University of Southampton

Modelling Demand For Tourism In Italy

Manuela Pulina

Doctor of Philosophy

Faculty of Social Sciences
Department of Economics

February 2002

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL SCIENCES

Doctor of Philosophy

MODELLING DEMAND FOR TOURISM IN ITALY

by

Manuela Pulina

The aim of this thesis is to construct and estimate the demand for tourism for the Italian Province of Sassari, in the Sardinian island, and Italy as a whole. Several propositions are investigated. A systematic understanding is carried out for separating the domestic from the international demand of tourism to Sassari Province. The historic evolution of tourists' flows, the seasonality, trading-day effects and the empirical findings from the econometric investigation validate the separation of the two components. The sample period under estimation is from 1972 up to 1995. Three dynamic models are estimated at monthly, quarterly and annual frequencies for Sassari Province; similarities and differences are explored amongst the three models. On balance, the evidence indicates that the monthly and quarterly data models are superior to annual data models. However, one does not want to omit the annual estimation. Ideally, one should integrate and learn from each of the separate analysis.

Some of the recently developed econometric techniques are used. A pre-modelling data analysis is undertaken for the economic series of interest. Seasonal and long run unit roots tests have given insight on the properties of the variables under study. The Johansen cointegration analysis is used in order to examine possible long run relationships amongst variables integrated of order one. Dynamic estimations are run in terms of the number of tourists for Sassari Province and monthly data expenditure for Italy. The LSE general-to-specific methodology is followed and a full range of diagnostic tests is provided. Short and long run income elasticities, negativity and substitutability are tested in the light of economic theory and other empirical studies existing in the tourism literature.

LIST OF CONTENTS

CHAPTER 3 CHARACTERISTICS OF INTERNATIONAL AND DOMESTIC DEMAND FOR TOURISM IN THE ITALIAN PROVINCE OF SASSARI: A GENERAL INTRODUCTION

CHAPTER 4 INTERNATIONAL DEMAND FOR TOURISM IN THE NORTH OF SARDINIA

CHAPTER 5 THE DOMESTIC DEMAND FOR TOURISM IN THE NORTH OF SARDINIA

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

My first thought goes to my supervisor Mr Ray O'Brien. His insightful comments, his patience in reading all my drafts with admirable care, his support and encouragement have been a guide over these years.

I am also grateful to Mr Peter Smith for his supervision and helpful comments that have led to exploring new knowledge.

I would also like to thank Isabel Andrade for her time and support during my research; and to all those members of the Economic Department that, either directly or indirectly, have helped in achieving the aim.

Thanks must go to Sardinian Region for providing financial support for this research.

Finally, a special mention to Stefano Usai, from the Economic Department of Sassari University, for his valuable support.

ABBREVIATIONS

ADF - Augmented Dickey and Fuller unit root test

AIC - Akaike Information Criterion

ARDF - Augmented Rank Dickey-Fuller

CADF - Covariate Augmented Dickey Fuller

CI - Cointegrating Vector

D.E.I.S. - Dipartimento di Economia Istituzioni e Società, Sassari University

DF - Dickey and Fuller unit root test

ECM - Error Correction Models

EEC - European Economic Community

EPT - Ente Provinciale per il Turismo

ESIT - Ente Sardo Industrie Turistiche, Sardinia (Italy)

EU - European Union

GDP - Gross Domestic Product

GNP - Gross National Product

HQ - Hannan and Quinn information criterion

IFS - International Financial Statistics

ISTAT - Istituto Centrale di Statistica, Italy

LR - Likelihood Ratio

LSE - London School of Economics general-to-specific methodology

MDV - Mean of the Dependent Variable

M-TAR - Momentum Threshold AutoRegressive model

OECD - Organisation for Economic Co-operation and Development

OLS - Ordinary Least Squares

PPP - Purchasing Power Parity

PVAR - Parsimonious Vector AutoRegressive model

RDF - Rank Dickey-Fuller

RSS - Residual Sum of Squares

SC - Schwarz information Criterion

SEM - Structual Econometric Model

SER - Standard Error of the Regression

SSE - Sum of the Squared Errors

SSEL - Sum of the Squared Errors for the Linear specification

SSELL - Sum of the Squared Errors for the Log-Linear specification

TAR - Threshold AutoRegressive model

UVAR - Unrestricted Vector AutoRegressive model

VAR - Vector AutoRegressive models

WTO - World Tourism Organisation

CHAPTER 1.

INTRODUCTION

1.1 TOURISM AND ECONOMETRIC ANALYSIS

Tourism is a heterogeneous activity. Its *sui generis* nature involves multiple aspects which interact with geographical, environmental, political, sociological and economic elements. Thus, different disciplines have analysed tourism as a phenomenon because of its importance and impact which has been growing throughout the world over the recent decades.

Since the Fifties studies on tourism demand have been undertaken; however, the dawn of a systematic economic analysis of tourism has been seen with Gray (1966). In the Seventies, an increased number of empirical studies appeared in the tourism literature. The determinants of international demand for tourism started to be analysed by applying economic concepts, econometric methodologies and forecasting tools (see for example Artus, (1972); Archer, (1976)). Crouch (1994) and Lim (1997) provide a comprehensive literature review for more than one hundred empirical studies over three decades of international tourism demand. In these surveys, a detailed account is provided on the type of data used, methodologies adopted, dependent and explanatory variables employed. According to Lim (1997) and Sinclair (1998), extensive econometric effort still needs to be done in the study of international tourism demand. Small sample sizes, lack of discussion of the appropriate functional forms, failure in including the full range of diagnostic tests are pointed out as some of the main deficiencies in empirical tourism demand studies. One also might point out that more advanced econometric approaches, which include amongst others Hendry's methodology, seasonal and long run unit roots and cointegration analysis are still much neglected in the tourism literature (examples in this direction are Lanza and Urga, 1995; Syriopoulos, 1995; Vogt and Wittayakorn, 1998; Song *et al.*, 2000; Kulendran and Witt, 2001). Little attention is also paid to the analysis of the determinants of domestic demand for tourism, which represents a great quota of tourism demand in developed countries (see Seddighi and Shearing, 1997).

This thesis analyses and models tourism demand in Italy by making use of recent econometric methodologies. In particular, the focus on tourism in the Northern Province of Sardinia will be taken.

1.2 AIM AND PROPOSITIONS OF THE THESIS

The aim of this thesis is to formulate and validate an economic model of tourism for Italy and Sardinia. It anticipates that this study will concentrate on the demand of tourism (both foreign and domestic) in the Northern Province of Sardinia. This Province sees the major quota of tourist flows in the island for the sample period under analysis. While, the demand for tourism to the north of Sardinia will be modelled in terms of numbers, the Italian tourism will be modelled in terms of tourism receipts given the existing availability of data.

Song *et al.* (2000) have shown that more sophisticated econometric approaches have given significant results in analysing, modelling and testing economic theory. Kulendran and Witt (2001) have also shown that the forecasts obtained by using more recent econometric methodologies are more accurate than those obtained by least squares regression. Hence, the proposition of this thesis is the following: can advanced econometric approaches give more insight in modelling and understanding tourism demand in Sardinia and Italy? The major questions arising from this proposition are the following:

- Are there any differences between domestic and international tourism? The majority of the studies focus on the analysis and modelling of international demand for tourism. In general, there has been little attention in understanding the validity in differentiating the domestic from international demand for tourism. One of the aims of this study is to give foundations for modelling the two components separately. For this purpose, graphical and econometric tools will be employed.
- Are there common findings by using different data frequencies (*i.e.* annual, quarterly and monthly)? One of the suggestions given by Witt and Witt (1992) for further research is to estimate tourism models at different data frequencies. They write: "First, only annual data have been used to estimate the models and forecast

tourism demand. This is by no means uncommon, in that almost all the studies concerned with international tourism demand forecasting employ annual data. However, the use of monthly and quarterly data would allow for more precise estimation and examination of lags. It would also be interesting to see whether the results established for annual data hold for monthly and quarterly data" (p. 171). The lack of research in this area is also pointed out by Uysal and Roubi (1999) "the use of different data periods is one of the areas that would need further research in tourism demand and forecasting studies" (p.116). The scope of this thesis is to investigate this proposition.

- Are the economic propositions always satisfied? This thesis makes use of economic concepts that are commonly applied in the tourism literature. The aim is to test the theory by using dynamic econometric modelling. Short and long run effects of changes in income and relative prices on the demand for tourism in Italy and Sassari Province will be investigated. Propositions such as negativity and substitutability will be tested. Hence, one will consider: a) the sensitivity of the demand for tourism to changes in the prices for goods and services in the destination country relative to prices in the source countries; b) the sensitivity of tourism demand to changes in the prices of tourist goods and services relative to prices in other competitor destinations.
- Is there any conflict between economic theory and econometric results? The role of econometrics in falsifying economic theory and/or adding new knowledge is still the object of major debate amongst academics; see Hylleberg and Paldam (1991) for a discussion of different schools of thoughts. The scope of this thesis is not that of assessing new economic knowledge from the conflict between data and priors. Instead, the aim is to use the guidance and help of the existing theoretical framework in interpreting and co-ordinating the results from the econometric analysis. Hence, econometrics is employed as a tool for testing *a priori* theoretical propositions making use of several models and time series that are new with respect to other empirical studies available in the tourism literature.

This thesis answers these questions by making use of distinct research steps, as follows:

a) literature review that focuses on the aspects and characteristics of the tourism phenomenon having in mind *a priori* economic assumptions;

b) data collection and modelling the demand of tourism using particular case studies;

c) use of more advanced and recent econometric tools to test the theory;

d) feedback to economic theory.

1.3 OUTLINE OF THE CHAPTERS IN THE THESIS

An outline of each chapter of the thesis is given below.

- **Chapter 2: Methodology**

This chapter is dedicated to the methodology adopted in the thesis and links with the literature review on tourism economy covered in Chapter 1. This allows a focus on the aspects and characteristics of the tourism phenomenon having in mind *a priori* economic assumptions. The next step links together the theory with the empirical practise that requires data collection to be undertaken. From the raw data, variables of interest are calculated in monthly, quarterly and annual frequencies. Such variables are tested for both possible seasonal and long run unit roots. Once the status of the variables of interest is established, further testing for possible cointegration is carried out whenever necessary. Hence, the LSE general-to-specific methodology is used in modelling the demand for tourism. The empirical results obtained are compared with economic theory.

- **Chapter 3: Characteristics of International and Domestic Demand for Tourism in the Italian Province of Sassari: A General Introduction**

Chapter 3 is a general introduction to the main characteristics of tourism demand in Sassari Province of Italy. An account is given of the differences between the international and domestic demand based on the evolution of the tourist flows and on the seasonal distribution. In accordance with economic theory, the possible determinants that might have a role in explaining the demand for tourism are described.

- **Chapter 4: International Demand for Tourism in the North of Sardinia**

The aim of this chapter is to examine the economic factors affecting the demand for international tourism in the Province of Sassari. An econometric model is developed for a short run and long run analysis. Elements like the "trading-days" factor and the Easter effect are examined and included into the equation. The relationship between the short run, long run income and price elasticities is investigated by making use of different time frequencies.

- **Chapter 5: The Domestic Demand for Tourism in the North of Sardinia**

Chapter 5 is dedicated to the analysis of the domestic demand. It is possible to find out other distinctive characteristics that differentiate this component from the international demand. Further investigation is carried out to assess the possible validity of the correction of the dependent variable (*i.e.* the number of domestic arrivals) for the number of weekends in a month. Relationships between short run, long run income and price elasticities are also explored and a comparison with other empirical findings is given. Different data frequency models are estimated. A section is dedicated to test and establish that the Italian production index can be considered as a valid proxy of the personal disposable income.

- **Chapter 6: Sassari Province Competitors: Substitute Price and Exchange Rate**

In Chapter 6, the inclusion of the exchange rate for the main competitors in the Mediterranean area is considered. A careful investigation is carried out to include either the aggregated substitute price variable adjusted for the exchange rate or a disaggregated real substitute price for each of the competitor countries. The analysis is undertaken for both the domestic and international demand for tourism in the north of Sardinia.

- **Chapter 7: Italian Tourism: Seasonality, Numbers and Expenditure**

Chapter 7 gives an in-depth analysis on tourism in Italy as a whole. A graphical analysis identifies possible differences between domestic and international demand for tourism. An analysis on the seasonal pattern of the major origin countries, that is Belgium, France, Germany, Japan, Sweden, Switzerland, United Kingdom and United

States is undertaken.

Tourism demand can be expressed in terms of the number of tourists in registered accommodation and in terms of tourist expenditure. Hence, a comparison is made amongst the number of tourists' arrivals, nights of stay, nominal and real tourism expenditure for the period from 1972 to 1995.

- **Chapter 8: Estimating Italian Tourism**

Chapter 8 gives an account of the findings in modelling the demand for tourism in Italy as a whole. In this case, monthly expenditure data (*i.e.* tourist receipts from the balance of payments) are used as the dependent variable. One of the aims is to model the real tourism receipts, commonly used in time-series empirical studies on tourism. The second variable is the real aggregated budget share for the main source countries, commonly used in cross-section studies. Monthly data are used for the period 1972:1 up to 1990:12.

- **Chapter 9: General Discussion**

In Chapter 9, details are given on the main contributions of the present thesis to the tourism literature. The initial propositions are investigated in the light of the findings obtained from this empirical analysis. For the first proposition, an understanding is given on whether more advanced econometric approaches are able to give insight to modelling and estimating the demand for tourism. For the second proposition, it is reported whether it is appropriate to separate domestic from international tourism in terms of evolution of tourists' flows, seasonality, statistics and econometric findings. Under the third proposition, similarities and/or differences that have been encountered in estimating tourism demand at different time frequencies are underpinned. An analysis of the empirical findings in terms of economic theory is carried out. Finally, for the last proposition to be investigated, it is assessed whether any conflict emerges between theory and econometric findings from this analysis.

- **Chapter 10: Conclusions**

Chapter 10 gives concluding remarks.

CHAPTER 2.

METHODOLOGY

Aim of the Chapter:

To introduce the methodological steps followed in this thesis.

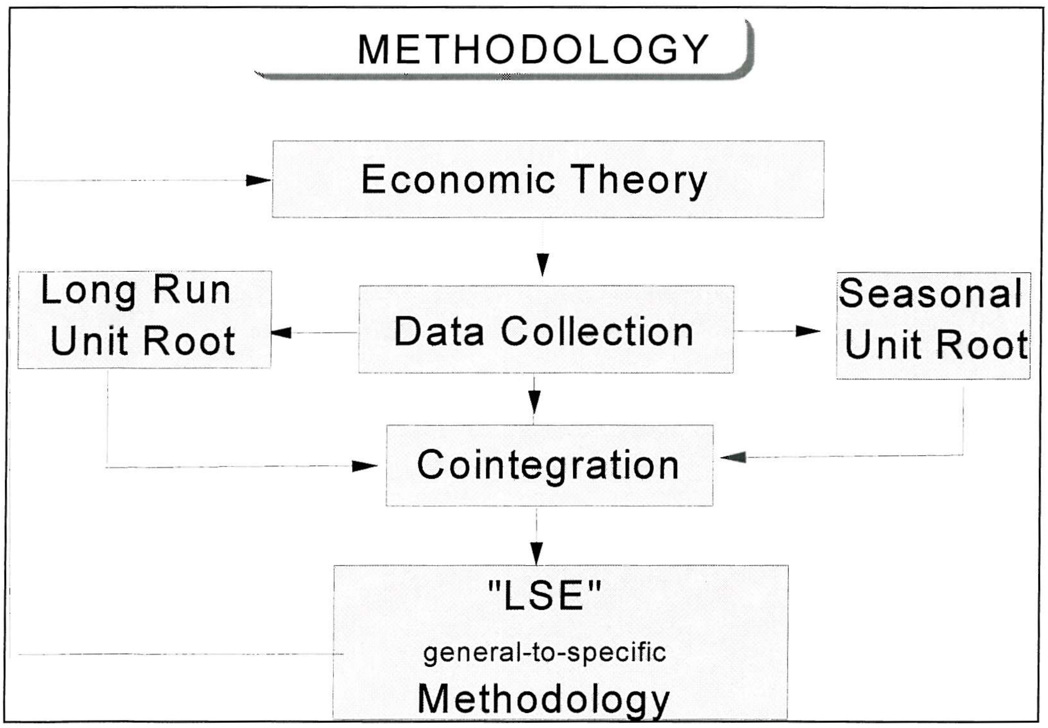
2.1 INTRODUCTION

This chapter will introduce the main methodological steps followed in this thesis. The aim is to use some recent developments in econometric methodology to test the theory and analyse the demand of tourism in Italy and in the Sardinian Province of Sassari. The following sections are dedicated to the topics which this work is based on.

2.2 METHODOLOGY

The distinct research steps for this thesis are shown in Figure 2.1.

Figure 2. 1 Methodology of the Thesis



Economic theory is derived from a literature review on tourism economy. The next step consists of linking together the theory with the empirical practice. For this specific study, data collection is obtained both in the local *Ente Provinciale per il Turismo* (EPT of Sassari) and in the official statistical sources, e.g. Istituto Centrale di Statistica (ISTAT), Datastream and Bank of Italy. From the raw data, approximations to the variables of interest are calculated on a monthly, quarterly and annual frequency. Such variables are tested for both possible seasonal and long run unit roots. Once the status of the variables of interest has been established, further testing for possible cointegration is carried out whenever necessary. Hence, the LSE general-to-specific methodology is used in modelling the demand for tourism. Finally, the empirical results obtained are compared with economic theory.

In the following sections, a more detailed analysis of each of the steps shown in Figure 2.1 is given.

2.3 ECONOMIC THEORY

Economic analysis of tourism involves modelling the supply and/or the demand side. This thesis focuses on the demand side; however, some consideration of the supply side will be given in Chapters 4 and 5. Very few studies exist on the demand for tourism in Sassari Province (Solinas, 1992; D.E.I.S., 1995; Contu, 1997) and none of them makes use of the most recent econometric methodology.

The aim of the thesis is to analyse the most significant determinants of the demand for tourism in the north of Sardinia and in Italy. According to neo-classical consumer demand theory, a tourist is a consumer who derives utility from a vector of goods and services that range from food through to travel and recreation. Consumer theory also suggests that an individual consumer maximises his/her utility subject to a budget constraint. Thus, by setting up the utility maximisation condition subject to a budget constraint, one can derive the tourism demand equation by solving using the Lagrange multiplier (see Var *et al.*, 1990). By means of this conceptual model one can understand the main factors which influence the international and domestic demand of tourism for the north of Sardinia and Italy. The most relevant determinants are the following: the personal disposable income level of the potential tourists; the price of the commodities and services of tourism; the price of substitutes; the exchange rate on the grounds that some consumers may be more aware of exchange rates than destination costs of living for tourists; the tastes and preferences of the potential tourists (Archer, 1976). The generic demand equation for tourism can be written in the following manner:

$$D = f\,(IN,\ RP,\ EX,\ SP,\ DM)$$

D = Dependent variable. It can be defined as the demand for tourist goods and services. In several empirical studies, different proxies are used as the dependent variable. In some of the studies, the level of expenditure and receipts on goods and services is used as a measure of consumption (Smeral, 1988; Di Matteo and Di Matteo, 1993; Gonzàles and Moral, 1996; Uysal and Roubi, 1999). In many other studies, the number of arrivals is taken as the dependent variable (Martin and Witt, 1989; Makridakis *et al.*, 1989; Witt and Witt, 1992; Carraro and Manente, 1996). Which is the variable to best approximate the demand for tourism? The answer depends on the aim of the analysis. As Sheldon (1993) points out, "measures of international tourism volume and international tourism expenditures are both important for a destination" (p.13). Forecasts of the demand for tourism, using tourist expenditure as the dependent variable, are needed to assess the economic impact of tourism. Forecasts of the demand for tourism, using the number of arrivals as the dependent variable, are important for private tourism businesses and for governments in planning their activities in terms of investments and infrastructure needs. In analysing the demand for tourism in the north of Sardinia, the number of arrivals of tourists is taken as the dependent variable. Even if the number of arrivals is the variable one wishes to model, using economic theory for expenditure involves an approximation. This choice is constrained by the availability of the data. Figures on tourist expenditure are, in fact, not available for the Province of Sassari. One of the limitations of the Bank of Italy reports in terms of tourism receipts is that there is no availability of disaggregated data by region and/or province (see Ballatori and Vaccaro, 1992). On the other hand, the analysis of Italian tourism will involve the use of expenditure as the dependent variable.

IN = Income is considered as one of the main relevant explanatory variables in the analysis of tourism. Tourism, in fact, is defined as a consumption good. International tourism activity might be regarded as a luxury good whereas domestic tourism may on estimation appear as a necessity good. Income, according to economic theory as well as empirical findings, constitutes one of the main indicators of an origin country, and individual wealth. One would expect that an increase in income level leads to an expansion in tourism. Empirical studies have shown that international demand for Italy presents values for income elasticities within a range of 1.00 (Germany) to 2.40 (for UK) (Syriopoulos, 1995). Note that these elasticities are obtained by estimating the foreign demand for tourism in Italy in terms of expenditure. Malacarni (1991), in estimating the demand for tourism in terms of numbers, finds a value of income elasticity of 1.49 for the foreign demand and a value of 0.92 for the domestic demand. The latter results show foreigners seem to consider the Italian destination as a luxury good more than the locals. An increase in income level leads to a tourism flow expansion. The expected sign of this variable is positive (Summary, 1987).

RP = Relative Price. Another important determinant of the demand for tourism is the "own price" of goods and services. Two types of costs can be considered: living cost in the destination country and travel cost. As Sinclair (1998) and Lim (1999) report, transport costs have appeared to be statistically not significant and in the majority of single equation studies have been excluded. Price effects can be substantial. Many studies confirm that the elasticity of tourism commodities with respect to a unitary change in the holiday price is negative and sometimes more than unity (Grasselli, 1982; Gardini, 1984; Syriopoulos, 1995). In Syriopoulos (1995) for example, the range of elasticities is within the range -0.38 for United States to -1.61 for Germany, when Italy is considered as the destination country. One of the aims of this thesis is to test these findings for the Province of Sassari.

EX = Exchange Rate. In the empirical tourism literature, the exchange rate is included as an explanatory variable. It is used either as a proxy for the tourist price index or together with the relative price. The main evidence for this combination is that the international visitor will consider the exchange rate before going to a certain destination country. Hence, the exchange rate is seen as a good approximation of the holiday cost. "Prices are seldom completely known in advance by travellers so that the price level foreseen by the potential traveller will depend predominantly upon the rate of exchange of his domestic currency.... The rate of exchange can be expected to be a prime indicator of expected prices" (Gray, 1966, p.86). Including the exchange rate alone as a proxy of the tourist price index can lead to biased results by not taking into account the inflation rate of the destination country. Hence "though the exchange rate in a destination may become more favourable this could be counter balanced by a relatively high inflation rate" (Witt and Witt, 1992, p.19). In this study the relative price and exchange rate will be included separately.

SP = Substitute Price. Economic theory suggests that the price of substitutes are another relevant determinant of demand. In the tourism literature, a common specification is to consider the substitution price between tourist visits to the foreign destination under consideration and domestic tourism (Gray, 1966; Martin and Witt, 1988). On the other hand, Sinclair and Stabler (1997) point out "occasionally, the prices and exchange rates of other competing destinations have been also incorporated" (p. 42). For this argument, amongst the other studies, see Gonzàles and Moral (1995), Syriopoulos (1995) and Lee *et al.* (1996). Chapter 6 will evaluate which variable best explains tourism demand.

DM = Extra Economic Variables. In this thesis, the models will include other determinants which are assumed to have an impact on tourism demand. Amongst these are a weather variable, an "Easter" dummy and seasonal dummies. The construction of these variables will be discussed later in the thesis.

2.4 DATA COLLECTION

Bearing in mind the main determinants for tourism demand, the next step consists in the collection of relevant data. In this way, it is possible to analyse and build a model for the demand of tourism.

One of the aims of this thesis is to consider whether one can reach common findings using data at different frequencies (*i.e.* annual, quarterly and monthly). The number of tourist arrivals in Sassari Province are available in a monthly frequency and are supplied by the government agency EPT of Sassari. These data are defined in terms of the number of tourists' arrivals in all registered accommodation in the north of Sardinia. Such data omit tourist movement in private and non-registered accommodation. According to Solinas' study (1992), registered accommodation supply represents almost 1/3 of the total accommodation supply. However, such an omission is common to many tourist statistical sources, as pointed out by Lickorish (1997).

The data used to create the other explanatory variables are obtained from several statistical sources. The Bank of Italy is the source for exchange rates and tourist receipts data. The International Financial Statistics (IFS) Datastream is the source for the industrial production index, consumer price index and private consumption. The Organisation for Economic Co-operation and Development (OECD) is the source for the number of tourists' arrivals for the competitors. The Central Institute of Statistics in Italy (ISTAT) is the source for the consumer price index in Sassari and the annual Italian personal disposable income. Of the non-economic variables, the weather variable is supplied by the Agricultural University of Sassari. This variable is expressed in terms of the monthly average temperature in Sassari.

2.4.1 SPECIFICATION FORM

The collected data will be transformed to the appropriate variables, as economic theory suggests. In the next chapters, some experiments will be carried out in adopting either a linear or a logarithmic specification form. The majority of the empirical studies on tourism demand employ the logarithmic specification (see Gray, 1966; Quayson and Var, 1982; Martin and Witt, 1988; Lee, *et al.*, 1996). However, as Qiu and Zhang (1995) highlight, the superiority of the logarithm or the linear form has to be supported by the data. In order to test the right specification form the Box and Cox (1964) test will be adopted. In running the test, Griffiths *et al.* (1993) description of the Box-Cox procedure is followed for choosing between linear and log-linear functional forms. The procedure is as follows. A model is estimated both in a linear and log-linear specification. The sum of the squared errors of the two specifications are saved (SSE_L and SSE_{LL}). The aim is to test the null hypothesis that the two models are empirically equivalent. If the null is rejected one needs to establish which specification fits the data better. In doing this, one calculates the SSE for the linear model with $(Y/\bar{Y}_G)$ as the dependent variable. Note that $\bar{Y}_G$ is the geometric mean defined as follows:

$$\bar{Y}_G = exp\left\{\frac{1}{T}\sum_{t=1}^{T} lnY_t\right\}$$

Hence, the sum of the squared errors for the latter model are equivalent to $(SSE_L/(\bar{Y}_G)^2)$. The next step is to calculate the χ^2 given by the following formula:

$$l = \frac{T}{2}\left| ln\left(\frac{SSE_L / \overline{YG}^2}{SSE_{LL}}\right)\right| \sim \chi^2_{(1)} \qquad (2.4.1.1)$$

where T are the number of observations. As stated before, if the calculated value is greater than the critical value the null hypothesis fails to be accepted and the two models have to be considered as empirically different. Moreover, if the $SSE_L/(\overline{Y}G)^2$ is higher than SSE_{LL}, one concludes that the logarithmic specification form fits the data better than the linear model.

2.5 LONG RUN UNIT ROOTS

In this thesis, as previously mentioned, one will deal with monthly, quarterly and annual series. After having transformed the data to the appropriate variables of interest, one would test for the possible existence of unit roots. Hence, the next step of the methodology used in this thesis is to test for possible long run unit roots.

Dickey and Fuller's (1981) framework will be used. The theory suggests as a series can be non-stationary in the level. In particular, a series whose growth does not depend by a positive trend is defined as a random walk. To test the latter, one can make use of the such-called Augmented Dickey and Fuller (ADF) unit root test. The ADF test consists in running equation (2.5.1):

$$\Delta Y_t = \alpha + \beta T + (\rho - 1) Y_{t-1} + \sum_{i=1}^{p} \lambda_i \Delta Y_{t-1} + \eta D + \varepsilon_t \qquad (2.5.1)$$

where a constant, the first lag of the series, the lagged difference terms, a time trend (T) and seasonal dummies (D) are included. The augmentation is set to the first statistically significant lag, testing downwards and upon white residuals. Note that the ADF without any augmentation corresponds to the Dickey-Fuller (DF) test. In the next chapters, results of the ADF test will be given for each of the possible combination: equation (2.5.1) with the inclusion of the constant term, the constant and the trend, the constant and seasonals and, finally, the constant, the trend and the seasonals. Given the generic model (2.5.1), the ADF test consists in running a t-test on the coefficient of the first lag of the dependent variable. Hence, the null hypothesis is $\rho = 1$; when failing to reject the null one treats the dependent variable as non-stationary. Secondly, one can apply a joint F-test testing whether the restriction for $\alpha = 0$, $\beta = 0$ and $\rho = 1$ holds. If the F-statistic value is smaller than the correspondent critical value, one has to treat the variable of interest as a random walk. Some of the critical values can be found in Dickey and Fuller (1991, p.1063). The critical values when including the seasonal dummies are provided by the package used to run the test, that is PcGive 9.0 package (see Doornik and Hendry, 1996, pp.93-95).

The use of the unit roots test enables one to assess the status of the economic series of interest. If a series is found to be non-stationary in the level, whenever the null hypothesis fails to be rejected, it has to be differenced. For example, if Y_t in (2.5.1) has been found to be a random walk it needs to be differenced. In particular, Y_t is said to be integrated of order one if the first difference ΔY_t is stationary but Y_t is not (*i.e.* I(1)). More in general, a series can be integrated of higher order if the series differenced d times is stationary but the series differenced d-1 times is not (*i.e.* I(d)). Note also that a I(0) series is stationary in the level.

2.6 QUARTERLY AND MONTHLY SEASONAL UNIT ROOTS

Recently, many studies have involved the investigation of seasonal variation. This development is due to the realisation that the seasonal components can be the main cause for the variations in many economic time series, and that the seasonal variation in many time series is often irregular. Thus, the seasonal pattern of many economic time series cannot be described by deterministic seasonal dummies, *i.e.* it cannot be represented by a model which assumes that the seasonal components are regular and non-changing. As Hyllerberg points out (see Hargreaves, 1994, pp.153-177), there are many different causes for seasonal variation. As far as tourism is concerned, the change in tourists' preferences (e.g. winter holidays being preferred to summer holidays) or the change in the timing of vacations by institutions and/or employers can cause a shift in the seasonal pattern. The possibility of an irregular seasonal pattern can be tested by means of investigating the possible existence of seasonal unit roots. As Hylleberg *et al.* (1990) point out, in order to test for unit roots in quarterly time series one has to estimate the auxiliary equation (2.6.1). "There will be no seasonal unit roots if π_2 and either π_3 or π_4 are different from zero, which therefore requires the rejection of both a test for π_2 and a joint test for π_3 and π_4" (p. 223). The auxiliary equation is given by:

$$\phi^*(B)y_{4,t} = \mu_t + \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1} + \varepsilon_t \qquad (2.6.1)$$

where $\phi^*(B)$ is a polynomial function in B, and where:

$$y_{1,t} = y_t + y_{t-1} + y_{t-2} + y_{t-3}$$

$$y_{2,t} = - y_t + y_{t-1} - y_{t-2} + y_{t-3}$$

$$y_{3,t} = -y_t + y_{t-2}$$

$$y_{4,t} = y_t - y_{t-4}$$

moreover, μ_t represents the deterministic part, and in this particular study consists of a constant, a trend and 3 seasonal dummies. Equation (2.6.1) is fitted by OLS. Critical values are provided in Hylleberg *et al.* (1990, p.226-227).

One needs also to test for possible seasonal unit roots at a monthly frequency. As Franses (1991a) points out "testing for unit roots in monthly time series is equivalent to testing for the significance of the parameters in the auxiliary regression:

$$\phi^*(B)\ y_{8,t} = \pi_1\ y_{1,t-1} + \pi_2\ y_{2,t-1} + \pi_3\ y_{3,t-1} + \pi_4\ y_{3,t-2} + \pi_5\ y_{4,t-1} + \\ + \pi_6\ y_{4,t-2} + \pi_7\ y_{5,t-1} + \pi_8\ y_{5,t-2} + \pi_9\ y_{6,t-1} + \pi_{10}\ y_{6,t-2} + \\ + \pi_{11}\ y_{7,t-1} + \pi_{12}\ y_{7,t-2} + \mu_t + \varepsilon_t \qquad (2.6.2)$$

where $\phi^*(B)$ is some polynomial function of B..." (p.202), and where:

$$y_{1,t} = y_t + y_{t-1} + y_{t-2} + y_{t-3} + y_{t-4} + y_{t-5} + y_{t-6} + y_{t-7} + y_{t-8} + y_{t-9} + y_{t-10} + y_{t-11}$$
$$= \sum_{S=0}^{11} y_{t-s}$$

$$y_{2,t} = - y_t + y_{t-1} - y_{t-2} + y_{t-3} - y_{t-4} + y_{t-5} - y_{t-6} + y_{t-7} - y_{t-8} + y_{t-9} - y_{t-10} + y_{t-11}$$

$$y_{3,t} = - y_t + y_{t-2} - y_{t-4} + y_{t-6} - y_{t-8} + y_{t-10}$$

$$y_{4,t} = - y_t + \sqrt{3}\ y_{t-1} - 2\, y_{t-2} + \sqrt{3}\ y_{t-3} - y_{t-4} + y_{t-6} - \sqrt{3}\, y_{t-7} + 2\, y_{t-8} - \sqrt{3}\, y_{t-9} + y_{t-10}$$

$$y_{5,t} = - y_t - \sqrt{3}\ y_{t-1} - 2\, y_{t-2} - \sqrt{3}\, y_{t-3} - y_{t-4} + y_{t-6} + \sqrt{3}\, y_{t-7} + 2\, y_{t-8} + \sqrt{3}\, y_{t-9} + y_{t-10}$$

$$y_{6,t} = - y_t + y_{t-1} - y_{t-3} + y_{t-4} - y_{t-6} + y_{t-7} - y_{t-9} + y_{t-10}$$

$$y_{7,t} = - y_t - y_{t-1} + y_{t-3} + y_{t-4} - y_{t-6} - y_{t-7} + y_{t-9} + y_{t-10}$$

$$y_{8,t} = y_t - y_{t-12}$$

and where μ_t, which represents the deterministic part, consists of a constant, a trend and 11 seasonal dummies. Equation (2.6.2) is fitted by Ordinary Least Squares (OLS), for each of the time series of interest. One can test the null hypothesis of unit root both running a *t*-test of the separate π's, as well as the joint *F*-test of the pairs, and the π's in the interval π_3 ... π_{12}. Critical values for the seasonal unit roots test are given in

Franses (1991b, pp. 161-165). If the null hypothesis is rejected one can treat the variable of interest as stationary.

Note that both the tests for quarterly and monthly seasonal unit roots are run in Microfit 4.0 package (Pesaran and Pesaran, 1997).

2.7 COINTEGRATION

As shown in Figure 2.1, the findings from the seasonal and long run unit roots test lead to a possible cointegration of the variables of interest.

As Johansen (1995) points out "it is important that one allows the components of a vector process to be integrated of different orders. The reason for this is that when analysing economic data the variables are chosen for their economic importance and not for their statistical properties. Hence, one should be able to analyse for instance I(0) as well as I(1) variables in the same model, in order to be able to describe the long-run relation as well as the short-run adjustments" (p. 34). If one has a vector of time series X_t, that achieves stationarity after differencing and a linear combination $\beta 'X_t$ is stationary, the time series X_t are said to be co-integrated with co-integrating vector β. Given two generic series y_t and x_t, and the components of the vector $X_t = (y_t, x_t)'$ are both I(1), then the equilibrium error, if it exists, would be I(0) (Engle and Granger, 1987).

Many estimators of long run coefficients exist in the literature (see Hargreaves, 1994, pp. 87-131 for a more detailed review). In investigating cointegrating relations Engle and Granger (1987), for example, suggest using the Cointegrating Regression Durbin Watson approach. Amongst the others is the Johansen Vector AutoRegressive (VAR) maximum likelihood estimator. Hargreaves' (1994) Monte Carlo simulations suggest that the Johansen estimator is best, amongst other five estimators of cointegrating relations (*i.e.* OLS, Augmented OLS, Fully-Modified, Three-Step, and Box-Tiao), as long as the sample is reasonably large (about 100 observations) and the model is accurately specified.

In this thesis, cointegration testing in single equations as well as the Johansen cointegration procedure will be used.

2.7.1 Cointegration Analysis In Single Equations

One of the approaches used for cointegration in single equations is the (Augmented) Dickey-Fuller test. Assume one has two generic variables y_t and x_t that are both stationary in the first difference. From the estimated long run relationship, that is from the estimation of the following static model:

$$y_t = \beta\, x_t + u_t \qquad (2.7.1.1)$$

one would expect that $\hat{u}_t \sim \mathrm{I}(0)$ and, therefore, the two variables to be cointegrated of order CI(1,1). Whereas, the null hypothesis of no cointegration implies that $\hat{u}_t \sim \mathrm{I}(1)$. Thus, to test that y_t and x_t are not cointegrated, one has to test whether $\hat{u}_t \sim \mathrm{I}(1)$ against the alternative that $\hat{u}_t \sim \mathrm{I}(0)$. For this purpose, the ADF test is used and it takes one of the following specifications:

$$\Delta\hat{u}_t = \rho\ \hat{u}_{t-1} + \sum_{i=1}^{p} \rho_i\, \Delta\, \hat{u}_{t-1} + \varepsilon_t \qquad (2.7.1.2)$$

$$\Delta\hat{u}_t = \mu + \rho\ \hat{u}_{t-1} + \sum_{i=1}^{p} \rho_i\, \Delta\, \hat{u}_{t-1} + \varepsilon_t \qquad (2.7.1.3)$$

$$\Delta\hat{u}_t = \mu + \delta t + \rho\ \hat{u}_{t-1} + \sum_{i=1}^{p} \rho_i\, \Delta\, \hat{u}_{t-1} + \varepsilon_t \qquad (2.7.1.4)$$

If a constant term is included in (2.7.1.1) and model (2.7.1.2) is used, it will be equivalent to using model (2.7.1.3). Whereas, if a constant and a time trend are added to (2.7.1.1) and model (2.7.1.2) is used, it will be as using model (2.7.1.4). Thus, a test with just a constant implies model (2.7.1.3) with no constant in the cointegrating regression (2.7.1.1). The null hypothesis of no-cointegration is based on a *t*-test with a non-normal distribution. However, as Harris (1995) points out the standard Dickey-Fuller distribution would tend to over-reject the null. Moreover, the number of regressors included in (2.7.1.1) affect the distribution of the test statistic under the null (see pp. 52-57). The critical values have been calculated from MacKinnon's table provided by Banerjee *et al.* (1993) using the following relation:

$$C(p) = \phi_\infty + \phi_1 T^{-1} + \phi_2 T^{-2} \qquad (2.7.1.5)$$

where $C(p)$ is the p per cent upper-quantile estimate and T are the number of observations.

Note that the analysis has to be run also by regressing x_t on y_t.

2.7.2 Johansen Cointegration Procedure

One of the main limitations of the single equation model is that of the difficulties caused by the introduction of more than two variables. The Johansen cointegration analysis is thus more general. It also avoids potential conflicts when regressing y_t on x_t then x_t on y_t. It uses a simultaneous approach, which involves the interdependencies of the variables under study.

One starts analysing the cointegration relation with a p-dimensional VAR system for the series of interest. One can consider a generic p-variate vector autoregression of order k, which is specified as follows:

$$X_t = \Phi D_t + \Pi_1 X_{t-1} + \ldots\ldots + \Pi_k X_{t-k} + \varepsilon_t \quad (\mathrm{t} = 1,\ldots\ldots\mathrm{T}) \qquad (2.7.2.1)$$

where ε_t are *iid* $N_p\ (0,\Omega)$. X_t is a vector (px1) of variables and each Π is a (pxp) matrix of parameters. The vector D_t is a matrix of deterministic components, possibly containing a constant term, time trend, impulse and seasonal dummies. Since this process is non-stationary, as it includes I(1) variables, one rewrites the model in first difference terms, *i.e.*:

$$\Delta X_t = \Phi D_t + \Gamma_1 \Delta X_{t-1} + \ldots\ldots + \Gamma_{k-1} \Delta X_{t-k+1} + \Pi X_{t-k} + \varepsilon_t \qquad (2.7.2.2)$$

where,

$$\Gamma_i = -I + \Pi_1 + \ldots\ldots + \Pi_i \qquad (\mathrm{i} = 1,\ldots\ldots\mathrm{k}\text{-}1)$$

and

$$\Pi = -(I - \Pi_1 - \ldots\ldots - \Pi_k)$$

This system contains information on both the short and long run adjustments to changes in the vector X_t via the estimates of $\hat{\Gamma}_i$ and $\hat{\Pi}_i$, respectively. Note that model (2.7.2.2) is called an Error Correction Model (Engle and Granger, 1987) or Equilibrium Correction model (Mizon, 1996).

Once formulated the vector autoregressive model, the hypothesis of cointegration has to be tested. Let the rank of Π be r. Three different cases can be taken into account:

a) if Π has full rank $(r = p)$ then the vector X_t is stationary;

b) if the rank of Π is such that $r < p$ then Π can be written as the product of two r x p matrices, α and β, *i.e.* $\Pi = \alpha\ \beta'$; β is a matrix of long run coefficients and α is a matrix of weights which represent the speed of adjustment to the equilibrium.

c) if the rank of Π is zero, the Π will be a zero matrix and no cointegration exists.

One is interested in case b), thus in testing the null hypothesis of reduced rank, *i.e.*:

$$Ho : rank(\Pi) = r < p$$

To test the cointegration hypothesis Johansen (1988) has introduced the likelihood ratio test statistic for the hypothesis that there are at most r cointegrating vectors. In Lütkepohl (1991, p.384) notation $Ho: r = r_0$; one tests a specific cointegration rank $r = r_0$ versus the alternative $H_1: r_0 < r \leq p$. One uses:

$$\lambda_{trace} = -2 \ln (Q) = - T \sum_{i=r_0+1}^{p} \ln (1 - \hat{\lambda}_i) \qquad (r_0 = 0,1,2,.......,p-2,p-1)$$

where Q is given by the ratio between the restricted maximum likelihood and the unrestricted maximum likelihood, and $\hat{\lambda}_i$ are the eigenvalues of a particular matrix. Similarly, the likelihood ratio test for testing $Ho: r=r_0$ versus the alternative $H_1: r=r_0+1$ (that is testing for the existence of r_0 cointegrating vectors against the alternative that r_0+1 cointegrating vectors) is given by:

$$\lambda_{max} = - T \ln (1 - \hat{\lambda}_{r_0+1}) \qquad (r_0 = 0,1,2,.......,p-2,p-1)$$

Critical values are provided by Johansen and Juselius (1990) and Osterwald-Lenum (1992). However, these values refer to the case in which the constant term is included either unrestrictedly or restrictedly in the cointegrating space. Critical values for the inclusion of other deterministic variables, such as seasonals, are not available.

2.8 LSE GENERAL-TO-SPECIFIC METHODOLOGY

In this thesis, the so-called LSE methodology is adopted in modelling time series data. In this section, a brief outlook of the main components of the LSE econometric modelling is given, based on the comprehensive survey provided by Mizon (1996).

The LSE methodology can be considered as the *in medio* between the extreme methodologies of the Structural Econometric Models (SEMs) and Vector AutoRegressive models (VARs) by Sims (1980). The first can be seen as a theoretical methodology. It is based on *a priori* economic theory that defines both the exogeneity status of the variables and the restrictions for the identification of the structural parameters. The second can be seen as an empirical methodology. The modelling describes the dynamic structure of the relationships between variables. Hendry and

Mizon (1990) propose: SEMs can be derived from an underlying congruent VAR representation of the data via a sequence of reductions. Hence, in Hendry and Mizon methodology both VARs and SEMs play an important role. Starting with a general Unrestricted Vector AutoRegressive model (UVAR), one obtains a Parsimonious VAR (PVAR) via a series of valid statistical reductions; hence, a congruent SEM is achieved. This final congruent SEM can be viewed as an important source for achieving meaningful and interpretable hypotheses in terms of economic theory; moreover, the capability of these hypotheses to encompass the PVAR can be statistically tested.

The main central concepts, on which the LSE methodology is based, follow. An econometric model needs to be congruent and encompassing.

1. Which are the information for which a model can be recognised as congruent?

- Economic Theory.

One of the requirements for an econometric model is to be founded upon economic theory. The theory is useful in choosing the variables to include in the model as well as the functional form to characterise the relationship between them. It is desirable for the econometric model to be coherent with economic theory. However, as Hendry (1993) points out, the findings can be in contrast with economic theory. The divergence between theory and empirical evidence is the first step for the development of new theories. The main argument is the theory cannot be considered as endowed with veracity *a priory* and the theory has to be proven by evidence. This proposition is still the object of major debates amongst academics.

- Relative Past, Present and Future Sample Information.

a) A model is not congruent with the past sample information if the errors are correlated with their lagged values. This type of congruence can be tested by the serial correlation test. It represents a way for checking the adequacy of the dynamic specification of the model.

b) Several tests can be used to test for model congruence with the present sample information. These are tests of homoscedasticity, omitted variables and normality in the error distribution.

c) A model is considered as congruent to the future sample information whenever the parameter estimates are approximately constant across varying estimation periods.

Tests are available to find out such a property: e.g. Chow (1960) prediction test statistics.

- Measurement System.

This is another property required for a congruent model. This means that the variables included in the model have to be transformed in an admissible way. One can extend the concept to other characteristics of the data such as stationarity, deterministic non-stationarity, integratedness and seasonality.

- Rival Models.

Encompassing is an essential property for a model to be congruent. For this, one requires a model to dominate other rival models. Moreover, a model that encompasses the general model and is data admissible will be the dominant model. Hence, the parsimonious encompassing refers to a model that is an acceptable reduction of the congruent embedding model.

2. What is the information for which a model can be recognised as encompassing?

- Encompassing is the other property of the LSE methodology. Previously, a definition of parsimonious encompassing has been given. However, there is the possibility that further information is available after having completed the modelling and having reached a parsimonious encompassing model with respect to the "old" information set. In this case, it is necessary to incorporate the new information and find out if the original model is still robust for this new set of information. As Mizon (1996) writes, "each model is evaluated with respect to an information set more general than the minimum one required for its own implementation, thus achieving robustness to extensions of its information set in directions relevant for competing models" (p.122).

The strategy of "general-to-specific" is recognised to be the best strategy within the LSE methodology. Starting with a very general model, it is possible via a testing down procedure to reach a congruent and encompassing model, which might also validate *a priori* economic theory.

In estimating an autoregressive distributed lag model the choice of the lag length is of extreme importance. In choosing the lag length one might use the statistical tests: Wald test and likelihood ratio test (LR). These tests allow one to test whether it is statistically significant to reduce the lag length by one. The lag length of a model can be also chosen by making use of information criteria: Hannan-Quinn, denoted as the HQ criterion; Schwarz, denoted as the SC criterion; finally, Akaike, denoted as the AIC criterion. The information criteria are defined as follows:

$$HQ = -2 \log L/T + 2n \log(\log(T))/T$$

$$SC = -2 \log L/T + n \log(T)/T$$

$$AIC = -2 \log L/T + 2n/T$$

where L is the maximised likelihood, T is the sample size and n is the number of parameters. The estimated information criteria are chosen so that they are minimised.

The final step of the methodology used in this thesis is the feedback to economic theory (Figure 2.1). The results obtained from the congruent and encompassing model are compared with the theory. The initial propositions with respect to income and price elasticities and, in general, the capability of the independent variables to explain the dependent variable will be examined.

2.8.1 Developments Of The General-To-Specific Procedure

Recent developments have been made in adopting the general-to-specific approach. Hoover and Perez (1999) find that Hendry and Doornik's computer-automated *PcGets* performs well in evaluating econometric model selection strategies by simulation. A further improvement has been reached by Hendry and Krolzig (1999a and 1999b). They implement *PcGets* by introducing concepts from the LSE methodology such as: tests for pre-selection and encompassing tests for choosing between multiple models which are found to be congruent with the information set. A unique model is reached either from the combination of congruent contenders when the algorithm terminates or by using an information criterion. This new "data mining reconsidered" needs still further improvements as Hendry and Krolzig (1999b) point out. Problems appear in the appropriate parameterisation, functional forms, variable choice, and inclusion of seasonals and dummies which requires a careful prior analysis. Moreover, issues as the role of structural breaks and, for example, first difference constraints on the lags of stationary variables, have still to be investigated further. It is worth noting that *PcGets* in Hendry and Krolzig (1999a and 1999b) is employed to reconsider existing empirical estimations (*i.e.* UK money demand and the

US narrow-money demand). In their re-analysis, Hendry and Krolzig (1999a) write "it remains to stress that these cases benefit from "fore-knowledge (e.g. of dummies, lag length etc.), some of which took the initial investigators time to find" (p.21). Moreover, Monte Carlo investigation of model selection, which is only possible because *PcGets* is automatic, suggests that the pre-test bias is quite manageable. That is, using 5% level tests gives a reasonable probability of finding the right model, rather than always finding something, simply because one has done so much testing.

2.9 OTHER METHODS

The following subsections are dedicated to give an account of other methods that are used in this thesis.

2.9.1 Simultaneity

In Chapters 4 and 5, the problem of simultaneity will be examined by including some supply variables, that is the number of boats and the number of flight arrivals in the north of Sardinia. One will employ the Durbin-Wu-Hausman's simultaneity test when using annual and monthly data (see Pindyck and Rubinfeld, pp. 303-305 for more details). Such a test will assess the lack of correlation between a right-hand side variable and the error term. The null hypothesis of no simultaneity implies that the variable of interest will be treated as predetermined; the alternative hypothesis is that such a variable can be treated as endogenous.

A brief note, in terms of terminology used, is due. According to the *Cowles Foundation approach* the classification of variables into "exogenous" and "endogenous" (in the case they are determined outside or within the model) and the causal structure of the model are given *a priori* and are untestable. This approach has been criticised on several grounds (see Maddala, 1992, p.389). In particular, Leamer (1985) suggests re-defining the concept of exogeneity. Two concepts of exogeneity are distinguished:

1) *Predeterminedness*. A variable is predetermined in a particular equation if it is independent of the contemporaneous and future errors in that equation.

2) *Strict exogeneity*. A variable is strictly exogenous if it is independent of the contemporaneous, future and past errors in the relevant equation.

Engle, Hendry and Richard (1983) extend such concepts in order to "seek conditions which validate treating one subset of variables as given when analysing others" (Hendry, 1995, p.156) and the new notions of weak exogeneity, strong exogeneity and super-exogeneity are given. One can argue that the notion of weak exogeneity involves more than a lack of correlation between a right-hand side variable and the error term.

In this study, the Durbin-Wu-Hausman's test is employed. This test assesses whether the variable under study is predetermined. One can assume that the generic demand and supply models can be expressed as follows:

$$y_t = \alpha_0 + \alpha_1 x_t + \varepsilon_{1,t} \qquad (2.9.1.1)$$

$$x_t = \beta_0 + \beta_1 y_t + \beta_2 z_t + \varepsilon_{2,t} \qquad (2.9.1.2)$$

where:

y_t = a generic explanatory variable;

x_t = variable which is thought to be determined by the level of y_t;

z_t = variable treated as the instrument.

