610.1684	0.000	-460.1868	0.0019	-1037.023	0.000	-3048.275	0.000	-1938.617	0.000
Day 6	-2852.812	0.000	-2780.346	0.000	-6438.353	0.000	-12786.79	0.000	-10770.64	0.000
Day 7	-6590.405	0.000	-6595.013	0.000	-11532.88	0.000	-17137.72	0.000	-13748.35	0.000
Precipitation	5.836322	0.498	10.52018	0.2172	-12.86234	0.2681	-72.24721	0.0342	-87.35759	0.0006
Max Temperature	12.95743	0.3115	17.89974	0.1576	102.1165	0.000	223.2956	0.000	220.1263	0.000
DIHP tourists	0.005105	0.000	0.00681	0.000	0.002061	0.0044	0.021628	0.000	0.023659	0.000
DIHP residents	0.00803	0.0012	0.012671	0.000	0.012797	0.000	0.052246	0.000	0.064663	0.000
Volume (-1)	0.621346	0.000	0.540219	0.000	0.13109	0.000	0.384725	0.000	0.322376	0.000
Holidays										
January 6	-7590.646	0.000	-7898.894	0.000	-12693.81	0.000	-20209	0.000	-19471.96	0.000
January 20	-4767.02	0.000	-5385.632	0.000	-8721.864	0.000	-16756.76	0.000	-17769.5	0.000
March 1	-6456.237	0.000	-6446.847	0.000	-12449.25	0.000	-17542.62	0.000	-15947.2	0.000
April 17	-5769.361	0.000	-5519.196	0.000	-12305.7	0.000	-11475.64	0.000	-8482.284	0.000
May 1	-4974.031	0.000	-4724.776	0.000	-12354.24	0.000	-10856.26	0.000	-9348.526	0.000
August 15	-4599.606	0.000	-4395.17	0.000	-10142.26	0.000	-13409.8	0.000	-10194.93	0.000
October 12	-6293.444	0.000	-6141.311	0.000	-11229.92	0.000	-16720.67	0.000	-15361.39	0.000
November 1	-6829.242	0.000	-7090.178	0.000	-12426.07	0.000	-19836.52	0.000	-16347.18	0.000
December 6	-5727.227	0.000	-5913.239	0.000	-10365.41	0.000	-16593.43	0.000	-15125.75	0.000
December 8	-5695.087	0.000	-5961.291	0.000	-12041.77	0.000	-17586.03	0.000	-16709.12	0.000
December 25	-7111.256	0.000	-6783.755	0.000	-7464.228	0.000	-10606.14	0.0104	-11083.62	0.0003
December 26	-4716.502	0.000	-4874.891	0.000	-12284.02	0.000	-16341.79	0.0001	-15104.03	0.000
Good Friday	-4852.254	0.000	-4754.157	0.000	-12883.72	0.000	-15375.75	0.000	-14542.57	0.000
Equation Statistics										
R-squared	0.888434		0.890238		0.916204		0.857917		0.89685	
Mean dep. Var.	18405.69		17180.44		23217.39		58558.17		55813.35	
S.D. dep. var	3884.939		3871.476		4810.92		10756.17		9413.449	
No. of observations	1081		1079		1055		1006		1001	
AIC	17.21979		17.19663		17.36224		19.50171		18.91504	
Period	Jan '04 - Dec, '06									

Through the signs of the estimated parameters for the day-of-the-week dummies it can be observed that the most active day of the week in terms of traffic is Friday. This contrasts with evidence presented in the Highway Capacity Manual (TRB, 2000), according to which Sunday is the most active day for recreational routes. However, the same evidence shows that there is a peak in traffic activity on Fridays, when it is higher than the other weekdays and Saturday as well. It is also important to note that Inca cannot be defined as a recreational road, yet it also presents higher volumes on Fridays. Besides the day of the week, the inclusion of dummies to account for national and local holidays showed that, during these days traffic volumes are lower.

Regarding the weather variables, it can be seen that precipitation is a significant predictor only in the case of Llucmajor, where the sign follows the

expected relationship, indicating that lower traffic volumes can be associated to an increase in the amount of rainfall. As for the temperature, it turned out to be a significant predictor in three cases (Inca, Lluçmajor 1 and Lluçmajor 2), indicating that maximum temperatures are positively related with traffic volumes, a result that has also been found in similar studies.

Both the population stock and the tourist stock variables were significant determinants of traffic volumes in all cases, indicating a positive relationship between population and traffic volume. The distinction between tourist and resident population shows, in a direct way, the impact of tourists in determining traffic volumes.

2.4.2. Results for speed

Results for speed are presented in Table 2.2. Again, because of the high level of first order autocorrelation shown in the static estimation, a dynamic estimation including the first lag of the independent variable was carried out, allowing rejecting the hypothesis of more complex lag structures in view of the statistical tests mentioned in the case of volume.

For traffic speed, evidence regarding the day-of-the-week dummies is mixed, suggesting that the lowest speeds occur during Fridays only in the case of Lluçmajor 1 and during Mondays in the cases of Inca and Lluçmajor 2. Meanwhile, from the results for the stations of Gènova, since most of the day-of-the-week dummies are not significant, it can only be said that speeds are lower on Fridays than on Saturdays. It is also possible to observe that, higher speeds can be expected on the specific holidays considered.

EMPIRICAL CHAPTERS

Table 2.2 Estimations for average daily speed

Counter	Gènova 1		Gènova 2		Inca		Llucmajor 1		Llucmajor 2	
	Coefficient	P-value								
Intercept	43.021	0.000	64.95212	0.000	34.81725	0.000	68.7994	0.000	57.106	0.000
Day 1	0.170013	0.207	0.400569	0.0035	-0.4144	0.0059	0.518023	0.0614	-1.846353	0.000
Day 2	-0.019148	0.8851	0.200337	0.1406	0.213962	0.1333	1.061931	0.0001	0.669272	0.0913
Day 3	0.082827	0.5302	0.214015	0.1148	0.160633	0.2589	0.993393	0.0002	1.386871	0.0005
Day 4	-0.032146	0.8081	0.199873	0.142	0.138691	0.3312	1.166189	0.000	1.349613	0.0007
Day 6	0.581015	0.000	0.364942	0.0071	1.228832	0.000	3.601073	0.000	7.28537	0.000
Day 7	0.801025	0.000	0.181804	0.1808	1.496761	0.000	3.152866	0.000	3.25264	0.000
Wind Speed	-0.012169	0.0001	-0.010202	0.0011	-0.00425	0.1944	-0.009846	0.1074	-0.014942	0.1032
Precipitation	-0.105596	0.000	-0.115855	0.000	-0.05939	0.000	-0.188692	0.000	-0.191041	0.000
Max. Temperature	0.020956	0.0631	0.066043	0.000	0.036746	0.0029	-0.025418	0.2579	-0.071881	0.0334
DIHP tourists	-1.2×10^{-6}	0.0466	-1.7×10^{-6}	0.0055	-2.1×10^{-6}	0.0013	-4.2×10^{-6}	0.0004	-8.1×10^{-6}	0.000
DIHP residents	3.1×10^{-6}	0.1423	7.2×10^{-6}	0.001	-5.4×10^{-6}	0.0217	-1.0×10^{-5}	0.0166	-3.0×10^{-5}	0.000
Speed (-1)	0.468432	0.000	0.202184	0.000	0.447057	0.000	0.364996	0.000	0.63291	0.000
Holidays										
January 6	1.388933	0.04	0.049169	0.9436	1.942789	0.0077	2.993025	0.0252	5.069651	0.0114
April 17	2.777116	0.0001	2.043023	0.004	2.768712	0.0002	5.286527	0.0001	6.996877	0.0007
Holy Thursday	1.235013	0.0681	0.688284	0.3225	1.926046	0.0084	3.378385	0.0118	7.819002	0.0001
Good Friday	0.887073	0.1905	0.811405	0.2437	1.965616	0.0073	2.252143	0.0928	4.454049	0.0268
Equation Statistics										
R-squared	0.461126		0.315322		0.422107		0.451601		0.717758	
Mean dep. Var.	84.98992		89.07254		56.99477		95.88491		89.76083	
S.D. dep. var	1.565948		1.428848		1.631079		3.067998		6.407998	
No. of observations	1091		1089		1089		1060		1062	
AIC	3.146833		3.203111		3.298296		4.510308		5.319055	
Period	Jan '04 - Dec, '06									

In the case of weather related variables, the expected relationships were found for precipitation, indicating that speed is reduced on rainy days. This relationship is also negative in the case of wind speed; however, this variable is not significant in three cases (Inca, Llucmajor 1 and Llucmajor 2). Unlike precipitation and wind speed, the signs for the variable “maximum temperature” are contradictory across estimations. The indicators of the number of people were significant as well. In the case of tourist population, the coefficients are significant and their signs indicate that speeds decrease as tourists increase. In the case of residents, however, the signs are contradictory across estimations.

2.4.3. Results for congestion and hyper-congestion

Results for the congestion indicator are presented in Table 2.3. Using the same considerations for the lag structure of the residuals mentioned for the above cases, the first lag of the dependent variable was used as an independent variable,

although for Inca and Lluçmajor 2 the introduction of lags were not necessary, making it possible to keep the static equations.