To test for the existence of simultaneity one follows two steps. Firstly, the reduced form is obtained by regressing x_t on the variables included into equation (2.9.1.2) and the instrument variable z_t. Hence, in the second phase, the residuals obtained from this regression (say $res = \hat{w}$) are added into equation (2.9.1.1). The null hypothesis of no simultaneity would be rejected at the 10% level when using the two-tailed t-test. Hence, if the null is rejected the variable (x_t) can be treated as endogenous.

2.9.2 Testing For Structural Breaks

In Chapter 5, one will investigate the possible existence of structural breaks in the seasonal pattern. The presence of seasonal unit roots at some frequencies on one hand, and problems of specification form appeared in the residuals on the other, can be considered as a symptom of possible non-stationarity which needs to be examined. Adopting the Chow test (1967) the null hypothesis of no change is tested, that is the seasonal coefficients have remained the same within the period under investigation. An F statistic is calculated and this value will be compared with the correspondent critical value from the conventional tables. The rejection of the null implies that a structural break is present. However, such a test assumes the knowledge of the change point. In the case in which the change point is unknown, one needs to make use of non-standard distribution. Andrews (1993) has investigated and provided critical values which one will use in testing for structural changes.

2.9.3 Non-linear Model

In Chapter 8, a non-linear model in modelling real tourism expenditure for Italy is used. The weighted real industrial production for the main origin countries has been found to be I(1) from the ADF test. This implies that the income proxy is non-stationary in the level, whereas the tourist expenditure is stationary in the level. Hence, one might want to investigate a non-linear specification for the real income proxy, in order to reconcile an I(1) variable as the explanatory variable with an I(0) variable as the dependent. The non-linear model is run with TSP 4.3A. The approach involves n iterations whose objective is to minimise the residuals sum of the squares and Gauss's method is used. Firstly, derivatives of the equation residual with respect to each parameter are computed. Then, TSP regresses the current residual on the derivatives, and the resulting regression coefficients are used as the changes in the parameters. If a local minimum of the residual sum of squares is achieved the final coefficients in the artificial regression will be zero. After convergence is achieved, non-linear estimation statistics have a standard OLS interpretation, although the underlying theory is only asymptotic.

2.10 CONCLUSION

This chapter has given a description of the main methodological steps adopted in the thesis. Firstly, a tourism literature review is undertaken which gives the economic theoretic assumptions. One of the aims of the thesis is to test the theory using advanced econometric modelling. An extensive data collection exercise is undertaken for achieving the purpose. Data are collected on different time frequencies and appropriate variables are constructed. In Section 2.6, an account of Franses (1991a, and 1991b) and Hylleberg *et al.* (1990) methodologies is given; these methods will be used to test for the existence of possible monthly and quarterly seasonal unit roots, respectively. Hence, the ADF test for testing the existence of long run unit roots is described. Once established the integration status of the economic series of interest, a cointegration analysis is undertaken both using single equations and Johansen cointegration analyses. These cointegration tests have been described in Sections 2.7.1 and 2.7.2. Hence, an account of the LSE general-to-specific methodology is given as provided in Mizon (1996).

Other topics used in the thesis have been discussed in Section 2.9. In Section 2.9.1, an account is given of the Durbin-Wu-Hausman's simultaneity test when including supply variables into the model. In Section 2.9.2, it has been discussed the use of Andrews' (1993) critical values in the case of unknown structural break point. Finally, Section 2.9.3 has given an account on non-linear modelling which will be one of the objectives of Chapter 8. Further details for such methods will be given when they are used.

CHAPTER 3.

CHARACTERISTICS OF INTERNATIONAL AND DOMESTIC DEMAND FOR TOURISM IN THE ITALIAN PROVINCE OF SASSARI: A GENERAL INTRODUCTION

Aim of the Chapter:

To examine and identify possible differences between domestic and international tourism in the north of Sardinia.

3.1 THE DEVELOPMENT OF TOURISM IN THE NORTH OF SARDINIA

According to Crenos (2000), the Sardinian economic system has seen structural changes since the '50s. Its traditional specialisation sectors, that is agriculture and mining, have declined due to public policies of industrialisation. High capital industries, that is chemicals, energy and construction, have driven the economy to a brief period of expansion until the first half of the '70s. However, along the decades, this transformation has led to a progressive worsening of Sardinian performance in terms of internal productivity capacity (GDP from labour and GDP *per capita*) relative to other Italian regions. Crenos points out the failure of the actual productivity system in creating jobs opportunities. There are several causes: the small size of the industries, low levels of highly specialised human capital and institutional inefficiencies.

This study is concerned with tourism that can be considered as a key sector in the economic development strategy of this Mediterranean island. Sardinia represents one of the important tourist areas in the Mediterranean Sea. Its main attraction for tourists consists in the quantity and quality of its natural resources. Its "isolation", its physical distance from the main Italian tourist destinations (e.g. *Costa Adriatica*) and investment projects such as the well known *Consorzio Costa Smeralda* (North-East Coast) in the early Sixties are all elements that increase the image of this island as an "exclusive and elite holiday" destination.

In this study, particular attention is given to the flows of foreign and domestic tourists into the north of Sardinia, known as Sassari Province. This Province sees 54% of the overall flows of tourism to this Italian island (see Confcommercio, 1994). Since the Fifties the north of Sardinia has been the first Sardinian Province to invest in tourism. The first investment projects were seen in Alghero (North West Coast), followed by the Consorzio of Costa Smeralda by Prince Aga Khan. In the Seventies, other coastal centres were developed such as: Stintino, La Maddalena, Palau and Santa Teresa di Gallura. Another element that encouraged tourist flows to Sassari Province is the presence of two of the most important airports and harbours in Sardinia that serve both national and international traffic to the Western and Eastern Coast.

3.2 FOREIGN VERSUS DOMESTIC DEMAND

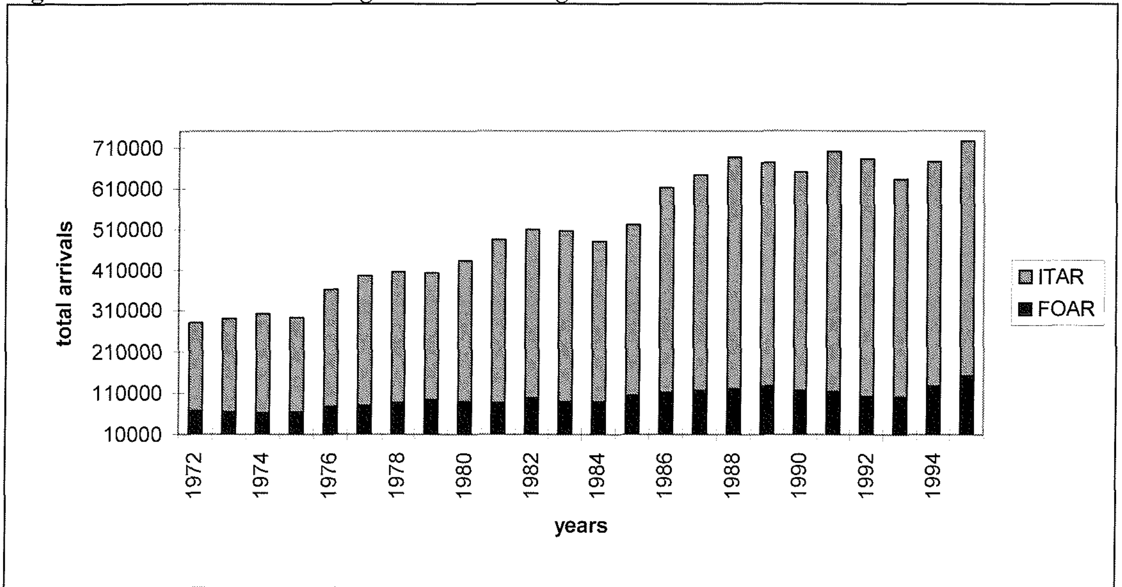
In the tourism literature, domestic and foreign flows of tourists are investigated separately, due to their different characteristics. The majority of empirical studies are dedicated to the analysis of international demand for tourism. Only a few consider the domestic component (see for example Raeside *et al.*, 1997; Seddighi and Shearing, 1997) or both the components (see Malacarni, 1991).

The following sections are dedicated to capturing the characteristics and possible differences of the two components in Sassari Province.

3.2.1 Tourist Flows And Their Evolution

The first main difference between domestic and international flows can be noticed in terms of percentages; the domestic flows of tourists count for the 81% against a 19% of foreigners. These percentages are calculated in terms of the number of arrivals of tourists in registered accommodation, averaged for the period between 1972-1995. Figure 3.1 shows the time series from 1972 to 1995 for the arrivals of foreign and domestic tourists, and gives us a better understanding of the historical evolution of the two components.

Figure 3. 1 Domestic and Foreign Arrivals in Registered Accommodation for Sassari Province



Source: Figures based on EPT (Ente Provinciale per il Turismo) Sassari. Key words: ITAR (Italian tourists' arrivals); FOAR (Foreign tourists' arrivals).

The evolution of tourism in the north of Sardinia seems to be dependent on the economic events that have occurred both at the national and international level. In the second half of the Seventies an increase of both the foreign and domestic flows of tourism can be seen, due to a general European economic prosperity (Sesto Rapporto sul Turismo Italiano, 1995). The increasing of individual and aggregate income has a positive influence on the demand for tourism, which is considered as a normal good.

The years between 1976 and 1979, as far as the foreign demand is concerned, show an increase of the arrivals also encouraged by the devaluation of the lira. The Eighties are characterised by deep economic changes. The first half of this decade saw monetary restrictions in both the USA and UK to diminish the level of inflation, which might have implied a decreasing number of foreign arrivals. Note that United States and British tourists in the same period counted for 20.4% of the international flow of tourism towards the north of Sardinia. The second half of the Eighties was characterised by a positive economic performance together with a general optimism for a possible integration for the EEC countries, which saw a new expansion of the arrivals of foreign tourists. The monetary restrictions adopted by the major industrialised countries in 1989, together with an average reduction in GDP growth from 2.1% to 0.2% in 1991 (Sesto Rapporto sul Turismo Italiano, 1995), have produced a negative effect on tourism (see Figure 3.1, 1990-1992). The subsequent expansion of tourism flows in 1994 and 1995 may be due to different causes: the depreciation of the lira at the end of 1992, political crises in some of the main competitor countries in the Mediterranean area (e.g. ex-Yugoslavia), the stabilisation of prices for tourist goods and the intensification of promotional policies by private firms and public agencies in Sardinia.

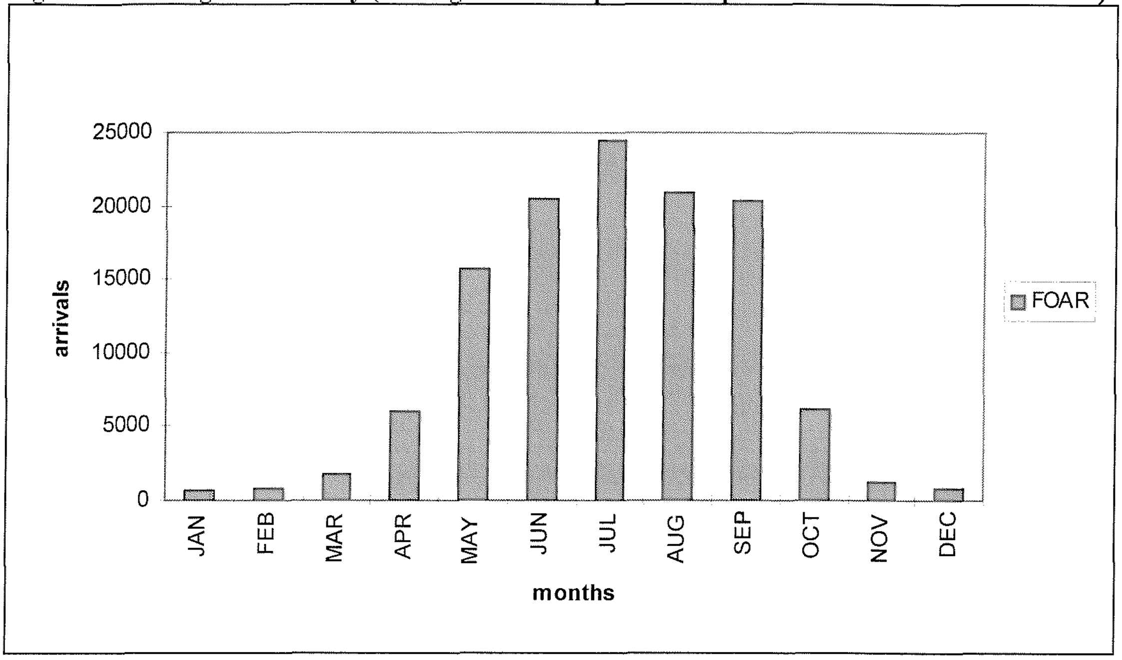
As far as domestic arrivals are concerned, an upward trend of arrivals can be seen within the period under study, with the exception of 1984, 1989, 1990 and 1993; the latter shows the lowest performance since the second half of the 80's.

3.2.2 Seasonality

Other aspects distinguishing the domestic to the foreign demand for tourism have to be highlighted. One of the main characteristics of tourism is its seasonality. The distribution of public holidays and school vacations, as well as the climatic conditions, have a strong impact on tourism. For example, in the western industrialised countries (e.g. Italy and France) the major periods of public holidays are either in the summer months or at Christmas or Easter. Thus, many resorts and tourist regions experience "overcrowded" and the "holiday-rush" seasons (Hartmann, 1986). This strong seasonality appears to be problematic for the tourist service system in the north of Sardinia. There are a variety of problems, including: an uneven utilisation of tourism facilities such as hotels, holiday villages, beaches and entertainment, whereas a more uniform utilisation might lead to a reduction in average prices and an improved profitability. In the low season, the work force is under-utilised with the consequence of increased uncertainty for labour conditions. The public sector, on the other hand, faces problems related to the optimal scale of the public services and infrastructure. In recent years, both private and public agencies (such as the *Ente Sardo Industrie Turistiche* (ESIT)) have undertaken several projects aimed at the promotion and marketing of specific products for the period of low season (between October and May). These could cause changes in the seasonal pattern. It should be noted that Sardinia's mild climate, with average temperature between 9^0 C degrees in the colder months and 25^0 C degrees in the hotter months[1], allows for an extension of the tourist season, in particular for the foreign markets.

Figures 3.2 and 3.3 show the comparative seasonality of foreign and domestic tourism in the Nineties, calculated as average arrivals per each equivalent month within the period January 1990 - December 1995. The two figures show a difference in the seasonal behaviour of the foreign and domestic flows.

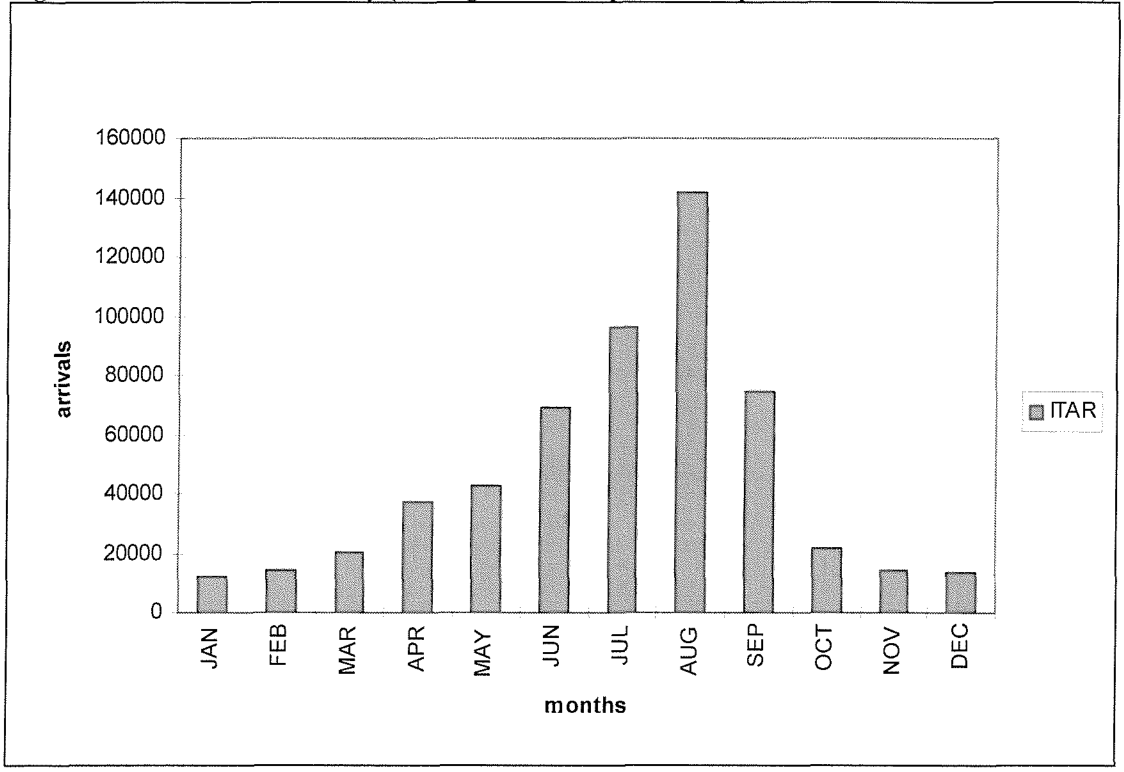
Figure 3. 2 Foreign Seasonality (Averages Arrivals per each Equivalent Month 1990:1 - 1995:12)



Source: Figures based on EPT (*Ente Provinciale per il Turismo*) Sassari. Key word: FOAR (Foreign tourists' arrivals).

[1] Note that these values are calculated in average with respect to the period under study: 1972:1 up to 1995:12 for the Province of Sassari (author's own calculation on data of: *Agronomia e Coltivazioni Erbacee dell'Università di Agraria di Sassari*)

Figure 3. 3 Domestic Seasonality (Averages Arrivals per each Equivalent Month 1990:1-1995:12)



Source: Figures based on EPT Sassari. Key word: ITAR (Italian tourists' arrivals).

The seasonality of arrivals of foreign tourists shows overall smaller variations for the months between June and September, with the highest value in July. There is some evidence that the seasonal pattern for foreign tourism is flattening out at the end of the period, which may be due to the success of various promotional projects in Sardinia. On the other hand, the seasonality of the Italian flows of tourism presents an irregular distribution, with the highest peak in August.

One can conclude that the two seasonal distributions are characterised by different aspects. Firstly, the weather conditions during the year seem to influence Italians and foreign tourists differently. Foreigners seem to try to avoid the hot months whereas domestic tourists still choose August as month to leave on holidays. This different choice depends also on "institutional" elements such as the timing of holidays and school vacations. As already stated, Italians, and in particular Italian families with school children, are more constrained than foreigners in choosing their holidays in the winter and summer.

Another component for choosing the holiday destination and/or time could be the price differential between the high and low season. May, June and September, as low season months, are characterised by lower prices for tourist accommodation. From the previous graphical comparison, it seems that foreign tourists are more sensitive than Italians to price changes. Such a hypothesis has been investigated by Malacarni (1991) in a study on Italian tourism. Malacarni has derived a price elasticity of -0.83 for the international tourism demand in Italy and a price elasticity of -0.16 for the domestic demand. One of the aims is to consider whether these results can be extended to Sassari Province case.

So far, the main differences in the evolution of tourist flows as well as in the seasonal characteristics suggest modelling foreign and domestic demand for tourism separately. However, such *a priori* assumption needs to be investigated further. In the next chapters the aim will consist in validating such a distinction within an econometric and statistical frame.

3.3 POSSIBLE DETERMINANTS OF TOURISM IN NORTH SARDINIA

In the following subsections, an account will be given of the main determinants that are expected to have an impact on the demand for tourism in the north of Sardinia in accordance with economic theory.

3.3.1 The Dependent Variable

In analysing the demand for tourism in the north of Sardinia the number of arrivals of tourists is taken as the dependent variable. As already noted in Section 2.3 (Chapter 2), this choice is constrained by the availability of the data.

Given the complexity of the tourism phenomenon, difficulties appear in finding a comprehensive and agreeable definition. In the tourism literature, as Masberg (1998) writes, dissimilar definitions appear. In this thesis, one accepts the notion provided by the United Nations Conference on International Travel and Tourism of 1963 (see Sinclair, 1998) for which: a tourist is a "temporary visitor who spends more than 24 hours in destinations other than their normal place of residence, whose journey is for the purpose of holiday-making, recreation, health, study, religion, sport, visiting family or friends, business or meetings. Those who spend less than 24 hours in their destinations are defined as excursionists" (p.4). Hence, when one talks about tourists' flows one does not separate the flows for recreation purposes from those for business.

However, the latter component can be considered as a small quota in the overall tourist flows to the north of Sardinia. Note, in fact, that the most industrialised province is that of Cagliari in the south of the region, with its own airport and harbour. More generally, it is worth noting that very few surveys exist that make a differentiation between the two components, see for example the Swedish Tourism and Travel Data Base (Nordström, 1996 gives a detailed discussion). In Italy, such a differentiation has only been carried out since 1995 by means of surveys, and, therefore, the sample size available is quite small.

As Baron (1989) points out the "trading-day factors" might be important in the analysis of monthly data, that take into account the effects of four of five Saturdays (or Sundays) in a particular month. As far as the demand for international tourism is concerned, Saturday has been chosen as the starting day of the holiday. The majority of the charter flights and boat trips to the north of Sardinia occur, in fact, on a Saturday. Therefore, the dependent variable has been adjusted in order to take into account the number of Saturdays in each month for the period under consideration. It is worth noting that the latter normalisation cannot be considered arbitrary. A previous investigation has been carried out using the raw series of the foreign arrivals (see Appendix A for a complete discussion). Such analysis encountered problems of non-normality and heteroscedasticity (at the 1% level), which have been corrected with the adjustment of the dependent variable for the number of Saturdays in a month.

In the analysis for the domestic demand of tourism, Sunday has been chosen as the starting day of the holiday. Such a choice will be supported by the results obtained from two separate models. The first model includes the series of arrivals normalised for the number of Sundays in a month as the dependent variable. The second model has the dependent variable normalised for the number of Saturdays in a month. A complete discussion is provided in Chapter 5. "It is evident that the distribution of public holidays, especially school vacations, has a strong impact on the individual timing of the vacation and travel days within the annual cycle", (Hartmann, 1986, p. 26). It is plausible to believe that a great part of arrivals of domestic clients tend to occur on Sunday, as, in Italy, for the majority of private and public activity Saturdays are trading-days. However, any day of the week is likely to be the starting day of the holidays for Italians and this point will emerge from the econometric analysis later in this thesis.

In the present study, one will also consider whether the current tourist consumption is influenced by the past demand. For this purpose, one will include the lagged dependent variable in order to determine whether domestic and foreign tourists can be defined as "psychocentric" or "allocentric"[2] (Sinclair and Stabler, 1997).

3.3.2 Economic Determinants

As already stated in the methodology chapter (Section 2.3), income, relative price, exchange rate and substitute price are considered as the main relevant explanatory variables in the analysis of tourism.

a) *Income*. The best variable to construct the income index is the personal disposable income (see Witt and Witt, 1992). However, such a variable is not available with a monthly frequency. Thus, as suggested by some authors (see Gonzàlez and Moral, 1995; Garcìa-Ferrer and Queralt, 1997) one will use the weighted average of the industrial production (see Appendix B for the construction of this variable), for the main clients of the north of Sardinia, as a proxy of income. The index of industrial production in Italy is taken as a proxy of income for the analysis of the domestic demand for tourism. From a VAR analysis, as given in Chapter 5, there is statistical evidence that the Italian industrial production index is a valid proxy of the Italian disposable income *per capita*.

b) *Own price and exchange rate*. According to economic theory another important determinant of the demand for tourism is the price of goods and services. However, there are many problems in determining a tourist consumer price index. The tourist good, in fact, is heterogeneous including costs of transportation, accommodation resorts, entertainment, souvenirs, etc. Because of the lack of such an index, many researchers use as its proxy the consumer price index. As Sinclair and Stabler (1997) point out, Martin and Witt (1987) study is virtually the only study to investigate its

2 "Phychocentric" is a tourist who prefers a familiar destination other than a new destination. In this case, a positive coefficient sign is expected between current and past demand, and the coefficient for the long lags is expected to be statistically significant. "Allocentric" is a tourist who prefers new destination for his/her own holidays. A negative coefficient sign is hence expected together with a rapid adjustment. Note that the negative sign for the lagged dependent variable could be also due to undesiderable characteristic of a particular destination.

suitability as a proxy. As many other authors, the present study will base on their empirical findings employing the consumer price index as a valid proxy.

Furthermore, in analysing the international demand for tourism in the north of Sardinia, the exchange rate is included on its own, in addition to the relative price. In particular, the exchange rate has been constructed using a weighting system based on market shares of the major clients (*i.e.* Belgium, France, Germany, Sweden, Switzerland, UK and USA) for the north of Sardinia (see Appendix B for a more detailed analysis).

c) *Substitute price*. One of the debates in the tourism literature is related to which is the most appropriate specification for the substitute price. Sometimes substitute prices and exchange rates have been included as separate explanatory variables, whereas in other studies a real substitute price was included.

The analysis will be articulated in the following manner. In Chapters 4 and 5, one will assume that the nominal substitute price could be thought to be the main determinant in explaining the demand for tourism in the north of Sardinia. In Chapter 6, one will examine the possibility to include either the exchange rate for the competitors as a separate variable or the real substitute price.

d) *Supply variables*. So far one has taken into consideration the variables linked with the demand side. "Tourism is traditionally conceived in terms of the demand-side or consumer characteristics (e.g. the duration and geographic extent of a trip taken by a traveller), which places the industry in a difficult political and statistical position" (Smith, 1995, p.34). The tourism activity, in fact, can be considered *sui generis*, since it is not defined in terms of its products but in terms of its consumers.

It is important to classify the components of the supply of tourism, which is a composite of activities, services, and industries that facilitate travel and activity away from one's usual environment. The main components of tourism supply are: natural resources which include climate, natural beauty, flora, fauna, beaches and many others; infrastructures, such as harbours, airports, roads, bus and train station facilities, hotels, restaurants, entertainments and similar structures; transportation that includes aeroplanes, boats, trains, taxis, and other facilities; hospitality and cultural resources, such as the attitude of the residents toward tourists, arts, history, traditions, sports and many others (McIntosh, Goeldner and Ritchie, 1995).

It might be relevant, therefore, to consider variables that express both the quantity and the quality of the supply services. The amount of accommodation and for example: beds, flights, trains or/and boats can define the extent of the quantity of supply; while the proportion of 3-5 star hotels and/or the ratio (beds/toilets) could be considered as a good proxy for the quality of the supply.

In order to take into consideration a component of the quantity of the supply services, in estimating the annual foreign and domestic demand for tourism in the north of Sardinia, one will include two variables that is: the number of arrivals of boats in the two main ports (Porto Torres and Olbia) and the number of arrivals of flights in the two main airports (Fertilia and Olbia)[3]. The results from estimating the international and domestic demand for tourism using annual data will suggest a further investigation by adopting the simultaneity Durbin-Wu-Hausman's test.

Providing an ample tourism supply to match anticipated demand in the long run is a challenge for the planners. The evolution of the supply of tourism is strictly linked with the evolution of the demand. In developed countries, in recent years, there has been an increase both in the standard required in, and the total demand for, tourism. Not only wealthier individuals look for leisure time but also larger and larger proportions of the total population. Demand can change quite rapidly. Supply components (e.g. infrastructures and services) are more rigid.

In a short run context it is important to detect the seasonal fluctuations of the demand and supply levels. Tourism, unlike many other products, is a composite product and has a perishable nature, since unfilled airline seats and unused hotel rooms cannot be stockpiled. "If firms selling tangible goods can deal with demand fluctuation through the inventory process, this option is not available to firms providing travel services. In the travel industry, an effort must be made to reduce seasonal fluctuations as much as possible" (McIntosh *et al.*, 1995, p.291).

The intent is to illustrate the supply situation associated with fluctuating demand levels for the Province of Sassari. In order to understand the extent of utilisation of accommodation by the total number of tourists in Province of Sassari, one introduces the following definition:

$$RU = (P * 100) / (B * D)$$

3 The source is the "*Annuario Statistico Italiano*" (1972-1996).

where:

RU = actual rate of utilisation expressed in percentage;

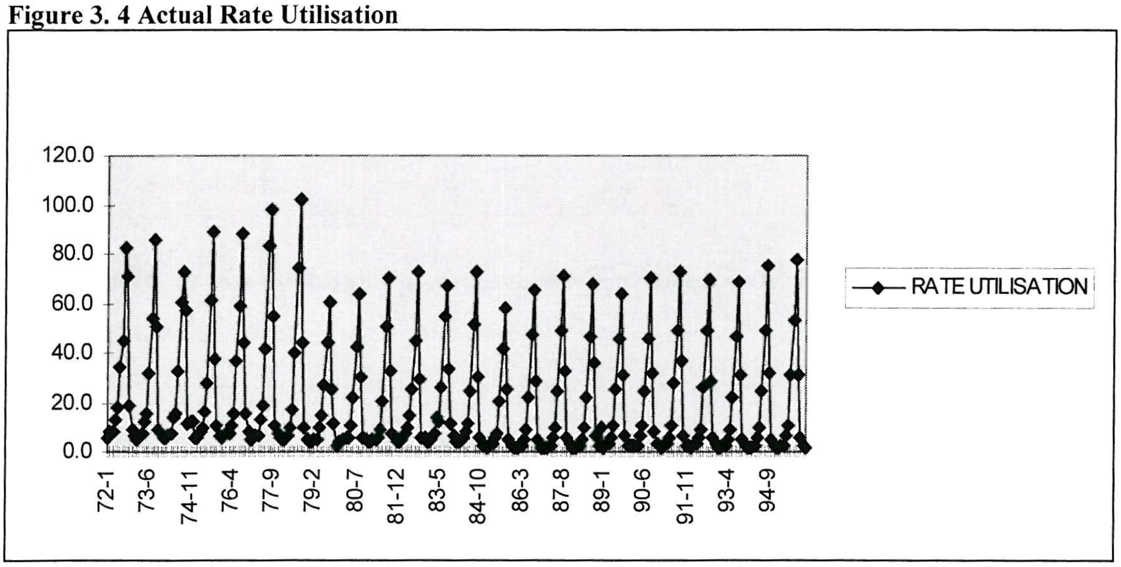
P = number of nights spent by all tourists in a month;

B = number of beds available;

D = Number of days per month.

The rate of utilisation is the main determinant of the profitability of the hospitality resources. Figure 3.4 provides a graphical analysis of the seasonal actual rate of utilisation for the period 1972(1) to 1995(12). In particular, one can notice that the low season (October - May) presents a low rate of utilisation, an average of 6% (calculated for the all period under consideration), whereas the period between June and September presents an average rate of utilisation of 40%. August, as the peak month, sees an average rate of utilisation of 71%.

Figure 3. 4 Actual Rate Utilisation



Source: Figures based on EPT Sassari data.

Therefore, during the low season accommodation will suffer from low occupancy levels, with the consequence of a loss in terms of profitability. Note also that 1978(8) sees a rate of utilisation equals 102% with the consequence of an overcrowded month. However, it probably reflects a measurement error in the data. The main conclusion is, therefore, that there is no evidence that supply acts as a constraint on demand, except in this one month.

3.3.3 Extra Economic Determinants

Tastes and preferences of tourists, and environment are considered as the main factors affecting the demand for tourism. Dummy variables can be introduced in the econometric model to take some of these qualitative determinants into account.

In this study some extra-economic determinants have been taken into consideration: a) the Easter effect which has major importance for the performance of the tourism in the north of Sardinia. Some mis-specifications have arisen in the model by not taking into account the former variable. b) The weather conditions that can also be considered as a natural resource and, therefore, regarded as a component of the tourism supply, as mentioned in the previous subsection (McIntosh *et al.*, 1995). c) Two main streams of thought are in the literature. First, a time trend is included in the model in order to pick up possible changes in consumers' tastes for a specific destination over time. Second, the time trend variable is recognised as hiding problems of multicollinearity, *i.e.* as possibly highly correlated to other economic variables such as income. In this study, a time trend will be included in the final restricted model upon a statistically significant coefficient.

3.4 CONCLUSION

So far, one has given a general introduction on the main characteristics of the tourism demand in Sassari Province of Italy. An account has been given of the differences between the international and domestic demand based on the evolution of the tourist flows and on the seasonal distributions. In accordance with economic theory, the possible determinants that might have a role in explaining the demand for tourism in the north of Sardinia have been described.

The plan of the next three chapters is the following. Chapter 4 will be dedicated to the analysis of the international demand. In Chapter 5, a model for the domestic tourism demand will be estimated. In Chapter 6, a deeper investigation for the inclusion of the nominal substitute price and exchange rate for the main competitors will be carried out.

CHAPTER 4.

INTERNATIONAL DEMAND FOR TOURISM IN THE NORTH OF SARDINIA

Aim of the Chapter:

To model and estimate the international demand for tourism in Sassari Province of Sardinia (north of Sardinia), Italy.

4.1 INTRODUCTION

The aim of this chapter is to examine the economic factors affecting the demand for international tourism in the Province of Sassari in the island of Sardinia with the view of developing an econometric model for the short run and long run analysis. As already pointed out, given the noticeable differences in the seasonal pattern of the domestic and foreign demand, particular attention will be given in studying the determinants of foreign demand and of the domestic demand separately.

In many empirical studies the demand for tourism has been investigated using annual or quarterly data (Summary, 1987; Martin and Witt, 1988; Smeral, 1988; Kulendran, 1995; Seddighi and Shearing, 1997; Song *et al.*, 2000; Kulendran and Witt, 2001), whereas very few studies deal with monthly data (Rugg, 1973; Bond, 1979; Gonzàles and Moral, 1996; Lim and McAleer, 2001). In the present study particular emphasis is given to the use of monthly time series (from January 1972 until December 1995). Such a time interval allows for the investigation of seasonal variation in the time series under consideration. It is possible, for example, that changes in the tastes and preferences of tourists can affect the seasonal pattern. Other influences on demand will also be examined. In particular, the importance of variations in the date of Easter will be highlighted, where Easter Sunday varies between March 26 and April 22 as far as the period under modelling is concerned. In addition, the so called "trading-days" effect will be carefully studied (Baron, 1989). This involves allowing for the effect of four or five weekends in a particular month. The study includes also a comparison between the use of monthly, quarterly and annual data.

Use is made of the Johansen multivariate cointegration procedure in order to identify stable long run relationships amongst the variables under consideration. In the literature, there are just a few studies which apply the notion of cointegration to modelling demand for tourism. See Sinclair (1998) for a review. Lanza and Urga (1995), for example, apply the Johansen procedure to annual tourism expenditure for 13 European countries from 1975 to 1992. Vogt and Wittayakorn (1998), apply a test for stationarity and cointegration to variables with an annual frequency in estimating the determinants of the demand for Thailand's exports of tourism. Notably, Song *et al.* (2000) apply the LSE methodology in estimating the outbound tourism demand in the UK; annual expenditure data from 1965 to 1994 have been employed.

In the present study, the total arrivals of foreign tourists in all registered tourist accommodation in the north of Sardinia will be analysed and modelled. "Tourist arrivals" have been used in many empirical studies and are considered a good proxy for the demand for tourist goods and services (Crouch, 1994, gives a detailed review).

This study assumes that the supply of tourism services and in particular the supply of accommodation in the north of Sardinia is perfectly elastic in the short run, *i.e.* not acting as a constraint on demand. The support for this assumption has already been given in the general introduction (Chapter 3). As a reminder, the only exception was for August 1978.

In terms of economic analysis, emphasis will be given in exploring the relationship between short and long run income and prices elasticities.

Section 4.2 will be dedicated to the analysis and estimation of the international demand for tourism in the north of Sardinia employing monthly data. In the subsections the analysis will be concentrated on the following: a time series analysis for the international demand; the notion of cointegration will be introduced for the variables which present long run unit roots *(i.e.* relative price and exchange rate); a monthly frequency model will be estimated, where short run and long run relationships will be taken into consideration. In Section 4.3, an annual model will be estimated. The subsections will be dedicated to simultaneity problems when supply variables, such as the number of international number of flights and the total number of boat arrivals in Sassari Province will be considered. Section 4.4 will be dedicated to the estimation of a model when employing quarterly data. A summary and conclusions will be given as final sections.

4.2 INTERNATIONAL DEMAND FOR TOURISM USING MONTHLY DATA

4.2.1 A Time Series Analysis

One needs to take into account seasonality and non-stationarity, and test for cointegration that will lead to the framework of the Error Correction Models (ECM). A "pre-modelling" analysis is of particular value in order to assess the properties of the variables under study without which the quality of the empirical results may be questionable (see Song *et al.* 2000; Lim and McAleer, 2000).

Seasonality and Dickey-Fuller unit root tests

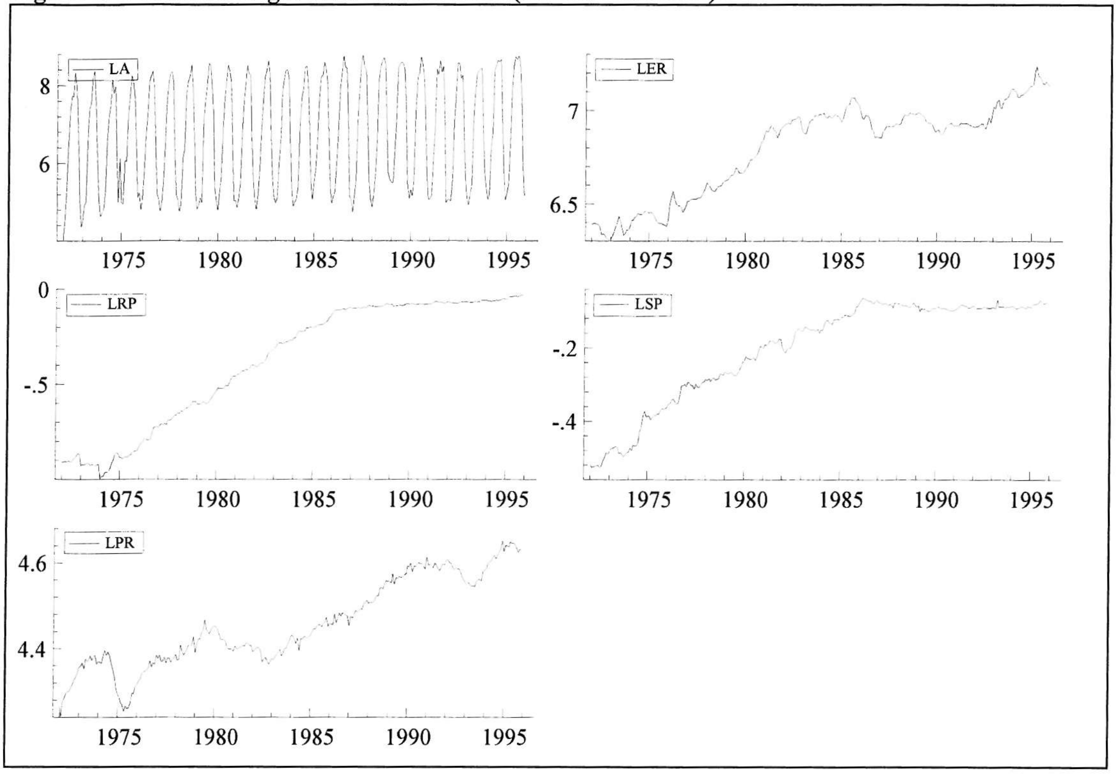
In this study the method described in Franses (1991a, 1991b) to test for seasonal unit roots with monthly data is used. This is applied to five series: *i.e.* the adjusted foreign arrivals of tourists[4], *LA*, for the period January 1972 up to December 1995; the exchange rate, *LER*, (1972:1 to 1995:12); a relative price index, *LRP*, is taken into consideration for the period from January 1972 to December 1995 and a substitution price index, *LSP*; and finally, the index of industrial production (1990=100) from 1972:1 to 1995:12, *LPR*. See Appendix B for a detailed discussion of the determination of these variables. In each case the natural logarithm of the variables is used. Graphs of each series are provided in Figure 4.1.

4 The tourist arrivals data are collected by the *Ministero del Turismo e dello Spettacolo* and *ENIT* through the *Enti Provinciali per il Turismo (EPT) e le Stazioni di cura, di soggiorno e di turismo.* Data are collected all over the national territory for all the accommodation infrastructures (hotels, *pensioni, locande,* youth hotels, camp sites, tourist residences, houses for holidays, *rifugi, pensionati, colonie,* religious institutions, private residences, villas, flats or rooms rented for holidays).

The data collection is done on the basis of the daily declarations from the providers of accommodation, of the clients' arrivals and of total nights spent in the particular accommodation.

Note that this variable has been created dividing the monthly number of tourists' arrivals by the number of Saturdays in a month (as stated in Chapter 3).

Figure 4. 1 Natural Logarithm of the Series (1972:1 - 1995:12)



One now gives an account of the main results obtained by fitting the equation (2.6.2) by OLS, for each of the five time series mentioned above [5]. Note that μ_t, which represents the deterministic part, in this particular case, consists of a constant, a trend and 11 seasonal dummies. The results are reported in Table 4.1.

5 The auxiliary regression (2.6.2) is run using Microfit 4.0 package.

Table 4. 1 Testing for Seasonal Unit Roots

t-statistics	Variable				
	LA	***LRP***	***LER***	***LRP***	***LSP***
π1	-3.502 ***	-3.081	-1.659	0.071	-1.152
π2	-4.715 ***	-4.627 ***	-4.914 ***	-4.262 ***	-4.509 ***
π3	1.966	-1.467	-6.812 ***	-6.056 ***	-5.623 ***
π4	-6.907 ***	-6.133 ***	-3.357 *	-3.537 **	-3.843 ***
π5	-6.679 ***	-6.944 ***	-7.508 ***	-7.550 ***	-6.581 ***
π6	-7.387 ***	-6.790 ***	-6.332 ***	-7.841 ***	-6.584 ***
π7	2.516	-2.757 ***	-1.776 ***	-2.977 ***	-2.544 ***
π8	-4.627 ***	-1.197	-1.221	-0.356	-0.671
π9	-2.117	-4.428 ***	-6.675 ***	-5.479 ***	-5.705 ***
π10	-6.017 ***	-7.998 ***	-2.716	-5.969 ***	-6.266 ***
π11	1.606	-3.364 ***	-4.932 ***	-3.498 ***	-4.016 ***
π12	-4.919 ***	-3.215 *	-1.650	-3.114	-2.712
F-statistics	***LA***	***LPR***	***LER***	***LRP***	***LSP***
π3, π4	26.594 ***	20.207 ***	27.867 ***	26.654 ***	19.847 ***
π5, π6	27.506 ***	25.696 ***	26.822 ***	32.300 ***	25.126 ***
π7, π8	16.902 ***	33.898 ***	24.417 ***	22.591 ***	22.531 ***
π9, π10	18.848 ***	32.151 ***	26.385 ***	22.831 ***	21.111 ***
π11, π12	12.723 ***	24.846 ***	30.397 ***	25.052 ***	36.727 ***
π3, ...,π12	24.184 ***	208.198 ***	94.600 ***	186.018 ***	150.043 ***

Notes: The three, two and one asterisks indicate that the seasonal unit root null hypothesis is rejected at the 1%, 5% and 10% level, respectively.

In the case of testing for the presence of seasonal unit roots with respect to the (log) foreign arrivals, the null hypothesis cannot be accepted at the 1% level of significance[6], both running the *t*-tests of the separate π's (except for π_3, π_7, π_9, π_{11} where the null hypothesis cannot be rejected) and the *F*-test of the pairs of π's, as well as the joint *F*-test of $\pi_3=...=\pi_{12}=0$. The main conclusion is that the arrivals of foreign tourists, *LA*, can be considered as having a deterministic seasonal pattern and, furthermore, the null hypothesis of a long run unit root is not accepted (*i.e.* H_0: $\pi_1 = 0$) so this variable can be modelled as a stationary process, *i.e.* I(0). The latter result has also been confirmed running the ADF test where the null of the presence of a unit root cannot be accepted at a 1% level of significance (Table 4.2)[7].

Running the auxiliary regression (2.6.2) for the log index of industrial production, *LPR*, the *t*-tests of the separate $\pi_2,..,\pi_{12} = 0$ cannot be accepted, in general, at a 1% level, nor can the *F*-test of the pairs of π's and the joint *F*-test of $\pi_3= ...=\pi_{12}=0$, with the exception for π_3 and π_8. Therefore, there appears to be no evidence

6 The critical values for the seasonal unit roots test are provided in Franses (1991a, pp.161-165).

7 The augmented Dickey-Fuller unit root test has been run using the PcGive 9.0 package (Doornik and Hendry, 1996, pp.93-95).

for the presence of seasonal unit roots. On the other hand, the null hypothesis $\pi_1 = 0$ cannot be accepted at a 20% level of significance, but can at the 10% level. The latter result is investigated further by an ADF test. As shown in Table 4.2, one can see that the level of this series is stationary as the null hypothesis of the presence of a unit root cannot be accepted at a 5% level where a constant and a trend have been included. Thus, it can be concluded that *LPR* is stationary in the level about a trend. However, the income proxy (*LPR*) can be considered non-stationary in the level, but stationary in the first difference when including just the constant (and seasonal dummies). This result suggests that deviations from linear trend might have a role in explaining the international demand for tourism.

The main result of running equation (2.6.2) for the relative price, (*LRP*), is that the null of the presence of seasonal unit roots cannot be accepted, taking into account the outcomes for the *t*-tests (with the exception for π_8 and π_{12}) the *F*-tests for pairs of π's, and for the joint *F*-test, at a 1% level of significance. Note also that the null hypothesis $\pi_1 = 0$ cannot be rejected at a 5% level. From the ADF unit root test (Table 4.2) one concludes that *LRP* is integrated of order one, when a constant, constant and trend, and constant, trend and seasonal dummies are included.

The seasonal unit roots test has been run for the exchange rate (*LER*) for the same period. Running the auxiliary regression (2.6.2), one can conclude that there appears to be no evidence for the presence of seasonal unit roots, denoting a regular seasonal pattern. The null hypothesis has to be accepted for π_8, π_{10} and π_{12}, as well for the long run frequency. As in the previous case, the null hypothesis of non-stationarity cannot be rejected at a 5% level of significance, testing π_1=0 using the *t*-test. This result is confirmed also by the ADF test (see Table 4.2). Thus, *LER* has to be considered an I(1) process.

The last variable under consideration is the substitute price (*LSP*). The main result from the OLS regression is that the presence of seasonal unit roots cannot be accepted at a general 1% level, both performing the *t*-tests of the separate π_s (except for π_8 and π_{12}) and the *F*-test of the pairs of π's, as well as the joint *F*-test of π_3=...= π_{12}=0. Whereas, the H_0: π_1=0 cannot be rejected at a 5% level; this result seems to indicate that this series is non-stationary. To test further the latter finding, an ADF test is run. From this test the null of the presence of a unit root cannot be accepted at a 5%

level (Table 4.2), but can, marginally, at the 1% level of significance, when including either a constant or a constant and seasonals. Thus, one might treat *LSP* as a stationary process, *i.e.* I(0). At this point a brief observation is due with respect to the substitute price. Experiments have shown as the substitute price can be considered stationary in the level when a constant is included, whereas such a variable seems to be integrated of order one when a constant and a time trend are included (Table 4.2). The choice of including just a constant in performing the ADF test is supported by the following assumptions. Firstly, the inclusion of a trend implies the presence of unit root plus a quadratic trend. Secondly, as can be seen in Figure 4.1, the data show an adjustment to a stable situation, given the zones for exchange rate stability in the European Union (EU). As the competitors included are EU (France, Greece, Spain and Portugal), it is difficult to accept long run non-stationarity in this variable.

Table 4. 2 Augmented Dickey-Fuller Unit Root Test

Series	*ADF*(1)	*LAG*(2)
LA(c)	- 3.87 **	9
LA(c,t)	- 4.36 **	10
LA(c,s)	- 4.06 **	2
LA(c,s,t)	- 6.33 **	2
LPR(c)	- 0.46	3
DLPR(c)	- 11.77 **	2
LPR(c,t)	- 3.84 *	8
LPR(c,s)	- 0.45	3
DLPR(c,s)	- 7.73 **	2
LPR(c,t,s)	- 3.77 *	8
LRP(c)	- 2.16	12
DLRP(c)	- 3.11 *	11
LRP(c,t)	- 0.64	12
DLRP(c,t)	- 3.78 *	11
LRP(c,s)	- 2.89 *	0
LRP(c,t,s)	- 0.47	12
DLRP(c,t,s)	- 3.76 *	11
LER(c)	- 1.68	1
DLER(c)	- 10.70 **	1
LER(c,t)	- 2.20	1
DLER(c,t)	- 10.72 **	1
LER(c,s)	- 1.61	1
DLER(c,s)	- 10.51 **	1
LER(c,t,s)	- 2.11	1
DLER(c,t,s)	- 10.54 **	1
LSP(c)	- 3.03 *	0
LSP(c,t)	- 1.24	0
DLSP(c,t)	- 15.36 **	0
LSP(c,s)	- 3.16 *	0
LSP(c,t,s)	- 1.18	0
DLSP(c,t,s)	- 15.13 **	0

Notes: The one and two asterisks indicate that the unit root null hypothesis is rejected at the 5% and 1% level, respectively. The capital letter *D* denotes the first-difference operator defined, in a general notation, by $Dx_t = x_t - x_{t-1}$.
(1) Augmented Dickey-Fuller statistics with constant (*c*) critical values = -2.872 at 5% and -3.455 at 1% level; with constant and trend (*c, t*) c.v.= -3.428 at 5% and -3.995 at 1% level; with constant and seasonals (*i.e. c, s*) c.v. = -2.872 at 5% and -3.456 at 1% level; with constant, trend and seasonals (*i.e. c, t, s*) c.v. = -3.428 at 5% and -3.995 at 1% level;
(2) Number of lags set to the first statistically significant lag, testing downward and upon white residuals. Note that ADF(0) corresponds to the DF test.

The main findings are that *LA*, *LPR* and *LSP* have to be considered as stationary, *i.e.* I(0), whereas *LRP* and *LER* are integrated of order d=1, denoted $X_t \sim I(1)$; therefore, the latter series have to be differenced one time to become stationary. Moreover, all the series present a regular and non-changing seasonal pattern.

4.2.2. Cointegration Analysis

In this study one makes use of the notion of cointegration. As already stated, the relative price (*LRP*) and exchange rate (*LER*) have been found to be stationary in the first difference. Hence, given that the components of the vector $X_t=(LRP,LER)'$ are both I(1), then the equilibrium error, if it exists, would be I(0) (Engle and Granger, 1987). In investigating the cointegration relation, two approaches can be used: the single equation cointegration approach (see Section 2.7.1), whose results for *LRP* and *LER* are reported in Appendix C; the Johansen VAR maximum likelihood estimator is also used.

One starts analysing the cointegration relation with a 2-dimensional VAR system for the series *LRP* and *LER*. A bivariate vector autoregression of order $k=13$ can be specified as in as follows:

$$X_t = \Phi D_t + \Pi_1 X_{t-1} + \ldots + \Pi_{13} X_{t-13} + \varepsilon_t \qquad (t = 1,\ldots,T)$$

In this case, the vector D_t contains a constant term and 11 seasonal dummies, both included unrestrictedly. A preliminary inspection of the residuals suggests the need for a 0-1 dummy *i.e.* *i1974p1* which it is possibly picking up the first oil shock. A VAR (13), as above specified, has been re-estimated with the impulse dummy *i1974p1* included unrestrictedly. However, problems in terms of diagnostic tests still persist such as non-normality, though largely reduced, conditional heteroscedasticity, non-homoscedasticity and serial correlation. From the joint *F*-test it is possible to reduce the system to a VAR (3). Whereas, the SC and HQ criteria suggest to run a system with 2 lags (Table 4.3); however, similar results have been obtained.

Table 4. 3 System Reduction

system	T	p		log-likelihood	SC	HQ	AIC
1	275	30	COINT	2512.4789	-17.660	-17.896	-18.273
2	275	34	COINT	2534.9070	***-17.741***	***-18.009***	-18.436
3	275	38	COINT	2540.6785	-17.702	-18.001	-18.478
4	275	42	COINT	2545.7655	-17.657	-17.988	-18.515
5	275	46	COINT	2550.6216	-17.610	-17.973	-18.550
6	275	50	COINT	2552.9521	-17.546	-17.939	-18.567
7	275	54	COINT	2556.9897	-17.493	-17.919	-18.596
8	275	58	COINT	2560.6230	-17.438	-17.895	-18.623
9	275	62	COINT	2563.3856	-17.376	-17.865	-18.643
10	275	66	COINT	2565.5167	-17.310	-17.830	-18.658
11	275	70	COINT	2566.2845	-17.234	-17.785	-18.664
12	275	74	COINT	2566.7857	-17.156	-17.739	-18.668
13	275	78	COINT	2568.0470	-17.084	-17.698	***-18.677***

System 13 --> System 12: F(4, 470) = 0.54016 [0.7063]
System 12 --> System 11: F(4, 474) = 0.21616 [0.9294]
....
System 4 --> System 3: F(4, 506) = 2.3618 [0.0523]
System 3 --> System 2: F(4, 510) = 2.7041 [0.0298] *

This lag gives a satisfactory portmanteau test statistic for serial correlation in PcFiml; however, non-normality as well as heteroscedasticity problems have not been eliminated.

Note also that in the case under study each equation is fitted with $kp+m=19$ parameters leaving 248[8] degrees of freedom for the variance. To test the cointegration hypothesis one makes use of the procedure reported in Section 2.7.2 (Chapter 2). The results of the eigenvalue and eigenvector calculations are given in Table 4.4[9].

Table 4. 4 The Eigenvalues $\hat{\lambda}$, Eigenvectors $\hat{\beta}$, and the Weights α

Eigenvalues $\hat{\lambda}$
(0.0737 0.0063)

Standardized β' eigenvectors		Standardized $\hat{a}$ coefficients		
LRP	***LER***	***LRP***	-0.02	0.019
1.00	-1.09	***LER***	0.01	-0.017
-0.43	1.00			

Table 4.5 reports the results of the tests for reduced rank. The test statistics are the maximal eigenvalue (λ_{max}) and the trace statistics (λ_{trace}), as previously described.

---

8 This is calculated with the following formula: $T - (pk + m)$ where T is the sample size, p is the number of variables, k is the number of lags, m consists of the constant, the dummies and the trend, when included.

9 All the results concerning with the cointegration testing are obtained using the PcGive and PcFiml modules of PcGive 9.0 (see Doornik and Hendry, 1996).

Table 4. 5 Johansen Tests for the Number of Cointegrating Vectors

Ho	H_1	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
r=0	r=1	21.82**	21.36**	14.1	23.62**	23.13**	15.4
r=1	r=2	1.81	1.77	3.8	1.81	1.77	3.8

Notes:
(1) Adjusted by the degrees of freedom (see, Reimers, 1992)
(2) Critical values at a 5% level of confidence (see Osterward-Lenum, 1992).
* and ** denotes rejection of the null (*i.e.* non-cointegration) at a 5% and 1% level, respectively.

From Table 4.5, therefore, one can reject the hypothesis that $r=0$, at least at the 1% level, concluding that there is one cointegrating relationship. Such a result, acceptance of one cointegrating vector, is against the results obtained using the Cointegrating Regression Durbin Watson approach as given in Appendix C.

The coefficient estimates of the cointegrating relation are found in Table 4.4 as the first row of β' matrix, and the equivalent error correction mechanism is the following:

$$CI = LRP - 1.0893LER \quad (4.2.2.1)$$

The coefficient for the (log) weighted exchange rate has been tested for the following restriction: $\beta=-1$ and such a restriction has been accepted at the 1% level from the χ^2 test[10]. Hence, in the long run, there are no main price differentials amongst the origin countries under analysis. In this way one can model the following cointegrating vector:

$$CI = LRP - 1 * LER$$

"It is natural to give the coefficients of α an economic meaning in terms of the average speed of adjustment towards the estimated equilibrium state, such that a low coefficient indicates slow adjustment while a high coefficient indicates rapid adjustment" (Johansen and Juselius, 1990, p.183). Note that in this case the average speed of adjustment is approximately 0.01 in modulus (see Table 4.4). As Johansen (1995 p.41) points out, when the cumulated disturbances push the economic variables around in the attractor space, the agents (that are assumed to be rational and identical) tend to react to these forces and create economic variables that react to the disequilibrium errors through the adjustment coefficients α and are forced back towards the attractor set. Alternatively, in matrix notation one can express the adjustment towards the estimated equilibrium state as follows:

10 The results for the restriction test on the coefficient is: $\chi^2(1) = 0.64365$ [0.4224]

$$\Delta X_t = \Pi X_{t-1} + u_t \text{ or } \qquad (4.2.2.2)$$
$$= a\,\beta' X_{t-1} + u_t;$$

one multiplies each side by β':

$$\beta' X_t = \beta' X_{t-1} + (\beta' a)\,\beta' X_{t-1} + \beta' u_t$$

let $\beta' X_t = Z_t$:

$$Z_t = (I + \beta' a)\, Z_{t-1} + \beta' u_t$$

The process Z_t is stationary if the matrix $(I + \beta' \alpha)$ has its eigenvalues inside the unit circle. When the roots are near to the value of zero a fast adjustment will occur, whereas roots close to one imply a slow adjustment.

From Tables 4.4 and 4.5, one has $I = 1$ and:

$$(1 + \beta' \alpha) = 1 + \begin{bmatrix} 1 & -1.0893 \end{bmatrix} \begin{bmatrix} -0.02 \\ 0.01 \end{bmatrix}$$

the root equals 0.97 thus a rather slow adjustment will occur. As 0.97^{12}=0.686, only the 31% of the adjustment to equilibrium occurs in the first year and 53% in the first two years. This calculation is only an approximation, as the lags on ΔX_t (in equation 4.2.2.2) have not been included. See page 64 for a further discussion on this issue.

The main finding is the existence of a cointegrating relationship between the relative price (*LRP*) and the (weighted) exchange rate (*LER*). This result validates, statistically, that there is evidence to separate prices and exchange rate in the short run. Whereas, a real effective exchange rate should be used in the long run in accordance to the Purchasing Power Parity (PPP).

4.2.3 Model Specification Using Monthly Data

In this section the relationship amongst the I(0) variables is estimated, *i.e.* the adjusted arrivals of foreign tourists (*LA*), the substitute price (*LSP*), the industrial production (*LPR*), the first difference of the exchange rate (*DLER*) and of the relative price (*DLRP*) and, finally, the cointegrating vector (*CI*) defined by the following relationship after having imposed the restriction on the coefficient for *LER*:

$$CI = LRP - 1 * LER$$

The main purpose is to estimate whether these variables are able to explain the foreign demand for tourism in the north of Sardinia, with respect to the period from 1972:1 to 1995:12. One starts with an unrestricted system[11] that includes 13 lags[12], the one lagged error correction mechanism (CI_{t-1})[13], a constant, a trend that might pick up the deviations of the (log) industrial production from the trend (as in Section 4.2.2., *LPR* is found to be stationary about a trend), 11 seasonal dummies, an "Easter" dummy and four impulse dummies (*i1974p12*, *i1979p3*, *i1985p3* and *i1991p11*). See Appendix D, Table D.1[14].

The Easter dummy was introduced into the model in order to capture the Easter holiday effect. This effect, in fact, "cannot be captured by the seasonal components due to its mobility so it has to be modeled separately" (Gonzàles and Moral, 1996, p. 748). As far as the period under modelling is concerned, Easter is between the 26th March and the 22th April. Thus, the dummy variable "Easter" has been constructed giving the value one in the Easter month and zero otherwise. Note also that the Saturday before Easter has been considered as the first day of the holiday, in the case when the Easter period is split into March and April. For example, in 1972 Easter Sunday was the second of April, therefore the value of one is given to the April month instead of the March month. This worked better empirically than giving a value 0.5 in each month as experimented by Gonzàles and Moral, 1996.

The four impulse dummies are constructed in order to avoid non-normality problems in the residuals. However, such dummies are difficult to interpret. Possible factors for outliers could be related to particular events, such as strikes for boats or

---

11 These results are obtained using the PcGive and PcFiml modulus of PcGive 9.0 (see Doornik and Hendry, 1996).

12 A system with 13 lags and 12 lags, respectively, was initially estimated, but according to the test of system reduction, as provided in PcGive 9.0, the restricted system cannot be accepted at the 1% level. Note that the HQ criterion leads to the same conclusion.

system	T	p		log-likelihood	SC	HQ	AIC
12	274	395	OLS	5906.7627	-35.023	-38.141	-41.115
13	274	420	OLS	5965.6730	-34.941	-38.256	-40.545

System 13 lags > System 12 lags: $F(25, 692) = 3.3991$ [0.0000] **

13 One could put in the first lag of the cointegrating vector and the free lags of *DLRP* and *DLER*, as in this case; either free lags of the cointegrating vector and *DLRP*, or free lags of the cointegrating vector and *DLER*.

14 It is worth noting that in a previous stage a "weather" variable (*LW*) with the average temperatures for the Province of Sassari was included into the system. However, such a variable does not seem to have any particular effect in determining the foreign demand for tourism; hence, it has been excluded at an early stage.

planes, or particular discounts for holidays package in Sardinia. Particular sport events could also be thought to have positively effected the demand for tourism such as rallies, cycle races and so on.

After a general-to-specific simplification, as "an efficient way to find a congruent encompassing model" (Mizon, 1996, p.123), one has obtained a parsimonious model as reported in Table D.2. The unrestricted model could be reduced by using both the joint *F*-test and the SC criterion.

Considering Table D.2., the estimates of the parameter coefficients of the short run variables are significant, in general, at the 5% level. The R^2 explains 98% of the variance of the dependent variable. Moreover, as the relevant *F*-statistic indicates, the overall significance of the regression is satisfactory. Looking at the diagnostic tests the model specification has to be accepted, as well as the conditions of no serial correlation, conditional homoscedasticity, normality and homoscedasticity.

The impulse dummies as well as the "Easter" dummy are statistically significant as in the unrestricted model case. As far as the seasonal component is concerned, it makes evident the concentration of tourist arrivals in the period between May and September as suggested by the data analysis given in Section 3.2 (Chapter 3). Note that December is the variable omitted.

As one can notice the first and the second lag of the substitute price present coefficients almost of the same size but opposite sign (*i.e.* $\gamma_1 \approx \gamma_2$). Hence, an *F*-test is run to test if the null hypothesis H_0: $\gamma_1+\gamma_2=0$ holds. The appropriate *F* statistic is 1.90 with q=1 degrees of freedom in the numerator and N-K=246 in the denominator. This value is smaller than the critical value of the *F* distribution at a 5% level (*i.e.* smaller than 3.84), thus failing to reject the null hypothesis one concludes that the restriction holds. The same conclusion is reached from the SC criterion (*i.e.* -2.79541) greater in absolute value than the one for the unrestricted model (*i.e.* -2.78261). Note that LSP_1-LSP_2 is called *RLSP*.

A restriction has also been tested on the coefficient for LSP_{11} and LSP_{12}, however, it has not be accepted at the 5% level. The same result has been achieved when testing for the restriction on the third and seventh lag of the industrial production: the null hypothesis has to be rejected at the 5% level.

As a further experiment a restriction on the seasonal dummies has been imposed. In particular, the seasonal *May*, *Jun* and *Jul*[15] present coefficients almost of the same size and same sign (*i.e.* $\mu_5 \approx \mu_6 \approx \mu_7$, Model 1) confirming the seasonal pattern shown in Figure 3.2. A new dummy (say *MJJ*, Model 2) was created giving the value 1 to May, June and July months and 0 otherwise. An F-test is run in order to test Model 1 versus Model 2 (where the RSS equals 9.6393). The appropriate F statistic is 0.155 with q=2 degrees of freedom in the numerator and N-K=247 in the denominator. This value is smaller than the critical value of the F distribution at a 5% level (*i.e.* 3.00), thus failing to reject the null hypothesis the restriction holds. Note that the SC criterion suggests accepting the restriction as in Model 1 it is equal to -2.79541 and in Model 2 it equals -2.83513.

Given the previous analysis, the final parsimonious encompassing model is reported in Table 4.6.

[15] Note that *May* is created giving the value 1 in May and 0 otherwise; *Jun* takes the value 1 in June and 0 otherwise; finally, *Jul* takes the value 1 in July and 0 otherwise.

Table 4. 6 Results from the Restricted Parsimonious Model for the Foreign Demand of Tourism

EQ(3) Modelling LA by OLS (using For.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-2.7541	1.6850	-1.634	0.1034	0.0106
LA_1	0.12464	0.047236	2.639	0.0089	0.0272
LA_2	0.10335	0.040793	2.534	0.0119	0.0251
LA_3	0.13435	0.043241	3.107	0.0021	0.0373
LA_11	0.10844	0.049556	2.188	0.0296	0.0189
LPR_3	2.5695	0.62997	4.079	0.0001	0.0626
LPR_7	-2.0073	0.62160	-3.229	0.0014	0.0402
RLSP	-4.7452	1.8776	-2.527	0.0121	0.0250
LSP_11	5.2114	1.8885	2.760	0.0062	0.0297
LSP_12	-4.5311	1.8806	-2.409	0.0167	0.0228
CI_1	-0.33807	0.15560	-2.173	0.0308	0.0186
easter	0.42625	0.071951	5.924	0.0000	0.1235
i1974p12	1.5557	0.20568	7.564	0.0000	0.1868
i1979p3	-0.57510	0.20496	-2.806	0.0054	0.0306
i1985p3	0.67531	0.20578	3.282	0.0012	0.0415
i1991p11	-0.59522	0.20399	-2.918	0.0038	0.0331
JA	0.25746	0.079375	3.244	0.0013	0.0405
FE	0.63111	0.11875	5.315	0.0000	0.1019
MAR	1.0567	0.15893	6.649	0.0000	0.1508
AP	1.7281	0.21368	8.087	0.0000	0.2080
MJJ	2.7638	0.23245	11.890	0.0000	0.3622
AU	2.5070	0.21165	11.846	0.0000	0.3604
SE	2.3545	0.16297	14.448	0.0000	0.4560
OT	1.1904	0.12748	9.338	0.0000	0.2593
NO	0.017192	0.084658	0.203	0.8392	0.0002

R^2 = 0.981794 F(24,249) = 559.5 [0.0000] sigma = 0.196754
DW = 1.91 RSS = 9.639300629 for 25 variables and 274 observations

```
AR 1- 7 F( 7,242) =      1.0461 [0.3995]
ARCH 7  F( 7,235) =     0.42111 [0.8886]
Normality Chi^2(2)=      3.8046 [0.1492]
Xi^2    F(34,214) =       1.646 [0.0187] *
RESET   F( 1,248) =      1.5331 [0.2168]
Tests of parameter constancy over: 1995 (5) to 1995 (12)
Forecast Chi^2( 8)=      10.082 [0.2593]
Chow    F( 8,241) =      1.1653 [0.3209]
```

where:

LA = (log) normalised series of foreign arrivals for the number of weekends (*i.e.* Saturdays) in a month.

LPR = (log) weighted average industrial production index for the origin countries.

RLSP = difference between the coefficients of the first and second lag of the (log) substitute price.

LSP = (log) substitute price (consumer price index in Sassari by the weighted average consumer price index in other destinations in the Mediterranean area).

CI = cointegrating vector.

EASTER= dummy 0 -1 with respect to the Easter holiday.

MJJ = seasonal dummy giving the value of 1 to May, June and July months and 0 otherwise.

In terms of statistical tests, one can say the same results achieved for the model reported in Table D.2., hold for the final parsimonious model which, therefore, can be considered as data congruent. Further tests for parameter constancy over the last eight observations of the sample are reported. The null hypothesis of constancy fails to be rejected. Moreover, it has to be noted that the null hypothesis of White homoscedasticity for the residuals is marginally rejected at the 5% level. In this case the "ordinary least-squares parameter estimators are unbiased and consistent, but they are not efficient; *i.e.* the variances of the estimated parameters are not the minimum variances" (Pindyck and Rubinfeld, 1991, p.128). A White correction for heteroscedasticity has been used for the standard errors[16] as reported in Table 4.7.

16 Such a correction has been run using Microfit 4.0.

Table 4. 7 Results after Correcting for Heteroscedasticity

Ordinary Least Squares Estimation
Based on White's Heteroscedasticity adjusted S.E.'s

Dependent variable is LA
274 observations used for estimation from 1973M3 to 1995M12

Regressor	Coefficient	Standard Error	T-Ratio[Prob]
CONSTANT	-2.7543	1.9648	-1.4019[.162]
LA(-1)	.12465	.051951	2.3994[.017]
LA(-2)	.10333	.036099	2.8624[.005]
LA(-3)	.13435	.037856	3.5491[.000]
LA(-11)	.10844	.049199	2.2041[.028]
LPR(-3)	2.5695	.65517	3.9218[.000]
LPR(-7)	-2.0073	.62053	-3.2348[.001]
RLSP	-4.7461	1.7801	-2.6662[.008]
LSP(-11)	5.2137	2.2075	2.3618[.019]
LSP(-12)	-4.5334	2.1908	-2.0693[.040]
CI(-1)	-.33808	.16315	-2.0722[.039]
EASTER	.42625	.070251	6.0676[.000]
i1974P12	1.5556	.060414	25.7492[.000]
i979P3	-.57510	.071683	-8.0228[.000]
i985P3	.67534	.070258	9.6123[.000]
i991P11	-.59523	.045902	-12.9675[.000]
JAN	.25741	.079456	3.2397[.001]
FEB	.63108	.11713	5.3877[.000]
MAR	1.0567	.16644	6.3485[.000]
APR	1.7280	.21773	7.9366[.000]
MJJ	2.7638	.24273	11.3865[.000]
AUG	2.5070	.22332	11.2260[.000]
SEP	2.3545	.17169	13.7138[.000]
OCT	1.1904	.14359	8.2899[.000]
NOV	.017208	.080258	.21441[.830]

R-Squared	0.9818	R-Bar-Squared	0.98004
S.E. of Regression	0.19675	F-stat. F(24, 249)	559.5285[.000]
Mean of Dependent Variable	6.7907	S.D. of Dependent Variable	1.3926
Residual Sum of Squares	9.6389	Equation Log-likelihood	69.7942
Akaike Info. Criterion	44.7942	Schwarz Bayesian Criterion	-0.36989
DW-statistic	1.9079		

The R-squared adjusted, and the ratio between the standard error of the regression (*SER*) and the mean of the dependent variable (*MDV*), that is equal to 0.03, indicate that the variables included are significant determinant of the international demand for tourism. In terms of signs of the coefficients they are as expected. The lags coefficients of the foreign arrivals, as explanatory variables, present a positive sign. This indicates that foreign tourists are possibly “psychocentric” and that the Province of Sassari is viewed as a desirable destination area (see definition in Section 3.3.1, Chapter 3). This is also consistent with the adjustment of the dependent variable to changes in the right hand side variables. The foreign demand for tourism shows a rather strong dependence, on the index of industrial production (used as an indicator of the main clients' income). The latter presents a positive short run coefficient, and on average a positive sign in the long run (*t*-value 2.62 in Table 4.8), confirming that the higher the income of the nations the higher the demand for leisure. An increase in the substitution price will have a positive impact on the demand for tourism in the short run. This seems to be in contradiction with the expectation that when prices in the north of Sardinia become higher, *ceteris paribus*, the demand for tourism decreases.

The coefficient of the cointegrating vector (*CI*) presents a negative sign. This denotes that if the *CI* increases, by a deviation either in *LRP* (*i.e.* the relative price) or in *LER* (weighted exchange rate) from the respective long run relations, the foreign demand for tourism decreases in the short run.

The long run dynamics and the long run standard errors are reported in Table 4.8. One notices the long run multipliers and the standard errors are, in general, well-specified. Moreover, they are statistically significant and present the expected signs, with the only exception for the substitute price (*i.e. LSP*). However, it may appear that imposing the restriction on *RLSP* overstates the precision of the *LSP* coefficient, but on investigation (*i.e.* including all the four coefficients for *LSP*) the effect was found to be very marginal. Hence, another possibility is the presence of measurement errors. This issue will be discussed further in Chapter 6.

Table 4. 8 Solved Static Long Run Equation

```
Solved Static Long Run equation
          LA =     -5.204              -0.6388 CI             +0.8054 easter
(SE)          (     3.283)           (   0.3027)            (   0.1627)
                  +1.062 LPR           -8.967 RLSP            +1.286 LSP
              (   0.4049)            (    3.638)            (   0.3014)
                   +5.223 MJJ
              (   0.4346)
                   +2.94 i1974p12      -1.087 i1979p3         +1.276 i1985p3
              (   0.4894)            (   0.3961)            (   0.4052)
                   -1.125 i1991p11     +0.4865 JA              +1.193 FE
              (   0.3993)            (   0.1585)            (   0.2453)
                   +1.997 MAR           +3.265 AP              +4.737 AU
              (   0.2989)            (   0.3852)            (   0.3275)
                   +4.449 SE            +2.249 OT             +0.03249 NO
              (   0.3497)            (   0.2634)            (   0.1591)


ECM = LA + 5.2041 + 0.638808*CI - 0.805439*easter - 2.93956*i1974p12
     + 1.08671*i1979p3 - 1.27607*i1985p3 + 1.12472*i1991p11 - 0.486492*JA
     - 1.19253*FE - 1.99677*MAR - 3.26534*AP - 4.73731*AU - 4.44912*SE
     - 2.24937*OT - 0.0324867*NO - 1.06219*LPR + 8.96657*RLSP - 1.28553*LSP
     - 5.22252*MGL;

WALD test Chi^2(18) = 635.14 [0.0000] **
```

Thus, with few exceptions, one can conclude that the model gives satisfactory results in the short run as well in the long run, in terms both of statistical tests and economic theory.

4.2.4 Linear Versus Logarithmic Specification

In this paragraph, an account is given of the experiments carried out in assessing whether the logarithmic specification form is more appropriate than the linear form, in estimating the demand for international tourism.

The integration status of the variables expressed in a linear specification will be tested using: *A* (number of arrivals modified for number of weekends), *PR* (income proxy, as a weighted average for the source countries), *RP* (relative price, Sassari/origin countries), *SP* (substitute price, Sassari/other destinations) and *ER* (weighted average exchange rate for the origin countries). At this point, it is worth noting that one cannot directly compare the results from the ADF for the logarithmic form with those obtained using the linear specification for the variables of interest. As Granger and Hallman (1991) point out, a unit root test invariant to the transformation has to be used. Such tests are the Rank Dickey-Fuller (RDF) and Augmented RDF (ARDF). It is not in the scope of this thesis to investigate the possibility that the logarithmic transformation used in the present study could lead to an over-rejection of the null hypothesis. However, this can be thought as further work. From the ADF test with 13 lags the following results have been obtained: *A*, *PR* and *SP* can be considered as I(0). The income proxy is stationary in the level when the time trend is included in the ADF test, whereas the substitute price is stationary in the level when the constant, and a constant and seasonals are included. On the other hand, *RP* and *ER* are I(1). As for the logarithmic case, the cointegration status for the coefficients of the latter variables has been investigated. The Johansen analysis, has shown the existence of a cointegration relationship between the coefficients of the relative price and exchange rate. In particular, an initial 13 lag system, used to run the Johansen analysis, which also includes a constant and seasonal dummies could be reduced to a VAR(3). The cointegration relationship for the linear specification is defined as follows:

$$ECL = RP - 0.000913631\ ER$$

In order to run the Box and Cox test, one estimates an unrestricted 13 lag tourism demand equation expressed both in a logarithmic and linear form, where the independent variables are defined as before.

1) logarithmic form

$$LA_t = a_1 + a_2 LA_{t-1}... + a_3 LPR_t... + a_4 LSP_t.. + a_5 DLER_t.. + a_6 DLRP_t.. + a_7 CI_{t-1} + \\ + a_8 Easter + a_9 Seas + a_{10} Dummies + e_t$$

and

2) linear form

$$A_t = a_1 + a_2 A_{t-1}... + a_3 PR_t.. + a_4 SP_t.. + a_5 DER_t.. + a_6 DRP_t.. + a_7 ECL_{t-1} + \\ + a_8 Easter + a_9 Seas + a_{10} Dummies + e_t$$

The sum of the squared errors from the logarithmic form (SSE_{LL}) is equal to 1.90, whereas the sum of the squared errors for the linear form (SSE_L) equals 34301958.14. To test whether the two models are empirically equivalent and find out which of the two models fits better the data. In doing this, the sum of the squared errors needs to be calculated for the linear model with $(A/\overline{A}_G)$ as the dependent variable. Note that $\overline{A}_G$ is the geometric mean defined as follows:

$$\overline{A}_G = exp\left\{\frac{1}{T}\sum_{t=1}^{T} lnA_t\right\}$$

For the latter model, the sum of the squared errors (that is $SSE_L/(\overline{A}_G)^2$) equals 46.4971. The Box-Cox test (see formula 2.4.1.1, Chapter 2) indicates that the two models are empirically different as the calculated χ^2 is 438.06 and the correspondent tabulated value is 3.84 at the 5% significance level. Moreover, the $SSE_L/(\overline{A}_G)^2$ is higher than SSE_{LL}, hence one can conclude that the logarithmic specification form fits the data better than the linear model.

Note that it would be possible to compare different specification in linear and logarithmic forms, but the rejection of the linear model is so clear that this has not been pursued.

4.3 THE MODEL SPECIFICATION USING ANNUAL DATA ANALYSIS

From the analysis so far, several advantages have emerged in the use of monthly data. They give the possibility to identify the short run characteristics of the demand for tourism. One can study carefully the seasonal pattern that seems to be of extreme importance for operators in the tourism activity. Moreover, the relatively large number of observations available in the previous analysis (288 observations) has allowed one to test for the possible presence of seasonal unit roots as well as for long run unit roots.

At this stage it could be interesting to assess the characteristics of the international demand for tourism also in the long run by making use of annual and quarterly data. The annual and quarterly analysis should be broadly consistent with the monthly results. Further, the availability of extra variables, only obtainable on a quarterly and annual basis, may enrich the analysis or remove the need to use proxy variables. On the other hand, one of the main limitations when dealing with annual tourism data could be the relative small number of observations available.

One starts estimating an annual model for the period 1972-1995. In order to obtain homogeneous results and comparisons between models, the same time series as for the monthly case will be used. A "pre-modelling" analysis is carried out. Table 4.9 reports the results from running an ADF test for each of the economic series. However, as one can notice, such a table can be interpreted as illustrative of the problems of using ADF tests in small samples (*i.e.* $T=24$) rather than being informative as to the integration status of the variables under study.

Table 4. 9 Augmented Dickey-Fuller Unit Roots Test using Annual Data

Series	*ADF*(1)	*LAG*(2)
LA(c)	- 1.23	0
DLA(c)	- 3.35 *	0
LA(c,t)	- 3.35	1
DLA(c,t)	- 3.18	0
LPR(c)	- 0.16	0
DLPR(c)	- 6.30 **	0
LPR(c,t)	- 3.01	0
DLPR(c,t)	- 6.02 **	0
LRP(c)	- 4.32 **	1
LRP(c,t)	- 0.82	1
DLRP(c,t)	- 3.89 *	0
LER(c)	- 1.19	0
DLER(c)	- 3.43 *	0
LER(c,t)	- 2.25	0
DLER(c,t)	- 2.70	0
DDLER(c,t)	- 5.12	0
LSP(c)	- 5.79 **	0
LSP(c,t)	- 1.38	1
DLSP(c,t)	- 3.98 *	0

Notes: The one and two asterisks indicate that the unit root null hypothesis is rejected at the 5% and 1% level, respectively. The capital letter *D* denotes the first-difference operator defined, in a general notation, by $Dx_t = x_t - x_{t-1}$.
(1) Augmented Dickey-Fuller statistics with with constant and trend (*c, t*) c.v.= -3.735 at 5% and -4.671 at 1% level; with constant c.v. = -3.066 at 5% and -3.93 at 1% level.
(2) Number of lags set to the first statistically significant lag, testing downward (starting with a maximum of 2 lags) and upon white residuals.

A comparison with Table 4.2 suggests major differences in the results. These findings suggest possible mis-specification in determining whether a variable is stationary in the level when using annual data with a short sample size. Note, therefore, that one will consider the above variables as having the same integration status as suggested by the ADF test when using monthly data; the modified series of foreign arrivals[17] (*LA*) will be treated as I(0), as well as the income proxy (*LPR*) and the substitute price (*LSP*). Whereas, the relative price (*LRP*) and the exchange rate (*LER*) will be treated as I(1).

A Johansen cointegration analysis is undertaken in order to check a possible cointegration relationship between *LRP* and *LER*. An initial bivariate VAR with k=3 is run which includes the unrestricted constant. A further reduction to a VAR of order one is carried out, as suggested by the system reduction test and by the SC and HQ criteria[18]. From the diagnostic tests the null hypothesis of homoscedasticity fails to be

17 In the annual data case, the modification of the number of arrivals of foreign tourists has been done by dividing the annual figures by the average number of weekends (*i.e.* Saturdays) in a year.

18

accepted at the 5% level. This finding seems to confirm those obtained for the system using data with monthly frequency.

Table 4.10 reports the results of the tests for reduced rank. The test statistics, even when corrected by the degrees of freedom, suggest that the null hypothesis of the existence of one cointegrating vector cannot be rejected at the confidence level of 1%.

Table 4. 10 Johansen Tests for the Number of Cointegrating Vectors using Annual Data

Ho	H_1	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
r=0	r=1	32.44**	29.35**	14.1	33.51**	30.32**	15.4
r=1	r=2	1.06	0.96	3.8	1.06	0.96	3.8

Note: As in Table 4.5

The results of the eigenvalue and eigenvector calculations are given in Table 4.11.

Table 4. 11 The Eigenvalues $\hat{\lambda}$, Eigenvectors $\hat{\beta}$, and the Weights α

Eigenvalues $\hat{\lambda}$ (0.7867 0.0494)				
Standardized β' eigenvectors		Standardized $\hat{\alpha}$ coefficients		
LRP	***LER***			
1.00	-0.88	***LRP***	-0.22	0.0016
-0.54	1.00	***LER***	-0.06	-0.1485

Therefore, the equivalent error correction mechanism is the following:

$$CI = LRP - 0.88\ LER \qquad (4.3.1)$$

The coefficient for the (log) weighted exchange rate has been tested for the following restriction: β=-1 and such a restriction has been accepted at the 5% level from the χ^2 test[19]. In this way the following cointegrating vector can be modelled:

$$CI = LRP - 1\ LER$$

as has been done in the monthly case.

Again, the coefficients of α have an economic meaning in terms of the average speed of adjustment towards the estimated equilibrium state. Note that in this case the

system	T	p		log-likelihood	SC	HQ	AIC
1	21	6	OLS	149.58949	-13.377	-13.610	-14.247
2	21	10	OLS	153.96649	-13.214	-13.603	-14.663
3	21	14	OLS	154.86913	-12.720	-13.265	-13.749

System (2 lags) -> System (1 lag): F(4, 30) = 1.7381 [0.1677]
System (3 lags)-> System (2 lags): F(4, 26) = 0.2855 [0.8847]

Note that the AIC criterion suggests running a VAR k=2; however, the residuals show problems of serial correlations at the 1% level. Nevertheless, same results are obtained in terms of cointegration analysis.

19 The results for the restriction test on the coefficient is: $\chi^2(1)$ = 1.5489 [0.2133]

average speed of adjustment is approximately 0.14 in modulus (see Table 4.11). From Tables 4.10 and 4.11, the root is:

$$(1 + \beta' \alpha) = 1 + \begin{bmatrix} 1 & -0.88357 \end{bmatrix} \begin{bmatrix} -0.21641 \\ -0.06070 \end{bmatrix}$$

that equals 0.8372. After one year 16.28% of the adjustment has occurred. Further, as $0.8372^2 = 0.7009$, only 29.91% of the adjustment to equilibrium occurs in the first two years. This calculation is only an approximation, as the lags on ΔX_t are not included (see equation 4.2.2.2). If the lags are taken into account, the dynamics of the adjustment of the cointegrating vector *Z* cannot be completely isolated from the rest of the system. The dominant root becomes 0.9676, and the quartiles of the lag distribution for *Z* become 19, 31 and 32 months. One can conclude that this result seems to be in line with the findings of Dwyer *et al.* (2000). They argue that "One thing that is striking is that there are wide variations in destination price competitiveness. In short, tourism prices differ widely from country to country.... These observations are consistent with the more general observation that purchasing power parity does not hold across countries - even approximately. There are systematic differences in price levels, even between countries which trade intensively" (p.17). From the statistical analysis in this thesis, it appears that the adjustment to the equilibrium indeed occurs as expected from economic theory. However, this adjustment is relatively slow.

The annual data used in this study cover a period of 24 years (1972-1995). The initial model is estimated by regressing the logarithm of the modified series of arrivals (*LA*) on the logarithm of the following variables: the index of industrial production (*LPR*), the substitute price (*LSP*), the first lag of the cointegrating vector (CI_{t-1}), the first difference of the relative price index in Sassari (*DLRP*), the first difference of the exchange rate (*DLER*), the weather variable (*LW*), a time trend (*TREND*) is also included in order to take into consideration "possible changes in the popularity of the holiday over the period as a result of changing tastes" (Martin and Witt, 1988). Note that with this analysis the aim consists in replicating the monthly model, and, as far as possible, comparing the results with those obtained using monthly data.

A one lag structure is tested suggesting no problems in terms of diagnostic tests. The results from the final model, using annual data, obtained after a general-to-specific simplification, are reported in Table 4.12.

Table 4. 12 Final Model for the Foreign Demand of Tourism using Annual Data

```
EQ(1) Modelling LA by OLS  (using datiann.in7)
The present sample is:  3 to 24

Variable      Coefficient    Std.Error  t-value   t-prob PartR^2
Constant          -5.0875       4.1228   -1.234   0.2340  0.0822
LPR                2.3421      0.63582    3.684   0.0018  0.4439
LSP_1              1.5474      0.34846    4.441   0.0004  0.5370
CI_1             -0.72958      0.27059   -2.696   0.0153  0.2995
Trend           -0.021042     0.011740   -1.792   0.0909  0.1589

R^2 = 0.9079  F(4,17) = 41.895 [0.0000] sigma = 0.0745249  DW = 1.42
RSS = 0.09441744098 for 5 variables and 22 observations
AR 1- 1 F( 1, 16) =      1.0549 [0.3197]
ARCH 1  F( 1, 15) =     0.26527 [0.6140]
Normality Chi^2(2)=     0.19461 [0.9073]
Xi^2    F( 8,  8) =     0.25693 [0.9641]
Xi*Xj   F(14,  2) =     0.18641 [0.9814]
RESET   F( 1, 16) =      5.2965 [0.0351] *
Tests of parameter constancy over: 24 to 24
Forecast Chi^2( 1)=      11.297 [0.0008] **
Chow    F( 1, 16) =      5.3833 [0.0339] *
```

Such a model is overall statistically well-specified and constitutes an admissible reduction of the underlaying unrestricted model. However, it shows non-linearity problems at the 5% level using the *RESET* test, which might be detecting the absence of relevant explanatory variables. The forecasting ability of this model and its parameter constancy is also evaluated: both the statistics are statistically significant implying the coefficients are not constant over the sample period.

The inclusion of the lagged dependent variable turns out to be statistically not significant, suggesting that the domestic demand is not influenced by its own history. The final model obtained can, therefore, be considered as a static model. In terms of coefficients of the explanatory variables, the (log) index of industrial production shows a positive sign as in the monthly case. Thus, the higher the income of the tourists' countries the higher the demand for leisure. This finding seems to be consistent with the result obtained by Arbel and Ravid (1985) for which "the income is found to be the single most important determinant of long run recreation use" (p.981).

The (log) substitute price coefficient has a positive sign both in the monthly and annual models, whereas one would expect a negative sign. Note, also, that in this model the time trend coefficient is statistically significant and shows a negative sign denoting a decreasing popularity for the north of Sardinia as a destination.

The cointegrating vector enters the equation with a negative sign, as is the case when using monthly data. The (log) weather variable once again does not play a role in explaining the domestic demand for tourism.

A couple of further comments are due. Experimenting with the unadjusted series of foreign arrivals of tourists gives almost equivalent results both in terms of coefficients, significance and diagnostic statistics.

Furthermore, a brief account on the appropriate functional form to use is given. The investigation of the integration status of the economic series of interest, by using the ADF test suggests: *A*, *PR* and *ER* to be non-stationary in the level, and statistical evidence is found for the substitute price and the relative price to be stationary in the level. However, these results, as already mentioned for the logarithmic specification, have to be considered with some caution given that the number of observations is quite small (24 in total). Hence, one proceeds as for the linear monthly case, treating only *RP* and *ER* as stationary in the first difference, which leads to a cointegration analysis for these variables. By adopting the Johansen analysis, evidence is found for the existence of one cointegrating vector. An initial unrestricted k=3 VAR, which includes a constant and a trend unrestrictedly, can be reduced to a one lag system. The resulting cointegrating vector for the linear specification is the following:

$$ECL = RP - 0.00083024ER$$

In order to run the Box and Cox test, an unrestricted 1 lag annual tourism demand equation is estimated, and expressed both in a logarithmic and linear form, where the explanatory variables are defined as before.

1) logarithmic form

$$LA_t = a_1 + a_2\, LA_{t-1}... + a_3\, LPR_t.. + a_4\, LSP_t..+ a_5\, DLER_t..+ a_6\, DLRP_t. + a_7\, CI_{t-1}+$$
$$+\ a_8\, LW + a_9\, Trend + e_t$$

and

2) linear form

$$A_t = a_1 + a_2\, A_{t-1}... + a_3\, PR_t.. + a_4\, SP_t..+ a_5\, DER_t..+ a_6\, DRP_t.. + a_7\, ECL_{t-1}+$$
$$+\ a_8\, W + a_9\, Trend + e_t$$

The sum of the squared errors from the logarithmic form (*SSELL*) is equal to 0.06381525754, whereas the sum of the squared errors for the linear form (*SSEL*) equals 52722686.58. The aim is to test whether the null hypothesis that the two models are empirically equivalent and find out which of the two models fits better the data. Using formula 2.4.1.1 (Chapter 2)[20], the calculated χ^2 equals 5.95 that is greater than the tabulated critical value, 3.84, at the 5% level; hence, the null fails to be accepted and the two models are empirically different. Moreover, one infers that the logarithmic specification is better than the linear specification as the $SSE_L/(\bar{A}\,G)^2$ is higher than SSE_{LL}.

4.3.1 The Model Specification Using Annual Data Analysis And Supply Components

As already stated, an advantage when dealing with annual data is the possibility of including variables which are just available with an annual frequency. Such variables might be able, in fact, to explain much better the variation of the dependent variable under study.

In estimating the annual foreign demand for tourism in the north of Sardinia, a component of the quantity of the supply services can be taken into consideration. For this purpose one includes two more variables, that is: the (log) number of total boats arrived (*LB*) in the two main ports (Porto Torres and Olbia) and the (log) number of international flights (*LAE*) in the two main airports (Fertilia and Olbia)[21].

An initial unrestricted one lag model including the explanatory variables as mentioned above (*i.e.* *LA*, *LPR*, *LSP*, *DLRP*, *DLER*, CI_{t-1}, *LW*), the two new variables (*i.e.* *LAE* and *LB* as above defined), a constant and a time trend is run. The final results are shown in Table 4.13.

---

20 Note that $(SSEL\,/\,(\bar{A}\,G)^2)$ equals 0.109696.

21 The source is the "*Annuario Statistico Italiano*" (1972-1996). These figures are available with an annual frequency. In particular, the total international flights and total boat arrivals are considered. Though, in the statistical sources, international arrivals of boats are reported separately from domestic arrivals, nevertheless, one has considered the total: *i.e.* international plus domestic arrivals of boats. One can argue, in fact, that foreign tourists (in particular Germans that represent the highest percentage) are more likely to use Genova, Livorno or Civitavecchia harbours to reach the north of Sardinia.

Table 4. 13 Static Model for Foreign Demand of Tourism with Supply Components

```
EQ(1) Modelling LA by OLS  (using datiann.in7)
The present sample is:  3 to 24
```

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-26.477	5.2151	-5.077	0.0005	0.7205
LSP_1	4.0415	0.56928	7.099	0.0000	0.8344
DLRP_1	-2.6862	0.64453	-4.168	0.0019	0.6346
DLER	2.6267	0.42678	6.155	0.0001	0.7911
DLER_1	0.54752	0.31078	1.762	0.1086	0.2369
CI_1	-2.0057	0.36824	-5.447	0.0003	0.7479
LW	1.4098	0.32061	4.397	0.0013	0.6591
LW_1	1.0208	0.27218	3.750	0.0038	0.5845
LB	1.2670	0.24993	5.070	0.0005	0.7199
LAE	0.70078	0.088519	7.917	0.0000	0.8624
LAE_1	0.30659	0.076114	4.028	0.0024	0.6187
Trend	-0.079658	0.013380	-5.954	0.0001	0.7800

```
R^2 = 0.981773  F(11,10) = 48.968 [0.0000] sigma = 0.0432263
DW = 2.12 RSS = 0.01868512734 for 12 variables and 22 observations

AR 1- 2 F( 2,   8) =      0.2682 [0.7714]
ARCH 1  F( 1,   8) =      1.3654 [0.2762]
Normality Chi^2(2)=      0.1875 [0.9105]
RESET   F( 1,   9) =      3.8546 [0.0812]
```

The results from the estimation using annual data show a satisfactory determination in terms of statistical significance of the coefficient for the total number of boat arrivals (*LB*) and international flights (*LAE*). This finding is in line with the expectation that an increase in the domestic demand is associated to an increase of the supply of means of transportation.

In comparison with model in Table 4.12, the overall performance of the model has improved in terms of diagnostic statistics that detect no problems. Further improvement has been obtained in the coefficient of determination.

The income proxy (*LPR*) does not influence the international demand for tourism and it is excluded. The substitute price presents the "usual" positive sign, as stated in the previous cases, and has a strong impact on the dependent variable. The time trend coefficient shows a downwards trend in popularity for Sassari Province, with a negative coefficient.

The first difference of the relative price (*DLRP*), is statistically significant and presents a negative sign. This indicates that a loss in terms of competitiveness between the origin and the destination country, *ceteris paribus*, is associated with a reduction of the arrivals of tourists. The exchange rate growth (*DLER*) is significant and shows that a depreciation of lira with respect to the origin countries currency is associated with an increase in demand for tourism. The coefficient for the cointegrating vector presents a negative sign confirming the results obtained so far.

The (log) weather variable, in this model, plays a role in explaining the international demand for tourism, differently from the previous models. The coefficient is statistically significant at the 5% level and it has a positive sign. It may well be that relative average high temperatures have a positive impact in the choice of clients that come from quite cold climates, such as Germans and Swiss tourists.

4.3.2 Testing For Simultaneity With Annual Data

The statistical significance of the coefficients for *LAE*, *LB* and *DLER* suggests a further issue in testing the presence of simultaneity for these explanatory variables. This test is carried out by adopting the Durbin-Wu-Hausman's procedure as discussed in Section 2.9.1 (Chapter 2).

It is assumed that the demand and supply models can be expressed as follows:

1) $LA_t = \alpha_0 + \alpha_1\, LSP_{t-1} + \alpha_2\, DLRP_{t-1} + \alpha_3\, DLER_t + \alpha_4\, DLER_{t-1} + \alpha_5\, CI_{t-1} + \alpha_6\, LW_t + \alpha_7\, LW_{t-1} + \alpha_8\, LB_t + \alpha_9\, LAE_t + \alpha_{10}\, LAE_{t-1} + \alpha_{11}\, TREND + \varepsilon_{1,t}$

2) $LAE_t = \beta_0 + \beta_1\, LA_t + \beta_2\, LPR_t + \varepsilon_{2,t}$

in accordance with the previous results obtained using OLS (Table 4.12), where:

a) *LA* = normalised series of foreign arrivals for the average number of weekends (*i.e.* Saturdays) in a year.

b) *LSP*= substitute price (consumer price index in Sassari by the weighted average consumer price index in other destinations in the Mediterranean area).

c) *DLRP*= relative price growth (Sassari - origin countries).

d) *DLER* = exchange rate growth.

e) *CI* = cointegrating vector.

f) *LW* = annual average temperatures in Sassari.

g) *LB* = total number of international and domestic arrivals of boats in the north of Sardinia.

h) *LAE* = total number of international flights in the north of Sardinia.

i) *LPR* = weighted average industrial production index for the main clients of foreign tourists. This variable is treated as instrument.

Firstly, one obtains the reduced form by regressing *LAE* on the variables included into equation 1) plus the instrument variable *LPR* (one assumes, in fact, that the industrial production index, used as the income proxy is able to explain the number of foreign arrivals as well as the number of international flights in the north of Sardinia). Hence, the residuals obtained from this regression (say $LAERES = \hat{w}$) are saved. The complete results are shown in Table 4.14.

Table 4. 14 Modelling Reduced Form for (log) Number of International Flights

```
Modelling Reduced Form for LAE
Variable      Coefficient    Std.Error  t-value
LPR                3.4288       1.1552    2.968
Constant           12.961       13.087    0.990
DLER_1            -1.5222      0.91445   -1.665
CI_1              0.57784      0.95444    0.605
LW                -1.0092      0.78151   -1.291
LW_1              -1.6318      0.71701   -2.276
LB                -1.1150      0.61021   -1.827
LSP_1             -2.3633       1.2952   -1.825
LAE_1            -0.32304      0.21358   -1.513
Trend            0.061291     0.028995    2.114
DLRP_1             3.2989       1.3770    2.396
DLER              -2.9171      0.81198   -3.593

R^2 = 0.944362  F(11,10) = 15.43 [0.0001]  RF sigma = 0.112596
DW = 2.27 RSS = 0.1267778188 for 12 variables and 22 observations

Modelling Reduced Form for LA
Variable      Coefficient    Std.Error  t-value
LPR                2.5314      0.88803    2.851
Constant          -18.110       10.060   -1.800
DLER_1           -0.56699      0.70296   -0.807
CI_1              -1.6437      0.73370   -2.240
LW                0.71708      0.60076    1.194
LW_1             -0.14934      0.55118   -0.271
LB                0.48131      0.46908    1.026
LSP_1              2.4340      0.99567    2.445
LAE_1            0.071125      0.16418    0.433
Trend           -0.038528     0.022289   -1.729
DLRP_1           -0.42126       1.0585   -0.398
DLER              0.59889      0.62419    0.959

R^2 = 0.926922  F(11,10) = 11.531 [0.0003]  RF sigma = 0.0865546
DW = 1.66 RSS = 0.07491706888 for 12 variables and 22 observations

EQ(1) Modelling LAE by IVE  (using datiann.in7)
The present sample is:  3 to 24

Variable      Coefficient    Std.Error  t-value  t-prob
LA               0.77853     0.41046    1.897  0.0732
LPR                1.2005      0.86835    1.383  0.1828
Constant          -6.4985       1.7582   -3.696  0.0015

Additional Instruments used:
   DLER_1     CI_1       LW      LW_1       LB    LSP_1
    LAE_1    Trend   DLRP_1      DLER

sigma = 0.170195  DW = 1.94
RSS = 0.5503625099 for 3 variables and 22 observations
2 endogenous and 2 exogenous variables with 12 instruments
Reduced Form sigma = 0.112596
Specification  Chi^2(9)   =      20.386 [0.0157] *
Testing beta=0:Chi^2(2)   =      56.68 [0.0000] **

AR 1- 2 F( 2, 17) =      0.2786 [0.7602]
ARCH 1  F( 1, 17) =      1.9823 [0.1772]
Normality Chi^2(2)=      3.7913 [0.1502]
Xi^2     F( 4, 14) =     0.8447 [0.5198]
Xi*Xj    F( 5, 13) =     0.7585 [0.5951]
```

The specification test $\chi^2(9)$ suggests the non validity of the instruments at the 5% level. The second test reported β=0 suggests that the coefficients excluding the constant term are jointly different from zero (it is analogue of the OLS *F*-test of R-squared).

The saved residuals (*LAERES*) have been included in the original regression to "correct" for simultaneity. The resulting results by OLS is given in Table 4.15.