Table 2.3 Estimations for congestion

Indicator Counter	Congestion Gènova 1		Congestion Gènova 2		Congestion Inca 1		Congestion + Hyper Congestion			
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Lluçmajor 1		Lluçmajor 2	
Intercept	1085.022	0.6027	1114.069	0.5948	-1398.824	0.3426	-51139.22	0.000	-10791.06	0.2018
Day 1	-404.0948	0.0191	-394.7182	0.0224	-257.5059	0.0029	-642.119	0.4308	-2956.833	0.000
Day 2	-677.6811	0.000	-676.3357	0.000	-288.6672	0.0008	-1462.595	0.0272	-2512.124	0.000
Day 3	-387.1982	0.0018	-385.3374	0.002	-269.652	0.0019	-2009.098	0.0028	-2477.725	0.000
Day 4	-312.2348	0.0103	-310.8098	0.011	-205.0592	0.0172	-1533.503	0.0215	-2260.856	0.000
Day 6	-1002.819	0.000	-1004.203	0.000	-1432.771	0.000	-9944.769	0.000	-5847.035	0.000
Day 7	-2982.215	0.000	-2968.245	0.000	-2677.783	0.000	-10448.67	0.000	-7639.903	0.000
Wind Speed	6.265066	0.0293	6.128149	0.0339	-1.915431	0.3405	31.51421	0.045	18.1351	0.1376
Precipitation	17.30728	0.0017	17.25524	0.0019	12.21044	0.0018	71.99745	0.0169	127.4883	0.0011
Max. Temperature	7.093483	0.5516	7.777572	0.516	-22.70271	0.0078	95.54606	0.1444	37.26724	0.4669
DIHP tourists	0.004408	0.000	0.004351	0.000	0.003078	0.000	0.035701	0.000	0.016631	0.000
DIHP residents	0.006046	0.0365	0.005968	0.0399	0.009337	0.000	0.085264	0.000	0.029705	0.0109
Congestion (-1)	0.110232	0.0149	0.114241	0.012	-	-	0.187555	0.0001	-	-
Holidays										
January 6	-3270.577	0.000	-3273.87	0.000	-2917.639	0.000	-10405.12	0.002	-7447.648	-3.016
January 20	-2076.879	0.0007	-1768.987	0.004	-1830.893	0.000	-8894.867	0.008	-5764.531	-2.348
March 1	-2205.327	0.0003	-2208.376	0.0003	-2738.577	0.000	-8483.631	0.0108	-3828.872	-1.572
April 17	-2597.363	0.0001	-2595.847	0.0001	-3318.801	0.000	-12769.87	0.0005	-9383.657	-3.047
May 1	-2006.695	0.001	-2009.554	0.001	-2775.855	0.000	-7690.292	0.0206	-4388.722	-1.805
October 12	-2592.521	0.000	-2588.879	0.000	-4446.403	0.000	-16976.65	0.000	-6744.346	-2.771
November 1	-2704.666	0.000	-2712.917	0.000	-2542.059	0.000	-13287.09	0.0001	-6363.951	-2.621
December 6		0.000	-2691.739	0.000	-2389.209	0.000	-10588.05	0.0016	-	-
December 25	-2934.913	0.000	-2940.059	0.000	-2604.054	0.000	-10970.91	0.001	-	-
December 26	-2486.765	0.0001	-2479.349	0.0001	-2647.773	0.000	-8379.872	0.0127	-	-
Good Friday	-2183.571	0.0003	-2185.336	0.0003	-3110.648	0.000	-11343.77	0.0007	-6966.584	-2.869
Equation Statistics										
R-squared	0.830891		0.828502		0.859282		0.852579		0.697917	
Mean dep. Var.	6679.571		6682.056		4682.223		22436.86		12827.11	
S.D. dep. var	1395.616		1392.633		1094.486		8210.4		4219.75	
No. of observations	357		357		361		359		314	
AIC	15.6745		15.68425		14.99761		19.08064		18.46009	
Period	Jan '06 - Dec, '06		Jan '06 - Dec, '06		Jan '06 - Dec, '06		Jan '06 - Dec, '06		Jan '06 - Dec, '06	

Results for the congestion indicators suggest that congestion levels are higher on Fridays than the rest of the week, while they are lower on specific holidays. The weather also plays a significant role in determining congestion levels and, while precipitation was significant in all cases and its sign would suggest that more rainfall can be associated to higher congestion, the signs of wind speed and temperature vary across locations. Thus, while higher temperatures and wind speeds could be associated to higher congestion levels on most recreational roads (Gènova and Lluçmajor), the relationship would be the opposite for Inca. This relationship is

EMPIRICAL CHAPTERS

straightforward in the sense that recreational roads are more congested when the weather conditions are most favorable for tourism.

In this case, again, the stock of residents and non-residents were significant and showed the expected sign, reflecting a positive relationship between congestion and both resident and tourist population. According to the results, the increase in resident and tourist population would lead to increases in the levels of road traffic congestion.

Table 2.4 Estimations for hyper-congestion

Counter	Llucmajor 1		Llucmajor 2	
	Coefficient	P-value	Coefficient	P-value
Intercept	-2309.056	0.0851	-12916.13	0.0213
Day 1	-102.6529	0.2177	-598.9674	0.0644
Day 2	-81.10834	0.3317	-544.9851	0.0863
Day 3	-119.2595	0.1535	-470.2396	0.1298
Day 4	-136.5321	0.1024	-561.1657	0.0749
Day 6	-548.5105	0.000	-1354.959	0.0006
Day 7	-446.0578	0.000	-1594.228	0.0003
Precipitation	3.63298	0.354	45.26264	0.0001
DIHP tourists	0.00204	0.000	0.006279	0.000
DIHP residents	0.00301	0.0841	0.014872	0.0396
Hyper-congestion (-1)	0.168558	0.008	-	-
Equation Statistics				
Adjusted R-squared	0.30934		0.104503	
Mean dep. Var.	273.5748		119.9838	
S.D. dep. var	426.0308		465.0414	
Left censored obs	95		293	
Uncensored obs	269		72	
AIC	11.32615		3.813998	
Period	Jan '06 - Dec, '06		Jan '06 - Dec, '06	

Finally results for the hyper-congestion indicator are presented in Table 2.4, where the only measurement points (Llucmajor 1 and Llucmajor 2) that showed a considerable number of hyper-congested observations are presented. Results show how weather does not play a significant role in determining hyper-congestion with the only exception of the amount of rainfall which is significant in Llucmajor 2. The weekly behavior obtained when analyzing the above indicators was also observed in

the case of hyper-congestion. However, the special holiday variables could not be included because no hyper-congestion observations were found in those special days. Finally, the expected relationship with respect to the stock of residents and non-residents was found.

2.5. Discussion and conclusions

The purpose of this work has been to incorporate a quantitative measure to account for the influence of tourism on the generation of a specific road transport related externality (congestion) in Mallorca, Spain. For this task proxies like flow, speed and an indicator developed to assign each observation along a speed-flow diagram its corresponding level of congestion, were related to a set of popular and previously used explanatory variables and two measures of the resident and tourist population, that accurately reflect the strong seasonal fluctuation in the amounts of residents and non-residents.

Using time series analysis, results showed how, in accordance with the previous applied literature, some weather related variables can be associated to traffic congestion levels. Therefore, it has provided additional evidence on how the amount of rainfall, wind speed and temperature could be related to higher levels of traffic congestion.

Nevertheless, in general terms, the evidence regarding how the weather affects the different proxies of congestion is rather mixed. In addition, it should be noted that, while the relevance of including “maximum temperature” as an explanatory variable is evident from the literature review, due to its high correlation with the indicator of tourism pressure, it turned out to be non-significant in most of the models.

Regarding the fluctuations in levels of traffic along the week, the results do not follow what the literature suggests in the sense that Sunday is not the most congested day of the week. Instead, Fridays are the days with higher congestion and

EMPIRICAL CHAPTERS

traffic volumes, a result that is not surprising for recreational roads where Friday is usually the second highest day of the week in terms of traffic activity. Although these results might differ from those for main recreational roads around the most important cities in the world, it should be noted that in the case under study leisure use is linked to tourists that stay for an average of 8 days in the Island and probably do not change their mobility behavior from weekdays to weekends. In any case, as in other regions, results suggest that congestion levels are lower and speeds are higher during holidays.

The results also show how tourism can be associated to different variables that account for road traffic congestion and confirmed the expected relationships between this externality and tourism, showing that increasing numbers of tourists can be associated to lower speeds, higher traffic volumes and more congestion. Moreover, the use of data from roads differentiated in terms of the type of traffic they carry allowed observing that, while the influence of residents on congestion is higher than that of tourists, the influence of the latter is higher in recreational roads (as compared to the non-recreational road). In concrete, the impact of residents on traffic volume is six times higher than that of tourists on the road with less tourist-related traffic, while on recreational roads, it is not bigger than three times. A similar behavior can be observed for congestion, where the ratio is three times as much for the non-recreational road and two times as much (in average) for recreational roads.

By providing information on how the tourist pressure variable affects different indicators of traffic congestion and hyper-congestion, for different monitoring stations, it is possible to classify the roads in terms of usage by tourists in order to anticipate the levels of traffic intensity, especially during peak periods. This can be useful for stakeholders at the destination level in order to assess how congestion might affect the destination's image, especially considering that this issue could affect the satisfaction of visitors, having a negative effect on a possible future visit or even lead such visitors to seek out alternative destinations.

The application of the model could be extended to other destination areas in order to assign tourism its responsibility in traffic congestion, bearing in mind that the dependent variables available for this empirical application are proxies of congestion, which constitutes a complex phenomenon. Therefore, whenever data are available, the extension of a study of this nature can be done by considering other measures of congestion.

Another way to improve this application would be to work with data from other traffic counters and weather related variables from each of those measuring points.