Table 4. 15 Testing for Simultaneity for *LAE*

EQ(2) Modelling LA by OLS (using datiann.in7)
The present sample is: 3 to 24

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-6.7854	5.7318	-1.184	0.2668	0.1347
LSP_1	1.5171	0.70222	2.160	0.0590	0.3415
DLRP_1	-0.84641	0.59551	-1.421	0.1889	0.1833
DLER	0.75461	0.52219	1.445	0.1823	0.1883
DLER_1	-0.15242	0.25538	-0.597	0.5653	0.0381
CI_1	-0.89911	0.35029	-2.567	0.0303	0.4226
LW	0.49673	0.29585	1.679	0.1275	0.2385
LW_1	0.10535	0.27719	0.380	0.7127	0.0158
LB	0.35399	0.26862	1.318	0.2201	0.1618
LAE	0.92344	0.076561	12.062	0.0000	0.9417
LAE_1	0.048428	0.077931	0.621	0.5497	0.0411
Trend	-0.031338	0.014270	-2.196	0.0557	0.3489
LAERES	-0.76665	0.18463	-4.152	0.0025	0.6570

R^2 = 0.993749 F(12,9) = 119.23 [0.0000] sigma = 0.0266843
DW = 2.01 RSS = 0.006408459165 for 13 variables and 22 observations

AR 1- 2 F(2, 7) = 0.15886 [0.8561]
ARCH 1 F(1, 7) = 0.69407 [0.4323]
Normality Chi^2(2)= 0.11222 [0.9454]
RESET F(1, 8) = 1.12270 [0.3203]

The results suggest that the coefficient for the residuals (*LAERES*) is statistically different from zero. Thus, the null hypothesis of no simultaneity fails to be accepted and this variable can be treated as endogenous. This finding seems to be likely since the number of planes and/or charters might be changed more promptly depending on the number of passengers booking for a place. Thus, the number of international flights could be thought to be endogenous depending on the actual number of arrivals.

The same analysis has been done for the number of total arrivals of boats in the north of Sardinia (*LB*). As a first step to test for simultaneity with respect to *LB*, one estimates the reduced form regressing *LB* on the variables previously mentioned and the industrial production treated as the instrument from which the residuals (say *LBRES*) are saved.

Table 4. 16 Modelling Reduced Form for Total Number of Boat Arrivals

```
Modelling Reduced Form for LB
Variable      Coefficient    Std.Error  t-value
Constant           14.233       4.1964    3.392
LPR               0.86102      0.65670    1.311
DLER              -1.3127      0.36297   -3.617
DLER_1           -0.45134      0.44119   -1.023
CI_1              0.88308      0.33489    2.637
LW               -0.25629      0.37001   -0.693
LW_1             -0.30176      0.38473   -0.784
LSP_1             -1.4852      0.47921   -3.099
LAE              -0.22449      0.12286   -1.827
LAE_1            -0.16660     0.092249   -1.806
Trend            0.032957     0.011674    2.823
DLRP_1             1.0014      0.70753    1.415

R^2 = 0.940376  F(11,10) = 14.338 [0.0001]  RF sigma = 0.0505226
DW = 2.15 RSS = 0.02552534634 for 12 variables and 22 observations

Modelling Reduced Form for LA
Variable      Coefficient    Std.Error  t-value
Constant          -9.3172       6.2568   -1.489
LPR                1.2973      0.97914    1.325
DLER              0.93757      0.54119    1.732
DLER_1           -0.10381      0.65781   -0.158
CI_1             -0.94701      0.49932   -1.897
LW                 1.0768      0.55169    1.952
LW_1              0.57041      0.57364    0.994
LSP_1              2.1915      0.71450    3.067
LAE               0.39256      0.18318    2.143
LAE_1            0.082608      0.13754    0.601
Trend           -0.039055     0.017406   -2.244
DLRP_1            -1.3965       1.0549   -1.324
R^2 = 0.944647  F(11,10) = 15.515 [0.0001]  RF sigma = 0.0753295
DW = 1.70 RSS = 0.05674537316 for 12 variables and 22 observations

EQ(1) Modelling LB by IVE  (using datiann.in7)
The present sample is:  3 to 24

Variable      Coefficient    Std.Error  t-value  t-prob
LA               -0.60116      0.21053   -2.855  0.0101
Constant           4.2690      0.94289    4.528  0.0002
LPR                2.2521      0.44963    5.009  0.0001

Additional Instruments used:
     DLER   DLER_1    CI_1      LW     LW_1    LSP_1
      LAE    LAE_1   Trend  DLRP_1

sigma = 0.0917932  DW = 1.07
RSS = 0.1600937483 for 3 variables and 22 observations
2 endogenous and 2 exogenous variables with 12 instruments
Reduced Form sigma = 0.0505226
Specification  Chi^2(9)   =     10.504 [0.3112]
Testing beta=0:Chi^2(2)   =     38.706 [0.0000] **
AR 1- 2 F( 2, 17) =    2.418   [0.1191]
ARCH 1  F( 1, 17) =    2.72    [0.1175]
Normality Chi^2(2)=    0.75372 [0.6860]
Xi^2    F( 4, 14) =    0.42341 [0.7893]
Xi*Xj   F( 5, 13) =    0.58006 [0.7150]
```

In this case, the specification test $\chi^2(9)$ suggests the validity of the instruments as the null cannot be rejected. The second test reported β=0 suggests that the coefficients excluding the constant term are jointly different from zero.

The residuals (*LBRES*) are included into the original model to "correct" for simultaneity. The results are provided in Table 4.17.

Table 4. 17 Testing for Simultaneity for *LB*

EQ(2) Modelling LA by OLS (using datiann.in7)
The present sample is: 3 to 24

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-19.805	7.6555	-2.587	0.0294	0.4265
LSP_1	3.4126	0.77473	4.405	0.0017	0.6831
DLRP_1	-2.3838	0.68335	-3.488	0.0068	0.5749
DLER	2.3832	0.46767	5.096	0.0006	0.7426
DLER_1	0.71794	0.33798	2.124	0.0626	0.3339
CI_1	-1.5892	0.50691	-3.135	0.0120	0.5220
LW	1.2136	0.35649	3.404	0.0078	0.5629
LW_1	1.0603	0.26934	3.937	0.0034	0.6326
LB	0.86487	0.42177	2.051	0.0706	0.3184
LAE	0.66761	0.091397	7.305	0.0000	0.8557
LAE_1	0.30648	0.074728	4.101	0.0027	0.6514
Trend	-0.066236	0.017425	-3.801	0.0042	0.6162
LBRES	0.29185	0.24895	1.172	0.2712	0.1325

R^2 = 0.984188 F(12,9) = 46.682 [0.0000] sigma = 0.0424392
DW = 2.38 RSS = 0.01620978162 for 13 variables and 22 observations

AR 1- 2 F(2, 7) = 0.56439 [0.5926]
ARCH 1 F(1, 7) = 0.49254 [0.5055]
Normality Chi^2(2)= 0.01228 [0.9939]
RESET F(1, 8) = 3.67 [0.0917]

The coefficient for the residuals (*LBRES*) is not statistically significant; the null hypothesis fails to be rejected, hence there is no simultaneity and *LB* can be treated as predetermined. One can argue that the number of boat arrivals is correlated with the capacity. Moreover, as far as the period under study is concerned (1972-1995), the number of boat arrivals is likely to be planned for the year (or years) ahead; thus, the number of boats cannot be adjusted promptly to the number of passengers requiring a place. Given this assumption, one can indeed treat such a variable (*LB*) as predetermined.

One might argue that it would be worth testing if *DLER* (the exchange rate growth) can be treated as predetermined, since the level of such a variable is statistically significant (see Table 4.13). Adopting Wu-Hausman's procedure, one obtains the reduced form regressing *DLER* on the variables previously mentioned and the industrial production as the instrument from which the residuals (say *RESDLER*) are saved.

Table 4. 18 Modelling Reduced Form for the Weighted Average Exchange Rate

Modelling Reduced Form for ***DLER***

Variable	Coefficient	Std.Error	t-value
Constant	7.1191	2.7176	2.620
LPR	0.54958	0.36879	1.490
LSP_1	-0.98811	0.22456	-4.400
DLRP_1	1.0232	0.30483	3.357
LAE	-0.19315	0.053765	-3.593
DLER_1	-0.14882	0.26172	-0.569
CI_1	0.45182	0.20522	2.202
LW	-0.31820	0.19251	-1.653
LW_1	-0.23057	0.21531	-1.071
LB	-0.43171	0.11937	-3.617
LAE_1	-0.063965	0.057466	-1.113
Trend	0.021902	0.0057072	3.838

R^2 = 0.869865 F(11,10) = 6.0767 [0.0040] RF sigma = 0.0289733
DW = 2.57 RSS = 0.008394546594 for 12 variables and 22 observations

Modelling Reduced Form for ***LA***

Variable	Coefficient	Std.Error	t-value
Constant	-8.6338	8.0049	-1.079
LPR	1.6415	1.0863	1.511
LSP_1	1.4956	0.66147	2.261
DLRP_1	-0.018281	0.89790	-0.020
LAE	0.17583	0.15837	1.110
DLER_1	0.070447	0.77092	0.091
CI_1	-0.87883	0.60449	-1.454
LW	0.58060	0.56706	1.024
LW_1	0.35311	0.63421	0.557
LB	0.12636	0.35162	0.359
LAE_1	0.12393	0.16927	0.732
Trend	-0.023676	0.016811	-1.408

R^2 = 0.928952 F(11,10) = 11.886 [0.0002] RF sigma = 0.0853437
DW = 1.81 RSS = 0.07283539944 for 12 variables and 22 observations

EQ(1) Modelling DLER by IVE (using datiann.in7)
The present sample is: 3 to 24

Variable	Coefficient	Std.Error	t-value	t-prob
LA	-0.26792	0.14295	-1.874	0.0764
Constant	0.75777	0.61560	1.231	0.2334
LPR	0.43985	0.30275	1.453	0.1626

Additional Instruments used:
LSP_1 DLRP_1 LAE DLER_1 CI_1 LW
LW_1 LB LAE_1 Trend

sigma = 0.0596325 DW = 1.59
RSS = 0.06756469577 for 3 variables and 22 observations
2 endogenous and 2 exogenous variables with 12 instruments
Reduced Form sigma = 0.0289733
Specification Chi^2(9) = 13.834 [0.1283]
Testing beta=0:Chi^2(2) = 3.8315 [0.1472]
AR 1- 1 F(1, 18) = 0.34706 [0.5631]
ARCH 1 F(1, 17) = 0.74569 [0.3999]
Normality Chi^2(2)= 1.1147 [0.5727]
Xi^2 F(4, 14) = 0.16079 [0.9547]
Xi*Xj F(5, 13) = 0.21733 [0.9488]

The residuals *RESDLER* are included into the original model which is run by OLS. The results are the following:

Table 4. 19 Testing for Simultaneity for *DLER*

EQ(2) Modelling LA by OLS (using datiann.in7)
The present sample is: 3 to 24

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
DLER	-0.80289	1.1278	-0.712	0.4945	0.0533
Constant	-8.0468	6.9479	-1.158	0.2766	0.1297
LSP_1	1.7765	0.82675	2.149	0.0602	0.3391
DLRP_1	-1.3511	0.62997	-2.145	0.0606	0.3382
LAE	0.41980	0.10962	3.829	0.0040	0.6197
DLER_1	0.61716	0.22653	2.724	0.0234	0.4520
CI_1	-0.72784	0.48439	-1.503	0.1672	0.2006
LW	0.65384	0.33351	1.960	0.0816	0.2993
LW_1	0.74090	0.21637	3.424	0.0076	0.5657
LB	0.68486	0.25838	2.651	0.0265	0.4384
LAE_1	0.20488	0.063900	3.206	0.0107	0.5332
Trend	-0.033503	0.017527	-1.912	0.0882	0.2888
RESDLER	2.2009	0.69592	3.163	0.0115	0.5264

R^2 = 0.991367 F(12,9) = 86.128 [0.0000] sigma = 0.0313582
DW = 2.60 RSS = 0.008850037787 for 13 variables and 22 observations

AR 1- 1 F(1, 8) = 1.4551 [0.2622]
ARCH 1 F(1, 7) = 0.082979 [0.7816]
Normality Chi^2(2)= 0.034143 [0.9831]
RESET F(1, 8) = 3.3803 [0.1033]

As one can notice the residuals from the reduced form are statistically significant, thus the null hypothesis of no simultaneity cannot be accepted. The conclusion is that one should consider *DLER* (exchange rate growth) as endogenous. However, it can be argued it is difficult to believe that the exchange rate, with respect to the main origin countries, can be determined by the model. Therefore, one rejects endogeneity *a priori*. Note that experiments with monthly data have confirmed that the weighted average exchange rate has to be considered predetermined, however, the results have not been included.

4.3.3 Testing For Simultaneity With Monthly Data When The Number Of Boat Arrivals Are Included

The acceptance of the null hypothesis for annual data might be caused by a shortage of data points. In this section one verifies whether the (log) number of international and domestic boat arrivals in the north of Sardinia (*LB*) is predetermined when using monthly data. Note that the annual boats figure for each year is kept constant along the year; it is assumed, in fact, that the capacity is fixed for the year in advance.

One starts with the most unrestricted model with 13 lags, as suggested by the joint *F*-test[22], where the usual variables are included (*i.e. LA*, *LPR*, *LSP*, *DLRP*, *DLER*, *EASTER,* the cointegrating vector, four impulse dummies, the time trend and the 11 seasonal dummies). No problems are detected in the residuals. A general-to-specific simplification is carried out according to the joint *F*-test statistic as well as to the SC criterion. The final parsimonious data congruent model is shown in Table 4.20.

Table 4. 20 Monthly Final Model when the Total Number of Boat Arrivals is Included

EQ(1) Modelling LA by OLS (using for.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-4.3618	2.3605	-1.848	0.0658	0.0135
LA_1	0.18618	0.043652	4.265	0.0000	0.0681
LA_3	0.18961	0.036370	5.213	0.0000	0.0984
LA_11	0.11843	0.049458	2.394	0.0174	0.0225
RLPR	2.2509	0.62079	3.626	0.0003	0.0502
LSP_1	-4.0493	1.8892	-2.143	0.0330	0.0181
LSP_2	5.0600	1.8865	2.682	0.0078	0.0281
RLSP1	4.2708	1.8729	2.280	0.0234	0.0205
CI_1	-0.52733	0.20185	-2.612	0.0095	0.0267
LB	0.32381	0.13785	2.349	0.0196	0.0217
RLB	1.5142	0.64903	2.333	0.0204	0.0214
i1974p12	1.4712	0.20459	7.191	0.0000	0.1720
i1979p3	-0.54071	0.20548	-2.631	0.0090	0.0271
i1985p3	0.68751	0.20543	3.347	0.0009	0.0430
i1991p11	-0.59885	0.20465	-2.926	0.0037	0.0332
easter	0.44111	0.071892	6.136	0.0000	0.1313
JA	0.19026	0.075547	2.518	0.0124	0.0248
FE	0.61605	0.11900	5.177	0.0000	0.0972
MAR	1.0174	0.15893	6.401	0.0000	0.1413
AP	1.6592	0.21328	7.780	0.0000	0.1955
MJJ	2.7017	0.23283	11.604	0.0000	0.3510
AU	2.4382	0.21215	11.493	0.0000	0.3466
SE	2.3180	0.16404	14.131	0.0000	0.4450
OT	1.1633	0.12875	9.035	0.0000	0.2469
NO	0.050349	0.083041	0.606	0.5449	0.0015

R^2 = 0.981731 F(24,249) = 557.51 [0.0000] sigma = 0.197098
DW = 2.02 RSS = 9.673094916 for 25 variables and 274 observations

AR 1- 7	F(7,242) =	0.82018 [0.5713]
ARCH 7	F(7,235) =	0.40581 [0.8982]
Normality	Chi^2(2)=	3.898 [0.1424]
Xi^2	F(34,214) =	1.4196 [0.0723]
RESET	F(1,248) =	1.7184 [0.1911]

where:

22 Note that the SC criterion suggests a further parameter reduction.

dep.var		T	k	df	RSS	sigma	Schwarz
12 lags: LA	OLS	274	95	179	6.81893	0.195178	-1.74727
13 lags: LA	OLS	274	101	173	6.18796	0.189126	-1.72145

Model 13 lags --> 12 lags: F(6, 173) = 2.9401 [0.0093] **

LA = (log) normalised series of foreign arrivals for the number of weekends (*i.e.* Saturdays) in a month.

RLPR = difference between the coefficients of the third and seventh lag of the (log) weighted average industrial production index for the origin countries. Such a restriction is suggested both from the joint *F*-test and the SC criterion[23].

RLSP1 = difference between the coefficients of the eleventh and twelfth lag of the (log) substitute price. One accepts such a restriction from the joint *F*-test and as suggested by the SC criterion[24].

LSP = (log) substitute price (consumer price index in Sassari by the weighted average consumer price index in other destinations in the Mediterranean area).

CI = cointegrating vector between *LRP* (relative price) and *LER* (exchange rate).

RLB = Difference between the coefficient of the fourth and fifth lag of the (log) total number of boat arrivals in north of Sardinia. Such a restriction has been suggested by the SC criterion and accepted by the joint *F*-test[25].

EASTER= dummy 0-1 with respect to the Easter holiday.

MJJ = seasonal dummy giving the value of 1 to May, June and July months and 0 otherwise. Such a restriction was possible given the results from the joint *F*-test and the SC criterion[26].

Considering the coefficients of *LB* the fourth lag has presented a positive sign that reflects as the number of foreign tourists is increasing given the number of boats determined at the beginning of the year (say in January); however, the fifth lag, that could occur in June, has shown a negative sign that reflects a decrease in the number

---

23 The restriction on the coefficients of the third and seventh lag is accepted at the 5% level from the *F*-test (1,246) as the calculated value (2.0527) is smaller than the critical value (3.84). Moreover, the SC criterion is minimised when the restriction is imposed; from -2.79465 to -2.80683, after imposing the restriction.

24 The restriction on the coefficients of the eleven and twelfth lag is accepted at the 5% level from the *F*-test (1,245) as the calculated value (0.23764) is smaller than the critical value (3.84). Moreover, the SC criterion is minimised when the restriction is imposed; from -2.77514 to -2.79465, after imposing the restriction.

25 The restriction on the coefficients of the fourth and fifth lag is accepted at the 5% level from the *F*-test (1,248) as the calculated value (3.515) is smaller than the critical value (3.84). Moreover, the SC criterion is minimised when the restriction is imposed; from -2.82522 to -2.83163, after imposing the restriction.

26 The restriction on the coefficients of May, June and July is accepted at the 5% level from the *F*-test (2,247) as the calculated value (0.56814) is smaller than the critical value (3.84). Moreover, the SC criterion is minimised when the restriction is imposed; from -2.80683 to -2.84321, after imposing the restriction.

of tourists' arrivals given the number of boats that has been planned to arrive to Sassari Province since the previous January. The oscillation of the boats supply (given by the coefficient of *RLB*) presents a positive sign and it is statistically significant, as well as the level of *LB*.

From Table 4.21, the long run multipliers and the standard errors are in general well-specified. Moreover, they are statistically significant and present the expected signs. The only exception is for the substitute price, though in the short run coefficient has turned out with the expected negative sign (Table 4.20). Again, one may argue that imposing the restriction on the coefficients of the eleventh and twelfth lag of the substitute price (*i.e. RLSP1*) overstates the precision of *LSP* effect. However, once more, the inclusion of the four *LSP* terms shows that the effect is very marginal.

Table 4. 21 Long Run Multipliers and Standard Errors

Solved Static Long Run equation

LA =	-8.624	+4.45 RLPR	+1.998 LSP
(SE)	(4.797)	(1.319)	(0.2972)
	+8.444 RLSP1	-1.043 CI	+0.6402 LB
	(3.81)	(0.4107)	(0.2715)
	+2.909 i1974p12	-1.069 i1979p3	+1.359 i1985p3
	(0.5136)	(0.416)	(0.4255)
	-1.184 i1991p11	+0.8721 easter	+0.3762 JA
	(0.4204)	(0.1738)	(0.1566)
	+1.218 FE	+2.011 MAR	+3.281 AP
	(0.2588)	(0.3147)	(0.4054)
	+5.342 MJJ	+4.821 AU	+4.583 SE
	(0.4624)	(0.3475)	(0.3739)
	+2.3 OT	+0.09955 NO	+2.994 RLB
	(0.2775)	(0.1613)	(1.345)

ECM = LA + 8.62377 - 4.45036*RLPR - 1.99824*LSP - 8.4438*RLSP1 + 1.04258*CI
- 0.640204*LB - 2.90874*i1974p12 + 1.06905*i1979p3 - 1.35928*i1985p3
+ 1.18399*i1991p11 - 0.872127*easter - 0.376165*JA - 1.21799*FE
- 2.01144*MAR - 3.2805*AP - 5.34163*MJJ - 4.82051*AU - 4.58286*SE
- 2.29998*OT - 0.0995451*NO - 2.9938*RLB;

WALD test Chi^2(20) = 604.07 [0.0000] **

The Durbin-Wu-Hausman's simultaneity test is employed in order to test whether *LB* can be treated as predetermined or endogenous. Note that the null hypothesis is of no simultaneity, that is predeterminedness. The demand and supply models can be expressed as follows:

1) $LA_t = \alpha_0 + \alpha_1 LA_{t-1} + \alpha_2 LA_{t-3} + \alpha_3 LA_{t-11} + \alpha_4 RLPR_t + \alpha_5 LSP_{t-1} + \alpha_6 LSP_{t-2} + \alpha_7 RLSP1_t + \alpha_8 LB_{t-4} + \alpha_9 LB_{t-5} + \alpha_{10} CI_{t-1} + \alpha_{11} EASTER + \alpha_{12} i1974p12 + \alpha_{13} i1979p3 + \alpha_{14} i1985p3 + \alpha_{15} i1991p11 + \alpha_{16} EASTER + \delta_{16} JAN + ... + \delta_{24} NOV + \varepsilon_{1,t}$

2) $LB_t = \beta_0 + \beta_1 LA_t + \beta_2 TREND_t + \varepsilon_{2,t}$

Note that the trend and the constant are used as instruments in the second equation.

Firstly, one obtains the reduced form by regressing *LB* on the variables included into equation 1) plus the instrument variable *TREND*; one can assume, in fact, that the time trend, used as a proxy in a possible change in the consumers' tastes, is able to explain the number of foreign arrivals as well as the number of boat arrivals in the north of Sardinia. Hence, the residuals obtained from this regression ($RESIDUAL = \hat{w}$) are saved. The results are reported in Table 4.22.

Table 4. 22 Modelling Reduced form for the Number of Boat Arrivals

Modelling Reduced Form for ***LB***

Variable	Coefficient	Std.Error	t-value
Constant	3.6605	0.64947	5.636
Trend	0.00044008	0.00012998	3.386
LA_1	0.0042628	0.0077787	0.548
LA_3	0.00058766	0.0064852	0.091
LA_11	0.0016118	0.0088222	0.183
RLPR	0.22839	0.11069	2.063
LSP_1	-0.40479	0.33739	-1.200
LSP_2	0.13825	0.34082	0.406
RLSP1	0.095369	0.33460	0.285
LB_4	0.80191	0.11803	6.794
LB_5	-0.061734	0.11822	-0.522
CI_1	0.23132	0.046003	5.028
i1974p12	0.037011	0.036533	1.013
i1979p3	-0.025451	0.036619	-0.695
i1985p3	-0.010772	0.036630	-0.294
i1991p11	0.016765	0.036561	0.459
easter	0.0034130	0.012822	0.266
JA	0.0061230	0.013476	0.454
FE	0.0085795	0.021237	0.404
MAR	0.0062162	0.028358	0.219
AP	-0.00031706	0.038063	-0.008
MJJ	-0.0081502	0.041574	-0.196
AU	-0.014919	0.037854	-0.394
SE	-0.013636	0.029254	-0.466
OT	-0.012154	0.022943	-0.530
NO	-0.0060506	0.014800	-0.409

R^2 = 0.940444 F(25,248) = 156.64 [0.0000] RF sigma = 0.0351398
DW = 0.522 RSS = 0.3062316914 for 26 variables and 274 observations

Modelling Reduced Form for ***LA***

Variable	Coefficient	Std.Error	t-value
Constant	-3.3508	3.6260	-0.924
Trend	0.00047273	0.00072569	0.651
LA_1	0.18428	0.043429	4.243
LA_3	0.18908	0.036207	5.222
LA_11	0.11632	0.049255	2.362
RLPR	2.3206	0.61797	3.755
LSP_1	-4.2018	1.8836	-2.231
LSP_2	4.9653	1.9028	2.609
RLSP1	4.4886	1.8681	2.403
LB_4	1.8089	0.65897	2.745
LB_5	-1.5728	0.66005	-2.383
CI_1	-0.47473	0.25684	-1.848
i1974p12	1.4839	0.20397	7.275
i1979p3	-0.54085	0.20445	-2.645
i1985p3	0.69427	0.20451	3.395
i1991p11	-0.59975	0.20412	-2.938
easter	0.44207	0.071588	6.175
JA	0.19183	0.075240	2.550
FE	0.61873	0.11857	5.219
MAR	1.0203	0.15832	6.445
AP	1.6659	0.21251	7.839
MJJ	2.7136	0.23211	11.691
AU	2.4502	0.21134	11.594
SE	2.3266	0.16333	14.245
OT	1.1680	0.12809	9.118
NO	0.053636	0.082632	0.649

R^2 = 0.981972 F(25,248) = 540.33 [0.0000] RF sigma = 0.196187
DW = 2.03 RSS = 9.545402109 for 26 variables and 274 observations

```
EQ(1) Modelling LB by IVE  (using For.in7)
The present sample is:  1973 (3) to 1995 (12)

Variable      Coefficient    Std.Error  t-value  t-prob
LA            -0.00049823    0.0045859   -0.109  0.9136
Constant           8.1349     0.032522  250.138  0.0000
Trend           0.0011413  7.9866e-005   14.290  0.0000

Additional Instruments used:
    LA_11      RLPR     LSP_1     LSP_2     RLSP1      LB_4
     LB_5      CI_1 i1974p12  i1979p3  i1985p3 i1991p11
   easter        JA        FE       MAR        AP        AU
       SE        OT        NO       MJJ      LA_1      LA_3

sigma = 0.103683  DW = 0.032
RSS = 2.913312047 for 3 variables and 274 observations
2 endogenous and 2 exogenous variables with 26 instruments
Reduced Form sigma = 0.0351398
Specification  Chi^2(23)  =      245.21 [0.0000] **
Testing beta=0:Chi^2(2) =        207.29 [0.0000] **

AR 1- 7 F( 7,264) =      859.13 [0.0000] **
ARCH 7  F( 7,257) =      389.38 [0.0000] **
Normality Chi^2(2)=      17.035 [0.0002] **
Xi^2    F( 4,266) =      4.9496 [0.0007] **
Xi*Xj   F( 5,265) =      3.9632 [0.0018] **
```

Note, however, that the reduced form presents problems in the residuals. Furthermore, the specification test $\chi^2(23)$ suggests the non validity of the instruments. The second test where β=0 suggests that the coefficients excluding the constant term are different from zero (it is analogue of the OLS *F*-test of R-squared). These mis-specifications suggest that a problem with the Hausman-type test is setting up a reasonable equation for *LB*.

The saved residuals (*RESIDUAL*) have been included in the original regression to "correct" for simultaneity.

Table 4. 23 Simultaneity Test for *LB* using Monthly Data

EQ(2) Modelling LA by OLS (using For.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-5.8184	2.3279	-2.499	0.0131	0.0246
LA_1	0.18615	0.043369	4.292	0.0000	0.0691
LA_3	0.18917	0.036154	5.232	0.0000	0.0994
LA_11	0.11748	0.049150	2.390	0.0176	0.0225
RLPR	2.3874	0.62210	3.838	0.0002	0.0561
LSP_1	-4.3059	1.8848	-2.285	0.0232	0.0206
LSP_2	5.1107	1.8754	2.725	0.0069	0.0291
RLSP1	4.5028	1.8650	2.414	0.0165	0.0230
LB_4	2.1219	0.69542	3.051	0.0025	0.0362
LB_5	-1.5539	0.64853	-2.396	0.0173	0.0226
CI_1	-0.44319	0.23560	-1.881	0.0611	0.0141
i1974p12	1.4932	0.20393	7.322	0.0000	0.1778
i1979p3	-0.55160	0.20421	-2.701	0.0074	0.0286
i1985p3	0.68875	0.20414	3.374	0.0009	0.0439
i1991p11	-0.59784	0.20365	-2.936	0.0036	0.0336
easter	0.44271	0.071479	6.194	0.0000	0.1340
JA	0.19316	0.075138	2.571	0.0107	0.0260
FE	0.62024	0.11837	5.240	0.0000	0.0997
MAR	1.0208	0.15803	6.460	0.0000	0.1440
AP	1.6631	0.21201	7.844	0.0000	0.1988
MJJ	2.7069	0.23136	11.700	0.0000	0.3557
AU	2.4418	0.21073	11.587	0.0000	0.3512
SE	2.3199	0.16294	14.238	0.0000	0.4498
OT	1.1630	0.12792	9.092	0.0000	0.2500
NO	0.051254	0.082529	0.621	0.5351	0.0016
RESIDUAL	-0.35678	0.33497	**-1.065**	0.2879	0.0046

R^2 = 0.982023 F(25,248) = 541.9 [0.0000] sigma = 0.195908
DW = 2.04 RSS = 9.518194855 for 26 variables and 274 observations

AR 1- 7 F(7,241) = 0.84238 [0.5531]
ARCH 7 F(7,234) = 0.4358 [0.8791]
Normality Chi^2(2)= 3.6614 [0.1603]
Xi^2 F(36,211) = 1.3694 [0.0909]
RESET F(1,247) = 1.0216 [0.3131]

From Table 4.23, the results suggest that the coefficient for the residuals (*RESIDUAL*) is not statistically significant. Thus, the null hypothesis of no simultaneity has to be accepted and this variable can be treated as predetermined, confirming the results from Table 4.17 employing annual data.

From the difficulty of interpreting the dynamic response, one can conclude that the inclusion of the total number of boats arriving in the north of Sardinia, in determining the international demand for tourism, gives evidence to believe that a spurious correlation might be present (see Table 4.20). It might be possible that this variable is picking up the effects of other components not explicitly included into the model.

So far the monthly model seems to give better results than the annual frequency model. The former has been able to give a better specification in terms of properties of the variables, and short run as well as long run elasticities. The next step is to run a model with quarterly time series.

4.4 THE MODEL SPECIFICATION USING QUARTERLY DATA ANALYSIS

Another aim of this analysis is to make a further comparison of monthly and annual data versus quarterly data.

The first interesting step is to test the series under study for the possible existence of seasonal unit roots, and compare these results with the ones obtained with the monthly data series. Hylleberg *et al.* (1990) methodology is followed as given in Chapter 2. The results reported in Table 4.24, are obtained by fitting the equation (2.6.1) in Chapter 2 by OLS, for each of the five time series above mentioned[27].

Table 4. 24 Testing for Seasonal Unit Roots

t-statistics	Variable									
	LA		***LPR***		***LSP***		***LER***		***LRP***	
$\pi 1$	-3.93	**	-3.23	*	-1.04		-1.70		0.05	
$\pi 2$	-3.32	***	-5.26	****	-6.05	****	-7.13	***	-5.54	****
$\pi 3$	-6.45	****	-4.17	****	-2.82		-3.32	*	-4.18	****
$\pi 4$	-2.19	**	-7.41	****	-5.87	****	-5.04	****	-6.34	****
F-statistics	***LA***		***LPR***		***LSP***		***LER***		***LRP***	
$\pi 3, \pi 4$	26.80	****	55.82	****	25.53	****	22.42	****	41.61	****

Notes: The four, three, two and one asterisks indicate that the seasonal unit root null hypothesis is rejected at the 1%, 5%, 10% and 20% level, respectively.

In the case of testing for the presence of seasonal unit roots with respect to the modified series of foreign arrivals[28] (*LA*), the null hypothesis cannot be accepted at a general 5% level of significance[29], both performing the *t*-tests of the separate π's (except for π_1 and π_4 where the null hypothesis fails to be accepted at the 10%) and the joint test for π_3 and π_4. Such a variable can be considered as having a deterministic seasonal pattern and to be stationary in the level. The last property has been tested further by an ADF test (Table 4.25). One can conclude that the level of this series is stationary, as the null hypothesis of the presence of a unit root cannot be accepted at a

27 The auxiliary regression (2.6.1) is run using Microfit 4.0 package.

28 The variable is normalised for the average number of Saturdays in each quarter of year.

29 The critical values for the quarterly seasonal unit roots test are provided in Hylleberg *et al.* (1990) pp. 226-227. Note that in this case one is taking into consideration the critical values for T=96 when intercept, trend and seasonal dummies are included.

1% level. These findings confirm the results obtained when using monthly data, but differ from the ones using annual data.

For the income proxy (*LPR*), the null hypothesis of seasonal unit root fails to be accepted at the general 1% level. The unit root test for the long run frequency suggests such a variable to be I(0) upon a trend. The latter result has been confirmed from the ADF test when including the constant term and a time trend, and the constant, time trend and quarterly seasonals. This finding confirms the monthly case (Table 4.2).

The seasonal unit roots test has been run for the (log) substitute price (*LSP*) for the same period (1972:1-1995:4). From estimating the auxiliary regression (2.6.1), one can conclude that there appears to be no evidence for the presence of seasonal unit roots, denoting a regular seasonal pattern. In fact, as Hylleberg (1990) suggests there is no seasonal unit roots if either π_3 or π_4 are different from zero, which requires a joint test. As one can see from the *F*-test one cannot accept the null hypothesis of seasonal unit root at the 1% level. However, the null hypothesis of non-stationarity cannot be rejected at a 5% level of significance, testing π_1=0 using the *t*-test. However, from the ADF test (see Table 4.25), one can conclude that this variable is stationary in the level, as found in the monthly case.

For the exchange rate (*LER*) the presence of seasonal unit roots cannot be accepted at a general 5% level. Whereas the null hypothesis of a long run unit root fails to be rejected. The ADF test confirms that *LER* is stationary in the first difference.

The last variable to be investigated is the relative price (*LRP*). As far as the Hylleberg's seasonal unit roots test is concerned, such a variable appears to show a deterministic seasonal pattern. However, the null hypothesis for π_1 cannot be rejected. Also from the ADF test one can treat *LRP* as stationary in the first difference.

Table 4. 25 Augmented Dickey-Fuller Unit Root Test with Quarterly Data

Series	*ADF*(1)	*LAG*(2)
LA(c)	- 3.39*	2
LA(c,t)	- 4.17**	5
LA(c,s)	- 4.29**	0
LA(c,t,s)	- 7.51**	0
LPR(c)	- 0.64	1
DLPR(c)	- 6.37**	0
LPR(c,t)	- 3.62*	2
LPR(c,s)	- 0.63	1
DLPR(c,s)	- 6.25**	0
LPR(c,t,s)	- 3.56*	2
LSP(c)	- 3.29*	2
LSP(c,t)	- 1.20	2
DLSP(c,t)	- 6.39**	4
LSP(c,s)	- 3.41*	2
LSP(c,t,s)	- 0.97	4
DLSP(c,t,s)	- 6.71**	3
LER(c)	- 1.16	2
DLER(c)	- 7.74**	1
LER(c,t)	- 1.40	2
DLER(c,t)	- 7.77**	1
LER(c,s)	- 1.15	2
DLER(c,s)	- 7.33**	1
LER(c,t,s)	- 1.42	2
DLER(c,t,s)	- 7.36**	1
LRP(c)	- 2.09	4
DLRP(c)	- 3.10*	3
LRP(c,t)	- 0.62	4
DLRP(c,t)	- 3.88*	3
LRP(c,s)	- 2.05	4
DLRP(c,s)	- 3.17*	3
LRP(c,t,s)	- 0.44	4
DLRP(c,t,s)	- 4.00*	3

Notes: * and ** asterisks indicate that the unit root null hypothesis is rejected at the 5% and 1% level, respectively. The capital letter *D* denotes the first-difference operator defined, in a general notation, by $Dx_t = x_t - x_{t-1}$.

(1) Augmented Dickey-Fuller statistics with constant critical values = -2.893 at 5% and -3.503 at 1% level; with constant and trend c.v.= -3.458 at 5% and -4.059 at 1% level; with constant and seasonals c.v. = -2.894 at the 5% and -3.505 at 1%; with constant, trend and seasonals critical values = -3.46 at 5% and -4.062 at 1% level.

(2) Number of lags set to the first statistically significant lag, testing downward and upon white residuals. Note that ADF(0) corresponds to the Dickey-Fuller test.

The main finding is that the results both from the quarterly seasonal unit roots test and ADF test lead to the same results as using monthly data, and both differ from the results obtained using annual data. Thus, one can treat *LA*, *LPR*, *LSP* as stationary in the level, and *LRP* and *LER* as stationary in the first difference.

As for the annual and monthly data cases, the possible cointegration between the two I(1) variables is tested. An initial unrestricted VAR with $k=5$ is first run which presented problems of non-normality and heteroscedasticity in the equation for the relative price. Two impulse dummies are added to pick up possibly the negative effects of the first oil shock (*i.e. i1974q1* and *i1975q1*). The first dummy is constructed giving the value of 1 to the first quarter of 1974 and 0 otherwise; the same construction holds for the second dummy. The VAR(5) is re-estimated with these two dummies, a constant and three quarterly seasonal dummies[30]. This system still shows problems in terms of serial correlation and heteroscedasticity for the relative price equation. However, it can be considered as the best system achievable. The poor statistical performance of the system seems to confirm the results obtained when using monthly data and annual data where non-homoscedasticity appeared. Note also that, once again, the coefficient of determinations for the first equation (*LRP*) is 0.99962; and, for the second equation (*LER*) it equals 0.99208.

The Johansen cointegration test has given the results reported in Table 4.26. The test statistics suggest as the null hypothesis of the existence of one cointegrating vector cannot be rejected at the confidence level of 1%.

Table 4. 26 Johansen Test for the Number of Cointegrating Vectors using Quarterly Data

Ho	H_1	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
r=0	r=1	53.49**	47.61**	14.1	55.35**	30.32**	15.4
r=1	r=2	1.86	1.65	3.8	1.86	1.65	3.8

Notes:
(1) Adjusted by the degrees of freedom (see, Reimers, 1992)
(2) Critical values at a 5% level of confidence (see Osterward-Lenum, 1992).
* and ** denotes rejection of the null (*i.e.* non-cointegration) at a 5% and 1% level, respectively.

The results of the eigenvalue and eigenvector calculations are given in Table 4.27.

Table 4. 27 The Eigenvalues $\hat{\lambda}$, Eigenvectors $\hat{\beta}$, and the Weights α

Eigenvalues $\hat{\lambda}$			
(0.7867 0.0494)			
Standardized β' eigenvectors		Standardized $\hat{a}$ coefficients	
LRP	***LER***	***LRP*** -0.07	0.0006
1.00	-0.89	***LER*** -0.03	-0.0600
-0.62	1.00		

The equivalent cointegrating vector is the following:

$$CI = LRP \ -0.89\ LER \qquad (4.4.2)$$

[30] The restricted system fails to be accepted at the 1% level; the same conclusion rises from the information criteria.

```
system    T     p          log-likelihood          SC          HQ          AIC
    4    91    28  OLS          748.00142     -15.052     -15.513     -16.440
    5    91    32  OLS          758.11087     -15.076     -15.602     -16.662
System (5 lags)> System (4 lags): F(4,148)= 4.3475 [0.0024]**
```

The coefficient for the (log) weighted exchange rate has been tested for the following restriction: β=-*1* and such a restriction has been accepted at the 5% level from the χ^2 test[31]. In this way one models the following cointegrating vector:

$$CI = LRP - 1\ LER$$

as for the monthly case and annual case.

The coefficients of α have an economic interpretation in terms of the average speed of adjustment towards the estimated equilibrium state. Note that in this case the average speed of adjustment is approximately 0.05 in modulus (see Table 4.27). In this specific case the root equals 0.952 that indicates a relatively slow adjustment to the equilibrium state, given by the solution of:

$$(1 + \beta'\alpha) = 1 + \begin{bmatrix} 1 & -0.89340 \end{bmatrix} \begin{bmatrix} -0.073354 \\ -0.028233 \end{bmatrix}$$

In particular, as $0.952^4 = 0.82$ only 18% of the adjstment to equilibrium occurs in the first year, and 33% in the first two years. As already stated, this calculation is only an approximation, as the lags on ΔX_t (in equation 4.2.2.2) have not been included (see p.64).

To model the foreign arrivals of tourists (*LA*) in the north of Sardinia, the sample period from 1972:1 to 1995:4 is used. The explanatory variables included in the model are: the income proxy (*LPR*), the substitute price (*LSP*), the first difference of the relative price (*DLRP*) and of the exchange rate (*DLER*), the cointegrating vector (CI_{t-1})[32], the (log) weather variable, a time trend and finally 3 quarterly seasonal dummies. An impulse dummy, *i1985q1*, is also added after inspecting the residuals; this dummy may detect the positive effects produced by the upturn in the economic performance of the EEC countries which started in the second half of the Eighties.

---

31 The results for the restriction test on the coefficient is: $\chi^2(1) = 2.2162$ [0.1366]

32 Note that one could include in the model the first lag of the cointegrating vector and free lags of *DLRP* and *DLER*; either free lags of the cointegrating vector and *DLER*, or free lags of the cointegrating vector and *DLER*.

An initial five quarters lag structure model is estimated. According to the joint F-test and SC criterion it can be reduced to a four lags well-specified model[33]. The unrestricted model is simplified in order to obtain a parsimonious, yet congruent, data characterisation. The final model is reported in Table 4.28.

Table 4. 28 Final Static Model for the International Demand using Quarterly Data

```
EQ(1) Modelling LA by OLS  (using Quadfor.in7)
The present sample is:  1973 (2) to 1995 (4)
```

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-0.47008	2.4693	-0.190	0.8495	0.0004
LPR	0.78767	0.31862	2.472	0.0155	0.0702
LSP	1.4237	0.23681	6.012	0.0000	0.3085
DLRP_4	-4.2241	1.1144	-3.791	0.0003	0.1507
CI_1	-0.42889	0.21231	-2.020	0.0467	0.0480
LW	0.55060	0.26390	2.086	0.0401	0.0510
I1985Q1	0.85333	0.16636	5.130	0.0000	0.2452
Seasonal	-0.64835	0.096951	-6.687	0.0000	0.3557
Seasonal_1	1.4939	0.074334	20.097	0.0000	0.8329
Seasonal_2	1.9135	0.15066	12.701	0.0000	0.6657

```
R^2 = 0.985802  F(9,81) = 624.87 [0.0000] sigma = 0.154922
DW = 1.74  RSS = 1.944063711 for 10 variables and 91 observations

AR 1- 7 F( 7, 74) =      0.88001 [0.5265]
ARCH 7  F( 7, 67) =      0.81989 [0.5742]
Normality Chi^2(2)=      0.48583 [0.7843]
Xi^2    F(14, 66) =      0.70677 [0.7598]
Xi*Xj   F(39, 41) =      0.84193 [0.7044]
RESET   F( 1, 80) =      0.23343 [0.6303]
Tests of parameter constancy over: 1995 (3) to 1995 (4)
Forecast Chi^2( 2)=       2.7672 [0.2507]
Chow    F( 2, 79) =       1.2475 [0.2928]
```

As in the annual case, one has arrived at a static model which manages to explain almost 99% of the variation in the number of foreign arrivals. Moreover, the ratio between *SER* and *MDV* equals 0.019363[34] which can be considered as satisfactory. The diagnostic statistics suggest no problems. In addition, the same model is re-estimated using 1995(3) to 1995(4) as forecasting sample data; the χ^2 prediction test statistic and the Chow prediction test statistic do not reject the null hypothesis of parameter constancy. As expected, the coefficient on the income proxy, *LPR*, is positive indicating that, other things being equal, the higher the income of the clients' countries the higher the demand for more trips. The substitute price (*LSP*) presents an anomalous positive sign as found in the other previous cases. It is also observed that an

33 The SC information criterion and the joint F-test.

dep.var		T	k	df	RSS	sigma	Schwarz
4: LA	OLS	90	32	58	1.42822	0.156922	-2.54345
5: LA	OLS	90	37	53	1.24058	0.152994	-2.43431

Model (5 lags) --> Model (4 lags): F(5, 53) = 1.6033 [0.1753]

34 SER/MDV = (0.154922/8.0011) = 0.019363

increase in the prices in Sassari Province, holding constant the prices in the origin countries, decreases the demand of tourism with a quite strong impact, with a negative elasticity of 4.22. However, the coefficient for *DLRP* shows a rather high standard error suggesting an imprecise estimation. The coefficient of the cointegrating vector (*CI*) shows a negative sign. This denotes that if the *CI* increases, by deviations either of the relative price or the exchange rate from the respective long run relations, the foreign demand for tourism decreases in the long run. The same result was obtained in the annual and monthly models. Turning to the climate variable, describing quarterly averaged temperatures, it turns out to be statistically significant. It appears that in the long run this variable has a positive impact on the international demand for tourism; whereas, in the monthly model case such a variable does not have any particular influence (see Table 4.7). Finally, the three seasonal dummy variables demonstrate that the foreign demand for tourism is rather highly influenced by seasonal factors, including statutory or religious holidays such as Christmas.

As for the monthly and annual models, a brief note on the use of the log-linear form is due. As a first step, one proceeds by testing the integration status of the variables expressed in a linear specification, that is: *A* (number of arrivals modified for the average number of Saturdays in a year), *PR* (income proxy, as a weighted average for source countries), *RP* (relative price, Sassari/origin countries), *SP* (substitute price, Sassari/other destinations) and *ER* (weighted average exchange rate for origin countries). Running the ADF test with an initial 5 lags, one infers the following results: *A*, *PR* and *SP* are I(0), and *RP* and *ER* are I(1). These results confirm those obtained in the linear monthly case.

Hence, a Johansen cointegration analysis is run for the relative price and exchange rate. An initial unrestricted $k=5$ VAR, which includes a constant, quarterly seasonals and a trend unrestrictedly, can be reduced to a two lag system. Statistical evidence is found for the existence of the following cointegrating vector:

$$ECL = RP - 0.0022679\ ER$$

In order to run the Box and Cox test, one runs an unrestricted 5 lag quarterly tourism demand equation expressed both in a logarithmic and linear form. The independent variables are defined as before.

1) logarithmic form

$$LA_t = a_1 + a_2\, LA_{t-1}\ldots + a_3\, LPR_t\ldots + a_4\, LSP_t\ldots + a_5\, DLER_t\ldots + a_6\, DLRP_t\ldots + a_7\, CI_{t-1} + a_8\, LW + a_9\, Trend + a_{10}\, Seas + e_t$$

and

2) linear form

$$A_t = a_1 + a_2\, A_{t-1}\ldots + a_3\, PR_t\ldots + a_4\, SP_t\ldots + a_5\, DER_t\ldots + a_6\, DRP_t\ldots + a_7\, ECL_{t-1} + a_8\, W + a_9\, Trend + a_{10}\, Seas + e_t$$

The *SSELL* from the logarithmic form equal 1.613690187, whereas the *SSEL* for the linear form equals 46091983.21. The aim is to test whether the null hypothesis that the two models are empirically equivalent and find out which of the two models fits the data better. Following the Box and Cox test procedure (see Section 2.4.1), the calculated χ^2 equals 53.03 that is greater than the tabulated critical value, 3.84, at the 5% level; hence, the two models are empirically different. Moreover, one infers that the logarithmic specification is "much better" than the linear specification as the *SSELL* (*i.e.* 1.613690187) is less than $SSEL/(\overline{A}\, G)^2$ (*i.e.* 5.243292).

4.5 SUMMARY

In this section the main economic findings in terms of income and price elasticities are reported, considering both the short and long run behaviour. Particular emphasis will be given to the main differences in using the three different data frequencies. Table 4.29 summarises the findings.

Table 4. 29 Short Run and Long Run Elasticities for the International Demand of Tourism

Elasticities	Monthly Model (288 obs.) (Tables 4.7 - 4.8)	Annual Model (24 obs.) (Table 4.12)	Quarterly Model (96 obs.) (Table 4.28)
INCOME (long run)	1.06 (2.62)	2.34 (3.68)	0.79 (2.47)
INCOME (short run)	2.56 (3.92)	=	=
REL.PRICE (long run)	-	-	- 4.22 (-3.79)
REL.PRICE (short run)	-	-	=
EX. RATE (long run)	-	-	-
EX.RATE (short run)	-	-	-
CI (long run)	- 0.64 (-2.11)	- 0.73 (-2.70)	- 0.43 (-2.02)
CI (short run)	- 0.34 (-2.07)	=	=
SUB.PRICE(long run)	1.29 (4.27)	1.55 (4.44)	1.42 (6.01)
SUB.PRICE(short run)	5.21 (2.36)	=	=

Notes: (1) *t*-values are given in parenthesis.
(2) For the annual and quarterly model long run elasticities equal short run elasticities as dealing with a final static model.
(3) Note that the short run elasticity corresponds to the first significant lag in the model (see Pindyck and Rubinfeld, p. 377, 1991).

The long run income elasticity shows different values with respect to the data frequency which has been used. In the annual model, the high income elasticity value indicates that foreign tourists hold strong preferences for Sardinian tourism. However, the monthly model shows a value just above unity, which indicates no strong evidence for the previous hypothesis. According to the quarterly data model, the relatively low income elasticity seems to indicate that Sardinian tourism needs some changes in order to attract higher number of foreign tourists. The differences in the magnitude of the elasticities are also likely to reflect different types of behaviour. Consumers' decisions are likely to be taken either on a yearly basis, at the last minute or somewhere in between. This fact has been confirmed by a recent survey by Blackwood & Partners (1994). The foreign respondents assess when they took the decision to spend their holidays in Sardinia: the January of the same year or the June of the same year were common responses. On balance, one considers monthly data to be the appropriate frequency for tourism decisions. This frequency, in fact, can give more insight in the differences existing amongst consumers and their preferences.

Some comparisons might be of interest. One can compare the annual model value with the figures obtained in other empirical studies for Italy. Malacarni (1991), for example, finds an income elasticity of 1.49 in estimating the aggregated international demand for Italy, where 17 observations have been employed. Clauser (1991), in a disaggregated study for international demand in Italy by main origin countries (20 observations), finds a value in the range within 0.55 for Holland and 2.42 for Japan. Witt and Witt (1992), using 16 observations in total, found values of 1.23 for Germany and 2.57 for France. Note that in all these studies the number of tourist arrivals have been used as the dependent variable. However, a comparison with other empirical studies is difficult. The income elasticities and, in general, the explanatory power of the other independent variables are highly dependent on other elements such as the level of aggregation, the time periods and the measure of demand used in each empirical study. As Sinclair and Stabler (1997) also note, one of the main problems is related with elasticity inferences obtained from models which have not included a full range of statistical tests. For example, problems of heteroscedasticity, as incurred in this study, are ignored in the majority of the cases.

From the quarterly data model one infers that international tourism demand is highly negative dependent on the growth in the relative price. This fact may suggest a high degree of substitutability of Sardinian tourism for the source countries. As a reminder, Malacarni (1991) finds a price elasticity of –0.83 for the international demand of tourism in Italy. Again, one notes that a comparison is difficult. Firstly, an annual model is estimated rather than a quarterly model as in this case. Secondly, this study concerns Sassari Province rather than Italy as a whole.

Note also that, in general, the short run price changes were found not to play any important role in explaining the foreign demand for tourism. The same conclusion has been reached when considering the exchange rate. However, the cointegrating vector appears to be statistically significant in each of three models.

Contrasting results appear for the substitute price elasticities, which present a positive sign. As already pointed out, this result might be indicating bias in not having taken into consideration other explanatory variables such as the exchange rate for the main north of Sardinia competitors. The latter hypothesis will be investigated in Chapter 6.

4.6 CONCLUSION

In this study an empirical investigation of the international tourism demand to the north of Sardinia for the period between 1972 and 1995 has been presented. Different concepts have been utilised such as: seasonality, non-stationary, cointegration and Equilibrium Correction Models. Particular determinants of the demand for tourism, *i.e.* the relative prices and the exchange rate, are I(1) and cointegrated.

In this study one makes use of different data frequencies. Monthly, annual and quarterly data have been used in order to assess the characteristics of the demand for tourism in the short run as well as in the long run. The relatively large number of observations available in this study ($T=288$), when using monthly data, has allowed to test the possible presence of seasonal unit roots as well as long run unit roots. One can notice that monthly and quarterly series have given homogenous results in terms of seasonal and long run unit roots testing. Whereas, annual data have shown different and perhaps misleading results. One of the main problems when dealing with tourism annual data is the relative short number of observations available (24 observations in this case). Nevertheless, testing for cointegration has revealed similar findings using any of the three frequencies.

There are numerous advantages in using monthly data. Firstly, they reveal the short term characteristics of the demand for tourism as well as the long run dynamics. This separation is of particular importance in tourism since consumers' decisions are likely to be taken several months in advance or sometimes at the last minute in response of "special offers".

Secondly, one can study carefully the seasonal pattern that seems to be of extreme importance for the operators in the tourism sector. In particular, foreign tourists seem to appreciate less crowded and cheaper holidays, and months in which the weather temperatures are milder. In this respect the island of Sardinia represents an appealing destination in the off-season months (May, June, July and September). Furthermore, the empirical analysis reveals the particular importance of the Easter holiday in explaining the pattern of tourism. However, the weather conditions, given by the average temperatures in Sassari Province, seem not to have any particular effects in determining international demand.

One can conclude that the monthly data model is rather satisfactory in terms of coefficient of determination (98% of the variance of the dependent variable is explained), diagnostic statistics (with the exception for heteroscedasticity detected at the 5% level), coefficients significance and signs. Moreover, the long run responses show an overall good specification in terms of low standard errors and statistical significance.

The best results have been obtained by taking into account the effects of four or five weekends (*i.e.* the number of Saturdays) in each month with respect to the foreign arrivals. In this way, it has been possible to correct the presence of non-normality and heteroscedasticity (at the 1% level) that have arisen when using the unadjusted series (see Appendix A).

In the annual model, some problems of mis-specification in the functional form have appeared. Overall, the final static model shows a worse performance than the monthly model, denoted also by the 91% of explanatory power shown by the coefficient of determination. An advantage from using annual data is the possibility of using data only available with annual frequency, in this case supply components. In this way one can test problems of simultaneity for determinants such as the number of international flight arrivals.

The use of quarterly data determines a final well-specified model. There have been no problems in terms of diagnostic tests. The coefficients are statistically significant and in general present the correct signs. In this case, the final choice of a static model might suggest that adjustment to equilibrium is quite rapid.

Finally, the Box and Cox (1964) test has been applied to determine whether the logarithmic form is appropriate. The three models have given statistical proof that the log-linear specification fits the data better. Note also that the ADF test for the economic series in a linear specification has shown to be more robust in the monthly and quarterly models. Moreover, by using the Johansen analysis, a cointegrating relationship has been suggested for *LRP* (relative price) and *LER* (weighted exchange rate) using each of the three time frequencies. Hence, in the long run, the use of effective exchange rates has been validated statistically.

CHAPTER 5.

THE DOMESTIC DEMAND FOR TOURISM IN THE NORTH OF SARDINIA

Aim of the Chapter:

To model and estimate the domestic demand for tourism in the Sardinian Sassari Province, Italy.

5.1 INTRODUCTION

In this chapter an account is given of the main findings from the analysis of the demand for tourism by domestic clients in Sassari Province. As pointed out in Chapter 3, various characteristics distinguish domestic from foreign flows of tourism for the north of Sardinia. Amongst others the seasonal pattern (see Figures 3.2 and 3.3) shows noticeably different behaviour for the two types of clients.

By making use of Franses' seasonal unit roots test, one will find out possible non-stationarity in the domestic seasonal pattern. From a deeper investigation, structural breaks will be identified in the seasonal pattern.

In the following analysis, further investigation is carried out to assess the possible validity of the correction of the dependent variable (*i.e.* the number of domestic arrivals) for the number of weekends in a month. As already highlighted in Chapter 3, several experiments are carried out for the raw series of domestic arrivals, as well as for the series adjusted for either the number of Saturdays or Sundays in a month.

Relationships between short and long run income and price elasticities will also be explored, and a comparison with other empirical findings will be given. Different data frequencies models will be estimated. A full range of test statistics, such as serial correlation, heteroscedasticity and specification form will be included. In the majority of tourism empirical studies these tests have been ignored (see Sinclair, 1998). The dynamics of the models will be investigated by including lagged dependent variable in accordance with joint F-test statistics as well as information criteria and lagged independent variables that are rarely taken into consideration in many empirical studies for tourism.

The structure of this chapter is the following. Section 5.2 will be dedicated to the analysis and estimation of the domestic demand for tourism in the north of Sardinia, when using monthly data. In the subsections, the analysis will be concentrated on the following issues: a seasonal unit roots test and an ADF test will be run on the variables of interest; possible structural breaks will be investigated in the domestic seasonal pattern; the specification of a monthly model will be included, taking into account the short and long run dynamics. In Section 5.3, an annual data model will be estimated. The subsections will include the following: a test and establishment that the Italian production index can be considered as a valid proxy of the Italian personal disposable income; estimation of a domestic demand of tourism using annual data; supply variables such as the number of domestic flights and the number of domestic boat arrivals in Sassari Province will be considered in order to test for simultaneity. Section 5.4 will be dedicated to the estimation of a model when using quarterly data. A summary in terms of economic findings and conclusions are given in the last two sections.

5.2 DOMESTIC DEMAND FOR TOURISM USING MONTHLY DATA

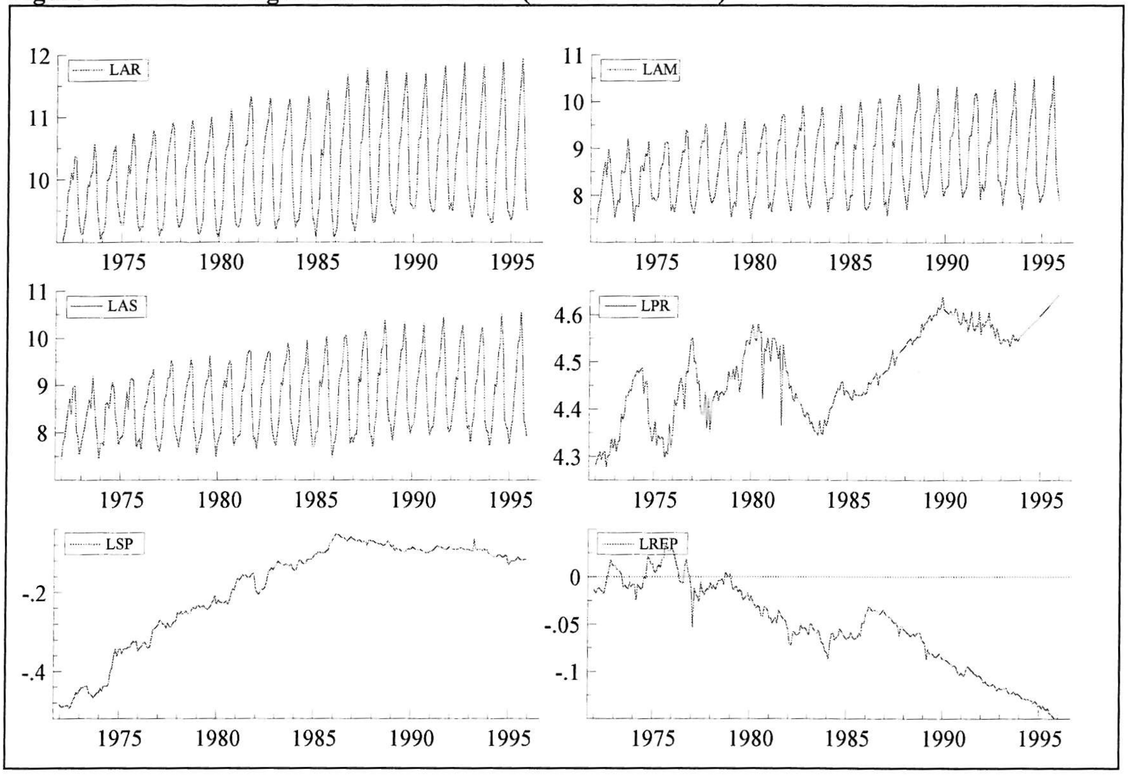
5.2.1 Seasonal Unit Roots Testing

In this thesis, the method of Franses (1991a and 1991b) to test for seasonal unit roots when dealing with monthly data is used (see Section 2.6, Chapter 2). Such a test is applied to the period from January 1972 until December 1995 to five series, *i.e.*: the (log) raw series of domestic arrivals[35] (*LAR*), the (log) modified series corrected for number of either Saturdays or Sundays in a month (*LAM* and *LAS*); the (log) index of industrial production (1990=100) in Italy[36] (*LPR*) as a proxy of income, as discussed later; a (log) relative price (1990=100) (*LREP*), defined as the difference between the (log) consumer price index in Sassari Province and the (log) consumer price index in

[35] The number of domestic tourist arrivals in all registered accommodation are collected by the EPT of Sassari.
[36] Source ISTAT.

Italy, as a proxy for the competitive price index for goods and services of tourism, and, finally, the (log) substitution price index (*LSP*)[37]. Graphs of each series are provided in Figure 5.1. One can notice that the income proxy (*LPR*) appears much volatile than one would expect income to be. Further work might involve the use of a moving average in order to make an attempt to smooth out the series.

Figure 5. 1 Natural Logarithms of the Series (1972:1 - 1995:12)



Equation (2.6.2) (see Section 2.6, Chapter 2) is fitted by OLS for each of the five series defined above. Note that in this case μ_t, which represents the deterministic part, includes a constant, a trend and 11 seasonal dummies. As Beaulieu and Miron (1993) have pointed out, "the loss of power that results from including seasonal dummies when unnecessary is insignificant compared to the bias that results from their omission when necessary" (p.318). The test results[38] are displayed in Table 5.1.

---

37 The nominal substitute price is defined as in Appendix B in formula (B.6), however, the weights α_i are defined as the quota of Italians (number of arrivals) choosing to spend their holidays in France, Greece, Portugal and Spain, respectively. The weights are allowed to change annually.

38 The critical values for the seasonal unit roots test are provided in Franses (1991) pp.161-165.

Table 5. 1 Testing for Seasonal Unit Roots

t-statistics	Variable					
	LAR	***LAM***	***LAS***	***LPR***	***LREP***	***LSP***
$\pi 1$	-1.008	-1.303	-1.098	-3.474 *	-2.164	-0.756
$\pi 2$	-3.407 ***	-4.466 ***	-3.995 ***	-4.318 ***	-4.264 ***	-4.657 ***
$\pi 3$	-1.467	1.115	-0.584	-3.012 ***	-2.895 ***	-4.803 ***
$\pi 4$	-4.178 ***	-5.683 ***	-4.423 ***	-6.020 ***	-6.567 ***	-5.336 ***
$\pi 5$	-6.120 ***	-5.516 ***	-4.309 ***	-5.720 ***	-6.637 ***	-6.613 ***
$\pi 6$	-6.679 ***	-6.098 ***	-5.727 ***	-5.688 ***	-7.020 ***	-6.768 ***
$\pi 7$	1.125	1.595	1.343	-0.343 **	0.097	-3.551 ***
$\pi 8$	-2.096	-2.335	-1.970	-3.151	-3.284 *	-0.405
$\pi 9$	-2.826 *	1.462	2.725	-4.374 ***	-5.489 ***	-6.045 ***
$\pi 10$	-6.155 ***	-4.169 ***	-3.720 **	-7.243 ***	-5.876 ***	-6.376 ***
$\pi 11$	0.437	2.378	1.491	-0.124	-2.700 ***	-1.850 ***
$\pi 12$	-2.965	-4.273 ***	-3.673 ***	-6.452 ***	-4.117 ***	-4.446 ***
F-statistics	***LAR***	***LAM***	***LAS***	***LPR***	***LREP***	***LSP***
$\pi 3, \pi 4$	9.937 ***	16.947 ***	9.980 ***	24.056 ***	27.387 ***	28.716 ***
$\pi 5, \pi 6$	22.590 ***	18.732 ***	17.608 ***	17.606 ***	25.436 ***	24.307 ***
$\pi 7, \pi 8$	3.187	3.104	2.214	33.028 ***	27.007 ***	19.994 ***
$\pi 9, \pi 10$	18.994 ***	19.116 ***	24.158 ***	26.776 ***	22.489 ***	27.049 ***
$\pi 11, \pi 12$	5.260 *	9.161 ***	6.848 **	32.454 ***	27.658 ***	24.246 ***
$\pi 3, \ldots\ldots, \pi 12$	13.454 ***	15.577 ***	14.401 ***	131.288 ***	106.458 ***	128.208 ***

Notes: One, two, and three asterisks indicate that the seasonal unit roots hypothesis is rejected at the 10%, 5% and 1% level, respectively.

As far as the series of *LPR*, *LREP* and *LSP* are concerned, the parameters in the auxiliary regression are significant in general at the 1% level, when one performs the *t*-test of the separate π's, with just a few exceptions, and the *F*-test of the pairs of π's as well as the joint *F*-test of $\pi_3=..=\pi_{12}=0$. Thus, there appears to be no noteworthy evidence for the presence of seasonal unit roots. The null hypothesis of a long run unit root in the *LPR* case cannot be rejected at the 5% level, but can at the 10% level. For the *LREP* and *LSP* cases, the null hypothesis of non-stationarity cannot be rejected at the 5% level. However, as Franses (1991b) points out "simulation evidence shows that the power of the test statistics may be low,...., and hence that significance levels of 10%, or even higher, may be more appropriate" (p.205). Therefore, the last findings are investigated further by performing an ADF test (see Table 5.2).

As far as the raw and the modified series of domestic arrivals are concerned (*i.e.* *LAR*, *LAM* and *LAS*), the null hypothesis of the presence of seasonal unit roots cannot in general be accepted at a 5% level of significance, both performing the *t*-test of the separate π's and the *F*-test of the pairs of π's as well as the joint *F*-test of $\pi_3=..=\pi_{12}=0$. However, some exceptions can be noticed from Table 5.1. The null hypothesis, in fact, cannot be rejected for some of the separate π's (*i.e.*: π_1, π_3, π_7, π_8, π_{11} and π_{12} for *lar*; π_1, π_3, π_7, π_8, π_9 and π_{11} for *LAM* and *LAS*) as well for one pair of π's, *i.e.* π_7, π_8 (*i.e.* at frequency $\pi/3$). In conclusion, for the unadjusted and adjusted series of domestic arrivals one rejects unit roots at most frequencies; thus, in accordance with the findings in Franses (1991b), and Beaulieu and Miron (1993), it appears that there is no strong evidence for the presence of seasonal unit roots. However, as Webb (1995) notices "other types of nonstationarity are also possible. An alternative...involves large, infrequent shocks" (p.277). The possibility of regime changes in the seasonal pattern is investigated in more detail in the next section. A further investigation is done in testing for the presence of a unit root at the zero frequency by running an ADF test. The results are reported in Table 5.2.

Table 5. 2 Augmented Dickey-Fuller Unit Root Test

Series	*ADF*(1)	*LAG*(2)
LAR(c)	- 4.39**	8
LAR(c,t)	- 5.48**	9
LAR(c,s)	- 9.28**	3
LAR(c,t,s)	- 4.31**	0
LAM(c)	- 3.47**	9
LAM(c,t)	- 3.63*	10
LAM(c,s)	- 4.28**	4
LAM(c,t,s)	- 4.07**	8
LAS(c)	- 3.43*	9
LAS(c,t)	- 3.75*	10
LAS(c,s)	- 4.19**	3
LAS(c,t,s)	- 3.84*	8
LPR(c)	- 2.95*	0
LPR(c,t)	- 3.62*	10
LPR(c,s)	- 2.19	5
DLPR(c,s)	- 4.67**	12
LPR(c,t,s)	- 3.96*	0
LSP(c)	- 3.06*	8
LSP(c,t)	- 0.80	8
DLSP(c,t)	- 7.19**	7
LSP(c,s)	- 3.28*	12
LSP(c,t,s)	- 0.52	12
DLSP(c,t,s)	- 6.34**	11
LREP(c)	- 0.03	1
DLREP(c)	- 9.18**	4
LREP(c,t)	- 2.33	12
DLREP(c,t)	- 9.20**	4
LREP(c,s)	- 0.01	5
DLREP(c,s)	- 8.73**	4
LREP(c,t,s)	- 2.01	4
DLREP(c,t,s)	- 8.75**	4

Notes: The one and two asterisks indicate that the unit root null hypothesis is rejected at the 5% and 1% level, respectively. The capital letter *D* denotes the first-difference operator defined, in a general notation, by $Dx_t = x_t - x_{t-1}$.

(1) Augmented Dickey-Fuller statistics with constant (*c*) critical values = -2.872 at 5% and -3.455 at 1% level; with constant and trend (*c, t*) c.v.= -3.428 at 5% and -3.995 at 1% level; with constant and seasonals (*i.e. c, s*) c.v. = -2.872 at 5% and -3.456 at 1% level; with constant, trend and seasonals (*i.e. c, t, s*) c.v. = -3.428 at 5% and -3.995 at 1% level;

(2) Number of lags set to the first statistically significant lag, testing downward and upon white residuals. Note that ADF(0) corresponds to the Dickey-Fuller test.

From the ADF test, one concludes that the raw and adjusted series of domestic arrivals (*i.e. LAR, LAM, LAS*), the level of the index of industrial production and the substitute price (*i.e. LPR* and *LSP*) are stationary, as one cannot accept the null hypothesis at the 1 and 5 percent levels, respectively. Furthermore, the relative price (*LREP*) is integrated of order one.

Experiments have also shown that the substitute price, *LSP*, can be considered stationary in the level when either a constant or a constant and seasonals are included.

On the other hand, such a variable is I(1) when either a constant and a time trend, or a constant, a trend and seasonals are included. The analysis will carry on treating *LSP* as I(0). It can be seen in Figure 5.1, the data show an adjustment to a stable situation. As the competitors included are EU (France, Greece, Spain and Portugal), it is difficult to accept long run non-stationarity in this variable.

5.2.2 The Model And Possible Regime Changes

Before investigating the possible existence of regime changes, a discussion of the explanatory variables included in the model is required. The three explanatory variables (*i.e. LPR, LREP* and *LSP*), for the period from January 1972 up to December 1995, show relatively high and negative contemporaneous cross-correlation coefficients for the pair *LREP* and *LSP*, and the pair *LPR* and *LREP*. Whereas, the pair *LPR* and *LSP* show a positive contemporaneous cross-correlation coefficient. In particular, the cross-correlation coefficients are the following: r(*LREP,LSP*)=-0.72, r(*LPR,LREP*)=-0.70 and r(*LPR,LSP*)=0.63. However, these values do not cause problems of multicollinearity.

Two other explanatory variables are included in the model. An "Easter" dummy (*E*) is included so as to take into account the possible "Easter effect". An experiment has been carried out for the "Easter" dummy, considering either the Thursday or Friday before Easter as the starting day of the holiday instead of the Saturday. The results have shown a better specification in taking the Saturday as the starting day of the holidays in domestic flows of tourism, as for the international analysis case. The (log) weather variable has also been included as monthly average temperature in degrees Celsius, recorded at the weather station in Sassari (*LW*).

The initial formulation of the equation for the total domestic arrivals can be expressed as:

$$LARRIVALS = f(LPR,\ LSP,\ DLREP,\ LW,\ E) \qquad (5.2.2.1)$$

Further investigations are needed in assessing the validity of a possible correction of the dependent variable for the number of Saturdays or Sundays in a month. The first step consists of running three different VARs where the first equation has either *LAR*, *LAM* or *LAS* as the dependent variable. The use of a VAR gives the possibility to better identify the lag size of the system. The first system for *LAR* includes a constant, 11 seasonal dummies, three impulse dummies (*i.e. i1974p11*, *i1992p3* and *i1993p3*) created after an inspection of the residuals in order to correct the presence of non-normality, the "Easter" dummy and the (log) weather variable (treated unrestrictedly)[39], plus 13 lags for each of the other explanatory variables and the dependent variable (treated as endogeneous). A system with 13 lags versus 12 lags is initially estimated. The period under study being from January 1972:1 to 1995:12. According to the joint F-test the restricted system with 12 lags cannot be accepted at the 1% level[40]. Therefore, a 13 lags system can be used. From Table 5.3, the diagnostic statistics show a good specification. The correlation of the actual and fitted values suggests that the equation explains the 99.4% of the variance of the dependent variable. No problems appear in terms of diagnostic tests.

Table 5. 3 Statistical Tests of the Equation for the Unadjusted Series of Domestic Arrivals (*LAR*)

σ = 0.0949295 RSS = 1.847379485		
correlation of actual and fitted		
LAR	0.99437	
LAR	:Portmanteau 12 lags=	9.8167
LAR	:AR 1- 7 F(7,198) =	1.009 [0.4261]
LAR	:Normality Chi^2(2)=	2.880 [0.2369]
LAR	:ARCH 7 F(7,191) =	0.320 [0.9443]
LAR	:Xi^2 F(104,100) =	0.553 [0.9985]

39 An experiment of including 13 lags of *LW*, as of the other explanatory variables has led to mis-specification problems. The final parsimonious model, for both the adjusted and unadjusted series of domestic arrivals, after having carefully dummied out the seasonals, has shown a worse fit and problems of functional specification. The same results have been observed when including a time trend in the models.

40 The reduction test, as given by PcFiml 9.0, is the following:

system	T	p		log-likelihood	SC	HQ	AIC
11	274	244	OLS	574.5094	-28.392	-30.318	-32.391
12	274	260	OLS	4669.4087	*-28.757*	*-30.809*	*-33.083*
13	274	276	OLS	4689.7316	-28.578	-30.756	-32.232
System 13 lags-> System 12 lags : F(16, 617) = 2.3270 [0.0024] **							

the restriction for twelve lags cannot be accepted at the 1% level by the joint F-test. On the other hand, the SC, HQ and AIC criteria suggest for a further parameter reduction. However, the most unrestricted system is chosen as using monthly data.

A second system for the modified series of domestic arrivals for the number of Saturdays (*LAM*) in a month is carried out. Such a system includes a constant, 11 seasonal dummies, an impulse dummy (*i1974p8*) in order to correct problems of non-normality in the residuals, the "Easter" dummy, the (log) weather variable and 13 lags[41] for each of the other explanatory variables and the dependent variable. The main statistical tests are reported in Table 5.4.

Table 5. 4 Statistical Tests of the Equation for the Modified Series of Domestic Arrivals for Number of Saturdays (*LAM*)

$\sigma = 0.140413$ $RSS = 4.081164656$			
correlation of actual and fitted			
LAM	0.98765		
LAM	:Portmanteau 12 lags=	20.836	
LAM	:AR 1- 7 F(7,200) =	3.650	[0.0010] **
LAM	:Normality Chi^2(2)=	1.434	[0.4881]
LAM	:ARCH 7 F(7,193)=	1.546	[0.1539]
LAM	:Xi^2 F(104,102) =	0.541	[0.9990]

From these statistical tests, the presence of serial correlation can be detected at the 1% level. Furthermore, the residual sum of squares (*RSS*) and the standard error show a greater value than the values for the unadjusted series. Equally, the value of the correlation of the actual and fitted values is smaller than the value in the equation for *LAR*, denoting a worse fit.

41 The test of system reductions are the following:

system	T	p		log-likelihood	SC	HQ	AIC
1	274	76	OLS	4254.5454	*-29.498*	-30.098	-31.055
2	274	92	OLS	4284.6632	-29.390	*-30.116*	-31.275
3	274	108	OLS	4302.4016	-29.192	-30.044	-31.404
...	....		......	......	......	......	
11	274	236	OLS	4462.4629	-27.738	-29.601	-31.573
12	274	252	OLS	4495.9475	-27.655	-29.644	-31.817
13	274	268	OLS	4518.3794	-27.491	-29.606	*-31.981*
System (13 lags) -> System (12 lags) : F(16, 623) = 2.1468 [0.0058] **							

the restriction for twelve lags cannot be accepted at the 1% level. Note that the SC and HQ criteria suggest for further reductions, and the AIC criterion is minimised for at least a 13 lag VAR.

The last system has been estimated taking into consideration the equation for the adjusted series of domestic arrivals for the number of Sundays (*LAS*) in a month. This system includes a constant, 11 seasonal dummies, two impulse dummies (*i.e. i1987p3* and *i1992p3*) created after inspecting the residuals in order to correct non-normality problems, the "Easter" dummy, the (log) weather variable and 13 lags[42] for the dependent and the other explanatory variables (*i.e. LPR*, *DLREP*, *LSP*). The statistical tests are given in Table 5.5.

Table 5. 5 Statistical Tests of the Equation for the Adjusted Series of Domestic Arrivals for Number of Sundays (*LAS*)

σ= 0.134399 *RSS* =3.721011797		
correlation of actual and fitted		
LAS	0.98869	
LAS	:Portmanteau 12 lags=	23.684
LAS	:AR 1- 7 F(7,199) =	3.169 [0.0034] **
LAS	:Normality Chi^2(2)=	1.6228 [0.4442]
LAS	:ARCH 7 F(7,192) =	0.745 [0.6339]
LAS	:Xi^2 F(104,101) =	0.568 [0.9977]

In terms of diagnostic tests, except for the presence of serial correlation at the 1% level, the equation satisfies the conditions of normality, conditional homoscedasticity and non-heteroscedasticity. The RSS and the standard error present smaller values than for the case in which the dependent variable has been corrected for the number of Saturdays in a month. However, these values denote a worse fit than for the unadjusted case.

The first finding is that one can consider the unadjusted series of domestic arrivals, *LAR*, and the adjusted series for the number of Sundays in a month, *LAS*, as the best specifications, and drop the series of arrivals normalised for number of Saturdays, *LAM*.

42 The test of system reductions are the following:

system	T	p		log-likelihood	SC	HQ	AIC
1	274	80	OLS	4247.3772	*-29.364*	-29.995	-31.003
2	274	96	OLS	4278.7413	-29.265	*-30.023*	-31.232
3	274	112	OLS	4294.2259	-29.050	-29.934	-31.345
....			.....	....	...		
10	274	224	OLS	4414.9237	-27.637	-29.405	-31.226
11	274	240	OLS	4443.8093	-27.520	-29.414	-31.437
12	274	256	OLS	4497.0521	-27.581	-29.602	-31.825
13	274	272	OLS	4520.7031	-27.426	-29.573	*-31.998*
System (13 lags) > System (12 lags): F(16, 620) = 2.2557 [0.0034] **							

the AIC criterion suggests to run at least a VAR(13). From the joint *F*-test the restriction for twelve lags cannot be accepted at the 1% level. However, the SC and HQ criteria suggest for a further reduction. Again, the most unrestricted system has been chosen as dealing with monthly data.

The next step is to investigate the possible existence of structural breaks, given the results obtained in Table 5.1. The existence of seasonal unit roots at some frequencies could be thought of as a symptom of non-stationarity which might be due to structural breaks. The first structural break test is carried out for the unadjusted series of domestic arrivals (*LAR*). Preliminary investigations of a structural break in all coefficients (*i.e.* the coefficients of the seasonal dummies, *LAR, LPR, DLREP* and *LSP* respectively) in the unrestricted model (see Table 5.6) do not clearly show the presence of coefficient changes. Running a Chow test (1967), the conventional F statistic indicates the presence of structural changes[43]. In Appendix E (Table E.1) the program for running the Chow structural break test for 64 restrictions is given[44]. The F statistic, in fact, is larger than the critical value of the F distribution with q=64 and $N+M-2K$=141 degrees of freedom, that is larger than 1.32 at the 5% level. In particular, the Chow test suggests that the main change has occurred between 1984/85, since the appropriate F statistic show the greatest value. However, given the multiple comparisons involved, one should use Andrews' (1993, p.840) critical values. Andrews' table do not go far enough for this case, but it appears the 5% critical value is around 2.05, that is the critical value obtained when Andrews' π=0.25[45] and 20 restrictions are considered. Hence, from the calculated values in Table 5.6, it appears that the null hypothesis of no structural change between the period 1984/85 cannot be rejected at the 5% level. A further investigation seems to be needed.

43 Note that the possible existence of a structural change is detected by moving the change point forward one year at a time.

44 The program is created with TSP Version 4.3A.

45 Note that π is given by the following formula: $\frac{1}{2}[OB/N]$, where OB is the total number of omitted observations from each period sides (130 in this case) and N is the total numer of observations (274 in this case).

Table 5. 6 *LAR*- RSS for Unrestricted and Restricted Models and *F* Statistic

73M3 95M12	**RRSS** =	1.84742			
		UNRSS			
73M3 78M12		1.17446	F(64, 141)	=	1.26
73M3 79M12		1.19130	F(64, 141)	=	1.21
73M3 80M12		1.22493	F(64, 141)	=	1.12
73M3 81M12		1.21965	F(64, 141)	=	1.33
73M3 82M12		1.19265	F(64, 141)	=	1.21
73M3 83M12		1.10303	F(64, 141)	=	1.49
73M3 84M12		1.05183	F(64, 141)	=	***1.67***
73M3 85M12		1.09844	F(64, 141)	=	1.50
73M3 86M12		1.15573	F(64, 141)	=	1.32
73M3 87M12		1.12813	F(64, 141)	=	1.40
73M3 88M12		1.10990	F(64, 141)	=	1.46
73M3 89M12		1.09675	F(64, 141)	=	1.51

The prior evidence from the seasonal unit roots test and visual inspection of the data (see Figure 5.1) suggests a possible change in the seasonal pattern around 1987/88 as far as the unadjusted series (*LAR*) is concerned. A preliminary investigation is performed running the model for *LAR* in which just a constant and the 11 seasonal dummies are included. In this case, the conventional F statistic (Table 5.7) with q=12 degrees of freedom in the numerator and $N+M-2K$=264 in the denominator, suggests that the main change in the seasonal pattern has occurred between 1980/81. This result is confirmed by Andrews' *tabulatum*. The critical value is obtained when π=0.25 and 12 restrictions are considered. The 5% critical value is, in fact, 2.41, smaller than the calculated value.

Table 5. 7 *LAR* - Chow Test for Different Sample Periods

72M1 - 78M12	35.2017
72M1 - 79M12	44.1433
72M1 - 80M12	**48.5646**
72M1 - 81M12	43.9208
72M1 - 82M12	38.4570
72M1 - 83M12	34.4241
72M1 - 84M12	35.4511
72M1 - 85M12	35.1059
72M1 - 86M12	27.2731
72M1 - 87M12	20.9503
72M1 - 88M12	14.5983
72M1 - 89M12	10.6042

However, the results have to be investigated further by running a full model in which just the seasonal coefficients are allowed to change. The results are reported in Table 5.8.

Table 5. 8 *LAR*-Chow Test for 12 Seasonal Coefficients

73M3 95M12 **RRSS** = 1.84742			
	UNRSS		
73M3 78M12	1.60884	F(12,193)=	2.38
73M3 79M12	1.61333	F(12,193)=	2.33
73M3 80M12	1.57379	F(12,193)=	2.80
73M3 81M12	1.58389	F(12,193)=	2.68
73M3 82M12	1.60435	F(12,193)=	2.44
73M3 83M12	1.51302	F(12,193)=	3.55
73M3 84M12	1.49247	F(12,193)=	**3.82**
73M3 85M12	1.53693	F(12,193)=	3.25
73M3 86M12	1.59043	F(12,193)=	2.60
73M3 87M12	1.60157	F(12,193)=	2.47
73M3 88M12	1.58476	F(12,193)=	2.67
73M3 89M12	1.56584	F(12,193)=	2.89
73M3 90M12	1.45724	F(12,193)=	**4.31**
73M3 91M12	1.48164	F(12,193)=	3.97
73M3 92M12	1.61556	F(12,193)=	2.31
73M3 93M12	1.74773	F(12,193)=	0.92
73M3 94M12	1.78196	F(12,193)=	0.59

The statistical values reported in Table 5.8 can be compared with the asymptotic critical values provided by Andrews (1993, p.840). For π=0.15[46], the critical value for 12 restrictions equals 2.51, at the 5% level. Hence, such a test suggests that the largest changes in the seasonal pattern have occurred between 1990/91 and 1984/85, respectively. The latter finding seems to confirm the result shown in Table 5.6. A better inspection of the changes of the seasonal pattern can be carried out graphically (see Figures 5.2 and 5.3). This investigation can be considered a rough comparison, fitting only one change in each case.

46 In this case π is given by: *½[OB/N]*, where *OB* equals 82 and *N* equals 274.

Figure 5. 2 Changes in Seasonal Pattern between 1990/91

Notes: *C1* represents the coefficient for the "non changing" December dummy; *Seas1* represents the coefficients of the various "non changing" seasonal dummies; *C1+C2* represents the sum of the coefficient for the "non changing" and "changing" December dummy, respectively; *Seas1+Seas2* represents the sum of the coefficients for the "non changing" and "changing" monthly seasonals, respectively.

Figure 5. 3 Changes in Seasonal Pattern between 1984/85

Note as for Figure 5.2.

The greatest changes in the seasonal pattern between 1990/91 seem to be in April, September and October with a decrease in the number of arrivals in the second period (*i.e.* from January 1991), in June and August with an increase in the number of domestic arrivals in the second period. For the structural break between 1984/85

(Figure 5.3), the main changes in the seasonal pattern seem to occur in June, July and August, with an overall increase of the number of domestic arrivals. Such assumptions have been investigated further (see the complete and final program for running the tests in Appendix E, Table E.2). Three separate dummies are fitted; for the whole period (*jan*, *feb*, *mar*, and so on), for 1985:1 onwards (*jan2*, *feb2*, *mar2*, and so on), and for 1991:1 onwards (*jan3*, *feb3*, and so on). Firstly, the structural break between 1984/85 has been considered. The F statistic (8,194) when the coefficients for the seasonals *apr2*, *jun2*, *jul2* and *aug2* are allowed to change between 1984 and 1985 has to be accepted at the 5% level. In fact, the F statistic (8,194) calculated equals 1.54 and it is smaller than the critical value (*i.e.* 1.94). Note that the null hypothesis when just the coefficients of *jun2*, *jul2* and *aug2* are allowed to change is not accepted at the 5% level.

Secondly, the structural seasonal change between 1990/1991 has been investigated by testing for possible restrictions on the coefficients of the seasonals: *c3*, *jan3*, *feb3*, *mar3*, *may3*, *jul3* and *nov3*; that is, allowing just the coefficients of the seasonals *apr3*, *jun3*, *aug3*, *sep3* and *oct3* to change. The F statistic (7,194), in such a case, equals 0.91 that is smaller than the conventional critical value at the 5% level (*i.e.* 2.01). Thus, the null hypothesis cannot be rejected.

The restricted seasonal changes with the unrestricted dummy model need to be compared. This is not a full specification search; in fact, when testing the 1984/1985 changes after restricting the 1990/91 changes the same results are not obtained. The restrictions on all the seasonal coefficients except for *apr2*, *jun2*, *jul2* and *aug2* cannot be accepted at a 5% level. That is the F statistic (8,188) equals 2.53 and this value is greater than 1.94 from the conventional tables. However, from a further investigation the coefficients for *jan2*, *apr2*, *jun2*, *jul2* and *aug2* seem to be changing in the second period, that is from 1985:1 until 1990:12. The F statistic (7,188) is equal to 1.96 smaller than the conventional critical value (2.01) at a 5% level.

The same investigation has been done for the structural change occurring between 1990 and 1991, after restricting the 1984/85 changes. The F statistic (7,188)[47] suggests that the restriction on all seasonal coefficients, with the exception for *apr3*, *jun3*, *aug3*, *sep3* and *oct3*, cannot be accepted at the 5% level. From a further analysis,

47 The F statistic (7,188) is equal to 2.48 greater than the critical value (2.01) at the 5% level.

however, the conclusion is that *apr3, may3, jun3, jul3, aug3, sep3* and *oct3* are changing between 1990 and 1991. The F statistic (5,188) calculated value is 2.02, smaller than 2.21 from the conventional table.

Once that the existence of a structural change in the seasonal pattern is accepted, the next interesting step is to understand whether just the seasonal coefficients or all the coefficients of the variables included in the full model are changing. The unrestricted model in which all the coefficients are changing can be compared with the restricted model in which just the coefficients for the seasonals are changing. This comparison can be done using the residual sum of squares reported in Tables 5.6 and 5.8, from which one can rewrite the results provided in Table 5.9.

Table 5. 9 *LAR* - Chow Test: 52 Coefficients vs 12 Seasonal Coefficients Changing

	RSS	UNRSS1			
73M3 78M12	1.60884	1.17446	F(52, 141)	=	1.00
73M3 79M12	1.61333	1.19130	F(52, 141)	=	0.96
73M3 80M12	1.57379	1.22493	F(52, 141)	=	0.77
73M3 81M12	1.58389	1.21965	F(52, 141)	=	0.81
73M3 82M12	1.60435	1.19265	F(52, 141)	=	0.94
73M3 83M12	1.51302	1.10303	F(52, 141)	=	1.01
73M3 84M12	1.49247	1.05183	F(52, 141)	=	1.14
73M3 85M12	1.53693	1.09844	F(52, 141)	=	1.08
73M3 86M12	1.59043	1.15573	F(52, 141)	=	1.02
73M3 87M12	1.60157	1.12813	F(52, 141)	=	1.14
73M3 88M12	1.58476	1.10990	F(52, 141)	=	**1.16**
73M3 89M12	1.56584	1.09675	F(52, 141)	=	1.15

The highest value of the F statistic seems to occur between 1988/89. The null hypothesis for which just the coefficients for the seasonals are changing cannot be rejected at the 5% level, if one uses the conventional F distribution. The calculated value, in fact, is smaller than the conventional critical value, 1.39, thus the restricted model holds. This result has also been confirmed by using Andrews' critical value, *i.e.* around 2.05 for π=0.25[48]. It appears, in fact, that the null hypothesis cannot be rejected at the 5% level. Hence, there is no evidence to believe that all the coefficients are changing.

The further step is to investigate the possible existence of regime changes in the series of domestic arrivals adjusted for the number of Sundays in a month (*LAS*). An unrestricted model in which the coefficients of the seasonal dummies, *LAS*, *LPR*, *DLREP* and *LSP* are allowed to change is estimated. The results from the Chow test

48 Note that π is given by: ½*[OB/N]* , where *OB* = 130 and *N* = 274.

for different sample periods are reported in Table 5.10. The F statistic with (64,142) degrees of freedom indicates that structural change occurs in the period between 1980/81. The F statistic, in fact, is larger than the critical value of the conventional F distribution, *i.e.* larger than 1.32 at a 5% level. However, a comparison of the calculated F statistic with Andrews' critical value (*i.e.* around 2.05) suggests that the null hypothesis of no regime change cannot be rejected at the 5% level. A deeper investigation is suggested.

Table 5. 10 *LAS*- RSS for the Unrestricted and Restricted Models and *F* Statistic

73M3 95M12	**RRSS** =3.71643		
	UNRSS		
73M3 78m12	2.27684	F(64,142)=	1.40
73M3 79m12	2.19587	F(64,142)=	1.53
73M3 80m12	2.19183	F(64,142)=	**1.54**
73M3 81m12	2.41427	F(64,142)=	1.20
73M3 82m12	2.56314	F(64,142)=	1.00
73M3 83m12	2.43634	F(64,142)=	1.17
73M3 84m12	2.33747	F(64,142)=	1.31
73M3 85m12	2.39871	F(64,142)=	1.22
73M3 86m12	2.43820	F(64,142)=	1.16
73M3 87m12	2.43215	F(64,142)=	1.17
73M3 88m12	2.42002	F(64,142)=	1.19
73M3 89m12	2.30524	F(64,142)=	1.36

The prior evidence from the seasonal unit roots test (see Table 5.1) and visual inspection of the data (Figure 5.1) suggests a possible change in the seasonal pattern around 1988/89 for *LAS*. A preliminary investigation is done by running the model for the adjusted series in which a constant and the 11 seasonal dummies are included. In this case, the appropriate F statistic (Table 5.11) with q=12 degrees of freedom in the numerator and $N+M-2K$=264 in the denominator, suggests that the largest change in the seasonal pattern has occurred between 1980/81, as in the case for the unadjusted series. The same result is confirmed also by using Andrews' tables. The 5% critical value equals 2.41 (for π=0.25 when 12 restrictions are considered), that is smaller than the calculated values.

Table 5. 11 *LAS* - Chow Test for Different Sample Periods

	LAS
72M1 - 78M12	27.4616
72M1 - 79M12	33.9360
72M1 - 80M12	**37.0588**
72M1 - 81M12	33.9318
72M1 - 82M12	29.9194
72M1 - 83M12	26.7712
72M1 - 84M12	28.0072
72M1 - 85M12	27.6995
72M1 - 86M12	22.6643
72M1 - 87M12	17.9130
72M1 - 88M12	12.5594
72M1 - 89M12	9.3017

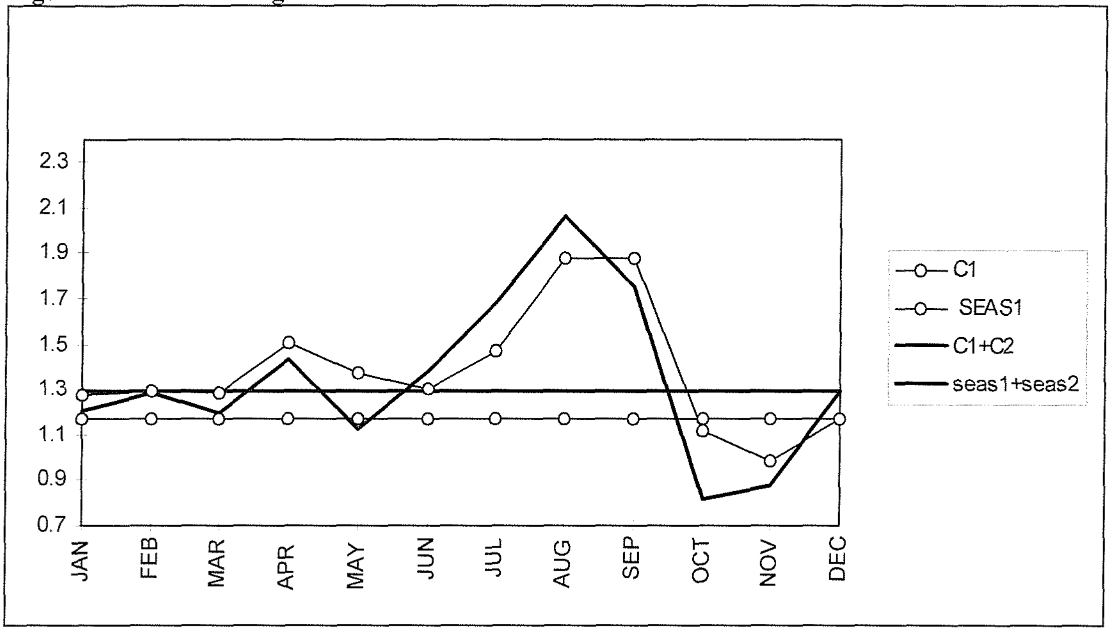
Again, these results have to be investigated further by running a full model in which all the variables that are assumed to effect the demand for tourism are included. However, just the seasonal coefficients are allowed to change. The results from running the Chow test for different sample periods are reported in Table 5.12.

Table 5. 12 *LAS* - Chow Test for 12 Seasonal Coefficients

72M3 95M12 **RRSS** = 3.71643			
	UNRSS		
73M3 78M12	3.33176	F(12,194)=	1.87
73M3 79M12	3.23401	F(12,194)=	2.41
73M3 80M12	3.07580	F(12,194)=	**3.37**
73M3 81M12	3.20463	F(12,194)=	2.58
73M3 82M12	3.28019	F(12,194)=	2.15
73M3 83M12	3.28008	F(12,194)=	2.15
73M3 84M12	3.16625	*F(12,194)=*	*2.81*
73M3 85M12	3.20330	F(12,194)=	2.59
73M3 86M12	3.20041	F(12,194)=	2.61
73M3 87M12	3.30221	F(12,194)=	2.03
73M3 88M12	3.29162	F(12,194)=	2.09
73M3 89M12	3.23856	F(12,194)=	2.38
73M3 90M12	3.081174	F(12,194)=	**3.33**
73M3 91M12	3.09827	F(12,194)=	3.22
73M3 92M12	3.22548	F(12,194)=	3.22
73M3 93M12	3.46374	F(12,194)=	1.18
73M3 94M12	3.57797	F(12,194)=	0.63

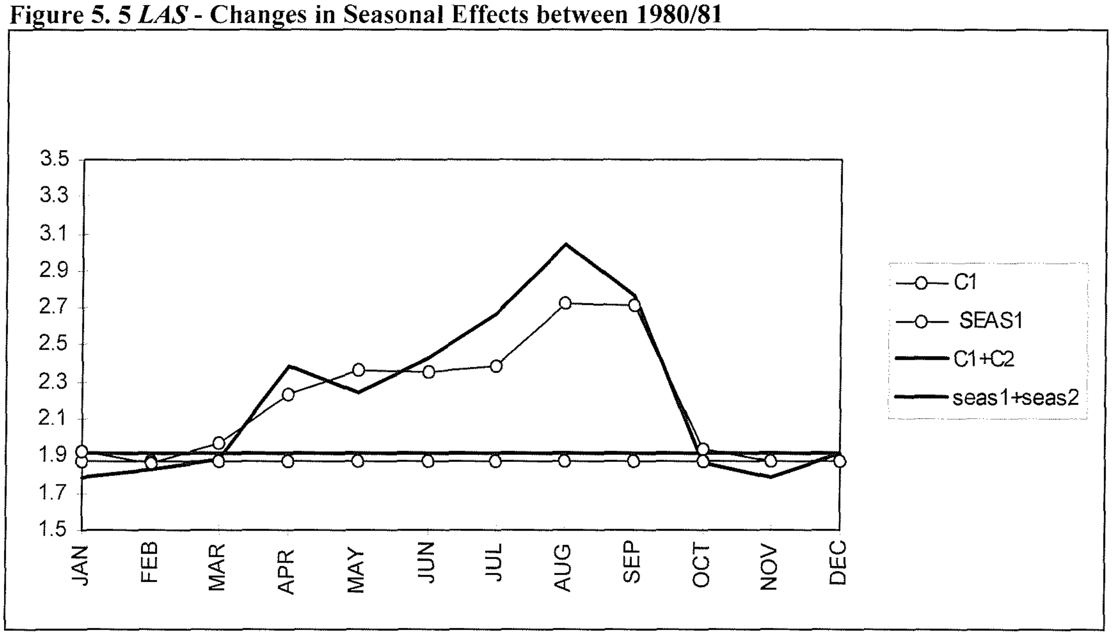
The statistical values reported in Table 5.12 can be compared with the asymptotic critical values provided by Andrews (1993 p.840). For π=0.15 the critical value for 12 restrictions equals 2.51, at the 5% level. Hence, such a test suggests that the largest changes in the seasonal pattern have occurred between 1990/91 and 1980/81, respectively. There is some evidence of an intermediate shift in 1984/85. One will model only two. Any serious mis-specification should be detected by the appropriate tests on the final model. One can roughly inspect the two seasonal changes in Figures 5.4 and 5.5.

Figure 5. 4 *LAS* - Changes in Seasonal Pattern between 1990/91



Note as Figure in 5.2

Figure 5. 5 *LAS* - Changes in Seasonal Effects between 1980/81



Note as in Figure 5.2

From Figure 5.4, it seems that the main changes in the seasonal pattern have occurred in May and October where the domestic demand for tourism in the second period decreases; in July and August where the domestic demand increases in the second period.

Taking into consideration the structural break between 1980 and 1981 (Figure 5.5), the changes in the seasonal pattern seem to occur in July, August and April that show an increasing demand for tourism in the second period.

These assumptions are investigated as follows. Three separate dummies are fitted; for the whole period (*jan, feb* and so on), for 1981:1 onwards (*jan2, feb2* and so on), and for 1991:1 onwards (*jan3* and so on). The first unrestricted model in which all seasonal coefficients are allowed to vary, with respect to the period between 1980/81 in which the structural break has occurred, has been compared with a restricted model in which just the coefficients for April, July and August are allowed to change. The F statistic (9,194) equals 1.70 that is smaller than the critical value, 1.88, at the 5% level. The main finding is that just the above mentioned seasonals are changing from January 1981.

The same type of analysis has been run for the structural break between 1990/91 where all seasonal coefficients are set to zero except for the coefficients for *may3, jul3, aug3* and *oct3*. The appropriate F statistic, with 8 degrees of freedom in the numerator and 194 in the denominator, equals 1.04. Such a value is smaller than the critical value, 1.94, at the 5% level, therefore, the restricted model can be accepted and, therefore, just the coefficients for May, July, August and October change from January 1991.

Also in this case, as for the unadjusted series, one has compared the restricted seasonal changes with the unrestricted dummy model. This is not a full specification search. Testing the 1980/81 change after restricting the 1990/91 seasonal coefficients gives slightly different results. That is just the coefficients for *apr2, jul2, aug2* and *sep2* change, as the calculated value, 1.56, from the F (8,190) is less than the critical value, 1.94. Note, in fact, that the F statistic (9,190), when just the coefficients on *apr2, jul2* and *aug2* are allowed for changes, is 1.97, greater than 1.88 at the 5% level from the conventional tables.

However, the same result has been obtained when testing for the 1990/91 change after imposing the restriction on the 1980/81 changes. The F statistic (8,190) equals 1.39 that is smaller than the correspondent critical value, 1.94. This finding confirms that just the coefficients of *may3, jul3, aug3* and *oct3* are changing.

As for the unadjusted series case, the next step is to investigate the possibility that all the coefficients of the variables included in the full model present a structural change. The RSS of the unrestricted model in which all the coefficients are changing (Table 5.10) can be compared with RSS of the restricted model in which just the coefficients for the seasonals are changing. From such a comparison one obtains the results presented in Table 5.13.