2.6. References

- Aguiló, E., Palmer, T., Rosselló, J. 2009. Road Tourism Transport: Demand Policies for the Encouraging the Public and Collective Alternatives. 11th Biennial Academy Conference of the International Academy for the Study of Tourism. Palma de Mallorca, Spain.
- Aguiló, E., Rosselló, J. 2005. Host Community Perceptions and Attitudes toward Tourism. A Cluster Analysis. *Annals of Tourism Research* 32, 925-941.
- Alegre, J., Cladera, M. 2006. Repeat Visitation in Mature Sun and Sand Holiday Destinations. *Journal of Travel Research* 44, 288-297.
- Argawal, M., Maze, T. H., Souleyrette, R. 2005. Impacts of Weather on Urban Freeway Traffic Flow Characteristics and Facility Capacity. Proceedings of the 2005 Mid-Continent Transportation Research Symposium, Ames, Iowa, USA.
- Bakhat, M., Rosselló, J., Sáenz-de-Miera, O. 2010. Developing a Daily Indicator for Evaluating the Impacts of Tourism in Isolated Regions. *European Journal of Tourism Research* 3, 114-118.
- Boarnet, M., Kim, E., Parkany, E. 1998. Measuring Traffic Congestion. Institute of Transportation Studies University of California, Irvine, U.S.A. Presented at the

EMPIRICAL CHAPTERS

- 77th Annual Meeting of the Transportation Research Board, Washington, D.C.
UCI-ITS-WP-98-3.
- Butler, S., Ringwood, J. V. 2007. Use of Weather Inputs in Traffic Volume Forecasting, Proceedings of the Irish Signals and Systems Conference, Derry.
- Cantos, P., Alvarez, O. 2009. El Valor del Tiempo con Congestión: el Caso de la Radial-3. *Revista de Economía Aplicada* 51, 55-80.
- Cools, M., Moons, E. Wets, G. 2007. Investigating Effect of Holidays on Daily Traffic Counts: Time Series Approach. *Transportation Research Record: Journal of the Transportation Research Board* 2019, 22-31.
- Cools, M., Moons, E. Wets, G. 2008. Assessing the Impact of Weather on Traffic Intensity. *Transportation Research Board Annual Meeting 2008 Paper #08-1903*.
- Cui, X., Ryan, C. 2010. Perceptions of Place, Modernity and the Impacts of Tourism - Differences among Rural and Urban Residents of Ankang, China: A Likelihood Ratio Analysis, *Tourism Management*, doi:10.1016/j.tourman.2010.05.012.
- Datla, S., Sharma, S. 2008. Impact of Cold and Snow on Temporal and Spatial Variations of Highway Traffic Volumes. *Journal of Transport Geography* 16, 358-372.
- Del Castillo, J. M., Benítez, F. G. 1995a. On the Functional Form of the Speed-Density Relationship-I: General Theory. *Transportation Research Part B* 29(5), 373-389.
- Del Castillo, J. M., Benítez, F. G. 1995b. On the Functional Form of the Speed-Density Relationship-II: Empirical Investigation. *Transportation Research Part B* 29 (5), 391-406.
- De Rus, G., Campos, J., Nombela, G. 2003. *Economía del Transporte*. Barcelona: A. Bosch.
- Dickinson, J., Robbins, D. 2008. Representations of Tourism Transport Problems in a Rural Destination. *Tourism Management* 29, 1110–1121.

- Dwyer, L., Forsyth, P. 1993. Assessing the Benefits and Costs of Inbound Tourism. *Annals of Tourism Research* 20, 751-68.
- Gujarati, D. N., 2004. *Gujarati: Basic Econometrics*, Fourth edition. The McGraw-Hill Companies.
- Haight, F. A. 1963. *Mathematical Theories of Traffic Flow*. Academic Press, New York.
- Haldrup, N., Hylleberg, S., Pons, G., Sansó, A. 2007. Common Periodic Correlation Features and the Interaction of Stocks and Flows in Daily Airport Data. *Journal of Business & Economic Statistics* 25 (1), 21-32.
- Hall, F. L., Allen, B. L., Gunter, M. A. 1986. Empirical Analysis of Freeway Flow-Density Relationships. *Transportation Research Part A* 20a (3), 197-210.
- Hassan, Y. A., Barker, J. 1999. The Impact of Unseasonable or Extreme Weather on Traffic Activity within Lothian Region, Scotland. *Journal of Transport Geography* 7, 209-213.
- Hau, T. 1998. Congestion Pricing and Road Investment. In J. Kenneth, K. Button & E. Verhoef (Eds.), *Road Pricing, Traffic Congestion and the Environment* (pp. 39-78). Cheltenham, England: Edward Elgar Publishing Limited.
- Hymel, K. 2009. Does Traffic Congestion Reduce Employment Growth? *Journal of Urban Economics* 65, 127-135.
- Ibrahim, A. T., Hall, F. L. 1994. Effect of Adverse Weather Conditions on Speed-Flow-Occupancy Relationships. *Transportation Research Record* 1457, 184-191.
- Key, K., Simmonds, I. 2005. The Association of Rainfall and Other Weather Variables with Road Traffic Volume in Melbourne, Australia. *Accident Analysis and Prevention* 37, 109-124.
- Key, K., Simmonds, I. 2006. Road Accidents and Rainfall in a Large Australian City. *Accident Analysis and Prevention* 38, 445-454.

EMPIRICAL CHAPTERS

- Koetse, M. J., Rietveld, P. 2009. The Impact of Climate Change and Weather on Transport: An Overview of Empirical Findings. *Transportation Research Part D* 14, 205–221.
- Kyte, M., Khatib, Z., Shannon, P., Kitchener, F. 2001. Effect of Weather on Free-Flow Speed. *Transportation Research Record* 1776, 60 –68.
- Lee, C. K., Kang, S., Long, P., Reisinger, Y. 2010. Residents' Perceptions of Casino Impacts: A Comparative Study. *Tourism Management* 31, 189-201
- Levine, N., Kim, K. E., Nitz, L. H. 1995. Daily Fluctuations in Honolulu Motor Vehicle Accidents. *Accident Analysis and Prevention* 27, 785-796.
- Liang, W.L., Kyte, M., Kitchener, F., Shannon, P. 1998. Effect of Environmental Factors on Driver Speed, *Transportation Research Record* 1635, Transportation Research Board, Washington, D.C.
- Lindsay, R., Verhoef, E. 2000. Congestion Modeling. In D. A. Hensher & K. J. Button (Eds.), *Handbook of Transport Modelling* (pp.377-397). Amsterdam, Netherlands: Pergamon.
- Li, M. Z. F. 2002. The Role of Speed–Flow Relationship in Congestion Pricing Implementation with an Application to Singapore. *Transportation Research Part B* 36, 731-754.
- Liu, Z., Sharma, S. 2006. Statistical Investigations of Statutory Holiday Effects on Traffic Volumes. *Transportation Research Record* 1945, 40-48.
- Lorde, T., Greenidge, D., Devonish, D. 2011. Local Residents' Perceptions of the Impacts of the ICC Cricket World Cup 2007 On Barbados: Comparisons of Pre- and Post-Games. *Tourism Management* 32 (2), 349-356.
- Nair, A. S., Liu, J-C., Rilett, L., Gupta, S. 2001. Intelligent Transportation Systems, *Proceedings. 2001 IEEE*, pp.: 681 – 685. Oakland, CA, USA.
- Oh, J. S., Shim, Y. U., Cho, Y. H. 2002. Effect of Weather Conditions to Traffic Flow on Freeway. *KSCE Journal of Civil Engineering*, 6 (4), 413-420.
- Palmer, T., Riera, A., Rosselló, J. 2007. Taxing Tourism: The Case of Rental Cars in Mallorca. *Tourism Management* 28, 271–279.

- Rendeiro, R., Ramírez, P. 2010. Ecological Footprint Analysis of Road Transport Related to Tourism Activity: The Case for Lanzarote Island. *Tourism Management* 31, 98–103.
- Riera, A., Mateu, J. 2007. Aproximación al Volumen de Turismo Residencial en la Comunidad Autónoma de las Illes Balears a Partir del Còmputo de la Carga Demogràfica Real. *Estudios Turísticos* 174, 59-71.
- Riera-Font, A., Aguiló-Pérez, E. 2009. Llibre Blanc del Turisme de les Illes Balears. Cap a una Nova Cultura Turística. Palma de Mallorca, Spain: Centre de Recerca Econòmica (UIB-“Sa Nostra”).
- Rosselló, J., Saenz-de-Miera, O. 2010. Road Accidents and Tourism: The Case of the Balearic Islands (Spain). *Accident Analysis and Prevention*, doi: 10.1016/j.aap.2010.10.011.
- Small, K. A., Chu, X. 2003. Hypercongestion. *Journal of Transport Economics and Policy* 37, 319-352.
- Smith, B. L., Byrne, K. G., Copperman, R. B., Hennessy, S. M., Goodall, N. J. 2003. An Investigation into the Impact of Rainfall on Freeway Traffic Flow. 83rd annual meeting of the Transportation Research Board, Washington D.C., USA.
- Smith, W. S., Hall, F. L., Montgomery, F. 1996. Comparing the Speed-Flow Relationship for Motorways with New Data from the M6. *Transportation Research Part A* 30 (2), 89-101.
- Taylor, M. A. P., Woolley, J. E., Zito, R. 2000. Integration of the Global Positioning System and Geographical Information Systems for Traffic Congestion Studies. *Transportation Research Part C* 8, 257-285.
- Teye, V. B. 1992. Land Transportation and Tourism in Bermuda. *Tourism Management* 13 (4), 395-405.
- Thomas, T., Weijermars, W.A.M., Berkum van E.C. 2008. Periodic Variations in Urban Traffic Flows. *European Journal of Transport and Infrastructure Research* 8 (3), 251-263.