Table 5. 13 *LAS* - Chow Test: 52 Coefficients vs 12 Seasonal Coefficients Changing

	RRSS	UNRSS		
73M3 78m12	3.33176	2.27684	F(52,142) =	1.26
73M3 79m12	3.23401	2.19587	F(52,142) =	**1.29**
73M3 80m12	3.07580	2.19183	F(52,142) =	1.10
73M3 81m12	3.20463	2.41427	F(52,142) =	0.89
73M3 82m12	3.28019	2.56314	F(52,142) =	0.76
73M3 83m12	3.28008	2.43634	F(52,142) =	0.95
73M3 84m12	3.16625	2.33747	F(52,142) =	0.97
73M3 85m12	3.20330	2.39871	F(52,142) =	0.92
73M3 86m12	3.20041	2.43820	F(52,142) =	0.85
73M3 87m12	3.30221	2.43215	F(52,142) =	0.98
73M3 88m12	3.29162	2.42002	F(52,142) =	0.98
73M3 89m12	3.23856	2.30524	F(52,142) =	1.11

The highest value of the F statistic occurs between 1979/80. Comparing this calculated value with the conventional critical values in the F distribution, the null hypothesis for which just the coefficients for the seasonals are changing cannot be rejected at the 5% level, thus the restricted model holds. Moreover, such a calculated value is smaller than Andrews' critical value (*i.e.* around 2.01) at the 5% level, when considering π=0.30 and 52 restrictions. Thus, no evidence appears that all the coefficients are changing.

So far, the main finding is the presence of a structural change in the seasonal pattern both for the unadjusted series and the adjusted series of domestic arrivals of tourists.

5.2.3 The Model Specification For The Unadjusted Series Of Arrivals Of Tourists

Further work is needed to assess which of the two series (*LAR* or *LAS*) produces the best specification. Starting with the unadjusted series of domestic arrivals (*LAR*) three sets of seasonal dummies are created. In this way, it is possible to allow the seasonal pattern to change.

The first set includes *c*, *jan*, *feb*, and so on. Note that *feb*, *mar* and *nov* present the same coefficients in all three periods. Note also that *jan* takes the value 1 in the first period (1972:1-1984:12) and in the third period (1991:1-1995:12), and zero in the second period (1985:1-1990:12). Whereas *may*, *sep* and *oct* take the value 1 in the first and in the second period, and zero in the last period. Moreover, the *apr*, *jun*, *jul* and *aug* dummies take the value 1 in the first period and zero in the second and third period.

The second set of seasonal dummies allows for a structural break in the second period and contains *jan2*, *apr2*, *jun2*, *jul2* and *aug2*. This set takes the value one in the second period and zero otherwise.

The third set of dummies, allowing for a structural break in the third period, contains *apr3*, *may3*, *jun3*, *jul3*, *aug3*, *sep3* and *oct3*. This set takes the value one in the third period and zero otherwise.

An initial unrestricted model has been fitted for *LAR* including 13 lags for the dependent variable as well as the (log) industrial production (*LPR*), the (log) first difference of relative price (Sassari Province-Italy) (*DLREP*) and the (log) substitute price (*LSP*). The model also includes the "Easter" dummy (*E*), the (log) weather variable (*LW*), three impulse dummies (*i1974p11*, *i1992p3* and *i1993p3*) created after inspecting the residuals in order to correct non-normality problems, and, finally, the above described seasonal dummies. The results for the equation for *LAR* are reported in the Appendix E, Table E.3. No problems appear in terms of diagnostic tests.

A general-to-specific simplification has been carried out, and a parsimonious model (see Table E.4, Appendix E) has been obtained with non-autocorrelated, normal residuals. The coefficients of the second and third lag of the industrial production is almost of the same size and opposite sign. The imposition of a restriction in such coefficients has been accepted at the 5% level[49]. No further restriction has been accepted. Note also that the relative price growth variable does not turn out to be statistically significant. Nevertheless, the exclusion of such a variable does not worsen the model noticeably, and its inclusion is suggested by the SC criterion as well as by a joint *F*-test on all 13 lags.

[49] The appropriate *F* statistic with $q=1$ and $N\text{-}K=237$, after imposing the restriction $\gamma_2 + \gamma_3 = 0$, is 0.34. Therefore, the restriction cannot be rejected since 0.34 is smaller than the critical value (3.84). Note that $LPR_2\text{-}LPR_3$ is called *RLPR*.

The final model (Table 5.14) is parsimoniosly well-specified in terms of reduction testing as well as in terms of the SC criterion. Moreover, the model does not present any problems in terms of diagnostic tests. The R-squared adjusted shows the good of fitness which is also supported by a satisfactory ratio between the *SER* and *DMV* that equals 0.007997[50]. Further tests for parameter constancy over the last 33 observations of the sample are also reported. The null hypothesis of parameter constancy cannot be rejected.

50 SER/MDV = (0.081626/10.146) = 0.007997.

Table 5. 14 Restricted Model for the Domestic Demand for Tourism

EQ(1) Modelling LAR by OLS (using vdomlar1.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	2.0709	0.58569	3.536	0.0005	0.0499
LAR_1	0.38762	0.042513	9.118	0.0000	0.2589
LAR_5	-0.080050	0.033139	-2.416	0.0165	0.0239
LAR_11	0.12137	0.046003	2.638	0.0089	0.0284
LAR_12	0.21460	0.053575	4.006	0.0001	0.0632
RLPR	-0.65314	0.18395	-3.551	0.0005	0.0503
LPR_11	0.31459	0.087485	3.596	0.0004	0.0515
LSP_6	1.7133	0.67707	2.530	0.0120	0.0262
LSP_7	-1.3387	0.67560	-1.981	0.0487	0.0162
jan	0.062585	0.029556	2.118	0.0353	0.0185
jan2	0.016667	0.041547	0.401	0.6887	0.0007
feb	0.052512	0.030513	1.721	0.0866	0.0123
mar	0.068376	0.055725	1.227	0.2210	0.0063
apr	0.30200	0.071873	4.202	0.0000	0.0691
apr2	0.44838	0.079360	5.650	0.0000	0.1183
apr3	0.26106	0.081491	3.204	0.0015	0.0413
may	0.26753	0.076261	3.508	0.0005	0.0492
may3	0.24610	0.086798	2.835	0.0050	0.0327
jun	0.30849	0.078161	3.947	0.0001	0.0614
jun2	0.47581	0.094223	5.050	0.0000	0.0968
jun3	0.53057	0.097691	5.431	0.0000	0.1103
jul	0.44813	0.084031	5.333	0.0000	0.1067
jul2	0.53537	0.10279	5.209	0.0000	0.1023
jul3	0.58224	0.10879	5.352	0.0000	0.1074
aug	0.57998	0.084264	6.883	0.0000	0.1660
aug2	0.75099	0.10190	7.370	0.0000	0.1858
aug3	0.84490	0.10815	7.812	0.0000	0.2041
sep	0.47825	0.082232	5.816	0.0000	0.1244
sep3	0.34631	0.10317	3.357	0.0009	0.0452
oct	-0.17855	0.064507	-2.768	0.0061	0.0312
oct3	-0.35791	0.075092	-4.766	0.0000	0.0871
nov	-0.091988	0.031107	-2.957	0.0034	0.0354
i1974p11	0.26292	0.085615	3.071	0.0024	0.0381
i1992p3	0.33224	0.085173	3.901	0.0001	0.0601
i1993p3	-0.20749	0.088385	-2.348	0.0197	0.0226
E	0.14997	0.030290	4.951	0.0000	0.0934

R^2 = 0.99036 F(35,238) = 698.62 [0.0000] sigma = 0.081626
DW = 2.04 RSS = 1.585748486 for 36 variables and 274 observations

AR 1- 7 F(7,231) = 0.68912 [0.6812]
ARCH 7 F(7,224) = 0.53527 [0.8073]
Normality Chi^2(2)= 0.95762 [0.6195]
Xi^2 F(43,194) = 0.74157 [0.8771]
RESET F(1,237) = 2.7144 [0.1008]
Tests of parameter constancy over: 1993 (4) to 1995 (12)
Forecast Chi^2(33)= 24.769 [0.8481]
Chow F(33,205) = 0.6736 [0.9117]

Analysing the coefficients of the explanatory variables, the sign of the industrial production index, as the income proxy, is expected to be positive. From Tables 5.14 and 5.15, the short and long run income elasticities are positive and statistically significant. Moreover, the magnitude of both the coefficients is less than unity. Hence, the hypothesis that domestic tourism is to be considered as a necessity good has been confirmed. This finding is also consistent with the result obtained by Malacarni (1991) when using annual data for the Italian case.

In general, in almost all economic transactions the principle that the demand and the price are inversely correlated holds. However, as in the case for the international demand for tourism, the coefficient for the substitute price, both in the short and long run, turns out to be positive. Once again, there might be problems in not taking explicitly into account the possible influence of the exchange rate for other competitors of the north of Sardinia. Chapter 6 is dedicated to a deeper investigation of modelling the domestic demand for tourism with the inclusion of the substitute price and exchange rate for the competitor countries.

From Table 5.14, the weather variable does not appear to have any effects on the determination of the domestic flows of tourists. The same results have been achieved in the international tourism demand case.

Turning to the "Easter" dummy, it is highly statistically significant in explaining domestic demand for tourism. This finding is in agreement with the international demand case.

In Table 5.15, the long run dynamics and standard errors (in parenthesis) are reported. In general, the standard errors are relatively low and and the long run coefficients are jointly significant as inferred from the Wald test. It can be concluded that the monthly model presents an overall good specification both in the short and long run.

Table 5. 15 Solved Static Long Run Equation

```
Solved Static Long Run equation
         LAR =      +5.81              +0.8825 LPR             +1.051 LSP
(SE)           (    1.039)           (  0.2269)             (   0.1724)
                  +0.1756 jan          +0.04676 jan2           +0.1473 feb
               (  0.08192)           (  0.1159)             (  0.08628)
                  +0.1918 mar          +0.8472 apr             +1.258 apr2
               (   0.1627)           (  0.2504)             (   0.3225)
                  +0.7324 apr3         +0.7505 may             +0.6904 may3
               (   0.2649)           (  0.2387)             (   0.2551)
                  +0.8654 jun          +1.335 jun2             +1.488 jun3
               (   0.2415)           (  0.3197)             (   0.3323)
                  +1.257 jul           +1.502 jul2             +1.633 jul3
               (    0.269)           (  0.3185)             (   0.3286)
                  +1.627 aug           +2.107 aug2             +2.37 aug3
               (   0.2902)           (  0.3635)             (   0.3871)
                  +1.342 sep           +0.9715 sep3            -0.5009 oct
               (   0.2462)           (  0.2695)             (   0.2227)
                  -1.004 oct3          -0.2581 nov             +0.7376 i1974p11
               (   0.3055)           (   0.101)             (   0.2727)
                  +0.9321 i1992p3      -0.5821 i1993p3         +0.4207 E
               (   0.2781)           (  0.2667)             (   0.1152)
                  -1.832 RLPR
               (   0.6097)
```

ECM = LAR - 5.8097 - 0.882537*LPR - 1.05084*LSP - 0.175575*jan - 0.0467577*jan2
- 0.147318*feb - 0.191821*mar - 0.847238*apr - 1.25789*apr2 - 0.732378*apr3
- 0.75052*may - 0.690403*may3 - 0.865441*jun - 1.33482*jun2 - 1.48845*jun3
- 1.25718*jul - 1.50193*jul2 - 1.6334*jul3 - 1.62708*aug - 2.10683*aug2
- 2.37028*aug3 - 1.34168*sep - 0.971534*sep3 + 0.500899*oct + 1.00409*oct3
+ 0.258061*nov - 0.737593*i1974p11 - 0.932072*i1992p3 + 0.582086*i1993p3
- 0.420728*E + 1.8323*RLPR;

WALD test Chi^2(30) = 373.09 [0.0000] **

5.2.4 Logarithmic Versus Linear Specification

In this section, an account will be given of whether a linear specification might be more appropriate in estimating the demand for domestic tourism than a logarithmic specification. As for the foreign demand of tourism, the Box and Cox (1964) test will be used (Chapter 2, Section 2.4.1).

Firstly, it is interesting to analyse the properties of the variables when expressed in a linear form. The ADF for testing the integration status of the variables of interest has given the following results: *AR* (number of domestic arrivals in registered accommodation), *PR* (income proxy) and *SP* (substitute price, Sassari/main competitors) have to be considered stationary in the level. Whereas, *REP* (relative price, Sassari/Italy) has been found to be I(1). Moreover, the seasonal unit roots test for the number of domestic arrivals (*AR*) suggests possible structural changes in the seasonal pattern, as with the log specification.

In order to run the Box-Cox test, one estimates the unrestricted 13 lag model for the demand equation as follows:

1) logarithmic form

$$LAR_t = a_1 + a_2\,LAR_{t-1\ldots} + a_3\,LPR_{t\ldots} + a_4\,LSP_{t\ldots} + a_5\,DLREP_{t\ldots} + a_6\,LW_t + a_7\,E + a_8\,Seas + a_9\,Dummies + e_t$$

and

2) linear form

$$AR_t = a_1 + a_2\,AR_{t-1\ldots} + a_3\,PR_{t\ldots} + a_4\,SP_{t\ldots} + a_5\,DREP_{t\ldots} + a_6\,W_t + a_7\,E + a_8\,Seas + a_9\,Dummies + e_t$$

where the explanatory variables are defined as before, and the seasonals are constructed so as to allow structural changes in the seasonal pattern. The impulse dummies are the following: *i1974p11*, *i1992p3* and *i1993p3*.

In particular, the sum of the squared errors from the logarithmic form (*SSELL*) equals 1.828891408, whereas the sum of the squared errors for the linear form (*SSEL*) equals 3626196276. The aim is to test whether the two models are empirically equivalent and find out which of the two models best fits the data. In doing this, one needs to calculate the sum of the squared errors for the linear model with ($AR/\bar{A}R\,G$) as the dependent variable and where $\bar{A}R\,G$ is the geometric mean. The sum of the squared errors for the latter model (*i.e.* $SSEL/(\bar{A}R\,G)^2$) equals 5.581458. The Box-Cox test indicates that the two models are empirically different as the calculated χ^2 is 152.86 and the correspondent tabulated value is 3.84 at the 5% level. Moreover, the $SSEL/(\bar{A}R\,G)^2$ is higher that *SSELL*, hence, it can be concluded the logarithmic specification form fits the data better than the linear model. The same result is obtained in the foreign model.

5.2.5 The Model Specification Using Monthly Data For The Adjusted Series Of Arrivals Of Tourists

The next step is to carry out an investigation for the adjusted series of domestic arrivals for the number of Sundays in a month (*LAS*). An unrestricted model is fitted for 13 lags of the dependent variable, *LPR*, *LSP*, *DLREP*, *E* and *LW* (as previously defined), and two impulse dummies (*i1987p3* and *i1992p3*) are included after having inspected the residuals in order to avoid non-normality problems. Three sets of seasonal dummies are created to allow the seasonal pattern to change, in accordance with the structural change findings as assessed in the previous section. The first set includes *c, jan, feb*, and so on. Note that *apr* and *sep* take the value 1 in the first period (1972:1-1980:12) and in the third period (1991:1-1995:12); whereas, *may* takes the value of one in the first and second period (1981:1-1990:12) and the value of zero in the third period. Note also that *jul* and *aug* dummies take the value 1 just in the first in the first period. The second set of seasonal dummies allows for the structural break in the second period and consists of *apr2, jul2, aug2* and *sep2*. This set of dummies takes the value of one in the second period and zero otherwise. Finally, the third set of dummies that allows for structural change in the third period consists of *may3, jul3, aug3* and *oct3*. In this case these dummies take the value of one in the third period and zero otherwise.

The results for the unrestricted model are shown in Appendix E (Table E.5) and no problems arise from the diagnostic tests. The residual sum of squares present, as pointed out in the VAR analysis, a greater value than the RSS for the unadjusted series. The results of the parsimonious model (see Appendix E, Table E.6) indicate problems in terms of specification (*RESET* test).

The main conclusion from this analysis is that, for the domestic demand of tourism, the adjustment of the dependent variable for the number of weekends (either Saturdays or Sundays in a month) does not appear to be acceptable. However, one can argue that such a finding seems plausible in that domestic tourists can easily arrive to the north of Sardinia any day of the week either by boat or plane. Foreign tourists, on the other hand, are much more constrained by the day of arrival as are more likely to use charter flights which occur mainly in the weekends, as far as the period under study is concerned.

5.3 THE MODEL SPECIFICATION USING ANNUAL DATA

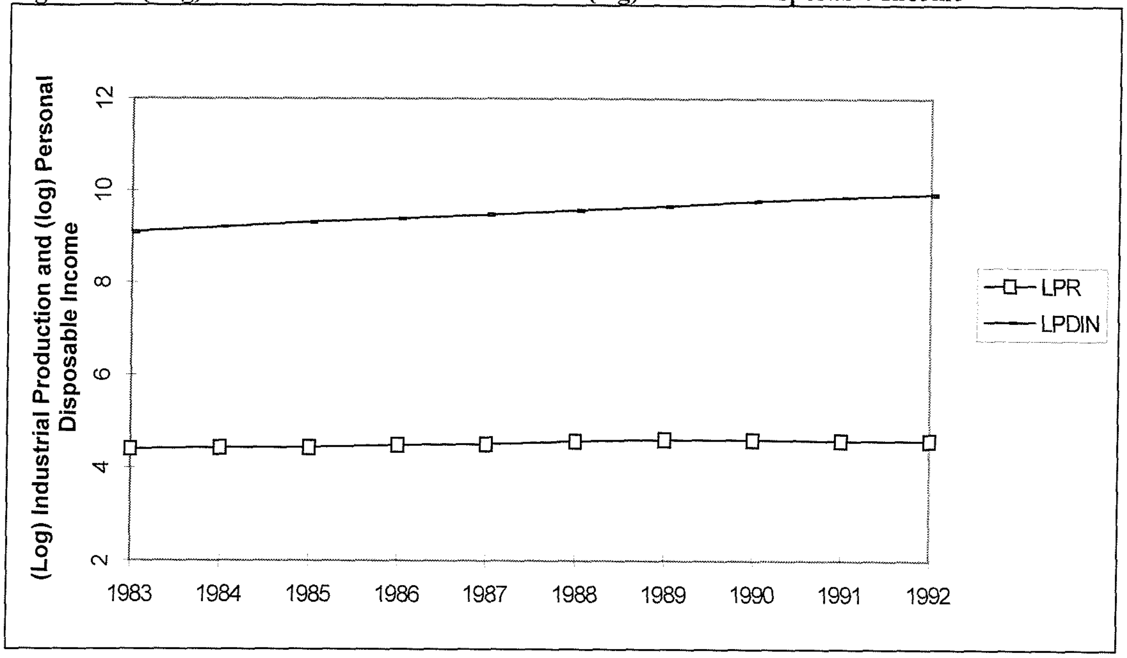
The aim of this section is twofold. Firstly, one will assess whether the Italian industrial production index (*LPR*) can indeed be thought to be a valid proxy for the Italian personal disposable income (*PDIN*).

According to Lim (1997), 62% of the total tourism studies have used annual data, whereas just a very limited number has made use of monthly data. One of the problems is the low precision of estimates associated with the use of annual data, as in many circumstances the sample sizes are very small and do not allow for specification of the dynamics of the demand for tourism (Witt and Witt, 1992; Lim, 1997). Hence, the second aim of this section is assessing the characteristics of the domestic demand for tourism in the long run by making use of annual data, and hence, comparing them with the short run characteristics.

5.3.1 Industrial Production Index As A Proxy For The Personal Disposable Income

It is of interest to find out whether the industrial production index can be considered a valid proxy of the personal disposable income in the Italian case. The analysis in this thesis refers to the period from 1983 until 1992. The total number of observations available for the Italian personal disposable income, *PDIN*, are ten (source ISTAT). In particular, the disposable income *per capita* consists of grants from public authorities, wages of self-employed and employees, earnings from different kind of investments minus income tax, health contributions, public and private contributions to pension funds. The Italian index of industrial production published annually (source IFS) is also used for the same period. The logarithm of the two series are shown in Figure 5.6.

Figure 5. 6 (Log) Industrial Production Index and (log) Personal Disposable Income



For the period 1983-1992 the contemporaneous cross-correlation coefficient is: r(*LPR*,*LPDIN*)=0.92 and the corresponding coefficient of determination equals 0.85. It can be stated that a relative high proportion of the total variance is explained by the other variable. In particular, when regressing the personal disposable income on the industrial production index, one obtains the following results with *t*-statistics in parentheses:

$$\underset{}{LPDIN} = \underset{(-2.28)}{-4.9101} + \underset{(6.69)}{3.1939}\ LPR$$

This finding can be thought to be one of the first arguments for using *LPR* as a proxy for *LPDIN*.

The next step is to test for the level of integration of the two variables. As already stated, an advantage in using monthly data is given by the relative large number of observations available (288 observations in this study) that allows one to test the possible presence of seasonal unit roots as well as determine the level of integration of the variables under study. On the other hand, a difficulty with ADF tests is that acceptance of the null hypothesis, and thus the presence of a unit root, may arise because the limited number of observations gives the test low power. In this specific case, one uses a Dickey-Fuller test (DF test) as one could not augment because of the lack of observations. In the previous monthly analysis, it has been found that *LPR* is stationary in the level when a constant and a time trend are added. However, when using annual data and just 10 observations, such a conclusion is not firm. Table 5.16 suggests also that it is not possible to establish the level of integration for the personal disposable income. Table 5.16 is more illustrative than conclusive. One could use appropriate critical values for relative small sample sets. However, this is out of the scope of this thesis.

Table 5. 16 DF Test for *LPR* and *LPDIN* (10 Obs. 1983-1992)

Series	*DF*(1)	*LAG*
LPR(c)	-1.70	0
DLPR(c)	-1.30	0
LPR(c,t)	-1.13	0
DLPR(c,t)	not calculated: too few observations for selected lag length and 4 parameters	
LPDIN(c)	-1.73	0
DLPDIN(c)	-2.48	0
LPDIN(c,t)	-1.84	0

(1) Dickey-Fuller statistics with constant and trend critical values= -4.082 at 5% and -5.478 at 1% level; with constant c.v. = -3.27 at 5% and -4.61 at 1% level.

A further experiment consists in running a VAR(1) (see Table E.7 in Appendix E) with the inclusion of a time trend. The results suggest that the industrial production index can be thought as an autoregressive process, since this variable seems to be dependent on its own past behaviour. In the first equation, almost 95% of the variance of the dependent variable is explained. From the second equation, one infers that the first lag of the industrial production index plays an important role in explaining the personal disposable income. Moreover, in this case, almost all the variance of the dependent variable is explained.

One can argue that more robust results might be given by using a higher number of observations. Nevertheless, the previous analysis has shown that the index of industrial production is highly statistically significant in explaining the Italian personal disposable income; hence, one can consider the former as a valid proxy for the latter.

5.3.2 Annual Data Analysis For The Domestic Arrivals (*LAR*)

The annual data used in this study cover a period of 24 years (1972-1995). The preliminary investigation carried out tests for the level of integration of the variables of interest. As already stated, a difficulty with ADF tests is that of acceptance of the presence of a unit root may arise because the limited number of observations which gives the test low power. Thus, it might be possible, to obtain different results from the monthly data analysis.

Table 5. 17 Testing for Long Run Unit Root with Annual Data (1972-1995)

Series	*ADF*(1)	*LAG*(2)
LAR(c)	- 2.29	0
DLAR(c)	- 3.87*	0
LAR(c,t)	- 1.95	0
DLAR(c,t)	- 3.62*	0
LPR(c)	- 1.27	0
DLPR(c)	- 5.06**	0
LPR(c,t)	- 2.32	0
DLPR(c,t)	- 4.97**	0
LREP(c)	- 0.21	0
DLREP(c)	- 4.07**	0
LREP(c,t)	- 2.21	0
DLREP(c,t)	- 4.13*	0
LSP(c)	- 4.39 **	0
LSP(c,t)	- 1.10	1
DLSP(c,t)	- 7.88 **	0

Notes: The one and two asterisks indicate that the unit root null hypothesis is rejected at the 5% and 1% level, respectively. The capital letter *D* denotes the first-difference operator defined, in a general notation, by $Dx_t = x_t - x_{t-1}$.

(1) Augmented Dickey-Fuller statistics with with constant and trend (*c,t*) c.v.= -3.735 at 5% and -4.671 at 1% level; with constant c.v. = -3.066 at 5% and -3.93 at 1% level.

(2) Lag is the length of the first significant lag. Note that ADF(0) corresponds to the Dickey-Fuller test; additional lags are included to whiten the residuals.

A comparison between Table 5.17 and Table 5.2 suggests major differences in the results. These findings imply possible mis-specification in determining whether a variable can be treated as stationary in the level when using annual data and a short sample size. Note, therefore, that one will consider the above variables as having the same integration status as suggested by the ADF test when using monthly data. The unadjusted series of domestic arrivals (*LAR*) will be treated as I(0), as well as the income proxy (*LPR*) and the substitute price (*LSP*), whereas, the relative price (*LREP*) will be treated as I(1).

The use of the logarithmic specification has been supported statistically as follows. First of all, the integration status of the variables expressed in a linear specification is tested. From the ADF test, *AR*, *PR* and *REP* are I(1), whereas *SP* is stationary in the level. Note that the results for *AR* and *PR* diverge from the linear monthly case (see Section 5.2.5). One argues that the ADF test run for the monthly series has higher robustness as a greater number of observations is involved in the estimation. Hence, one carries on treating *AR* and *PR* as stationary in the level. On this basis, one proceeds with the Box and Cox test. An unrestricted 1 lag domestic tourism equation is run both in a logarithmic and a linear form as follows:

1) logarithmic form

$LAR_t = a_1 + a_2\, LAR_{t-1}.. + a_3\, LPRa_t.. + a_4\, DLREP_t.. + a_5\, LSP_t\, .. + a_6\, LW_t\, .. + a_6\, Trend +$
$+\ e_t$

and

2) linear form

$AR_t = a_1 + a_2\, AR_{t-1}.. + a_3\, PRa_t.. + a_4\, DREP_t\, .. + a_5\, SP_t\, .. + a_5\, W_t\, .. + a_5\, Trend + \ e_t$

The *SSELL* is equal to 0.03665808659 and the *SSEL* equals 8.34E+09. One needs to test whether the null hypothesis, that the two models are empirically equivalent, holds. The calculated χ^2 equals 3.97 that is greater than the tabulated critical value, 3.84, at the 5% level; hence, the two models are empirically different. Moreover, it is inferred that the logarithmic specification is slightly better than the linear specification as the *SSELL* is smaller than $SSE_L/(\overline{\text{AR}}\, G)^2$ (*i.e.* 0.052569). Further specifications in linear and log-linear form could be compared; however, this is not pursued further as the above results confirm the finding for the monthly case.

The initial model is estimated regressing the logarithm of the unadjusted series of arrivals (*LAR*) on the logarithm of the following variables: the index of industrial production (*LPR*), the first difference of the relative price (*DLREP*), the substitute price (*LSP*), the weather variable (*LW*), a time trend (*TREND*) in order to take into consideration "possible changes in the popularity of the holiday over the period as a result of changing tastes" (Martin and Witt, 1988), and, finally, an impulse dummy variable (*D79*), after inspecting the residuals, which might picking up the effects of the second oil crises in the Seventies.

A one lag structure is tested suggesting no problems in terms of diagnostic tests. Following the joint *F*-test as well as the SC criterion, the initial unrestricted model has been reduced parsimoniously. The results from the final model, using annual data are reported in Table 5.18.

Table 5. 18 Regression Results for the Annual Domestic Demand Model (*LAR*)

```
EQ(1) Modelling LAR by OLS  (using datiann.in7)
The present sample is:  3 to 24
Variable       Coefficient    Std.Error  t-value   t-prob PartR^2
Constant            8.9790       1.0444    8.597   0.0000  0.8221
LPR                0.46779      0.18235    2.565   0.0207  0.2914
LSP                 1.4932      0.19786    7.547   0.0000  0.7807
LW                 0.69008      0.24934    2.768   0.0137  0.3238
Trend             0.015705    0.0037756    4.160   0.0007  0.5196
D79               -0.13515     0.048525   -2.785   0.0132  0.3265
R^2 = 0.981208  F(5,16) = 167.09 [0.0000] sigma = 0.0447279
DW = 1.97 RSS = 0.03200929317 for 6 var. and 22 obs.

AR 1- 7 F( 7,   9) =      1.9714 [0.1692]
ARCH 7  F( 7,   2) =     0.45772 [0.8169]
Normality Chi^2(2)=     0.27006 [0.8737]
Xi^2     F( 9,   6) =   0.30248 [0.9472]
RESET    F( 1,  15) =   0.31175 [0.5848]
Tests of parameter constancy over: 23 to 24
Forecast Chi^2( 2)=      4.5929 [0.1006]
Chow     F( 2,  14) =    2.0679 [0.1634]
```

Such a model is statistically well-specified and constitutes an admissible reduction of the underlying unrestricted model. The model is able to explain almost 98% of the variation of the dependent variable; moreover, the goodness of the fit is also suggested by the ratio of *SER* and *MDV* equal to 0.0035[51]. Note also that the null hypothesis of parameter constancy for the last two observations fails to be rejected.

The inclusion of the lagged dependent variable does not turn out to be statistically significant, suggesting that the model converges rapidly to the long run equilibrium. Moreover, this finding suggests that domestic tourists are likely to be "allocentric" in that they prefer new destinations for their own holiday trip. The (log) index of industrial production is found to be statistically significant. This finding seems to be consistent with the results obtained by Arbel and Ravid (1985) for which "the income is found to be the single most important determinant of long run recreation use" (p.981). Moreover, it confirms the results obtained in the monthly model; the coefficient is positive and less than unity suggesting that domestic tourism has to be regarded as a necessity good. Hence, the results obtained in Malacarni (1991) are confirmed. The (log) substitute price coefficient has a positive sign and is in contrast with economic theory. However, these findings restate the results from the

51 SER/MDV = (0.0447279/12.895) = 0.0035.

monthly case. The first difference of the (log) relative price has turned out not to be statistically significant as in the monthly case. The (log) weather variable plays a role in explaining the national demand for tourism. The coefficient is statistically significant and it presents a positive sign. Note that such a variable has not been found statistically significant in the monthly model. The time trend turns out to be highly statistically significant with a positive sign.

5.3.3 Supply Components And Simultaneity

In estimating the annual domestic demand for tourism in the north of Sardinia, a component of the quantity of the supply services is taken into consideration. Hence, two more variables have been included that is: the (log) number of domestic boat arrivals (*LB*) in the two main ports (Porto Torres and Olbia) and the (log) number of domestic flights (*LAE*) in the two main airports (Fertilia and Olbia)[52]. Note that the cross-correlation coefficients between the number of tourists' arrivals and the above mentioned variables are the following: r(*LAR*,*LB*)=0.69 and r(*LAR*,*LAE*)=0.66. The results from the estimation using annual data are presented in Table 5.19.

Table 5. 19 Final Model for Domestic Demand of Tourism using Supply Components

```
EQ(1) Modelling LAR by OLS  (using datiann.in7)
The present sample is:  3 to 24
```

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	9.1003	1.0609	8.578	0.0000	0.8214
LPR_1	0.64459	0.17732	3.635	0.0022	0.4523
LB	0.62007	0.15314	4.049	0.0009	0.5061
LB_1	-0.50664	0.16992	-2.982	0.0088	0.3572
LSP	1.2273	0.20421	6.010	0.0000	0.6930
Trend	0.017021	0.0045809	3.716	0.0019	0.4632

```
R^2 = 0.984848 F(5,16) = 207.99 [0.0000] sigma = 0.0401634
DW = 1.99 RSS = 0.02580961034 for 6 var. and 22 obs.

AR 1- 2 F( 2, 14) =     0.15833 [0.8551]
ARCH 1   F( 1, 14) =    0.03779 [0.8487]
Normality Chi^2(2)=     1.14330 [0.5646]
Xi^2     F(10,  5) =    0.18704 [0.9881]
RESET    F( 1, 15) =    1.89180 [0.1892]
```

The final parsimoniously well-specified model shows a statistical significance of the coefficient for just one of the supply component, *LB*. The coefficient of the (log) number of national boats arriving in the north of Sardinia is found to be statistically

52 The source is the "*Annuario Statistico Italiano*" (1972-1996).

significant with an average positive sign. It is expected, in fact, that an increase in the domestic demand is associated to an increase of the supply of transportation means.

The other coefficients can be compared with the coefficients reported in Table 5.18. In the present model, the income proxy (*LPR*) is statistically significant and shows a positive sign and an elasticity less than one. The substitute price shows a positive sign and it is highly statistically significant. The time trend has a positive sign denoting an upwards trend in the popularity for the north of Sardinia as a tourist destination.

Given the statistical significance of the level of the supply component, *LB*, a further aim is to test for possible simultaneity in the (log) number of domestic arrivals of boats (*LB*). Under the null hypothesis the tested variable will be treated as predetermined and, under the alternative, as endogenous. In accordance with the previous results obtained using OLS, it is assumed that the demand and supply models can be expressed as follows:

1) $LAR_t = \alpha_0 + \alpha_1 LPR_{t-1} + \alpha_2 LB_t + \alpha_3 LB_{t-1} + \alpha_4 LSP_t + \alpha_5 TREND + \varepsilon_{1,t}$

2) $LB_t = \beta_0 + \beta_1 LAR_t + \beta_2 LPR_t + \varepsilon_{2,t}$

where:

a) *LAR* = domestic arrivals of tourists in the north of Sardinia.

b) *LPR*= industrial production index in Italy.

c) *LSP*= substitute price (consumer price index in Sassari by the weighted average consumer price index in other destinations in the Mediterranean area).

d) *LB*= number of domestic arrivals of boats in the north of Sardinia.

e) *TREND* = time trend.

Firstly, one obtains the reduced form by regressing *LB* on the variables included in the first equation, treating the income proxy as an instrument. A number of alternative specifications have been investigated for the instrument set. However, no satisfactory instrument can be found for *LB*. The results obtained are given in Table 5.20.

Table 5. 20 Modelling Reduced Form for Arrivals of Domestic Boats (*LB*) in North of Sardinia

```
Modelling Reduced Form for LB
Variable      Coefficient     Std.Error  t-value
Constant           2.2775        1.6383    1.390
LPR             -0.074086       0.34156   -0.217
LB_1              0.87804       0.20232    4.340
LSP               0.30835       0.33519    0.920
Trend           0.0014308     0.0076183    0.188
LPR_1            -0.20277       0.31840   -0.637

R^2 = 0.873399  F(5,16) = 22.076 [0.0000]  RF sigma = 0.0654703
DW = 1.73 RSS = 0.06858179332 for 6 var. and 22 obs.

Modelling Reduced Form for LAR
Variable      Coefficient     Std.Error  t-value
Constant           10.435        1.4068    7.418
LPR               0.21425       0.29330    0.730
LB_1            -0.038087       0.17373   -0.219
LSP                1.3557       0.28782    4.710
Trend            0.019106     0.0065418    2.921
LPR_1             0.40780       0.27341    1.492

R^2 = 0.970313  F(5,16) = 104.59 [0.0000]  RF sigma = 0.0562189
DW = 1.29 RSS = 0.05056899133 for 6 var. and 22 obs.

EQ(1) Modelling LB by IVE  (using datiann.in7)
The present sample is:  3 to 24

Variable      Coefficient     Std.Error  t-value  t-prob
LAR               0.29251       0.10584    2.764  0.0124
Constant           1.1137        1.2595    0.884  0.3876
LPR               0.72687       0.38820    1.872  0.0766

Additional Instruments used:
     LB_1       LSP     Trend     LPR_1

sigma = 0.0975545  DW = 0.586
RSS = 0.1808206951 for 3 variables and 22 observations
2 endogenous and 2 exogenous variables with 6 instruments
Reduced Form sigma = 0.0654703
Specification  Chi^2(3)  =     16.207 [0.0010] **
Testing beta=0:Chi^2(2) =     35.719 [0.0000] **
AR 1- 2 F( 2, 17) =    13.211 [0.0003] **
ARCH 1  F( 1, 17) =     5.661 [0.0293] *
Normality Chi^2(2)=     7.580 [0.0226] *
Xi^2    F( 4, 14) =     0.860 [0.5118]
Xi*Xj   F( 5, 13) =     0.639 [0.6739]
```

As one can notice problems appear in the residuals. Moreover, the specification test $\chi^2(3)$ suggests that the null hypothesis, the validity of the instruments, cannot be accepted at the 1% level.

The saved residuals (*RELB*) from the reduced form are included into the original model to "correct" for simultaneity. The equation run by OLS is the following:

Table 5. 21 Simultaneity Test for *LB* (Domestic Boat Arrivals) using Annual Data

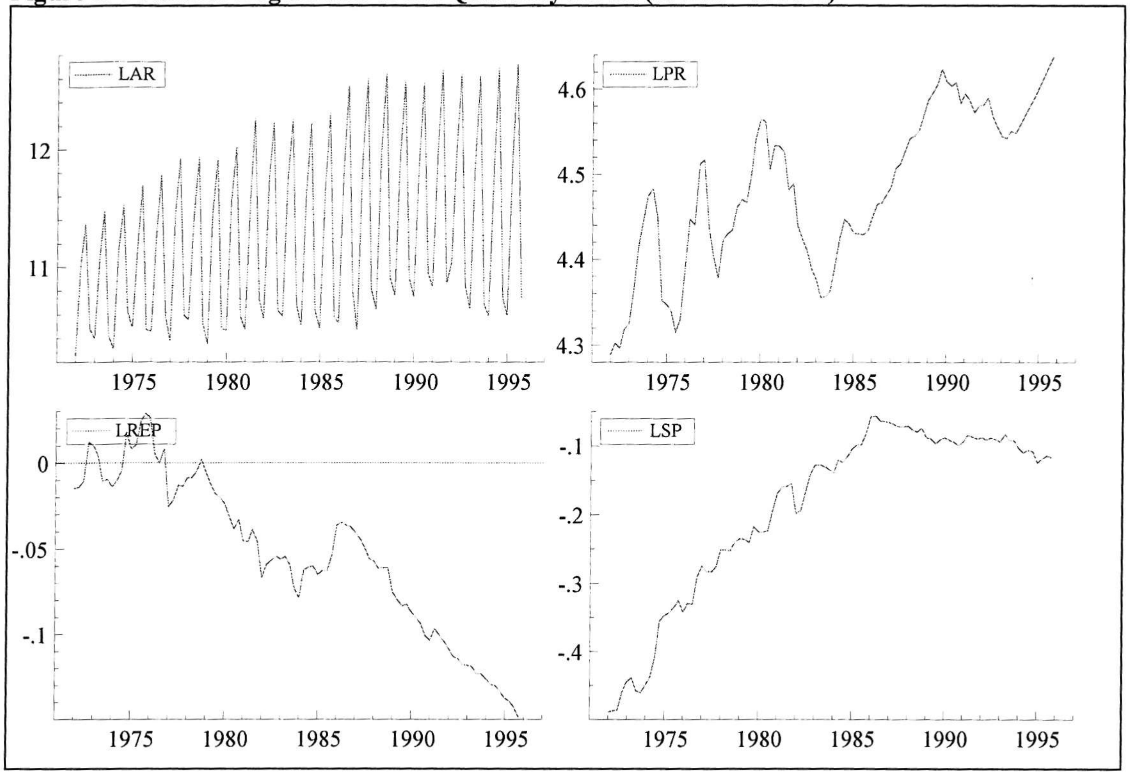
```
EQ(2) Modelling LAR by OLS  (using datiann.in7)
The present sample is:   3 to 24

Variable      Coefficient    Std.Error  t-value  t-prob PartR^2
LB                 1.0994      0.22351    4.919  0.0002  0.6173
Constant           6.8048       1.2542    5.425  0.0001  0.6624
LB_1             -0.55668      0.14618   -3.808  0.0017  0.4916
LSP               0.92023      0.20939    4.395  0.0005  0.5629
Trend            0.016069    0.0039243    4.095  0.0010  0.5278
LPR_1             0.36623      0.18431    1.987  0.0655  0.2084
RELB             -0.56571      0.21402   -2.643  0.0184  0.3178

R^2 = 0.989663  F(6,15) = 239.35 [0.0000] sigma = 0.0342615
DW = 1.93 RSS = 0.01760778592 for 7 variables and 22 observations

AR 1- 2 F( 2, 13) =     0.89447 [0.4326]
ARCH 1  F( 1, 13) =     0.24288 [0.6304]
Normality Chi^2(2)=     3.6818 [0.1587]
Xi^2     F(12,  2) =    0.080201 [0.9988]
RESET    F( 1, 14) =    1.0211 [0.3294]
```

The coefficient of *RELB* is statistically significant at the 5% level; as a result, the number of domestic arrival of boats can be treated as endogenous. This result does not confirm the finding obtained in the international demand for tourism, where the number of total (*i.e.* domestic and international) arrivals of boats has been found to be predetermined. It can be argued that the number of boats is more likely to be driven by the number of domestic rather than the number of foreign arrivals of tourists. Note, in fact, that the average share of domestic arrivals in northern Sardinia equals 81% for the period 1972-1995, whereas the average share of foreign arrivals is just 19%.

5.4 THE MODEL SPECIFICATION USING QUARTERLY DATA

Another aim of this analysis is to make a further comparison between monthly and annual data versus quarterly data. As a first step, the possible existence of quarterly seasonal unit roots as well as a long run unit roots is tested. The sample period under consideration is from 1972:1 to 1995:4. The variables under study are the unadjusted series of domestic arrivals (*LAR*), the industrial production index (*LPR*), the substitute price (*LSP*) and the relative price (*LREP*); all these variables are in logarithms. The graphs for the above variables are given in Figure 5.7.

Figure 5. 7 Natural Logarithms of the Quarterly Series (1972:1 - 1995:4)



To test for quarterly seasonal unit roots the auxiliary equation (2.6.1), given in Section 2.6, is run by OLS[53]. In this case μ_t consists of a constant, a trend and 3 quarterly seasonal dummies. An account is given of the main results obtained by fitting the equation, for each of the five time series above mentioned. The results are reported in Table 5.22.

Table 5. 22 Quarterly Seasonal Unit Roots

t-statistics	Variable			
	LAR	***LPR***	***LSP***	***LREP***
π1	-1.08	-3.51 *	- 0.37	-2.30
π2	-3.36 ***	-5.91 ****	2.93	-7.10 ****
π3	-2.37	-4.43 ****	-12.09 ****	-5.32 ****
π4	-0.78	-6.45 ****	3.32	-4.42 ****
F-statistics	***LAR***	***LPR***	***LSP***	***LREP***
π3, π4	3.14	45.80 ****	74.02 ****	32.79 ****

Notes: The four, three, two and one asterisks indicate that the seasonal unit root null hypothesis is rejected at the 1%, 5%, 10% and 20% level, respectively.

53 The auxiliary regression (2.6.1) is run using Microfit 4.0 package.

Testing for the presence of seasonal unit roots with respect to the series of domestic arrivals (*LAR*), the null hypothesis cannot be rejected at a general 5% level of significance[54], both running the *t*-tests for the separate π's (except π_2) and the joint test for π_3 and π_4. Such a variable does not appear to have a deterministic seasonal pattern and be stationary in the level. The last possibility has been tested further by an ADF test. From Table 5.23, it can be concluded that this series is stationary in the level, as the null hypothesis of the presence of a unit root cannot be accepted at a 1% level. It will be argued below that the apparent seasonal roots are in fact spurious, and it is structural changes in the seasonal pattern which determine this effect.

For the income proxy (*LPR*), the null hypothesis of seasonal unit root fails to be accepted at the general 1% level. The long run frequency unit root test suggests such a variable to be I(0). The latter result has been confirmed from the ADF test when including the constant term and a time trend. As in the monthly case, this variable can be considered as I(1) when either a constant or a constant and seasonals are included.

Even though π_4 is not statistically different from zero, the results from the joint *F*-test for the (log) substitute price (*LSP*) show that this variable does not present an irregular seasonal pattern. The null hypothesis of the presence of a long run unit root fails to be rejected at the 5% level. However, from the ADF test (see Table 5.23), *LSP* can be considered as a variable stationary in the level.

The seasonal unit roots test has been carried out for the (log) relative price (*LREP*) for the same period (1972:1-1995:4). Running the auxiliary regression (2.6.1), there appears no evidence for the presence of seasonal unit roots, denoting a regular seasonal pattern. However, the null hypothesis of non-stationarity cannot be rejected at a 5% level of significance, testing π_1=0 and using the *t*-test. From the ADF test (see Table 5.23), one can conclude that this variable is stationary in the first difference.

---

54 The critical values for the quarterly seasonal unit roots test are provided in Hylleberg *et al.* (1990) pp. 226-227. Note that in this case one takes into consideration the critical values for T=96 when the intercept, trend and seasonal dummies are included.

Table 5. 23 Augmented Dickey-Fuller Unit Root Test

Series	*ADF*(1)	*LAG*(2)
LAR(c)	-8.86**	0
LAR(c,t)	-9.53**	0
LAR(c,s)	-3.89**	0
LAR(c,t,s)	-7.11**	0
LPR(c)	- 2.19	1
DLPR(c)	- 6.79**	0
LPR(c,t)	- 3.59*	3
LPR(c,s)	- 1.88	1
DLPR(c,s)	- 6.57**	0
LPR(c,t,s)	- 3.58*	3
LSP(c)	- 3.16*	0
LSP(c,t)	- 0.72	3
DLSP(c,t)	- 7.22**	2
LSP(c,s)	- 3.03*	2
LSP(c,t,s)	- 0.43	4
DLSP(c,t,s)	- 7.18**	3
LREP(c)	-0.03	0
DLREP(c)	-9.42**	0
LREP(c,t)	-2.34	0
DLREP(c,t)	-9.49**	0
LREP(c,s)	-0.10	0
DLREP(c,s)	-8.83**	0
LREP(c,t,s)	-2.14	0
DLREP(c,t,s)	-8.91**	0

Notes: The one and two asterisks indicate that the unit roots null hypothesis is rejected at the 5% and 1% level, respectively. The capital letter *D* denotes the first-difference operator defined, in a general notation, by $Dx_t = x_t - x_{t-1}$.
(1) Augmented Dickey-Fuller statistics with constant, trend and seasonals (*i.e. c, t, s*) critical values = -3.46 at 5% and -4.062 at 1% level; with constant and trend c.v.= -3.458 at 5% and -4.059 at 1% level; with constant c.v. = -2.893 at 5% and -3.503 at 1% level.
(2) Lag is the length of the first significant lag. Note that ADF(0) corresponds to the Dickey-Fuller test; additional lags are included to whiten the residuals.

The previous analysis confirms the results obtained using monthly data. From the seasonal unit roots test one infers that all the variables under study denote a deterministic seasonal pattern, with the exception for the series of the domestic arrivals of tourists. Hence, the existence of a possible structural break requires investigation, as for the monthly case. From the long run unit roots test, it is confirmed that all the variables are stationary in the level, except the relative price which is I(1).

At this point, a brief note is due on the specification form chosen for the estimation. An ADF test is carried out for each of the economic series of interest. The results are the following: *AR* and *SP* are I(0); *PR* and *REP* are I(1). Once again, a contrast emerges for the income proxy *PR* with respect to the monthly case. However, the monthly ADF test gives more robust results as a greater number of observations are employed. Hence, one carries on using the same property for *PR* as for the monthly case (see Section 5.2.4). In order to run the Box and Cox test, an unrestricted five lag tourism demand equation is estimated, for the period 1972:1-1995:4, expressed both in a logarithmic and linear form. Given the previous definition for the explanatory variables, one estimates:

1) logarithmic form

$$LAR_t = a_1 + a_2\, LAR_{t-1}..+ a_3\, LPR_t..+ a_4\, LSP_t\, ..+ a_5\, DLREP_t\, ..+ a_6\, LW_t + a_7\, Trend + \\ + a_8\, Seas + e_t$$

and

2) linear form

$$AR_t = a_1 + a_2\, AR_{t-1}... + a_3\, PR_t... + a_4\, SP_t\, ..+ a_5\, DREP_t\, ..+ a_6\, W_t + a_7\, Trend + \\ + a_8\, Seas + e_t$$

Running the two equations, *SSELL* equals 0.4890330516 and *SSEL* equals 6.7E+09. The statistic χ^2 equals 36.29 and this value is greater than the correspondent critical value, 3.84, at the 5% level; hence, the null hypothesis cannot be accepted. Moreover, one infers that the logarithmic specification is better than the linear specification as $SSEL/(\overline{AR}\, G)^2$ (*i.e.* 1.095428) is greater than *SSEL*. One proceeds with the estimation of the quarterly domestic demand in a log-linear specification.

A generic unrestricted model is estimated for the domestic demand for tourism. This model includes the following variables: the (log) income proxy (*LPR*), the substitute price (*LSP*), the first difference of the relative price (*DLREP*), one impulse dummy (*i93q1*) which takes the value of one in the first quarter and zero otherwise, a trend in order to pick up possible changes in the consumers' tastes, the weather variable (*LW*), and, finally, 3 quarterly seasonal dummies.

Table 5. 24 *LAR* - Domestic Demand and the Unrestricted Quarterly Model

EQ(1) Modelling LAR by OLS (using quadr.in7)
The present sample is: 1973 (3) to 1995 (4)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	0.85122	1.9007	0.448	0.6558	0.0031
LAR_1	0.0064873	0.097673	0.066	0.9473	0.0001
LAR_2	-0.11113	0.097525	-1.139	0.2588	0.0199
LAR_3	-0.0070207	0.091499	-0.077	0.9391	0.0001
LAR_4	0.75186	0.089506	8.400	0.0000	0.5244
LPR	-0.040055	0.50119	-0.080	0.9365	0.0001
LPR_1	-0.37393	0.77933	-0.480	0.6330	0.0036
LPR_2	1.0246	0.77622	1.320	0.1915	0.0265
LPR_3	-0.18764	0.73873	-0.254	0.8003	0.0010
LPR_4	0.19763	0.45774	0.432	0.6674	0.0029
DLREP	1.1785	1.8187	0.648	0.5193	0.0065
DLREP_1	0.11851	1.7908	0.066	0.9474	0.0001
DLREP_2	1.2868	1.8564	0.693	0.4907	0.0075
DLREP_3	2.6670	1.7853	1.494	0.1401	0.0337
DLREP_4	-0.63122	1.4828	-0.426	0.6718	0.0028
LSP	0.33964	1.0469	0.324	0.7467	0.0016
LSP_1	-0.40238	1.6344	-0.246	0.8063	0.0009
LSP_2	0.47828	1.7217	0.278	0.7821	0.0012
LSP_3	0.080117	1.7395	0.046	0.9634	0.0000
LSP_4	0.12616	1.1879	0.106	0.9158	0.0002
Trend	-0.0012497	0.0012504	-0.999	0.3213	0.0154
Seasonal	0.12747	0.16117	0.791	0.4319	0.0097
Seasonal_1	0.13076	0.18043	0.725	0.4713	0.0081
Seasonal_2	0.19876	0.16854	1.179	0.2427	0.0213
LW	0.19250	0.16635	1.157	0.2515	0.0205
i93q1	-0.30297	0.095987	-3.156	0.0024	0.1347

R^2 = 0.990769 F(25,64) = 274.76 [0.0000] sigma = 0.0853171
DW =1.65 RSS = 0.465856358 for 26 variables and 90 observations

AR 1- 2 F(2, 62) =	2.7346 [0.0728]	
ARCH 1 F(1, 62) =	0.1123 [0.7387]	
Normality Chi^2(2)=	0.4137 [0.8131]	
Xi^2 F(46, 17) =	0.2850 [0.9996]	
RESET F(1, 63) =	6.0267 [0.0169]	*

An initial five quarters lag structure model is run which could be reduced to a four quarter lag structure, as suggested by the joint *F*-test and the SC criterion. However, such a model presents problems in terms of specification when testing at a 5% level with respect to the *RESET* test (see Table 5.24).