EMPIRICAL CHAPTERS

- TRB. 2000. Highway Capacity Manual. Washington D.C., USA: Transportation Research Board. National Research Council.
- UNWTO. 2008. Tourism Highlights 2008. Madrid, Spain: World Tourism Organization.
- UNWTO. 2009a. Facts & Figures Information, Analysis and Know-How. Tourism 2020 Vision. <http://unwto.org/facts/eng/vision.htm>. (Accessed 02.04.10)
- UNWTO. 2009b. Facts and Figures Information, Analysis and Know-How. Historical Perspective of World Tourism <http://unwto.org/facts/eng/historical.htm>. (Accessed 02.04.10).
- Verhoef, E. T. 2003. Inside the Queue: Hypercongestion and Road Pricing in a Continuous Time–Continuous Place Model of Traffic Congestion. *Journal of Urban Economics* 54 (3), 531-565.
- Verhoef, E., Rouwendal, J. 2004. A Behavioural Model of Traffic Congestion. Endogenizing Speed Choice, Traffic Safety and Time Losses. *Journal of Urban Economics* 56, 408-434.
- Verhoef, E., Rouwendal, J., Rietveld, P. 1999. Congestion Caused by Speed Differences *Journal of Urban Economics* 45, 533-556.
- Wang, C., Quddus, M. A., Ison, S. G. 2009. Impact of Traffic Congestion on Road Accidents: A Spatial Analysis of the M25 Motorway In England. *Accident Analysis and Prevention* 41, 798–808.
- Wooldridge, J.M., 2000. *Introductory Econometrics: A Modern Approach*, 2nd ed. South Western College Publishing, Thomson Learning.
- WTTC. 2007. Economic Data Search Tool. http://www.wttc.org/eng/Tourism_Research/Tourism_Impact_Data_and_Forecast_Tool/. (Accessed 12.04.10).

Chapter 3. Influence of tourism on air pollution: the case study of tropospheric ozone in Mallorca (Spain)

Abstract:

In this paper, the influence of tourism on daily concentration levels of tropospheric ozone is investigated using standard regression analysis and a Generalized Additive Model. Using the most popular air pollution determining variables (including meteorological conditions) and a measure of daily population pressure, the impact of both residents and tourists are considered using the case study of the Island of Mallorca (Spain), one of the most important tourist resort destinations in the Mediterranean. Using the daily stock measure of population it is possible to evaluate the impact of tourism from a joint perspective in contrast to the traditional economic sectorial approach. Results show that the daily stock of tourists is a significant predictor of concentration levels of tropospheric ozone.

3.1. Introduction

Transport activities are a key component of tourism and the development of destinations. Nevertheless, the transportation of passengers with holiday and leisure purposes has been identified in the literature as an important source of air pollution (Abeyratne, 1999; Becken, 2002; Johnson, 2002; Kelly et al., 2007). For this reason, during the last years there has been growing concern about the increased mobility of tourists within host countries or regions, which has a consequent rise in external costs including air pollution, that can be associated to means of transport (Palmer et al., 2007). This concern is motivated by the fact that potential environmental effects associated with energy consumption and emissions caused by tourist travel can be significant at local and global scales, especially in heavily visited destinations, where air pollution and ozone-related smog is often caused by unburned hydrocarbons and nitrogen oxides released from motorized vehicles (Kelly et al., 2007).

EMPIRICAL CHAPTERS

However, while a variety of specific economic activities like energy production (Mastral and Callen, 2000; Yi et al., 2006), manufacturing industries (Cole et al., 2008), mining (Kirchgessner et al., 1993) and livestock production systems (Cambra-Lopez et al., 2010) have been linked to air pollution, within the transport industry there has been a failure to consider the environmental impacts of transport from tourist activity (Peeters et al., 2007; Rendeiro-Martin-Cejas and Ramirez 2010), despite its growing impact for many destinations.

The present chapter tries to fill this gap by analyzing the concentration levels of tropospheric ozone (O_3) of the Island of Mallorca (Spain) with a daily population stock measure that exhibits a strong seasonal component due to the tourism pressure at such a typical mass summer destination. This is done with the objective of assigning the tourism sector its responsibility in the generation of social and environmental costs. Tropospheric ozone was chosen for this study because it is one of the main pollutants that can be associated with air pollution from transport (Maibach et al., 2008) and the availability of data. Tropospheric ozone has been recognized as the most important index substance of photochemical smog and is one of the main pollutants that degrade air quality. In addition, high ozone levels not only play a role in damage to plant species, various natural materials and manufactured goods, but also lead to the damage of lung tissues in humans (Abdul-Wahab et al., 2005). The Island of Mallorca is taken as case study because of (i) the relative importance of the tourism activity in the region, (with more than 8 million tourists yearly); (ii) the geographical nature of the islands, that makes it possible to thoroughly estimate the daily tourist stock from the ports and airports and (iii) the availability of the rest of the variables, required to carry out a study of this nature.

The remainder of the chapter proceeds as follows: Section 3.2 explains in detail, which are the most popular predictors of air pollution. Section 3.3 discusses the statistical methods for empirical applications and summarizes the main theoretical aspects of the alternatives that were used. Section 3.4 presents the empirical application describing the data employed in this study, showing empirical

estimations and presenting a brief discussion based on the results. Section 3.5 concludes.

3.2. Air pollution and tourism

The literature review shows a clearly identified set of variables that is frequently used to explain the concentration of the most important pollutants in the air, including tropospheric ozone, which consists mainly of weather conditions and spatial and temporal effects (Thompson et al., 2001) that can reflect human pressures (as, for instance, traffic variations).

Ozone variability, particularly variability in the exceedance of the ozone standards, can be substantially affected by meteorological variables and that is why there is a significant body of literature on the adjustment of ozone concentrations for meteorology. An important example of the inclusion of the weather can be found in Thompson et al. (2001), who revised different works considering a large set of meteorological variables showing that maximum surface temperature, wind speed, relative humidity, mixing height, and opaque cloud cover, as well as wind speed by temperature interaction were significant meteorological predictors over most major metropolitan areas of the United States.

More precisely, periods of high concentration levels are observed with slow-moving, high-pressure weather systems that result in sunshine, high temperatures, and stagnant air. Winds associated with high-pressure systems are typically light, thereby increasing the chance that pollutants will accumulate in the atmospheric boundary layer. And warm cloudless conditions associated with these systems are favorable to photochemical production of ozone (Thompson et al., 2001). Such conditions occur with a seasonal pattern (Gangoiti et al., 2001; Logan, 1985; Millán et al., 2002,) corresponding with high seasons characterized by incoming flows of tourists seeking sun and sand holidays in the northern hemisphere. In any case, while the literature review reveals a common set of meteorological variables, the choice of

EMPIRICAL CHAPTERS

the relevant ones for a specific study will depend on the purpose of the analysis, regional differences in meteorology and emission patterns, and data availability.

In accordance with the main objectives pursued when studying the concentration of diverse pollutants, many authors have included variables to account for a trend. This is the case of Reiss (2006), who included the number of days as a simple trend term.

Another category of determinants of air pollution that should be pointed out is related to traffic activity. Carslaw et al. (2007) includes vehicle flows of light and heavy duty vehicles as determinants of the concentration levels of various air pollutants. Traffic is also found within the set of explanatory variables in Aldrin and Hob ck-Haff (2005) and Shi and Harrison (1997). Reiss (2006) considers that different emissions profiles occur between weekend and weekdays, mostly due to differences in vehicle traffic patterns. Therefore, the author uses a model to examine the weekend–weekday concentrations of benzene and 1,3 butadiene.

With special interest for this study, Sandanaga et al. (2008) analyze the weekday–weekend differences of Ozone, NO_x (NO and NO₂) and Non-Methane Hydrocarbons in Tokyo and Osaka, Japan, from 2001 to 2004, in order to investigate what the authors call the Ozone Weekend Effect (OWE). According to them, the OWE is a common phenomenon of ozone behavior in the urban atmosphere: ozone concentrations on weekends are higher than those on weekdays despite lower concentrations of ozone precursors. The OWE was discovered in the 1970s in the United States and more recently, the same phenomenon was recognized in several other countries as well.

In summary, the relevance of accounting for the traffic and transportation activity in the determination of tropospheric ozone concentrations becomes evident from the multiple applications found in the literature review. Additionally it takes a key role in tourist destinations where the usual levels of commuter traffic can increase significantly during vacation periods and where environmental quality is a key factor for the tourism industry.

3.3. Methodology

Different statistical techniques have been regarded as useful to analyze the variables that explain and predict the concentrations of diverse pollutants in the air. Thompson et al. (2001) made a review of the methods used to study the meteorological variables behind the concentrations of tropospheric ozone, categorizing them into three broad statistical approaches: regression-based modeling, extreme value approaches, and space time models. According to Thompson et al. (2001), the most familiar of the methodologies employed in the literature is regression, which is normally used to achieve three main objectives:

- (a) Obtaining air quality forecasts. Forecasting extreme pollutant events in order to provide public health warnings may focus the analysis on investigating those observations exceeding a threshold and their association with readily predicted meteorology.
- (b) Investigating and estimating pollutant time trends. This will most frequently be done with a view to assessing the effect of changes in emissions.
- (c) Increasing scientific understanding of the underlying mechanisms.

Analytically, the formulation of the problem can be expressed as:

$$y_i = f(x_i^1, x_i^2, \dots, x_i^n) \quad (3.1)$$

Where the objective is to find the optimal set of n determinants, $x_i^1, x_i^2, \dots, x_i^n$, and the suitable functional form f to combine them and get the best estimated variable for the pollution indicator y_i . The simplest approximation to the problem is the use of linear regression models where the expression (3.1) is parameterized to:

$$E(y_i) = \beta_0 + \beta_1 x_i^1 + \beta_2 x_i^2 + \dots + \beta_n x_i^n, \quad (3.2)$$

assuming that the simple sum of the determinants will be able to reproduce the behavior of the dependent variable through the $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ coefficients of the observed values $x_i^1, x_i^2, \dots, x_i^n$ where $E(y_i)$ is the expected value of y_i . The basic formulation can be improved through the use of autoregressive terms or the application of principal component analysis before the estimation of the regression in order to find the best set of determining variables (Abdul-Wahab et al., 2005). The results from such linear estimation techniques can provide a first approximation to the relation between the explanatory variables and the dependent variable and their advantage is the easiness of coefficient interpretation. However, linear regression models are open to criticism since the underlying chemical and physical processes are unlikely to be linear and additive (Thompson et al., 2001).

In order to overcome this problem, artificial neural networks, which have become very widespread during the last years, constitute an alternative type of models that do consider the nonlinear associations among variables. Yi and Prybutok (1996), Kukkonen et al. (2003) and Schlink et al. (2003) have concluded that neural networks are superior to linear techniques in predicting PM_{10} , NO_2 , NO_x or ozone concentrations from several meteorological variables. However, although neural networks have evidenced certain higher performance on studies whose purpose is forecasting future levels of pollutant concentrations, it has been recognized that one of the difficulties using the neural network approach is that inference is difficult (Carslaw et al., 2007), and sometimes incoherent with theoretical underpinnings.

An alternative to both, linear regression models and artificial neural networks that has gained significant popularity in modeling air pollution during the last years is based on the General Additive Models (GAM) methodology (Hastie and Tibshirani,

1990). The GAM structure relaxes the assumption of linearity between the dependent and independent variables by fitting natural splines for each covariate (Reiss, 2006) while coherent inference between explanatory variables and dependent one remains. Thus, the technique offers the modeler an improvement with respect to linear models, which would be too restrictive because the relationships between explanatory variables are too complex to be represented as linear functions. GAMs have the advantage of allowing different associations between the variables and, supposedly, more sensitive estimates of ozone trends within regimes (Thompson et al., 2001).

Previous applications of GAM can be found in Carslaw et al. (2007) where daily concentrations of NO_x, NO₂, CO, benzene and 1,3-butadiene at a specific zone of London where concentrations are dominated by vehicles are modeled; in Holland et al. (2000), who modeled trends in sulphur and nitrogen species in the eastern United States and estimated spatial trends; Reiss (2006), who considered temporal trends in benzene and 1,3-butadiene concentrations in Houston; and Aldrin and Hob ck-Haff (2005), who determined the logarithm of hourly concentration of an air pollutant within the framework of the GAMs using traffic volume and several meteorological variables as the dependent ones.

Analytically a GAM structure can be written as:

$$g(\mu_t) = \beta_0 + s_1(x_t^1) + s_2(x_t^2) + \dots + s_n(x_t^n) \quad (3.3)$$

where $\mu_t \equiv E(y_t)$; g is a known monotonic differentiable 'link function'; β_0 is a parameter to be estimated; and s_1, s_2, \dots, s_n are unknown smooth functions that can be estimated using linear smoothers and backfitting. According to Aldrin and Hob ck-Haff (2005), the smoothness of each function s is controlled by a smoothness parameter, expressed by the number of degrees of freedom or effective

EMPIRICAL CHAPTERS

parameters for each function. This must be chosen before the function is estimated. As the number of degrees of freedom increases, the function becomes less smooth, but more flexible and it gives better fit to data. Hence, choosing the smoothness parameter is a trade-off between good fit and smoothness, or in other words, between bias and variance.

To understand the use and representation of smoothers, Wood and Augustin (2002) suggest that it is often desirable to include a smooth function like $h(x)$ into a model, without being very specific about the exact form of the function, whereby h has to be estimated. Estimation of h , using methods such as linear and generalized linear models, requires that h be represented in such a way that the model becomes linear, which can be done by choosing a *basis*. The task consists of choosing some basis functions, which will be treated as completely known: if $b_j(x)$ is the j^{th} such basis function, then h is assumed to have a representation

$$h(x) = \sum_{j=1}^m b_j(x) \alpha_j, \quad (3.4)$$

where the α_j are m unknown coefficients; $h(x)$ is made up of a linear combination of the basis functions $b_j(x)$; and estimating h is now equivalent to finding the α s (Wood and Augustin, 2002; Wood, 2006).

Cubic splines constitute popular basis functions because they can have good theoretic properties. A cubic spline is a curve made up of sections of cubic polynomial joined together so that they are continuous in value, as well as first and second derivatives. The points at which the sections join are known as the knots of the spline. For a conventional spline, knots occur wherever there is a datum, but for the regression splines of interest here, the locations of the knots must be chosen. Let the knot location be denoted by $\{x_j^* : j = 1, \dots, m-2\}$, a set of points in the range of x . If for instance, a rank 6 basis was used, meaning that $m=6$, there would be 4 knots.

The choice of degree of model smoothness would then be controlled by the basis dimension.

An alternative to controlling smoothness by altering the basis dimension is to keep the basis dimension fixed, at a size a little larger than it is believed could reasonably be necessary, but to control the model's smoothness by adding a "wiggleness" penalty to the least squares fitting objective. Thus, if we consider the model:

$$E(y_t) = \mu_t = \beta_0 + h_1(x_t^1) + h_2(x_t^2) + \dots + h_n(x_t^n) \quad y_t \sim N(\mu_t, \sigma^2) \quad (3.5)$$

Instead of fitting the model by minimizing

$$\sum_{t=1}^T (y_t - \beta_0 - h_1(x_t^1) - h_2(x_t^2) - \dots - h_n(x_t^n))^2, \quad (3.6)$$

it could be fitted by minimizing

$$\begin{aligned} & \sum_{t=1}^T (y_t - \beta_0 - h_1(x_t^1) - h_2(x_t^2) - \dots - h_n(x_t^n))^2 \\ & + \lambda_1 \int \left(\frac{\partial^2 h_1}{\partial [x^1]^2} \right) + \lambda_2 \int \left(\frac{\partial^2 h_2}{\partial [x^2]^2} \right) + \dots + \lambda_n \int \left(\frac{\partial^2 h_n}{\partial [x^n]^2} \right) \end{aligned} \quad (3.7)$$

where the integrated square of the second derivative penalizes models that are too "wiggly". The trade-off between model fit and model smoothness is controlled by the smoothing parameter, λ . Where $\lambda \rightarrow \infty$ leads to a straight line estimate for h , while $\lambda = 0$ results in an un-penalized regression spline estimate (Wood, 2006).

The next question that arises is which value of λ is best. A suggested technique is choosing the smoothing parameter λ through cross validation. According to Wood and Augustin (2002), the approach that is most consistent with

EMPIRICAL CHAPTERS

using GCV for smoothing parameter selection is to drop each term from the model in turn, and see if this reduces the GCV score relative to the full model. This approach could be used as the basis for a general backwards selection method. In practical modeling situations an *ad hoc* approach is sometimes useful. For this purpose, three questions need to be asked:

- 1) Are the estimated degrees of freedom for the term close to their lower limit (e.g. 1 for a univariate smooth with a second derivative based wigginess penalty)?
- 2) Does the confidence region for the smooth everywhere include zero?
- 3) Does the GCV score for the model go down if the term is removed from the model?

If the answer to all three of these is ‘yes’ then the term should be dropped. If the answer to 2 is ‘no’ then it probably should not be. Other cases will require judgment (Wood and Augustin, 2002).