A structural break analysis is, therefore, attempted as in the monthly data case. Preliminary investigations of a structural break in all coefficients (*i.e.* the coefficients of the seasonal dummies, *LAR*, *LPR*, *DLREP* and *LSP* respectively) in the unrestricted model are reported in Table 5.25. Running a Chow test (1967) the conventional *F* statistic indicates the absence of structural changes[55]. In particular, the Chow test

[55] Note that the possible existence of a structural change is detected by moving the change point forward one year at a time. The test is created with TSP Version 4.3A.

suggests that the main changes have occurred between 1978/79, since the appropriate F statistic show the greatest values. However, the F statistic, is smaller than the critical value of the F distribution with q=20 and $N+M-2K$=48 degrees of freedom, that is smaller than 1.84 at a 5% level. Given the multiple comparisons involved, one should use Andrews' (1993, p.840) critical values. At the 5% level the critical value equals 2.05, that is the critical value obtained when Andrews's π=0.25[56] and 20 restrictions are considered. Hence, from the calculated values in Table 5.25, it appears that the null hypothesis of no structural change between the period 1978/79 can be accepted at the 5% level.

Table 5. 25 *LAR*- RSS for Unrestricted and Restricted Models and *F-test* with Quarterly Data

73Q3 95Q4	**RRSS** =	0.471603			
		UNRSS			
73Q3 78Q4		0.28275	F(20, 48)	=	**1.57**
73Q3 79Q4		0.30211	F(20, 48)	=	1.32
73Q3 80Q4		0.31278	F(20, 48)	=	1.19
73Q3 81Q4		0.32012	F(20, 48)	=	1.11
73Q3 82Q4		0.33318	F(20, 48)	=	0.98
73Q3 83Q4		0.32622	F(20, 48)	=	1.05
73Q3 84Q4		0.32892	F(20, 48)	=	1.02
73Q3 85Q4		0.32605	F(20, 48)	=	1.05
73Q3 86Q4		0.36997	F(20, 48)	=	0.65
73Q3 87Q4		0.35739	F(20, 48)	=	0.75
73Q3 88Q4		0.30197	F(20, 48)	=	1.32
73Q3 89Q4		0.30737	F(20, 48)	=	1.26

However, the results have been investigated further by running a model in which just the seasonal coefficients are allowed to change, as in the monthly data case. The results are reported in Table 5.26.

56 Note that π is given by the following formula: *½[OB/N]*, where *OB* is the total number of omitted observations considering each time period (46 in this case) and *N* is the total number of observations (91 in this case).

Table 5. 26 *LAR*- Chow Test for 4 Seasonal Coefficients Changes, using Quarterly Data

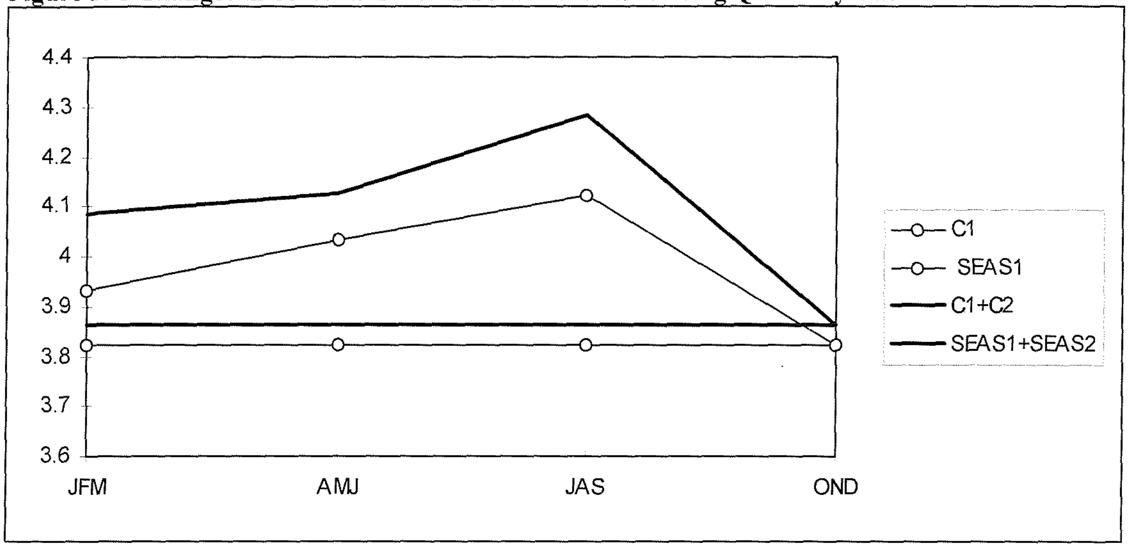
73Q3 95Q4	**RRSS=** 0.471603			
		UNRSS		
73Q3 78Q4		0.40621	F(4,64)=	2.53
73Q3 79Q4		0.41197	F(4,64)=	2.28
73Q3 80Q4		0.39835	F(4,64)=	2.90
73Q3 81Q4		0.39916	F(4,64)=	2.86
73Q3 82Q4		0.41125	F(4,64)=	2.31
73Q3 83Q4		0.40503	F(4,64)=	2.59
73Q3 84Q4		0.40658	F(4,64)=	2.52
73Q3 85Q4		0.39203	F(4,64)=	**3.20**
73Q3 86Q4		0.42737	F(4,64)=	1.63
73Q3 87Q4		0.41087	F(4,64)=	2.33
73Q3 88Q4		0.39724	F(4,64)=	2.95
73Q3 89Q4		0.40988	F(4,64)=	2.37
73Q3 90Q4		0.39531	F(4,64)=	**3.04**
73Q3 91Q4		0.41804	F(4,64)=	2.02
73Q3 92Q4		0.40087	F(4,64)=	2.78
73Q3 93Q4		0.45298	F(4,64)=	0.65
73Q3 94Q4		0.45763	F(4,64)=	0.48

The statistical values reported in Table 5.26 can be compared with the asymptotic critical values provided by Andrews (1993 p.840). For Andrews' π=0.20[57] the critical value for four restrictions equals 3.96 at the 5% level. The test does not suggest the presence of changes in the seasonal pattern. However, comparing the calculated values in Table 5.26 with the conventional critical value at the 5% (the tabulated value equals 2.53) one can suspect that a change in the seasonal pattern may have occurred between 1985/86 and 1990/91 respectively.

One can be encouraged to experiment further as the unrestricted model, without allowing for a seasonal change, has presented problems of form specification, as previously mentioned. A more detailed inspection of the changes in the seasonal pattern can be carried out graphically (see Figures 5.8 and 5.9). This investigation can be considered a rough comparison, fitting only one change in each case.

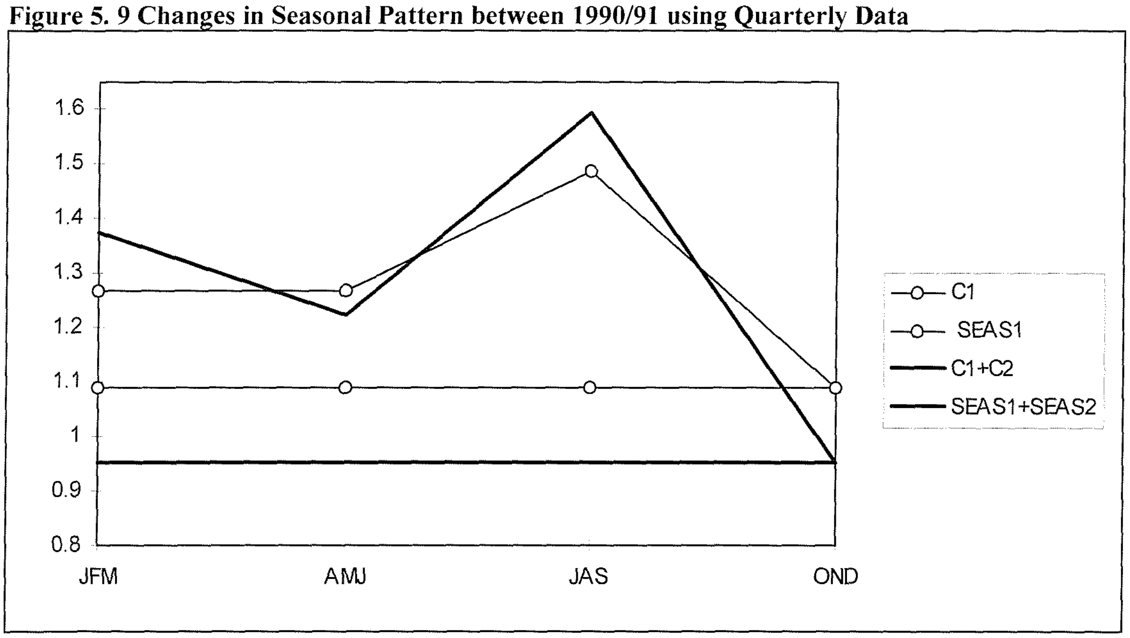
57 In this case π is given by: ½ *[OB/N]*, where *OB* equals 34 and *N* equals 91.

Figure 5. 8 Changes in Seasonal Pattern between 1985/86 using Quarterly Data



Notes: *C1* represents the coefficient for the intercept; *Seas1* represents the coefficients of the various "non changing" seasonal dummies, for quarters 1, 2 and 3; *C1+C2* represents the the sum of the coefficient for the "non changing" and "changing" intercept; *Seas1+Seas2* represents the sum of the coefficients for the "non changing" and "changing" seasonals for quarters 1, 2 and 3, respectively.

Figure 5. 9 Changes in Seasonal Pattern between 1990/91 using Quarterly Data



Note as in Figure 5.8.

Between 1985/86, the greatest changes in the seasonal pattern occur in January-February-March (first quarter, say *JFM*), April-May-June (second quarter, say *AMJ*) and July-August-September (third quarter, say *JAS*), with an increase in the number of arrivals in the second period (*i.e.* from the first quarter of 1986). Note that the arrivals in the third quarter double the domestic arrivals in the first quarter.

For the structural break between 1990/91, the main changes in the seasonals occur in the fourth quarter (October-November-December, say *C*) with a decrease in the number of domestic arrivals. On the other hand, the first quarter and third quarter show an overall increase of the number of domestic arrivals.

Such assumptions have been investigated further. Three separate dummies are fitted; for the whole period (*JFM*, *AMJ*, *JAS*, *C*), for 1986:1 onwards (*JFM2*, AMJ*2*, and so on), and for 1991:1 onwards (*JFM3*, *AMJ3*, and so on). Firstly, the structural break between 1985/86 has been considered. The F statistic (1,64), when the coefficients for the seasonals *JFM2*, *AMJ2* and *JAS2* are allowed to change between 1985 and 1986, has to be accepted at the 5% level. In fact, the F statistic (1,64) calculated equals 0.54 and it is smaller than the critical value (*i.e.* 4.00).

Secondly, the structural seasonal change between 1990/1991 has been investigated by testing for possible restrictions on the coefficients of the seasonals: *AMJ3* that is, allowing just the coefficients of the seasonals *C3, JFM3* and *JAS3* to change. The F statistic (1,64), in such a case, equals 1.51 that is smaller than the conventional critical value at the 5% level (*i.e.* 4.00). Thus, the null hypothesis cannot be rejected.

One has compared the restricted seasonal changes with the unrestricted dummy model. This is not a full specification search. However, the same results are obtained when testing the 1985/1986 changes after imposing the 1990/91 seasonal changes. The restriction on the coefficient of *C2* does still hold. That is the F statistic (1, 61) equals 1.57 and this value is smaller than 4.00 at the 5% level from the conventional table.

The same investigation has been done for the structural change occurring between 1990 and 1991, after imposing the changes in the seasonals from 1986:1 until 1990:4. The F statistic (1,62) suggests that the restriction on *AMJ3* holds. The calculated value equals 3.57 that is smaller than the critical value (*i.e.* 4.00) at the 5% level.

After assessing when and which seasonals are changing over the period under study, 1972:1-1995:4, three sets of seasonal dummies are created. The first set of dummies is the following: *C*, *JFM*, *AMJ* and *JAS* for the first period (1972:1-1985:4). Note that *JFM* and *JAS* take the value 1 in the first period (1972:1-1985:4) and zero otherwise. Whereas, *AMJ* takes the value 1 in the first and third period and zero otherwise. The second set of seasonal dummies allowing for the structural change in the second period (1986:1-1989:4) is the following: *JFM2* and *AMJ2*. Finally, the third set contains: *C3, JFM3* and *JAS3* for the third period (1990:1-1995:4).

An initial unrestricted model has been run with four lags for each of the independent variables, which has not presented any problems in terms of diagnostic statistics. The test of specification suggests as the changes in the seasonals had to be taken into consideration, since the null hypothesis for the *RESET* test has been accepted. The final parsimonious model, as suggested by the joint *F*-test and SC criterion, is given in Table 5.27.

Table 5. 27 *LAR* - Domestic Demand Final Restricted Model for Quarterly Data

```
EQ(2) Modelling LAR by OLS  (using quadr.in7)
The present sample is:  1973 (3) to 1995 (4)
```

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	2.9892	1.2980	2.303	0.0241	0.0677
LAR_1	0.25021	0.091740	2.727	0.0080	0.0925
LAR_2	-0.36846	0.096496	-3.818	0.0003	0.1665
LAR_3	0.28141	0.092796	3.033	0.0034	0.1119
LAR_4	0.38318	0.092833	4.128	0.0001	0.1892
LPR_3	0.44135	0.15085	2.926	0.0046	0.1050
LSP_1	0.53133	0.19982	2.659	0.0096	0.0883
i93q1	-0.27413	0.088507	-3.097	0.0028	0.1162
JFM	0.17345	0.12477	1.390	0.1687	0.0258
AMJ	0.22005	0.13621	1.615	0.1105	0.0345
JAS	0.61074	0.11868	5.146	0.0000	0.2662
JFM2	0.21725	0.14905	1.457	0.1493	0.0283
AMJ2	0.20812	0.15561	1.337	0.1852	0.0239
JAS2	0.73574	0.14586	5.044	0.0000	0.2585
C3	-0.10580	0.041729	-2.535	0.0134	0.0809
JFM3	0.41480	0.16725	2.480	0.0154	0.0777
JAS3	0.91124	0.16173	5.634	0.0000	0.3031

```
R^2 = 0.991676   F(16,73) = 543.55 [0.0000]   sigma = 0.0758582   DW =
2.11 RSS = 0.4200761679 for 17 var. and 90 obs.

AR 1- 5 F( 5, 68) =      1.9532 [0.0969]
ARCH 4   F( 4, 65) =    0.83848 [0.5058]
Normality Chi^2(2)=      3.3303 [0.1892]
Xi^2     F(22, 50) =    0.82229 [0.6853]
RESET    F( 1, 72) =     1.5844 [0.2122]
Tests of parameter constancy over: 1994 (1) to 1995 (4)
Forecast Chi^2( 8)=     14.272 [0.0749]
Chow     F( 8, 65) =    0.95023 [0.4823]
```

The *F*-test for the significance of the variables under study shows that all regressors have a significant explanatory role. The diagnostic statistics suggest that the estimated model is statistically well-specified and constitutes an admissible reduction of the underlying unrestricted model. In general, the model appears to be satisfactory as it is able to explain 99% of the dependent variable variation; the goodness of the fit is also indicated by the ratio of the *SER* and the *MDV* that is equal to 0.0067[58].

[58] SER/MDV=(0.0758582/11.267)=0.0067

Furthermore, the null hypothesis of parameter constancy over the last eight observations fails to be rejected.

The dependent variable seems to be highly influenced by its own history. The coefficient of the third quarter lag of *LPR* shows a positive sign indicating that the higher the income the higher the domestic flows of tourism. As in the monthly and annual analysis, the substitute price shows a positive sign that might be detecting a possible spurious correlation suggesting that it may be picking up other effects not expressly included in the model. The impulse dummy is statistically significant and presents a negative sign. Hence, it is possibly detecting the negative effects on the Italian economic recession in the early Nineties which caused a decline of the domestic demand for tourism, as it can be seen in Figure 3.1 (Chapter 3).

The coefficient of the growth of the relative price, the time trend and the "weather" dummy do not enter in the final restricted model. In general, the quarterly seasonal dummies, allowing for the seasonal pattern changes, appear to be statistically significant.

The long run dynamics are reported in Table 5.28 and can be compared with the results obtained in the previous analyses.

Table 5. 28 Solved Static Long Run Equation

LAR =	+6.589	+0.9729 LPR	+1.171 LSP
(SE)	(1.56)	(0.3347)	(0.2309)
	-0.6043 i93q1	+0.3823 JFM	+0.485 AMJ
	(0.2695)	(0.3044)	(0.3358)
	+1.346 JAS	+0.4789 JFM2	+0.4588 AMJ2
	(0.4467)	(0.3589)	(0.3658)
	+1.622 JAS2	-0.2332 C3	+0.9143 JFM3
	(0.5214)	(0.1298)	(0.4621)
	+2.009 JAS3		
	(0.6411)		

```
ECM = LAR - 0.972862*LPR - 1.1712*LSP + 0.604269*i93q1 -
0.382323*JFM - 0.485043*AMJ - 1.34625*JAS - 0.478869*JFM2 +
- 0.458753*AMJ2 - 1.62176*JAS2 + 0.233213*C3 - 0.914323*JFM3 +
- 2.00863*JAS3 - 6.58901;

WALD test Chi^2(12) = 161.82 [0.0000] **
```

The quarterly model shows a good specification also in the long run with relatively small standard errors, as given in parenthesis. The long run coefficients are statistically significant in general at the 5% level. Moreover, from the Wald test one cannot accept the null hypothesis, thus, the long run coefficients are jointly different from zero. The coefficient for the income proxy presents the positive correct sign.

However, once again, the substitute price shows a positive sign confirming the results obtained both in the monthly and annual models.

5.5 SUMMARY

This section is dedicated to a feedback of the main results into economic theory. The findings in terms of income and price elasticities are reported both in terms of short and long run dynamics. The analysis will be divided with respect to the data frequency of the model. The findings are summarised in Table 5.29.

Table 5. 29 Short and Long Run Elasticities for the Domestic Demand of Tourism

Elasticities	**Monthly Model (288 obs.) (Tab. 5.14 - 5.15)**	**Annual Model (24 obs.) (Tab.5.18)**	**Quarterly Model (96 obs.) (Tab. 5.27 - 5.28)**
INCOME (long run)	0.88 (3.89)	0.47 (2.57)	0.97 (2.91)
INCOME (short run)	0.31 (3.60)	=	0.44 (2.93)
REL.PRICE (long run)	-	-	-
REL.PRICE (short run)	-	-	-
SUB.PRICE(long run)	1.05 (6.10)	1.49 (7.55)	1.17 (5.07)
SUB.PRICE(short run)	1.71 (2.53)	=	0.53 (2.66)

Notes: (1) *t*-values are given in parenthesis.
(2) Note that the short run elasticity corresponds to the first significant lag in the model (see Pindyck and Rubinfeld, p. 377, 1991).

In terms of long run income elasticity the values are similar in the monthly and quarterly cases. However, in each of the three models their value is below unity. This means that Italians view domestic tourism as a necessity good. This finding is in line with the findings obtained in Malacarni (1991), who finds an income long run elasticity of 0.92, when estimating the domestic demand of tourism in Italy. As argued for the international case, it seems more appropriate to use monthly data in estimating the demand for tourism. This time frequency, in fact, is more consistent with the differences existing in the tourists' behaviour. Monthly data give more insight in understanding and differenciate the consumers' decision taking in the timing of their holiday, as well as their preferences for a certain destinations.

The price elasticity turned out to be statistically insignificant in each of the data frequency models. Hence, the findings obtained by Malacarni (1991) cannot be compared. In Malacarni's empirical work a price elasticity of –0.16 is estimated for Italian tourism demand using annual time series with 17 observations.

The substitute price elasticities are positive in both the long and short run in each of the three models. Possible problems might be derived from not having taken into consideration the exchange rate for the main north of Sardinia competitors as an explanatory variable. As already stated, such a hypothesis will be investigated in Chapter 6.

5.6 Conclusion

A phenomenon characterising almost all tourism is that of seasonality. Such a phenomenon is particular evident in countries or regions (such as Sardinia) in which tourists search for "sea and sun" holidays. From this research a different seasonal pattern has been detected between foreign and domestic tourist flows to Sassari Province. This finding has suggested modelling the two components separately.

Chapter 5 has been dedicated to the empirical study of the domestic demand for tourism to the north of Sardinia for the period between 1972 and 1995. An investigation of the possibility of an irregular seasonal pattern for the variables under study (*i.e.* domestic arrivals of tourists, the Italian index of industrial production, relative price - Sassari/Italy - and substitute price) has been carried out by means of testing for the possible existence of seasonal unit roots. Such a possibility has been rejected for all the series, with the only exception for the series of domestic arrivals. However, from a deeper investigation, the dependent variable has shown a structural break in the seasonal pattern. This finding has confirmed that the apparent seasonal roots from Franses' test are spurious and it is changes in the seasonal pattern which give rise to this effect. Calculated values have been compared both with the conventional tables and with Andrews' critical values, as a multiple comparison was involved. The integration status of the variables under study has been tested by using an ADF unit roots test. All the variables of interest have been found to be stationary in the level, with the only exception for the relative price, which is stationary in the first difference.

An important step has been investigating the effect of a correction of the domestic arrivals series for the number of weekends (Saturdays or Sundays) in a month. Such a normalisation has not been found to be appropriate as the best results are obtained with the unadjusted series. One can argue that domestic tourists are less constrained in the choice of the day of arrival, as they are likely to be able to catch a boat or a flight any day of the week. On the other hand, those foreign tourists who use charter flights or international boats are more likely to travel on Saturdays as the analysis in Chapter 4 has suggested.

Another aim of the chapter has been the use of different data frequencies: monthly, quarterly and annual data. The aim was to assess the characteristics of the domestic demand for tourism in the short and long run, as well as to assess the advantages and disadvantages in using each data frequency.

The advantages in using monthly data are as follows. Firstly, such a frequency has allowed to test for the possible presence of seasonal unit roots and to identify the integration status of the variables of interest, thanks to the relative large sample (T=288). However, it is interesting to note that monthly and quarterly data (T=96) have provided homogenous results in terms of seasonal and long run unit roots testing, confirming in this way the findings obtained from the international tourism demand. Moreover, the two data frequencies have given similar results in terms of testing for structural breaks. In both cases, a structural break in the seasonal pattern has been detected in the second half of the Eighties and in the first half of the Nineties.

Another advantage from the monthly series is the possibility to study the domestic demand dynamics. The final parsimonious model has given satisfactory results as 99% of the variance of the dependent variable has been explained. The inclusion of the "Easter" dummy has turned out to have an important role in explaining the pattern of tourism in the north of Sardinia, confirming the results in the international demand case. This model has allowed the study of short run as well as the long run dynamics. As an example, in the short and long run the Italian index of industrial production has presented a positive sign which is in line with economic theory and other empirical studies for Italy as a whole.

The model obtained using quarterly data has given satisfactory results. The coefficient of determination is 99% and the coefficients are overall well-specified in terms of statistical significance and signs. In both models the weather variable has turned out not to be statistically significant.

Another aim was to compare disaggregated and aggregate frequencies. One will refer to the advantages and disadvantages in using annual data. An important advantage that has revealed in using aggregated data has consisted in the possibility of making use of data that are often available only at an annual frequency. In this case, it has been possible to test and establish whether the Italian production index can indeed be considered as a valid proxy of the personal disposable income. Statistical evidence has been given from several statistics and tests such as the simple cross-correlation. From the single equation model, estimating *LPDIN* on *LPR*, the t-value for the coefficient of *LPR* is equal to 6.69. Finally, from the VAR analysis, the first lag of *LPR* plays an important role in explaining *LPDIN* and almost all its variance is explained. By using annual data, it has also been possible to take into consideration supply components such as the number of domestic boat and flight arrivals in the north of Sardinia. Evidence has been found for the arrivals of national boats to be an endogenous variable.

The disadvantages incurred are related to the power of testing with a small number of observations. No conclusive judgement could be made in terms of the integration status of the variables of interest. The same problem has occurred when using Wu-Hausman's test for testing the simultaneity of the supply component. However, the final parsimonious model does show an explanatory power of almost 98% and the model is well-specified in terms of diagnostic tests. Note, that this model converges rapidly to the long run equilibrium.

A further analysis has been carried out for giving statistical evidence in using the log-linear specification form. By using the Box and Cox (1964) test, it has been established that the logarithmic form is a better specification than the linear form. The same results have been achieved in each of the models (*i.e.* monthly, annual and quarterly). However, some divergence has been found in the properties of the economic variables of interest. The ADF test seems to give better results employing series at a monthly frequency.

On balance, one can conclude a better specification and findings seem to be given by using monthly or quarterly data. Nevertheless, one would not want to omit the findings obtained with aggregated annual data.

CHAPTER 6.

SASSARI PROVINCE COMPETITORS: REAL SUBSTITUTE PRICE

Aim of the Chapter:

To identify the impact of the real substitute price on domestic and international tourism demand in Sassari Province of Sardinia, Italy.

6.1 INTRODUCTION

As already stated, according to economy theory, one would have expected a negative sign for the substitute price which measures the price of tourism in Sassari Province relative to that in other destinations in the Mediterranean area. When the consumer price index in Sassari Province gets higher, *ceteris paribus*, one would expect a decrease in the number of arrivals. However, in the analysis in this thesis so far a positive sign has been found, which might be indicating problems of mis-specification.

So far, in both the domestic and international demand model, one has included the nominal substitute price as a possible determinant of the demand for tourism. However, one can argue that tourists are more aware of exchange rates than the cost of living in a foreign country. On this basis, Gray (1966), one of the pioneers in the tourism literature, includes the exchange rate alone as a proxy for the cost of living in the destination country. In many other empirical studies of tourism, prices and exchange rates have been combined and used separately as explanatory variables. Witt and Witt (1992), for example, argue that "consumer price index (either alone or together with the exchange rate) is a reasonable proxy for the cost of tourism. The exchange rate on its own, however, is not an acceptable proxy" (p.46). As Crouch (1994) writes, it does not appear clear cut in the literature which the best selection is: "How should changes in exchange rates be modelled? Would it be best to adjust prices for changes in exchange rates, or do tourists respond differently to exchange rate changes than they do to price changes?" (p.48). It seems clear that an answer cannot be unequivocal, as the statistical significance of these variables, used either separately or jointly, is highly dependent on the origin country and/or destination country under study (see Martin and Witt 1988; Witt and Witt, 1992). Some studies, amongst the others Gonzàles and Moral (1995), show that real relative prices and real substitute prices have an important role in explaining the international demand for tourism in Spain.

Another argument for including the exchange rate in the model is related to the Purchasing Power Parity. The theory of PPP assumes that, ignoring trade barriers and transportation costs, in the long run exchange rates should perfectly reflect the relative costs of living between countries. However, substantial deviations between prices and exchange rates are highly likely to occur in the short run (see among the others Lee, *et al.*, 1996; Garratt *et al.*, 1998). On this argument, statistical evidence has been given in Chapter 4.

The aim of this chapter is to carry out a further investigation of the characteristics of the real substitute price and its components. The plan of the chapter is as follows. The first section is dedicated to the analysis of the domestic demand for tourism in the north of Sardinia when taking into consideration the weighted average exchange rate for the other main destination countries in the Mediterranean area (*i.e.* France, Greece, Portugal and Spain). The aim is to determine whether the real substitute price, given by the difference between the (log) substitute price and the (log) weighted average exchange rate, plays a role in explaining the domestic demand for tourism. The second section is dedicated to the analysis of the characteristics of the individual components of the real substitute price. The main conclusion is that one can define and use a real substitute price for each of the main competitor countries of Sassari Province. In the third section, a model is estimated for the domestic demand for tourism in the north of Sardinia with the inclusion of the real substitute prices for each of the competitor countries under analysis. In the fourth section, an international demand model is estimated including the real substitute price for each of the main competitor countries. As will be assessed in the first section, the weighted average exchange rate includes the exchange rate lira/franc and the exchange rates for the other three countries. The latter are calculated as a ratio between lira/dollar and drachma/dollar, escudo/dollar and pesetas/dollar, respectively, as the series are readily available in the statistical sources. Hence, the weighted average exchange rate for the competitors is not employed in estimating the international demand for tourism in order to avoid possible problems of multicollinearity. Note, in fact, that this model includes the real substitute price for each of the individual source countries. The last section concludes the chapter.

6.2 REAL SUBSTITUTE PRICE FOR THE COMPETITORS AND DOMESTIC DEMAND

This section is dedicated to the investigation of the real substitute price as a possible determinant of the domestic demand for tourism in the north of Sardinia. So far, one has included only the nominal substitute price as an explanatory variable for the number of tourists' arrivals. However, one can argue that consumers are more aware of the exchange rates than the cost of living in a foreign country. The aim is to estimate a model in which the substitute price and exchange rate are expressed as weighted averages of the consumer price indices and exchange rates, respectively, for the main competitor countries in the Mediterranean area. The analysis is carried out with monthly data for the period between 1972:1 and 1995:12. The use of a monthly frequency takes into consideration seasonal effects, such as the "Easter" effect and structural changes in the seasonal pattern, which have been analysed in Chapter 5. Moreover, it gives more robust results in identifying the properties of the variables of interest, as one employs 288 observations.

Figure 6.1 shows the (log) weighted average exchange rate lira/currencies for the main destination countries (*LWTC*), the (log) nominal substitute price (*LSP*) and the difference between the two variables which gives the real substitute price (*LRSP*). Hence, *LRSP* can be defined as:

$$LRSP_t = LSP_t - LWTC_t$$

where:

SP_t = ratio between the consumer price index in Sassari Province and the weighted average consumer price index in the main competitor countries (*i.e.* France, Greece, Portugal and Spain).

WTC_t = weighted average exchange rate, lira/currencies for the competitor countries, calculated as follows:

$$WTC_t = \frac{\sum_{i=1}^{i=4} a_{i,t} * TC_{i,t}}{\sum_{i=1}^{i=4} a_{i,t}} \qquad (6.2.1)$$

where:

i = France, Greece, Portugal and Spain.

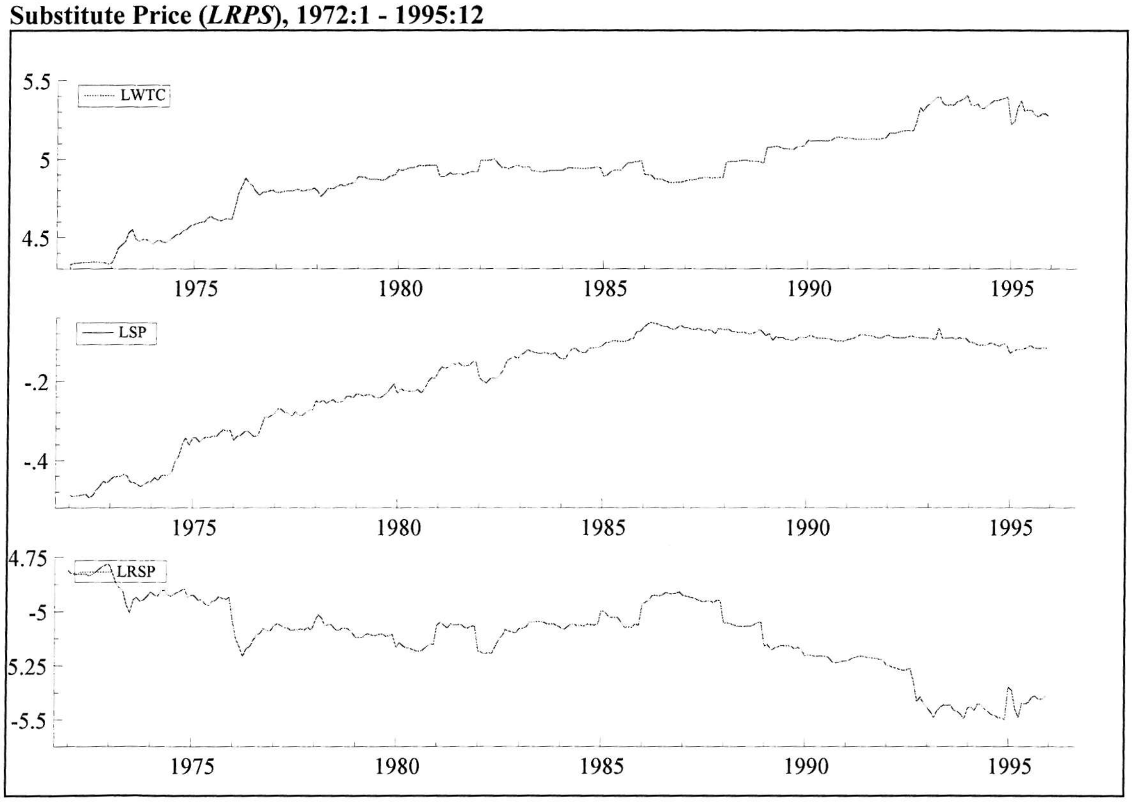
$TC_{i,t}$ = monthly exchange rate, lira/currency in country i and month t (Source: Banca d'Italia and IFS). Note that the exchange rates lira/drachma, lira/escudo and lira/pesetas are obtained from the ratio between lira/dollar and drachma/dollar (and so on).

$\alpha_{i,t}$ = weights are defined as

$$\alpha_{i,t} = \frac{AR_{i,t}}{\sum_{i=1}^{i=4} AR_{i,t}} \qquad (6.2.2)$$

where $AR_{i,t}$ are the number of Italian arrivals in each of the destination country (i) that is France, Greece, Portugal and Spain (Source: OECD and WTO). These weights are allowed to vary annually.

Figure 6. 1 Log: Weighted Average Exchange Rate (*LWTC*), Substitute Price (*LSP*) and Real Substitute Price (*LRPS*), 1972:1 - 1995:12



As already stated, the nominal substitute price (*LSP*) is stationary in the level if a constant, or a constant and seasonals are included in the unit root ADF test (see Table 5.2). Thus, one is interested in the integration status of the weighted average exchange rate and the real substitute price. An ADF unit root test with up to 13 lags is applied, as one is dealing with monthly data. Both the (log) weighted exchange rate and (log) real substitute price have been found to be I(1). This finding holds for each of the ADF cases, that is when including only the constant, the constant and seasonals and so on.

The next step is to use a Johansen cointegration analysis in order to test for a possible cointegration relationship between the (log) nominal substitute price and (log) weighted average exchange rate for the main competitors. Using the joint *F*-test, an initial unrestricted 13 lag system can be reduced to a 5 lag system when including only a constant term, and to a 7 lag system when including a constant and seasonals. Note that the information criteria SC, HQ and AIC are minimised when further coefficient restrictions are imposed in both the cases. From the cointegration analysis, using both the 7 and 5 lag system, there is no evidence for the two variables to be cointegrated. The same results have been obtained when including in the system the constant, and the constant and trend unrestrictedly.

The joint *F*-test and the other information criteria lead to the use of the first difference of the real substitute price, and thus consider only the short run adjustments. However, it is difficult to see how this variable can really be I(1); PPP must have some force, even if only in the long run. One does need to assume that consumers have a perfect knowledge of exchange rates as well as the cost of living in their own and competitor countries. Further, using only the first difference excludes consideration of the long run effect that is a strong assumption. Thus, one will include in the equation the level of the real substitute price and the level of the relative price (Sassari/Italy), *LREP* (see Chapter 5 for the definition).

The model for the domestic demand for tourism (*LAR*) contains the following determinants: the income proxy (*LPR*), the relative price (*LREP*), the real substitute price (*LRSP*), the "weather" variable (*LW*), the "Easter" dummy, the seasonal dummies that allow for structural breaks in the seasonal pattern in accordance with the findings in Chapter 5 and, finally, an impulse dummy (*i1992p3*) created in order to avoid non-normality problems in the residuals. As already mentioned, sometime it is difficult to give an economic interpretation of outliers when employing monthly series. Possibly, this impulse dummy is detecting some kind of effects determined by an unknown special event. The time trend, that can pick up possible changes in the tastes of the consumers, is not included as it has not been found statistically significant from a preliminary investigation.

An initial 13 lag model can be reduced to a 11 lag model, according to the joint F-test; however, this model presents problems of serial correlation (1% level) in the residuals; note also that the SC criterion is minimised for a further parameter reduction[59] that leads to further problems in the residuals. Hence, a 12 lag specification has been pursued since it has not presented problems in the residuals and it can be reduced parsimoniously as reported in Table 6.1.

59

Model statistics

dep.var		T	k	df	RSS	sigma	Schwarz
1: LAR	OLS	275	34	241	2.4386	0.100592	-4.03091
....		....	....	....	....	....	....
10: LAR	OLS	275	70	205	1.66182	0.0900357	-3.67913
11: LAR	OLS	275	74	201	1.53034	0.0872561	-3.67986
12: LAR	OLS	275	78	197	1.46048	0.0861023	-3.64488
13: LAR	OLS	275	82	193	1.43814	0.0863221	-3.5786

Model 13 --> 12: F(4, 193) = 0.74951 [0.5594]
Model 12 --> 11: F(4, 197) = 2.3557 [0.0551]
Model 11 --> 10: F(4, 201) = 4.3173 [0.0023] **

Table 6. 1 Final Domestic Demand Model with the Real Substitute Price

EQ(1) Modelling LAR by OLS (using vdomlar1.in7)
The present sample is: 1973 (2) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	0.53228	0.32402	1.643	0.1017	0.0110
LAR_1	0.43170	0.041543	10.392	0.0000	0.3077
LAR_11	0.17480	0.044687	3.912	0.0001	0.0592
LAR_12	0.20802	0.052758	3.943	0.0001	0.0601
RLPR	-0.66594	0.19159	-3.476	0.0006	0.0474
LPR_11	0.25561	0.081803	3.125	0.0020	0.0386
RLRSP	0.36688	0.13988	2.623	0.0093	0.0275
E	0.15638	0.031416	4.978	0.0000	0.0925
i1992p3	0.32633	0.088335	3.694	0.0003	0.0532
jan	0.035063	0.029328	1.196	0.2330	0.0058
jan2	0.0010724	0.039889	0.027	0.9786	0.0000
feb	0.049811	0.031068	1.603	0.1102	0.0105
mar	0.10379	0.045561	2.278	0.0236	0.0209
apr	0.35154	0.052659	6.676	0.0000	0.1550
apr2	0.52523	0.061779	8.502	0.0000	0.2293
apr3	0.33165	0.062832	5.278	0.0000	0.1029
may	0.30845	0.047316	6.519	0.0000	0.1489
may3	0.27609	0.063672	4.336	0.0000	0.0718
jun	0.31466	0.050897	6.182	0.0000	0.1359
jun2	0.51268	0.068053	7.534	0.0000	0.1893
jun3	0.53526	0.071802	7.455	0.0000	0.1861
jul	0.43281	0.058016	7.460	0.0000	0.1864
jul2	0.52693	0.076463	6.891	0.0000	0.1635
jul3	0.52461	0.081036	6.474	0.0000	0.1471
aug	0.55521	0.058066	9.562	0.0000	0.2734
aug2	0.73178	0.077664	9.422	0.0000	0.2676
aug3	0.80519	0.082661	9.741	0.0000	0.2808
sep	0.43160	0.069744	6.188	0.0000	0.1361
sep3	0.29939	0.091322	3.278	0.0012	0.0424
oct	-0.20430	0.058691	-3.481	0.0006	0.0475
oct3	-0.37688	0.070487	-5.347	0.0000	0.1053
nov	-0.077122	0.030106	-2.562	0.0110	0.0263

R^2 = 0.989327 F(31,243) = 726.58 [0.0000] sigma = 0.0852015
DW = 2.00 RSS = 1.76400701 for 32 var. and 275 obs.

AR 1- 5	F(5,238) =	1.244 [0.2891]
ARCH 4	F(4,235) =	0.584 [0.6743]
Normality	Chi^2(2)=	0.077 [0.9623]
Xi^2	F(37,205) =	0.737 [0.8652]
RESET	F(1,242) =	2.275 [0.1328]

where:

a) *LAR* = domestic arrivals of tourists in the north of Sardinia.

b) *LPR* = income proxy as industrial production index in Italy.

c) *RLPR* = Restriction on the coefficients for the second and third lag of the income proxy[60] .

[60] The restriction on the coefficients of the income proxy is accepted at the 5% level from the joint *F*-test (1,240). The calculated value, 2.13, is smaller than the critical value 3.84; the SC criterion is minimised when such a restriction is imposed: from -4.35821 to -4.36979 after imposing the restiction.

d) *RLRSP* = Restriction on the coefficients for the fifth and seventh lag of the real substitute price[61] .

e) *E* = "Easter" dummy.

The model is data congruent and well-specified. Restrictions on the coefficients can be imposed in accordance with the joint *F*-test and the SC criterion. The dependent variable is highly dependent on its own past behaviour. The income proxy shows a positive sign both in the short and long run. In the short run, the income elasticity is below unity; whereas, in the long run, the income elasticity is above unity with a *t*-value equal to 3.73 (see Table 6.2). The positive sign shows that the higher the income for Italians the higher the number of tourists in Sassari Province. Note also that the coefficient of the oscillation (*RLPR*) presents a negative sign, which can be picking up possible "over-time" effects. The relative price and the "weather" variable do not play any role in explaining the domestic demand for tourism. Both of them do not appear statistically significant in the final restricted model. Interestingly, the real substitute price enters the equation in its difference with a positive statistically significant coefficient. The "Easter" dummy once again appears to be highly significant. Moreover, the coefficients of the seasonal dummies, allowing for a changing seasonal pattern, are in general statistically significant.

The dynamics are reported in Table 6.2. The long run coefficients are, in general, well determined and the null hypothesis that they are all zero excluding the constant term is rejected at the 1% level.

---

61 The restiction on the coefficients of the real substitute price is accepted at the 5% level. The calculated *F*-test (1,242) is equal to 0.60 that is smaller than the critical value 3.84. The SC criterion is also minimised when the coefficient restriction is imposed: from -4.3776 to -4.3956 after imposing the restriction.

Table 6. 2 Long Run Dynamics for the Domestic Demand with Real Substitute Price

LAR =	+2.87	+1.378 LPR	+0.8431 E
(SE)	(1.655)	(0.3689)	(0.2271)
	+1.759 i1992p3	+0.189 jan	+0.005782 jan2
	(0.5391)	(0.1639)	(0.2151)
	+0.2686 feb	+0.5596 mar	+1.895 apr
	(0.1709)	(0.2401)	(0.3481)
	+2.832 apr2	+1.788 apr3	+1.663 may
	(0.4697)	(0.3573)	(0.1935)
	+1.489 may3	+1.696 jun	+2.764 jun2
	(0.2716)	(0.204)	(0.3345)
	+2.886 jun3	+2.334 jul	+2.841 jul2
	(0.3736)	(0.2573)	(0.2989)
	+2.828 jul3	+2.993 aug	+3.945 aug2
	(0.3172)	(0.324)	(0.4283)
	+4.341 aug3	+2.327 sep	+1.614 sep3
	(0.4992)	(0.3266)	(0.3955)
	-1.101 oct	-2.032 oct3	-0.4158 nov
	(0.4472)	(0.6279)	(0.1956)
	-3.59 RLPR	+1.978 RLRSP	
	(1.245)	(0.8322)	

ECM = LAR - 2.86981 - 1.3781*LPR - 0.843128*E - 1.75941*i1992p3 - 0.189043*jan
- 0.00578207*jan2 - 0.268558*feb - 0.559587*mar - 1.89536*apr
- 2.83181*apr2 - 1.78812*apr3 - 1.66303*may - 1.48856*may3 - 1.69648*jun
- 2.76411*jun2 - 2.88587*jun3 - 2.33351*jul - 2.84094*jul2 - 2.82845*jul3
- 2.99344*aug - 3.94541*aug2 - 4.3412*aug3 - 2.32698*sep - 1.61417*sep3
+ 1.10146*oct + 2.03195*oct3 + 0.415806*nov + 3.59042*RLPR - 1.97807*RLRSP;

WALD test Chi^2(28) = 466.12 [0.0000] **

The main finding is that problems still persist for the use of the substitute price adjusted for the weighted average exchange rate, in explaining the domestic demand for tourism in the north of Sardinia.

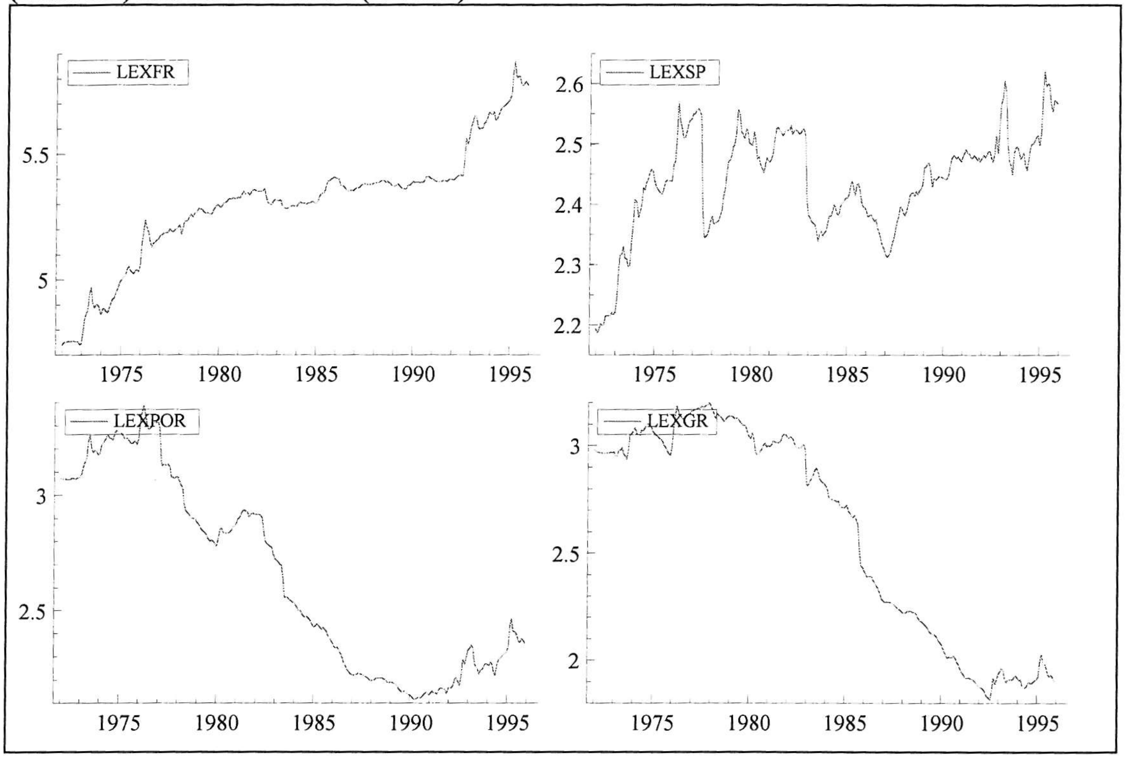
6.3 PRICES FOR COMPETITORS AND EXCHANGE RATES: A DISAGGREGATED STUDY

In tourism empirical studies, relative and/or substitute prices are, in general, used in an aggregated manner and weighted averages are included in models (see examples in Tremblay, 1989; Witt and Witt, 1992; Garcìa-Ferrer and Queralt, 1997). However, it might be that aggregation does not always lead to satisfactory results, as has been found from the previous analysis. The real substitute price, defined in terms of a weighted average, has given no conclusive results, and the positive sign has been confirmed.

It might be that the description and analysis of the individual components give a better understanding of the determinants that influence the demand for tourism. In this section, a description of the properties of the prices and exchanges rates for the main competitor countries of the Province of Sassari (*i.e.* France, Greece, Portugal and Spain) will be given.

Figure 6.2 represents the exchange rates: lira/franc, lira/pesetas, lira/escudo and lira/drachma, for the period from January 1972 up to December 1995.

Figure 6.2 (Log) Exchange Rates: Lira/Franc (*LEXFR*), Lira/Pesetas(*LEXSP*), Lira/Escudo (*LEXPOR*) and Lira/Drachma (*LEXGR*)

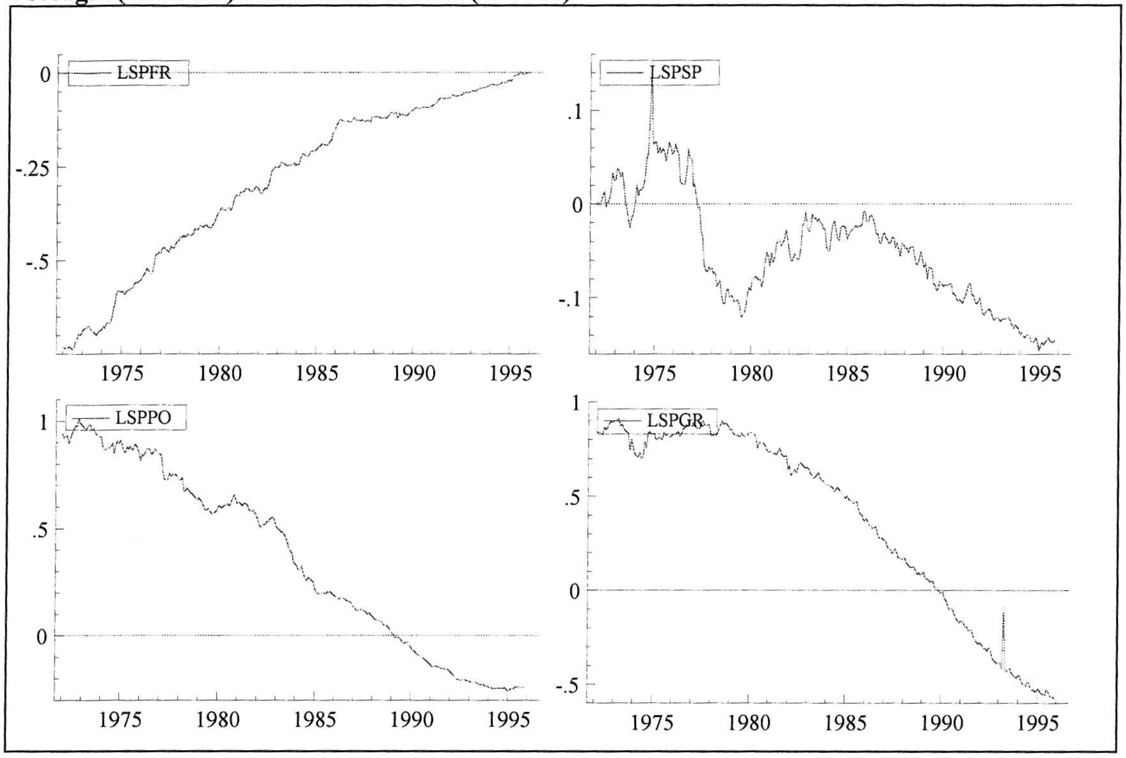


Again, the ADF unit root test with up to thirteen lags has been applied to each of the previous variables. It appears that the (log) exchange rate lira/franc, lira/drachma and lira/escudo are stationary in the first difference, whereas, the (log) exchange rate lira/pesetas is stationary in the level[62].