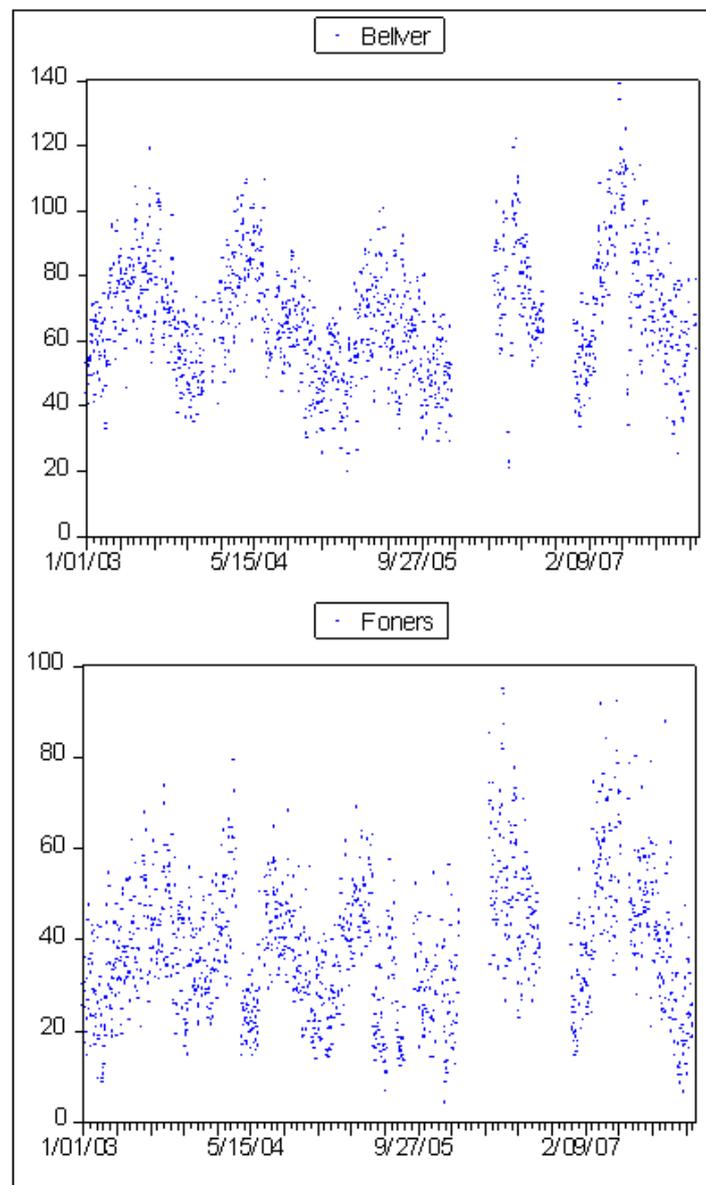
3.4. Empirical analysis

The tropospheric ozone variability in the Island of Mallorca (Spain) is used as a case study in order to analyze if the introduction of a tourist pressure variable improves both regression and GAM formulations using meteorological variables. Mallorca can be qualified as one of the most important tourist regions in the world, accounting for 8.7 million tourists with a population of 0.87 million inhabitants according to data from 2009. Since beaches and climate are often mentioned as the most central attractors, as in other tourist areas around the world, a high degree of seasonality characterizes the Island with a 75% of tourist arrivals registered during the May-October semester. The suitability of Mallorca as a case study is also supported by the location characteristics that makes easy to control, periodically, the number of people (tourists and residents) on the islands.

3.4.1. Data

Data on tropospheric ozone concentrations come from the *Consell de Mallorca* and are available for two monitoring stations, Bellver and Foners. The time series, with some periods of missing observations, range from January 1, 2003 to September 30, 2006 and can be seen in Figure 3.1.

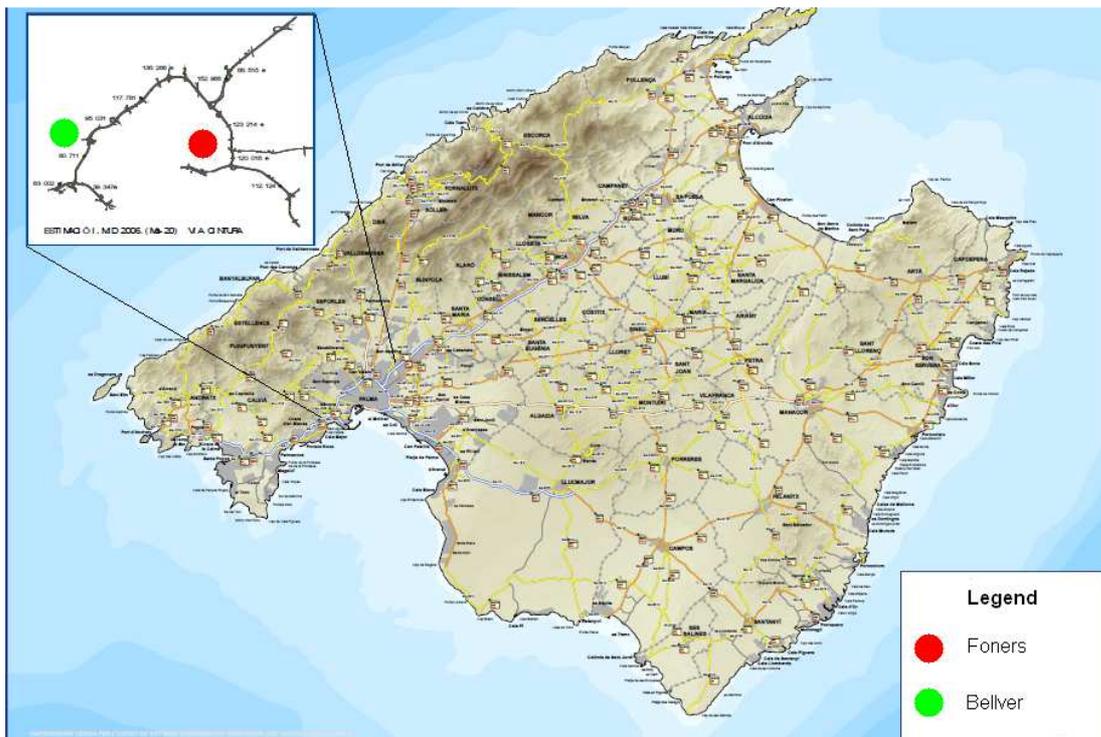
Figure 3.1 Daily Concentrations of O₃



EMPIRICAL CHAPTERS

The reasons for the specific location of both stations respond to the fact that they allow gathering information on emissions from traffic that flows not only on the metropolitan area of the capital city (Palma), but also to and from key tourist sites of the island. On the one hand, the station of Bellver would be able to reflect the influence of tourism on traffic fluctuations for it is located near the highways that link Palma (the main city with 400,000 inhabitants) with the main tourist resorts of western Mallorca such as Cala Major, Calvià, Andratx and Paguera. On the other hand, the station of Foners is an urban one situated inside Palma. Hence, it should be especially suitable to pick up the effect of residents and, to a lesser extent, that of tourists flowing to and from key tourists sites of eastern Mallorca like Platja de Palma, Can Pastilla and Coll de'n Rabassa (Figure 3.2).

Figure 3.2 Location of the Monitoring Stations



Source: Consell de Mallorca.

Meteorological variables are based on data provided by the *Instituto Nacional de Meteorología* of the Balearic Islands and are taken from the airport station. Meteorological variables collected initially include maximum daily temperature, minimum daily temperature, precipitation in mm, sunshine hours, mean humidity, wind speed and air pressure.

For the case of tourism pressure, the variable used is the Daily Indicator of Human Pressure (DIHP) for the Balearic Islands (Part I).

A point that is worth noting from such analysis is the fact that, on annual average, tourists account for a 25% of the total population pressure in Mallorca, even though for the high season months the amount of tourists is very close to that of residents. Although it should be admitted that the monthly average of the DIHP does not differ significantly from the daily airport (and port) arrival data, the DIHP can be used at daily level, avoiding the bias of the length of stay and the decomposition into tourists and residents provides further knowledge about the presence of both populations that can be useful for the design and implementation of policy measures.

3.4.2. Results

Following the methodological considerations mentioned above, the equations were estimated through the use of linear regression (Table 3.1) and then compared with estimations carried out within the framework of GAMs (Table 3.2), in order to observe which technique was the most suitable. Additionally, the relevance of including the indicator of tourist population (DIHP) was put to the test by trying alternative specifications, namely, one in which the variable was dropped from the equation and a second one where the DIHP was included as a parametric term in the GAM model.

Regarding the explanatory variables, the final estimation included: dummies, from Monday (D1) to Sunday (D7), to control for different levels of exposure

EMPIRICAL CHAPTERS

observed throughout the week; precipitation in liters per square meter; minimum and maximum daily temperature in Celsius; and the variables that account for the stock of resident population (DIHP residents) and tourist population (DIHP tourists). These variables are included both as linear terms in the case of the linear regression and as smooth functions in the case of the GAM estimations. However, it should be noted that the dummies are parametric in both cases; therefore the GAM estimation should more precisely be referred to as a semi-parametric model. In addition, the presence of autocorrelation in the time series suggested the estimation of dynamic models, hence the inclusion of lags of the dependent variable.

Table 3.1 Linear estimations of tropospheric Ozone (O₃)

Station	Belver	Foners
Intercept	-0.1857	9.373
Day 1	-4.579***	-0.1076***
Day 2	-2.033*	-5.477***
Day 3	-1.907	-5.621***
Day 4	-1.697	-5.835***
Day 5	-1.753	-6.019***
Day 6	-0.4753	-3.579***
Time	-3.1 x 10 ⁻³ ***	5.8 x 10 ⁻⁴
Precipitation	-0.2345***	-
Minimum Temperature	-4.201***	-
Maximum Temperature	2.843**	-
DIHP Residents	5.8 x 10 ⁻⁵ ***	3.2 x 10 ⁻⁶
DIHP Tourists	2.2 x 10 ⁻⁵ ***	4.8 x 10 ⁻⁶ **
O3(-1)	0.696***	0.7595***
Equation Statistics		
Adjusted R-squared	0.6095	0.6152
F-statistic	133.8	183.6
AIC	8377.016	8173.16

***, ** and * stand for statistical significance at the 1%, 5% and 10%, respectively

Bearing in mind that the reference observation is the tropospheric ozone concentration levels during Sunday, the results follow the expected presence of the OWE (Sandana et al., 2008), which means that concentrations are higher during

weekends. With regard to the weather, it could be seen that the amount of rainfall and the temperature are indeed significant determinants of the concentration levels of tropospheric ozone in the station of Bellver. The presence of a non-linear trend and its relevance as explanatory variable, as previous studies suggest, can also be seen in the results.

Table 3.2 GAM estimations of tropospheric Ozone (O₃)

Station	Bellver	Foners
Intercept	31.3508***	17.7726***
Day 1	-4.5051***	-10.2617***
Day 2	-2.3343**	-5.6352***
Day 3	-2.3501**	-5.8098***
Day 4	-2.0009*	-6.0762***
Day 5	-2.1171*	-6.2677***
Day 6	-0.9363	-3.857***
s(Time)	8.903***	8.206***
s(Precipitation)	2.122**	-
s(Minimum Temperature)	4.315***	-
s(Maximum Temperature)	1***	-
s(DIHP Residents)	1	1.983
s(DIHP Tourists)	6.211***	4.88**
O3(-1)	0.5526***	0.6643***
Equation Statistics		
Adjusted R-squared	0.645	0.635
No. of observations	1107	1143
AIC	8288.506	8123.819
CGV score	104.72	71.396

***, ** and * stand for statistical significance at the 1%, 5% and 10%, respectively

The stock of people, especially of non-residents, not only proved to be a significant predictor under both estimation techniques but its inclusion also played a significant role in obtaining the preferred equations. This can be seen on Table 3.3, which presents two statistics that allow comparisons between models: the AIC and the GCV Score. The AIC allows comparison both for linear regression and GAM models and, according to it, all the cases where the DIHP variable was dropped from the equation showed an increase in the value it takes. This means that, the

EMPIRICAL CHAPTERS

preferred models are the ones that include the variable DIHP-tourists. The CGV, on the other hand, confirms the same evidence in the case of the GAM estimations.

Table 3.3 Model comparison

Statistic Monitoring Station	Akaike Information Criterion (AIC)		GCV score	
	Bellver	Foners	Bellver	Foners
Model				
GAM	8288.506	8123.819	104.43	71.396
GAM with DIHP in parametric form	8291.854	8133.282	104.72	71.98
GAM excluding DIHP for tourists	8302.132	8133.576	105.69	71.997
Linear regression	8377.016	8173.16	-	-
Linear regression excluding DIHP for tourists	8390.328	8175.314	-	-

According to the AIC, it can also be seen that the GAM estimations are preferred to the linear regression estimations, which confirms that the equation had to be fitted by a model that was able to consider the non-linear relationships among variables.

Once it was seen that the GAM models are preferred, their own tool for deciding which GAM specification is best, the CGV score, was used. When removing the term (DIHP-Tourists), both in Bellver and Foners, the CGV score did not decrease (actually it increased), hence the variable should not be removed from the model. Furthermore, when introducing the variable in its parametric form (DIHP-Tourists), the CGV score increased too, suggesting that the preferred equation is the one that includes a smooth term of this variable.

3.4.3. Discussion

Results confirmed that the stock of people is a significant determinant of air pollution, especially for non-residents. Furthermore, beyond the significance of the stock of non-residents and the confirmation that it should be included in similar exercises, it is worth looking at its coefficient in order to get more detailed information about the relative importance of tourist activity as a determinant of air

pollution. In concrete, the higher value that it takes in the case of Bellver (as compared to Foners) confirms that tourism has a larger impact on the station that is situated near important resorts and is expected to present more intense traffic activity that can be directly attributed to tourism.

Regarding residents, in spite of the slight periodical decrease in their numbers that can be observed during summer, data on residents exhibit a rather uniform behavior throughout the year. However, the pattern of the series that should be considered to provide a better representation of the impact of residents is the one that reflects weekly traffic activity, but this effect is already being captured by day-of-the-week dummy variables.

Besides the way in which population has shown to influence ozone concentrations, it is important to examine what the results tell us about the relationship between weather conditions and air pollution. According to the signs of weather related variables, while higher temperatures can be associated to higher tropospheric ozone concentrations, rainy conditions can be associated to lower concentrations. Despite the facts that sunshine was not included due to its correlation with other variables (and especially with temperature) and that no data were available to directly account for cloudiness, the weather indicators that were used constitute fundamental features of the weather systems that are typically related to tropospheric ozone concentration. Moreover, the variables that were included in the final estimation allow confirming that high ozone concentrations are partly determined by seasonal weather conditions that combine high temperatures, clear skies and low precipitation levels.

3.5. Conclusions

Tourism is often pointed out as a strategic sector for economic and social development. However the increase in tourism numbers can be linked to the fact

EMPIRICAL CHAPTERS

that nowadays transport is considered responsible for an important share of air pollution. Then, it becomes essential to account, not only for the tourism benefits in economic terms but also to estimate the environmental impacts that tourism entails.

In this chapter, the influence of tourism on the concentration levels of tropospheric ozone has been analyzed using the case study of Mallorca (Spain), one of the most important tourist regions around the World. Linear regressions and GAMs have been estimated showing how the expected results for the meteorological variables are obtained. These have shown how the expected signs of the most important determining variables have been obtained; the approximation from the linear regressions seems too simple to describe reality and the GAM presented a better statistical behavior in terms of explaining the relationship between meteorological conditions and tropospheric ozone concentrations; and the OWE was obtained, confirming that concentration levels are higher during the weekends.

The results have shown how a daily indicator of tourism pressure was an important factor to be included in the quantification of the relationship between the concentration tropospheric ozone and its determinants, regardless of the applied approach (linear regression and GAM). Thus, it has been evidenced how tourism is a key determinant of air pollution from mobile sources, especially at a destination with such characteristics as Mallorca. This conclusion it based not only on the significance of the tourism variable but also on the evidence that better estimations were achieved thanks to its inclusion. Again, in line with the expectations, based on the literature review, it was seen how the equations estimated through the Generalized Additive Models technique were preferred to their linear counterparts.

In this study, the difficulty of assigning tourism its responsibility in the generation of certain external costs imposed onto society has been overcome. The application of the model could be extended to similar destination areas and other types of pollutants as well. Moreover, the consideration of other pollutants would make it possible to enrich the application by accounting for the precursors of tropospheric ozone.

As tourism and transport are essential to many countries and regions, one should strive for a balance between economy and ecology. Future investigations will address proposed measures for reducing inputs of atmospheric contaminants into ecosystems even if tourism around the World increases.

3.6. References

- Abdul-Wahab, S. A., Bakheit, C. S., Al-Alawi, S. M. 2005. Principal Component and Multiple Regression Analysis in Modelling Of Ground-Level Ozone and Factors Affecting Its Concentrations. *Environmental Modelling & Software* 20, 1263–1271.
- Abeyratne, R. 1999. Management of the Environmental Impact of Tourism and Air Transport on Small Island Developing States. *Journal of Air Transport Management* 5, 31-37.
- Aldrin, M., Hob ck-Haff, I. 2005. Generalised Additive Modelling of Air Pollution, Traffic Volume and Meteorology. *Atmospheric Environment* 39, 2145–2155.
- Bakhat, M., Rossell , J., S enz-de-Miera, O. 2010. Developing a Daily Indicator for Evaluating the Impacts of Tourism in Isolated Regions. *European Journal of Tourism Research* 3, 114-118.
- Becken, S. 2002. Analysing International Tourist Flows to Estimate Energy Use Associated with Air Travel. *Journal of Sustainable Tourism* 10 (2), 114-131.
- Cambra-Lopez, M., Aarnink, A. J. A., Zhao, Y., Calvet, S., Torres, A. G. 2010. Airborne Particulate Matter from Livestock Production Systems: A Review of an Air Pollution Problem. *Environmental Pollution* 158, 1–17.
- Carslaw, D.C., Beevers, S.D., Tate, J. E. 2007. Modelling and Assessing Trends in Traffic-Related Emissions Using a Generalised Additive Modelling Approach. *Atmospheric Environment* 41, 5289–5299.

EMPIRICAL CHAPTERS

- Cole, M. A., Elliot, R. J. R., Shanshan, W. 2008. Industrial Activity and the Environment in China: An Industry-Level Analysis. *China Economic Review* 19, 393–408
- Gangoiti, G., Millán, M. M., Salvador, R., Mantilla, E. 2001. Long-Range Transport and Re-circulation of Pollutants in the Western Mediterranean during the Project Regional Cycles of Air Pollution in the West-Central Mediterranean Area. *Atmospheric Environment* 35, 6267–6276.
- Hastie, T., Tibshirani, R. 1990. Generalized Additive Models. *Statistical Science* 1 (3), 297-318.
- Holland, D. M., De Oliveira, V., Cox, L. H., Smith, R. L. 2000. Estimation of Regional Trends in Sulfur Dioxide over the Eastern United States. *Environmetrics* 11, 373-393.
- Johnson, D. 2002. Environmentally Sustainable Cruise Tourism: a Reality Check. *Marine Policy* 26, 261-270.
- Kelly, J., Haider, W., Williams, P. W. 2007. A Behavioral Assessment of Tourism Transportation Options for Reducing Energy Consumption and Greenhouse Gases. *Journal of Travel Research* 45, 297-309.
- Kirchgessner, D.A., Piccot, S.D., Chadha, A. 1993. Estimation of Methane Emissions from a Surface Coal Mine Using Open-Path FTIR Spectroscopy and Modeling Techniques. *Chemosphere* 26, 23–44.
- Kukkonen, J., Pratanen, L., Karppinen, A., Ruuskanen, J., Junninen, H., Kolehmainen, M., Niska, H., Dorling, S., Chatterton, T., Foxall, R., Cawley, G. 2003. Extensive Evaluation of Neural Network Models for the Prediction of NO₂ and PM₁₀ Concentrations, Compared with a Deterministic Modelling System and Measurements in Central Helsinki. *Atmospheric Environment* 37, 4539–4550.
- Logan, J. A. 1985. Tropospheric Ozone: Seasonal Behavior, Trends, and Anthropogenic Influence. *Journal of Geophysical Research* 90, 10463-10482.

- Maibach, M., Schreyer, C., Sutter, D., van Essen, H.P., Boon, B.H., Smokers, R., Schroten A., Doll C., Pawlowska, B., Bak. M. 2008. Handbook on Estimation of External Costs in the Transport Sector. European Commission.
- Mastral, A. M., Callen, M. S. 2000. A Review on Polycyclic Aromatic Hydrocarbon (PAH) Emissions from Energy Generation. *Environmental Science & Technology* 34 (15), 3051-3057.
- Millán, M. M., Sanz, M. J. Salvador R., Mantilla, E. 2002. Atmospheric Dynamics and Ozone Cycles Related to Nitrogen Deposition in the Western Mediterranean. *Environmental Pollution* 118, 167–186.
- Palmer, T., Riera, A., Rosello, J. 2007. Taxing Tourism: The Case of Rental Cars in Mallorca. *Tourism Management* 28, 271- 279.
- Peeters, P., Szimba, E., Duijnisveld, M. 2007. Major Environmental Impacts of European Tourist Transport. *Journal of Transport Geography* 15, 83–93.
- Reiss, R. 2006. Temporal Trends and Weekend–Weekday Differences for Benzene and 1,3-Butadiene in Houston, Texas. *Atmospheric Environment* 40, 4711-4724.
- Rendeiro-Martin-Cejas, R., Ramirez, P. 2010. Ecological Footprint Analysis of Road Transport Related to Tourism Activity: The Case for Lanzarote Island. *Tourism Management* 31, 98–103.
- Riera, A., Mateu, J. 2007. Aproximación al Volumen de Turismo Residencial en la Comunidad Autónoma de las Illes Balears a Partir del Cómputo de la Carga Demográfica Real. *Estudios Turísticos* 174, 59-71.
- Sandanaga, Y., Shibata, S., Hamana, M., Takenaka, N., Bandow, H. 2008. Weekday/weekend Difference of Ozone and its Precursors in Urban Areas of Japan, Focusing on Nitrogen Oxides and Hydrocarbons. *Atmospheric Environment* 42, 4708–4723.
- Schlink, U., Dorling, S., Pelikan, E., Nunnari, G., Cawley, G., Junninen, H., Greig, A., Foxall, R., Eben, K., Chatterton, T., Vondracek, J., Richter, M., Dostal, M., Bertucco, L., Kolehmainen, M., Doyle, M. 2003. A Rigorous Intercomparison

EMPIRICAL CHAPTERS

- of Ground-level Ozone Predictions. *Atmospheric Environment* 37, 3237–3253.
- Shi, J. P., Harrison, R. M. 1997. Regression Modelling of Hourly NO_x and NO₂ Concentrations in Urban Air in London. *Atmospheric Environment* 31, 4081-4094.
- Thompson, M. L., Reynolds, J., Cox, L. H., Guttorp, P., Sampson, P. D. 2001. A Review of Statistical Methods for the Meteorological Adjustment of Tropospheric Ozone. *Atmospheric Environment* 35, 617-630.
- Wood, S. N., Augustin, N. H. 2002. GAMs with Integrated Model Selection Using Penalized Regression Splines and Applications to Environmental Modelling. *Ecological Modelling* 157, 157-177.
- Wood, S. N. 2006. *Generalized Additive Models. An Introduction with R*. Boca Raton: Chapman and Hall–CRC.
- Yi, H., Hao, J., Duan, L., Li, X., Guo, X. 2006. Characteristics of Inhalable Particulate Matter Concentration and Size Distribution from Power Plants in China. *Air & Waste Management Association* 56, 1243-1251.
- Yi, J., Prybutok, V. R. 1996. A Neural Network Model Forecasting for Prediction of Daily Maximum Ozone Concentration in an Industrialized Urban Area. *Environmental Pollution* 92 (3), 349-357.

III. Conclusions

Throughout this thesis, the relevance of accounting for the external costs of tourism has been constantly emphasized because there is still much to be done to achieve the imperative task of addressing tourism related issues from a holistic perspective. The literature has shown that there exist well established techniques to assess tourism benefits, but only recently have researchers started to develop tools to also incorporate tourism's negative impacts in their analyses. Furthermore, it has been observed how, for some of the economic sub sectors that compose the "so called" tourism industry, there has been a failure to account for the external costs that such industry imposes onto diverse sectors of society. With these considerations in mind, the empirical chapters have shown how this gap found in the literature can be overcome through the use of an indicator of tourist pressure.

In Chapter 1, the focus was on road traffic accidents, which were pointed out not only as one of the most important inland transport externalities but also as a key area of concern in the travel medicine. The challenge in this chapter was to find a suitable technique to deal with data on accidents, which are discrete and non-negative and usually follow a Poisson distribution that poses the additional problem of mean-variance equality. For this reason, Negative Binomial regressions were carried out, where the number of accidents was determined by a series of popular variables and a quantitative indicator of tourist pressure that had not been included in similar studies. Hence the objective was to show how tourism is a factor that increases accidents in a region where the population fluctuates drastically due to the massive influx of tourists.

The main conclusion of Chapter 1 was that the indicator of tourist pressure is a significant determinant of road traffic accidents; and thanks to such indicator the impact of tourism on accidents can be measured. In fact, a simulation carried out with data from the Balearic Islands allowed measuring how the level of accidents would decrease under a hypothetical scenario where the number of tourists was zero.

CONCLUSIONS

In Chapter 2, a similar empirical application was carried out but in this case the focus was on congestion, which is perhaps the most visible road transport externality and one of the major liabilities of modern life. In this case, despite the long history of efforts devoted to studying congestion, the inherent complexity of such phenomenon posed some methodological drawbacks. In concrete, the lack of surveys or probe vehicle technology to generate data for an adequate measure of congestion, motivated the use of proxies and the own elaboration of a variable that could account for this externality in a more accurate manner. Therefore, traffic volume, speed and a measure of congestion built upon classic speed-flow diagrams, were used as dependent variables in regressions where tourism was included as a determinant, along with other popular explanatory variables.

Conclusions from Chapter 2 reinforce the evidence found in chapter 1 in the sense that tourism is a significant determinant of road transport externalities, in this case congestion, in a popular tourist destination. With respect to the previous application, this one has the advantage of analyzing different roadways which in turn allows observing how tourism has a higher impact on the given externality levels on recreational roads. In addition, thanks to the need to overcome limitations in terms of data resources, this empirical application provided an alternative way to measure traffic congestion in a comprehensive manner.

In the last empirical chapter, the indicator of tourist population pressure is associated to air pollution. The literature review of this chapter evidenced that, so far the influence of road transport for tourists on air pollution has not been considered, which differs from other externalities like accidents and congestion, where such impact has already been approximated through dummy variables. Regarding the methodological details of this application, the fact that relations between variables are unlikely to be linear motivated the use of non-linear regression techniques. More specifically, a Generalized Additive Model was used to explain the concentration levels of a specific pollutant, through a series of variables and the indicator of tourist pressure.

The conclusion of this chapter is that the increase in tourist numbers can be associated to higher levels in the analyzed externality (air pollution). Also, the results of Chapter 3 reinforce existing evidence regarding the existence of an Ozone Weekend Effect; and that non-linear estimation techniques are more suitable for modeling air pollution.

A key characteristic that was observed throughout the empirical chapters is the crucial role that weather conditions play in determining the levels of all three externalities. In this sense and thanks to data availability the chapters allowed confirming how the external costs of road transport are partly affected by the weather, which substantially enriches the models both in terms of theoretical insights and of explanatory power for empirical applications.

The influence of traffic activity was another essential element of the empirical applications. By including day-of-the-week dummy variables, on the hand, valuable information was obtained regarding periodical fluctuations of externality levels and how these relate to the degree of infrastructure utilization by tourists, in the case of congestion; and the photochemical processes that define concentration levels over the week, in the case of air pollution. On the other hand, in the case of accidents, this provided evidence to support the view that thanks the use of dummies, valid models can be constructed even if data for traffic exposure variables (vehicle miles traveled, fuel sales, etc.) are not available.

Overall this thesis followed the objective of assigning tourism, considered as a whole, its responsibility in the generation of external costs of road transport. Through the implementation of the indicator of tourist pressure and the evidence of its significance in the three different applications presented here, the main result has been that of providing a new way of analyzing non-economic impacts of tourism.

As a result, the models developed for all three externalities could serve as a basis not only for the assessment of the effect of development policies but also to evaluate policies of seasonal redistribution of tourism. Although all three applications refer to the case study of the Balearic Islands, this methodology could

CONCLUSIONS

be extended to similar “isolated” regions, through the use of the DIHP and to other types of destinations through alternative indicators. Likewise, while this thesis has focused on three specific externalities, future research could be aimed at developing models for other resources or externalities, including those derived from energy demand, water consumption, public health services or urban waste production.

Finally, the opportunities for improvement are related to specific limitations for each externality. In the case of accidents, future similar research should distinguish participants in accidents, between tourists and residents. Additionally, although the destination under study is a highly seasonal one, it would be convenient to make the simulations of how accidents decrease with reduced numbers of tourists under different scenarios, other than that of zero tourists.

Regarding congestion, no data on travel times were available for the empirical application. Hence, further studies would have to use improved congestion variables that can measure concepts like delay, actual travel times, free-flow travel times or proportion of stopped time on a trip segment. This empirical application could also be improved through the use of an enriched dataset with information from more monitoring stations and climate data corresponding with each of those stations in order to reflect punctual weather conditions.

For the chapter on air pollution, possible ways to improve the application include working with data for other pollutants, especially the precursors of tropospheric ozone; and carrying out the estimations in a broader context in order to include the influence of certain regional factors that might influence the formation of ozone.