The next step is to describe and analyse each pair of substitute prices. The nominal substitute prices are defined as the difference between the log consumer price index in Sassari (*CPIss*) and the log consumer price index in each of the competitor countries (for France *LSPFR*, for Spain *LSPSP*, for Portugal *LSPPO* and, finally, for Greece *LSPGR*). The graphs for each of the variables are shown in Figure 6.3.

62 This is true when including the only constant, constant and trend, constant and seasonals and, finally, constant, trend and seasonals.

Figure 6. 3 (Log) Substitute Prices: Sassari-France (*LSPFR*), Sassari-Spain (*LSPSP*), Sassari-Portugal (*LSPPOR*) and Sassari-Greece (*LSPGR*)



As it emerges from the *LSPGR* series there appears a blip in the data. This can be attributed to a measurement error in the consumer price index for Greece. As already mentioned, the data are collected from the IFS *datastream.*

A long run unit roots ADF test has been applied to the above variables. In this case, *LSPFR* (*i.e.* difference between the log consumer price index in Sassari and the log consumer price index in France) appears to be I(0) when a constant, and a constant and seasonals are included. The other series (*LSPSP*, *LSPPO*, *LSPGR*) appear to be I(1).

Before proceeding further, a Johansen cointegration analysis has been used in order to test for a possible cointegration relationship between each pair, that is the nominal substitute price and the exchange rate. One will start with the first pair for France. An initial unrestricted 13 lag system with unrestricted constant and seasonal dummies is run. The former can be reduced to a 12 lag system in accordance with the information criteria SC and HQ, as well as to the joint *F*-test and upon no serial correlation in the residuals. From the analysis, it does not appear that there is a strong evidence for the existence of a cointegration relationship. Only the trace statistics, in fact, rejects the null hypothesis of no cointegration at the 5% level. The same result is achieved when only the constant is included in the system.

The Johansen test is also applied to the second pair of variables for Spain. A 13 lag system, with the inclusion of constant, trend and seasonal dummies treated unrestrictedly, can be reduced to a 2 lag system in accordance with the SC and HQ criteria, and the joint F-test; this system exhibits no serial correlation in the residuals. From the maximal eigenvalue and trace statistics there appears to be evidence for the coefficients of the substitute price and exchange rate to be cointegrated.

The third pair tested is for Portugal. In this case, a final 9 lag system with constant, trend and seasonals can be run. The results from the Johansen analysis seems to give evidence for stationarity between the coefficients of the two variables under study. Analogous results are obtained when including a constant and a trend in the system. However, such results can be considered as misleading as both variables are stationary in the first difference, as derived from the ADF test.

The last pair under consideration is for Greece. An initial unrestricted system of 13 lags, including a constant, trend and seasonals could be reduced to 7 lag system, which does not present any problems of serial correlation in the residuals. From the Johansen cointegration analysis, there appears to be evidence for stationarity between the coefficients of the substitute price and exchange rate. Once again, the result appears to be misleading.

As one can see the results are neither in agreement with the ADF test nor with each other or economic theory. It is difficult to see how an economy's exchange rate and price level, would not be cointegrated, if they are I(1), with a coefficient of unity. Given the conflicting empirical evidence it seems reasonable to impose the cointegration assumption. The analysis is, therefore, continued considering the substitute price adjusted for the exchange rate. That is, four different variables have been created and expressed as follows:

$$LPSj,t = ln\left(\frac{CPIi,t}{CPIj,t} \right) - ln\, EXj,t \qquad (6.3.1)$$

where:

j = France, Greece, Portugal and Spain.

$CPIi,t$ = monthly consumer price index in Sassari (1990=100) (Source: ISTAT).

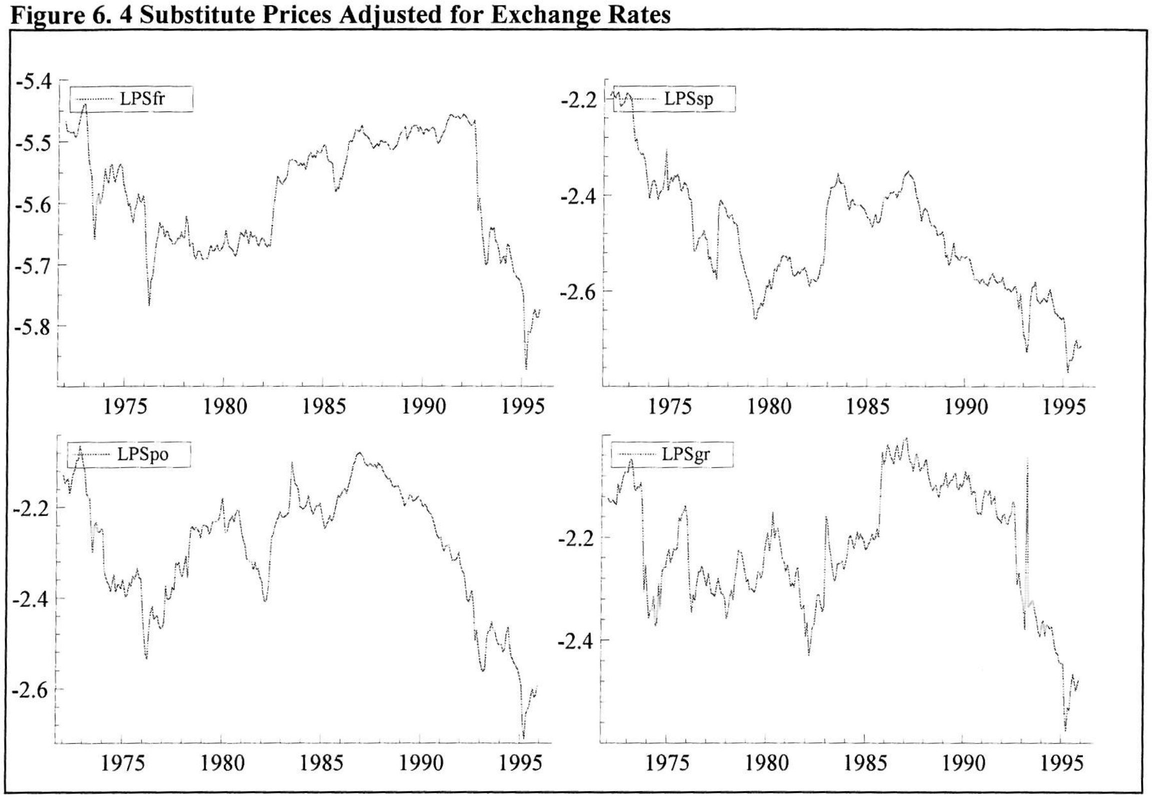
$CPIj,t$ = monthly consumer price index in country j (1990=100) (Source: IFS).

EXj,t = monthly exchange rate, lira per unit of currency of country j (elaborated on

IFS source).

The graphs for each of the variables created are given in Figure 6.4.

Figure 6. 4 Substitute Prices Adjusted for Exchange Rates



The noticeable characteristic of these new variables is the generally lower volatility in comparison to the individual components, as indicated in Table 6.3.

Table 6. 3 Standard Deviations

LEXFR	**LEXSP**	**LEXPOR**	**LEXGR**
0.23452	0.08560	0.41033	0.48823
LSPFR	**LSPSP**	**LSPPO**	**LSPGR**
0.22028	0.05876	0.41760	0.48760
LPSfr	**LPSsp**	**LPSpo**	**LPSgr**
0.09297	0.12303	0.14057	0.12107

An ADF test has been carried out for each of the series. All the series have been found to be stationary in the first difference. Thus, a Johansen cointegration analysis has been done. An initial 13 lag system has shown problems in the residuals (serial correlation, non-normality and heteroscedasticity). The inclusion of two impulse dummy variables *i1993p4* and *i1993p5* corrects for serial correlation and heteroscedasticity; however, though reduced, problems of non-normality are still present. A 13 lag system, which includes unrestrictedly the two impulse dummies, a constant, trend and seasonals is estimated according to the joint F-test as well as the

AIC criterion. From the analysis, the null hypothesis of no cointegration cannot be rejected. Analogous results have been obtained when including only a constant, or the constant and the time trend.

Overall, one can conclude the substitute prices adjusted for the exchange rates show a higher homogeneity in terms of properties and results than the individual components. Thus, the four real substitute prices will be used in running the model for the domestic and international demand for tourism in the north of Sardinia. However, there are reservations as to the order of integration of the constructed variables.

6.4 DOMESTIC DEMAND MODEL USING DISAGGREGATED REAL SUBSTITUTE PRICES

This section is dedicated to the specification of a model for the domestic demand for tourism in the north of Sardinia, that includes the four variables of substitute price adjusted for the exchange rate, as defined in the previous paragraph. It is possible that the heterogeneity of the properties of the individual components, that is the nominal substitute price and exchange rate for each of competitor countries, has created problems on an aggregated level (*i.e.* when using the real substitute price as a weighted average).

Firstly, one starts with the strong, but plausible assumption, that consumers have a perfect knowledge of the living costs in their own country as well as in the competitor countries. This assumption, can be thought to be true on the basis that tourists are well informed thanks to their own past experience, as well as from information received by other sources (newspapers, friends and family's experience). The assumption of perfect information on the exchange rates is widely accepted. In this way, one can use the level of those variables which appear to be stationary in the first difference, *i.e.* relative price Sassari-Italy (*LREP*), real substitute price for France (*LPSfr*), Greece (*LPSgr*), Portugal (*LPSpo*) and Spain (*LPSsp*).

The model for the domestic arrivals in Sassari Province (*LAR*) contains the following explanatory variables: the income proxy (*LPR*), the relative price (*LREP*), the real substitute prices for the competitor countries (*LPSfr*, *LPSgr*, *LPSpo* and *LPSsp*), the "weather" variable (*LW*), the "Easter" dummy variable (*E*), an impulse dummy (*i1992p3*), created after inspecting the residuals, in order to avoid non-normality problems, and, finally, the seasonal dummies allowing for structural changes in the seasonal pattern in accordance to the findings in Chapter 5.

The analysis is done for the period between January 1972 and December 1995. The first step consists of running an unrestricted model with 13 lags that can be reduced to a 12 lag model[63] . Such a model does not present any particular problems in the residuals. Note that further coefficient restrictions are suggested by using the SC criterion, however, problems appear in the residuals (serial correlation, heteroscedasticity and mis-specification). Hence, a 12 lag model is run and the final restricted model is obtained following both the *F*-test and the SC criterion. The final results are presented in Table 6.4.

[63] The reduction has been done according to the joint *F*-test and SC criterion as follows:

dep.var		T	k	df	RSS	sigma	Schwarz
1 lag : LAR	OLS	275	40	235	2.30404	0.0990172	-3.96512
....		....			......	....	.....
12 lags: LAR	OLS	275	117	158	1.01736	0.0802432	-3.20988
13 lags: LAR	OLS	275	124	151	0.999672	0.0813655	-3.08445

Model 13 -->12: $F(7, 151) = 0.3817$ [0.9120]
Model 12 -->11: $F(7, 158) = 2.9069$ [0.0069] **

Table 6. 4 Final Domestic Demand Model for Tourism with Inclusion of Real Substitute Prices

EQ(2) Modelling LAR by OLS (using vdomlar2.in7)
The present sample is: 1973 (2) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-0.71408	0.62081	-1.150	0.2512	0.0056
LAR_1	0.30081	0.052319	5.750	0.0000	0.1243
LAR_2	0.12476	0.043796	2.849	0.0048	0.0337
LAR_11	0.16391	0.042294	3.876	0.0001	0.0606
LAR_12	0.25527	0.050868	5.018	0.0000	0.0975
RLPR	-0.60252	0.17793	-3.386	0.0008	0.0469
LPR_11	0.12509	0.082981	1.507	0.1330	0.0097
LPSfr	0.47626	0.16778	2.839	0.0049	0.0334
LPSfr_3	-0.63965	0.30063	-2.128	0.0344	0.0191
LPSfr_5	0.97948	0.34391	2.848	0.0048	0.0336
LPSfr_7	-1.1636	0.22378	-5.200	0.0000	0.1040
RLPSfr	0.73741	0.18376	4.013	0.0001	0.0646
LPSgr_1	0.23833	0.088202	2.702	0.0074	0.0304
LPSpo_1	-0.59035	0.11109	-5.314	0.0000	0.1081
RLPSpo	0.94637	0.21582	4.385	0.0000	0.0762
LPSpo_11	0.54455	0.10020	5.435	0.0000	0.1125
RLPSsp	0.51039	0.14414	3.541	0.0005	0.0511
i1992p3	0.30929	0.083037	3.725	0.0002	0.0562
E	0.14841	0.029158	5.090	0.0000	0.1001
jan	0.069707	0.030899	2.256	0.0250	0.0214
jan2	0.012081	0.040118	0.301	0.7636	0.0004
feb	0.084199	0.034905	2.412	0.0166	0.0244
mar	0.15812	0.048823	3.239	0.0014	0.0431
apr	0.40881	0.060249	6.785	0.0000	0.1650
apr2	0.58327	0.068773	8.481	0.0000	0.2359
apr3	0.40203	0.071101	5.654	0.0000	0.1207
may	0.40135	0.061471	6.529	0.0000	0.1547
may3	0.33315	0.071521	4.658	0.0000	0.0852
jun	0.32198	0.053910	5.973	0.0000	0.1328
jun2	0.50841	0.065850	7.721	0.0000	0.2037
jun3	0.58221	0.070982	8.202	0.0000	0.2241
jul	0.42696	0.059183	7.214	0.0000	0.1826
jul2	0.55863	0.076659	7.287	0.0000	0.1856
jul3	0.60933	0.082882	7.352	0.0000	0.1883
aug	0.55415	0.061256	9.046	0.0000	0.2599
aug2	0.72581	0.078472	9.249	0.0000	0.2686
aug3	0.80408	0.084865	9.475	0.0000	0.2781
sep	0.41360	0.072871	5.676	0.0000	0.1215
sep3	0.32545	0.094063	3.460	0.0006	0.0489
oct	-0.24976	0.059455	-4.201	0.0000	0.0704
oct3	-0.42940	0.076369	-5.623	0.0000	0.1195
nov	-0.18503	0.044053	-4.200	0.0000	0.0704

R^2 = 0.991378 F(41,233) = 653.46 [0.0000] sigma = 0.0782021
DW = 1.88 RSS = 1.42492738 for 42 variables and 275 observations

AR 1- 7	F(7,226) =	1.8199 [0.0844]
ARCH 7	F(7,219) =	0.2696 [0.9652]
Normality	Chi^2(2)=	1.7118 [0.4249]
Xi^2	F(57,175) =	0.9393 [0.5993]
RESET	F(1,232) =	1.4712 [0.2264]

where:

a) *LAR* = domestic arrivals of tourists in the north of Sardinia.

b) *RLPR* = restriction on the coefficients for the second and third lag of the income proxy[64].

c) *LPR* = income proxy as industrial production index in Italy.

d) *LPSfr* = real substitute price between Sassari and France.

e) *RLPSfr* = restriction on the coefficients for the ninth and twelfth lag of the real substitute price for France[65].

f) *LPSgr* = real substitute price between Sassari and Greece.

g) *LPSpo* = real substitute price between Sassari and Portugal.

e) *RLPSpo*= restriction on the coefficients of the third and fifth lag[66].

f) *RLPSsp*= restriction on the coefficients of the seventh and tenth lag for the real substitute price between Sassari and Spain[67].

g) *E* = "Easter" dummy.

The final model is well-specified and data congruent. It seems to be interesting to give an explanation of the results obtained and compare them with the final model reached when using the aggregate real substitute price (see Table 6.1). From Table 6.4, the dependent variable shows a dependence on its own past behaviour and presents the expected positive sign. In terms of the income proxy, again, the restriction on the coefficients of this variable, picking up the short run dynamics, shows a negative sign. This confirms the findings presented in Table 6.1 where possible "over-time" effects might be acting. However, in both the cases, the short and long run elasticity are positive and less than one, confirming the results obtained in Chapter 5 and Malacarni's (1991) findings. Hence, domestic tourism turns out to be a necessity good. The relative price (Sassari-Italy) variable does not enter in the final

64 The restriction on the coefficients of the income proxy is accepted at the 5% level from the joint *F*-test (1,229). The calculated value, 1.41, is smaller than the critical value 3.84; the SC criterion is minimised when the restriction is imposed: from -4.33728 to -4.35156 after the restriction.

65 From the *F*-test (1,230) the calculated value is 0.33 that is smaller than 3.84 at the 5% level from the equivalent table. The SC is again minimised when the restriction is imposed The SC criterion is minimised when such a restriction is imposed: from -4.35156 to -4.37053 after imposing the restriction.

66 From the F-test (1,231) the calculated value is 0.01 that is smaller than 3.84 at the 5% level. The SC criterion is minimised when the restriction is imposed; it decreases from -4.37053 to -4.3909.

67 From the F-test (1,232) the calculated value, 1.52, is smaller than the critical value (3.84) at the 5% level. The coefficient restriction is suggested also by the SC criterion that decreases from -4.3909 to -4.40482.

model, confirming that such a variable does not play any role in explaining the domestic demand for tourism.

The results for the real substitute prices are quite heterogeneous. One can notice different effects with respect to each of the competitor countries. The real substitute prices for France and Portugal present an articulate lag structure. Analysing the long run effects in Table 6.5, a negative sign can be observed for the coefficient of the real substitute price for France and Portugal. This finding is in line with economic theory. Higher prices in the north of Sardinia relative to the competitors, everything else being equal, determine a change in the choice of the destination holiday, and a consequential decrease in the number of tourists' arrivals. However, if the long run coefficient is statistically significant for France, it does not appear to be statistically significant for Portugal. The short run elasticity for both of the former variables presents a negative sign and has statistically significant coefficients.

Differences can be noticed in the lag structure of the real substitute prices for Greece and Spain. Just the first lag of the real substitute price for Greece is found to be statistically significant and it presents a positive coefficient. The same positive sign appears in the restricted coefficient (*RLPSsp*) for the real substitute price for Spain.

The estimated coefficient for the "Easter" dummy is 0.148. The statistical significance of this variable in explaining the demand for tourism confirms the results obtained by Gonzàles and Moral (1996). The seasonal dummies that allow for changing pattern present statistically significant coefficients, with the highest percentage increase in the number of tourists' flows in August and July.

The long run dynamics are reported in Table 6.5.

Table 6. 5 Long Run Dynamics for the Model with the Real Substitute Prices

```
LAR =              -4.6                 +0.8058 LPR               -2.238 LPSfr
(SE)          (     4.52)          (    0.4481)           (     0.988)
                 +1.535 LPSgr          -0.295 LPSpo             +1.992 i1992p3
              (   0.6959)          (    0.5266)           (    0.6683)
                 +0.956 E              +0.449 jan             +0.07782 jan2
              (   0.2821)          (    0.2387)           (    0.2618)
                +0.5424 feb            +1.019 mar               +2.633 apr
              (   0.2669)          (    0.3722)           (    0.6474)
                 +3.757 apr2            +2.59 apr3              +2.585 may
              (   0.8714)          (      0.65)           (    0.5584)
                 +2.146 may3           +2.074 jun               +3.275 jun2
              (   0.4971)          (    0.3655)           (    0.5631)
                  +3.75 jun3            +2.75 jul               +3.598 jul2
              (   0.6605)          (    0.4386)           (    0.6006)
                 +3.925 jul3            +3.57 aug               +4.675 aug2
              (   0.6585)          (    0.5915)           (    0.7701)
                  +5.18 aug3           +2.664 sep               +2.096 sep3
              (   0.8602)          (    0.4961)           (    0.5161)
                 -1.609 oct            -2.766 oct3              -1.192 nov
              (   0.6569)          (    0.9801)           (    0.4614)
                 -3.881 RLPR             +4.75 RLPSfr           +6.096 RLPSpo
              (     1.47)          (     1.679)           (     2.011)
                 +3.288 RLPSsp
              (    1.157)

ECM = LAR + 4.59982 - 0.80579*LPR + 2.23829*LPSfr - 1.53526*LPSgr
     + 0.294997*LPSpo + 3.8812*RLPR - 4.75012*RLPSfr - 6.09617*RLPSpo
     - 3.28773*RLPSsp - 1.9923*i1992p3 - 0.956016*E - 0.449024*jan
     - 0.077818*jan2 - 0.542377*feb - 1.01852*mar - 2.6334*apr - 3.75717*apr2
     - 2.58972*apr3 - 2.5853*may - 2.14602*may3 - 2.07404*jun - 3.27496*jun2
     - 3.75037*jun3 - 2.75032*jul - 3.59844*jul2 - 3.92507*jul3 - 3.56959*aug
     - 4.67535*aug2 - 5.17959*aug3 - 2.66422*sep - 2.09641*sep3 + 1.60886*oct
     + 2.76605*oct3 + 1.19191*nov;

WALD test Chi^2(33) = 276.98 [0.0000] **
```

The Wald test suggests that the coefficients, except the constant, are jointly different from zero, as the null hypothesis is rejected at the 1% level. In particular, the long run coefficients are, in general, statistically significant, with the exception of the real substitute price for Portugal (*LPSpo*), as already stated.

6.5 MODEL FOR THE INTERNATIONAL TOURISM DEMAND USING THE REAL SUBSTITUTE PRICE IN A DISAGGREGATED MANNER

The aim of this section is to consider whether the real substitute price for the four competitor countries of the north of Sardinia (*i.e.* France, Greece, Portugal and Spain) show any importance in explaining the international demand for tourism. The real substitute prices are defined as in formula (6.3.1). As stated earlier in this chapter, one assumes that the consumers have a perfect knowledge of the living costs in the destination countries and that the same assumption holds for the exchange rates.

The model for the foreign arrivals of tourists in Sassari Province, normalised for the number of weekends (Saturdays) in a month (*LA*), contains the following explanatory variables: the income proxy (*LPR*), the first difference of the relative price (*DLRP*) and weighted average exchange rate (*DLER*), the first lag of the cointegrating vector (CI_{t-1}), the real substitute prices for the competitor countries (*LPSfr*, *LPSgr*, *LPSpo* and *LPSsp*), the "weather" variable (*LW*), the "Easter" dummy variable (*Easter*), a time trend (*Trend*), the seasonal dummies and, finally, five impulse dummies (*i1974p12*, *i1979p10*, *i1985p3*, *i1988p4* and *i1989p5*), created after inspecting the residuals, in order to avoid non-normality problems. These dummies might be picking up possible economic and non-economic events such as: strikes, particular exhibitions or sport meetings with a non cyclical pattern, special offers in accommodation in the north of Sardinia and so on.

The first step consists in running an unrestricted model with 13 lags which can be reduced to a 12 lag model. The parameter restrictions have been accepted by the joint *F*-test[68]. Such a model does not present any problems in the residuals. Hence, the unrestricted model can be reduced, according to the SC criterion and the joint *F*-test and it is reported in Table 6.6.

[68] The reduction has been done according to the joint *F*-test and the SC criterion as follows:

dep.var		T	k	df	RSS	sigma	Schwarz
1 lag : LA	OLS	274	28	246	11.6767	0.217867	-2.58193
......		....			.....	......	.....
12 lags: LA	OLS	274	124	150	5.7383	0.19559	-1.32572
13 lags: LA	OLS	274	132	142	5.3240	0.19363	-1.23677

Model 13 lags > 12 lags: F(8, 142) = 1.3812 [0.2096]
Model 12 lags > 11 lags: F(8, 150) = 2.9586 [0.0042] **

Table 6. 6 Final Model for Foreign Tourism Demand with Inclusion of Real Substitute Prices

EQ(2) Modelling LA by OLS (using For.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	4.0643	1.7509	2.321	0.0211	0.0214
LA_2	0.10550	0.049181	2.145	0.0329	0.0184
LA_3	0.18163	0.048831	3.720	0.0002	0.0532
LA_12	0.11742	0.048682	2.412	0.0166	0.0231
LPR_10	-0.85799	0.32970	-2.602	0.0098	0.0268
DLRP_12	-4.3347	1.3904	-3.117	0.0020	0.0380
LPSfr_3	1.3858	0.36841	3.762	0.0002	0.0544
LPSfr_6	-1.6715	0.37488	-4.459	0.0000	0.0748
RLPSsp	-1.3648	0.62515	-2.183	0.0300	0.0190
RLPSgr	1.0971	0.31891	3.440	0.0007	0.0459
Trend	0.0025721	0.00045652	5.634	0.0000	0.1143
Easter	0.42405	0.070613	6.005	0.0000	0.1279
i1974p12	1.5341	0.20072	7.643	0.0000	0.1919
i1979p10	0.65749	0.19982	3.290	0.0011	0.0422
i1985p3	0.93743	0.20323	4.613	0.0000	0.0796
i1988p4	-0.52929	0.19935	-2.655	0.0084	0.0279
i1989p5	0.38811	0.19969	1.944	0.0531	0.0151
JA	0.34255	0.10694	3.203	0.0015	0.0400
FE	0.87040	0.15960	5.454	0.0000	0.1079
MAR	1.2999	0.17808	7.300	0.0000	0.1780
AP	2.1134	0.20505	10.307	0.0000	0.3016
MAY	3.1630	0.20685	15.291	0.0000	0.4873
JUN	3.2228	0.20033	16.088	0.0000	0.5127
JUL	3.2074	0.19571	16.388	0.0000	0.5219
AU	2.9707	0.18733	15.858	0.0000	0.5055
SE	2.6550	0.18062	14.700	0.0000	0.4676
OT	1.4143	0.12823	11.030	0.0000	0.3309
NO	0.17214	0.087209	1.974	0.0495	0.0156

R^2 = 0.982701 F(27,246) = 517.56 [0.0000] sigma = 0.192961
DW = 1.81 RSS = 9.159507777 for 28 variables and 274 observations

AR 1- 7 F(7,239) =	2.4868 [0.0175]	*
ARCH 7 F(7,232) =	1.5517 [0.1509]	
Normality Chi^2(2)=	5.1804 [0.0750]	
Xi^2 F(37,208) =	1.2049 [0.2085]	
RESET F(1,245) =	0.1303 [0.7185]	

where:

a) *LA*= international arrivals of tourists in the north of Sardinia, normalised for the number of Saturdays in a month.

b) *LPR* = weighted average of the income proxy.

c) *DLRP* = first difference of the relative price.

d) *LPSfr* = real substitute price between Sassari and France.

e) *RLPSsp* = restriction on the coefficients for the third and fourth lag of the real substitute price (Sassari-Spain)[69] .

---

69 The calculated *F*-test (1,244) equals 0.04 that is smaller than the critical value at the 5% level. The restriction is also suggested by the SC criterion whose value decreases from -2.78429 to -2.8046

f) *RLPSgr* = restriction on the coefficients for the eleventh and twelfth lag of the real substitute price (Sassari-Greece)[70].

The model is able to explain a 98% of the variance of the dependent variable. However, the final model fails to accept the null hypothesis of non serial correlation in the residuals, at the 5% level. The past behaviour of the dependent variable plays a role in explaining the international demand for tourism in the north of Sardinia and has a positive sign coefficient. In terms of income proxy, only the tenth lag is found to be statistically significant with a negative sign. The cointegrating vector does not appear in the final restricted model, as well as the first difference of the weighted exchange rate. However, the relative price growth is statistically significant with the expected negative sign. Hence, the long run information does not have any importance in explaining the international demand.

The real substitute price for France, statistically significant in the long run, presents a negative sign coefficient, as reported in Table 6.7. However, the exchange rate and the consumer price index for France are also present in the weighted average exchange rate and consumer price index for the main origin countries. This might be indicating a substitute effect between home and Sardinian vacations for French tourists. The real substitute price for Portugal does not turn out to be statistically significant. Moreover, the real substitute prices for Spain and Greece enter in the equation as differences, since a coefficient restriction could be imposed. The first coefficient presents a negative sign, whereas the coefficient for Greece shows a positive sign.

The time trend, picking up changes in consumers' tastes, is highly significant and it has a positive sign. The "Easter" dummy, once again, plays an important role in explaining the international demand for tourism. The seasonal dummies show statistically significant coefficients, with the highest percentage increase in the number of foreign arrivals in June and July.

The long run dynamics are presented in Table 6.7. The Wald test suggests the joint significance of the long run coefficients.

---

70 From the *F*-test (1,246) the calculated value is 0.49 smaller than 3.84 at the 5% level from the conventional table. Again, the SC is minimised when this restriction is imposed, *i.e.* from -2.8046 to -2.8247.

Table 6. 7 Long Run Dynamics for Foreign Demand with the Inclusion of Real Substitute Prices

LA =	+6.826	-1.441 LPR	-7.28 DLRP
(SE)	(2.758)	(0.5715)	(2.585)
	-0.4798 LPSfr	-2.292 RLPSsp	+1.842 RLPSgr
	(0.2386)	(1.079)	(0.5869)
	+0.00432 Trend	+0.7121 easter	+2.576 i1974p12
	(0.0007019)	(0.1503)	(0.4954)
	+1.104 i1979p10	+1.574 i1985p3	-0.8889 i1988p4
	(0.3587)	(0.4031)	(0.3544)
	+0.6518 i1989p5	+0.5753 JA	+1.462 FE
	(0.3372)	(0.2396)	(0.4285)
	+2.183 MAR	+3.549 AP	+5.312 MAY
	(0.5279)	(0.6785)	(0.825)
	+5.412 JUN	+5.387 JUL	+4.989 AU
	(0.7477)	(0.6353)	(0.5167)
	+4.459 SE	+2.375 OT	+0.2891 NO
	(0.4272)	(0.2062)	(0.1295)

ECM = LA - 6.82563 + 1.44093*LPR + 7.27969*DLRP + 0.479773*LPSfr + 2.29215*RLPSsp - 1.8425*RLPSgr - 0.00431965*Trend - 0.712148*easter - 2.57637*i1974p12 - 1.1042*i1979p10 - 1.57434*i1985p3 + 0.888898*i1988p4 - 0.651793*i1989p5 - 0.575276*JA - 1.46175*FE - 2.18312*MAR - 3.54921*AP - 5.31192*MAY - 5.41249*JUN - 5.38657*JUL - 4.98905*AU - 4.45889*SE - 2.37527*OT - 0.289101*NO;

WALD test Chi^2(23) = 1847.7 [0.0000] **

6.6 SUMMARY

In this section, the main economic findings in terms of income and price elasticities are reported, considering both the short and long run behaviour. Table 6.8 summarises the findings.

Table 6. 8 Summary of Short and Long Run Elasticities

Elasticities	**Domestic Model** Aggregation LRSP (Tables 6.1 - 6.2)	**Domestic Model (Disaggregation:** ***LPSfr, LPSgr, LPSpo, LPSsp*****) (Tables 6.4- 6.5)**	**International Model** (Disaggregation $LPS_{i,t}$) **(Tables 6.6 - 6.7)**
INCOME (long run)	1.38 (3.73)	0.81 (1.80)	-1.44 (-2.52)
INCOME (short run)	0.26 (3.12)	0.13 (1.51)	-0.86 (-2.60)
REL.PRICE (long run)	-	-	-7.28 (-2.82)
REL.PRICE (short run)	-	-	-4.33 (-3.12)
EX. RATE (long run)	-	-	-
EX.RATE (short run)	-	-	-
R.SUB.PRICE (long run)	1.98 (2.38)	-	-
R.SUB.PRICE(short run)	0.37 (2.62)	-	-
SUB.PRICEfr(long run)	-	-2.24 (-2.26)	-0.48 (-2.01)
SUB.PRICEfr(short run)	-	-0.64 (-2.13)	1.39 (3.76)
SUB.PRICEgr(long run)	-	1.54 (2.21)	-
SUB.PRICEgr(short run)	-	0.24 (2.70)	-
SUB.PRICEpo(long run)	-	-0.53 (-0.56)	-
SUB.PRICEpo(short run)	-	-0.59 (-5.31)	-
SUB.PRICEsp(long run)	-	-	-
SUB.PRICEsp(short run)	-	-	-

Notes: (1) *t*-values are given in parenthesis.
(2) Note that the short run elasticity corresponds to the first significant lag in the model (see Pindyck and Rubinfeld, p. 377, 1991).

As already stated, there is a mix of evidence in terms of income and price elasticities. Note that the specifications for the domestic demand of tourism (namely the second and third column) the income proxy shows the correct sign in both the long and short run. However, in the second specification (third column) the income proxy presents a rather marginally statistically significant coefficient, though positive. The international model for tourism (fourth column) denotes problems in interpreting the income coefficients, as the sign is negative in contrast with the economic expectations. The negative coefficient could be due to over-time effects emphasised more by using the industrial production index as a proxy.

As already mentioned in Chapters 4 and 5, the differences in the magnitude of the elasticities are likely to reflect different types of behaviour, preferences and the time the decision is taken by the consumers.

There is mixed evidence that the inclusion of a substitute price for the competitors adjusted for the exchange rate gives better results than including just the nominal substitute price. In the second column, in fact, a positive sign appears for the

the aggregated real substitute price in its difference (namely *RLRSP*). The disaggregation of the real substitute price for each of the competitors has given better results for France and Portugal. These substitute prices, in fact, adjusted for the exchange rates show a long run coefficient with a negative sign. This finding is in line with economic theory. However, while the long run coefficient for France is statistically significant, the long run coefficient for Portugal is not. Both these real substitute prices present a distributed lag structure and reasonably short run dynamics, with significant and expected negative signs.

Different behaviour has been noticed for Spain and Greece. Both of them present a short run dynamic structure. Just the first lag of the real substitute price for Greece is found to be statistically significant with a positive sign. The same positive sign appears in the oscillation for the real substitute price for Spain (namely *RLPSsp* in Table 6.5).

Heterogeneous results have also been achieved in modelling the foreign demand for tourism. Only the long run coefficient of the real substitute price for France is found to be statistically significant, with the expected negative sign. The real substitute price for Portugal has turned out not to play any role in explaining the international demand. Moreover, just the oscillations (first differences) for Spain and Greece appear to be statistically significant in the final restricted model (see Table 6.7), the former with a negative sign and the latter with a positive sign.

On balance, the domestic model estimations give better results than the international demand model. This fact might also be suggesting a different choice of competitors for the source countries under analysis could be more appropriate. Interestingly, the only real substitute price for France enters in the final equation with the expected negative sign. One can think that France and the Corsican isle can be a substitute for example for Germans, British and Swiss. However, as one is dealing with an aggregated model for the origin countries, it may be possible that other countries can be thought to be competitors. For example, the British could think of Ireland or Holland as substitutes; Americans the Southern American countries; Germans their boundary countries and so on. It could be interesting to investigate these assumptions but is outside the scope of this thesis.

6.7 CONCLUSION

In the empirical tourism literature, the substitute price, used either with or without the exchange rate, has been identified as one of the main determinants of the demand for tourism, with a general expected negative sign. In the previous chapters, a positive sign for the nominal substitute price has been determined.

This chapter has been dedicated to investigating whether the exchange rate for the main competitor countries plays a role in explaining the demand for tourism in the north of Sardinia. The inclusion of the exchange rate is also supported by the Purchasing Power Parity theory, for which, in the short run, substantial deviations between prices and exchange rates are likely to occur.

In this chapter, the characteristics of the real substitute price in an aggregated and disaggregated manner have also been investigated for the main competitor countries (*i.e.* France, Greece, Portugal and Spain). Data with a monthly frequency, for the period between January 1972 and December 1995, have been used.

In the first section, evidence has been given that the use of a real substitute price used in an aggregated manner can lead to non-conclusive results. However, a deeper investigation of the characteristics and properties of the individual components of such variables has given more insight and conclusive results. Hence, four separate series have been considered, one for each of the competitor countries under study. According to the findings, the real substitute price for each of the destination countries has shown less variance and higher homogeneity in terms of properties, such as integration and cointegration status, than the individual components (*i.e.* nominal substitute price and weighted average exchange rate). These findings have suggested the use of four real substitute prices in modelling the domestic and international demand for tourism to Sassari Province.

From the results obtained, by fitting the models for the domestic and international demand for tourism, evidence has been found that there is a lack of homogeneity amongst the variables under analysis, both in terms of statistical significance as well as dynamic structure. On balance, France and Portugal have appeared to be the most likely substitute countries for the north of Sardinia tourism.

These findings seem to encourage a more careful investigation of the individual components of the determinants of tourism demand and problems of aggregation.

CHAPTER 7.

ITALIAN TOURISM: SEASONALITY, NUMBERS AND EXPENDITURE

Aim of the Chapter:

To introduce Italian tourism in terms of historic evolution of tourists' flows, seasonality, numbers and expenditure with an econometric analysis following in Chapter 8.

7.1 INTRODUCTION

So far, the demand for tourism in this thesis has been measured in terms of the number of tourists from particular origin countries, relative to their market share, to a certain destination that is the Italian Province of Sassari. The most significant determinants that influence the demand level of tourism in the north of Sardinia have been analysed. Chapters 7 and 8 will be dedicated to the study of Italian tourism as a whole.

Chapter 2 has covered the main debates in the tourism literature by assessing which variable best approximates the demand for tourism. The answer does not seem to be either unequivocal or conclusive. The first problem is the definition of tourism itself. A multitude of definitions can be found in the literature, and there is no common agreement as to what the constituents of tourism are. According to the World Tourism Organisation (WTO), tourism consumption should be defined as "the value of goods and services used by or for tourism units" (Nordström, 1996, p.15). Thus, the demand for tourism can be considered as a variegated "bundle" of goods and services. Given such a definition it does not seem to be clear cut which is the best variable to proxy the demand for tourism. In the majority of the current tourism literature, the number of arrivals is used as the dependent variable. However, there are some studies that analyse and/or forecast tourist expenditure as well as tourism arrivals (see Sheldon, 1993; Qiu and Zhang, 1995; Gonzàles and Moral, 1996). The aim of this chapter is to give a general introduction to Italian tourism with a particular emphasis on comparing expenditure and number of arrivals.

The chapter is divided in the following manner. Section 7.2 is dedicated to a general introduction to the domestic and international demand for tourism in Italy. Particular emphasis is given to the evolution of the flows of foreign tourists and to seasonality by country of origin. In the next section, a distinction is made between numbers (that is, number of arrivals of tourists and nights of stay in all registered accommodation) and expenditure. Several definitions are reported as given by the Bank of Italy. The last section concludes the chapter.

7.2 AN ANALYSIS OF ITALIAN TOURISM

Italy can be considered as one of the main tourist destinations amongst all European countries. As Papatheodorou (1999) points out, Italy together with Spain can be considered as the *core* of the six Mediterranean destinations that are examined in his study; on the other hand, Greece, Portugal, Turkey and Yugoslavia are defined as the *periphery*. He notes that the core has a share of almost 80% for the main source countries of tourism (Germany, France and UK) for the period 1957 to 1990. Indeed, Mediterranean tourism has experienced an Italian monopoly with a share of more than 75% in 1957. This share declined rapidly up to 1975 and then stabilised at lower levels. Baloglu and McCleary (1999) provide insight on the weaknesses and strengths of Italy with respect to two other Mediterranean competitors (Greece and Turkey) in the minds of U.S. visitor and non-visitor tourists. Italy is viewed as having superior quality accommodation provided, appealing local cuisine and high comfort for the whole travel experience. However, Italy has the minimum score in providing an unpolluted and unspoiled environment.

A description of the evolution of Italian tourism in terms of flows, expenditure and seasonality follows.

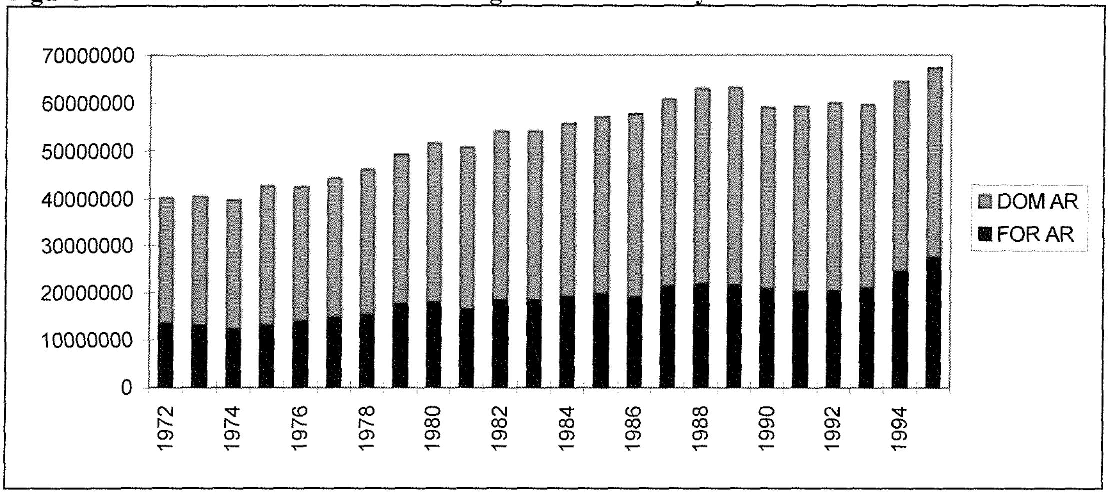
7.2.1 International Versus Domestic Flows

The characteristics of international and domestic flows in Italy are different, and a distinction between the two components is due. One notices, that domestic arrivals in all registered accommodation represent 65.7% against 34.3% of foreign arrivals; whereas domestic nights spent in registered accommodation represent 70.3% against 29.7% of foreigners. Note that such percentages are averages for the period between 1972 and 1995.

As far as the number of foreign arrivals are considered (see Figures 7.1 and 7.2), one can notice a general upward trend during the Seventies which sees the peak in 1979 and 1980. A similar pattern characterises the nights of stay in registered accommodation (Figure 7.2). The Eighties sees a period of maturity (see also Formica and Uysal, 1996). In these years Italian tourism has seen a loss of competitiveness with respect to other Mediterranean countries, e.g. Greece, Spain, Turkey and Yugoslavia. The causes are various: the high cost of living, the increasing congestion of most historical cities and the algae in the Adriatic sea which helped to spoil the image of the Italian beaches during 1988 and 1989. The Nineties (between 1990 and 1992) faced a decline, more evident in terms of nights spent in registered accommodation than arrivals. Some help for Italian tourism derived from the devaluation of Italian lira (September 1992) allowing a come-back in competitiveness since 1993.

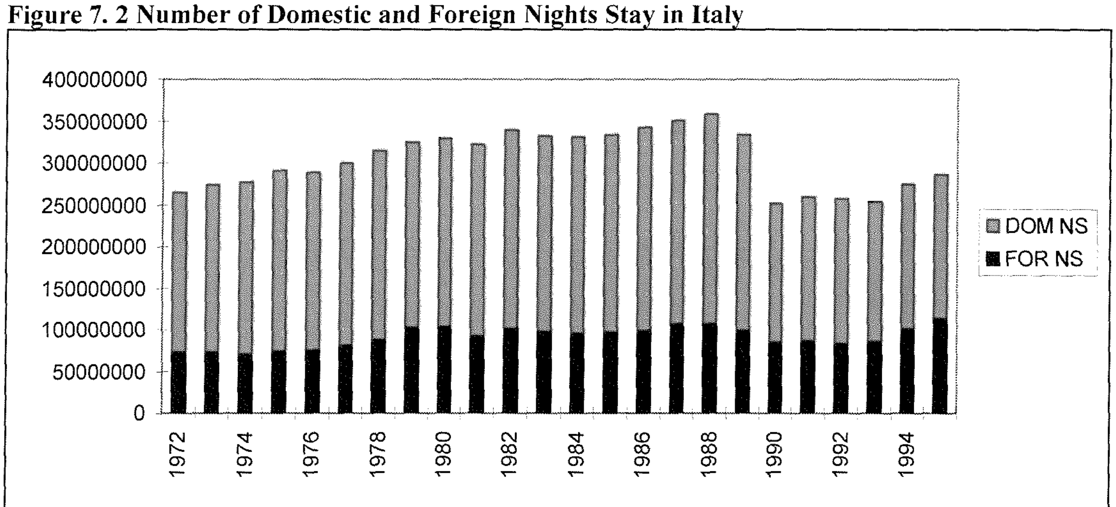
The domestic flows of tourists see a general upward trend in terms of arrivals during the Seventies and Eighties. One can also notice an upward trend in terms of number of nights of stay in the Seventies and a flattening of the trend in the Eighties (Figure 7.2). As is the case for foreign tourists, a decline can be seen since 1989, that is particularly evident in terms of nights of stay in registered accommodation. This decline might be caused by the growth of the "outgoing" phenomenon amongst Italian tourists, as well as by the increase of world-wide competition. As Formica and Uysal (1996) point out, "a one-week stay in the Seychelles (air ticket included) for an Italian resident is less expensive than the same amount of time spent in an Italian resort of equal quality" (p.327).

Figure 7. 1 Number of Domestic and Foreign Arrivals in Italy



Source: Figure based on ISTAT. Key words: DOM AR (Domestic tourists' Arrivals); FOR AR (Foreign tourists' arrivals).

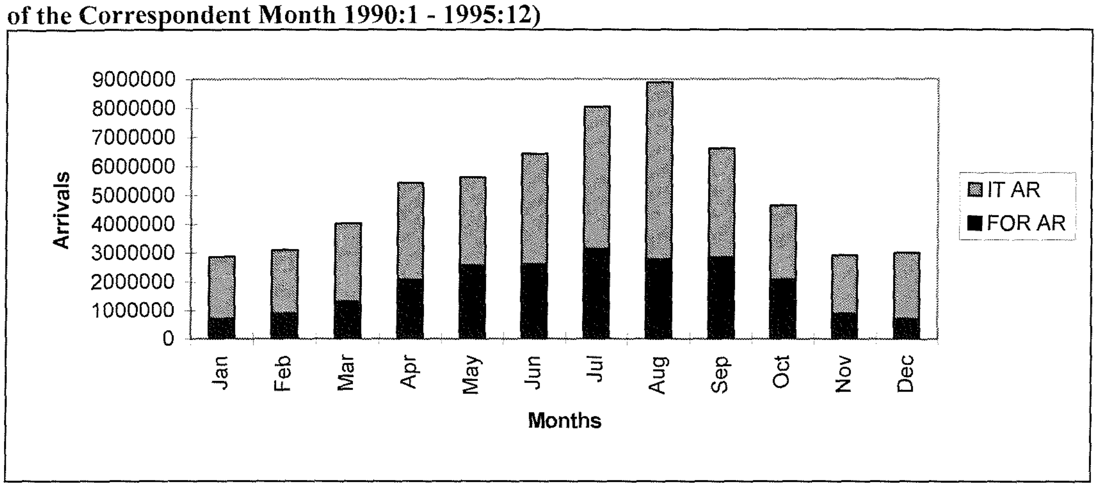
Figure 7. 2 Number of Domestic and Foreign Nights Stay in Italy



Source: Figure based on ISTAT. Key words: DOM NS (Domestic tourists' nights of stay); FOR NS (Foreign tourists' nights of stay).

A difference in the behaviour of domestic and foreign holiday-makers can also be detected by considering the seasonality of the number of arrivals in Italian accommodation. Figures 7.3 and 7.4 show a comparative analysis. One can see that the seasonality of arrivals of foreigners shows overall smaller variations for the months between April and October, with July showing the highest number of foreign arrivals. On the other hand, Italians prefer August. Overall, the domestic seasonal distribution exhibits larger variations.

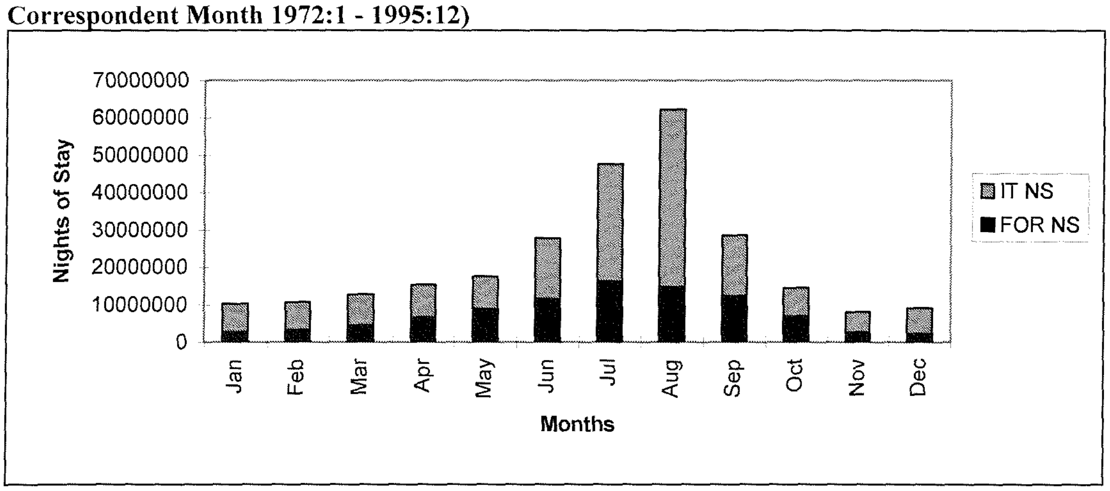
Figure 7. 3 Number of Arrivals of Foreign and Domestic Tourists: Seasonality (Averages for Each of the Correspondent Month 1990:1 - 1995:12)



Source: Figure based on ISTAT. Key words: IT AR (Italian tourist arrivals); FOR AR (Foreign tourist arrivals).

In terms of length of stay, foreigners prefer July and August, followed by September and June. The latter months are in general characterised by lower prices for tourist goods and services, milder temperatures and less congestion. Once again, the seasonality for domestic nights of stay presents a more irregular distribution with the highest peak in August.

Figure 7. 4 Nights Spent by Foreign and Italian Tourists: Seasonality (Averages for Each of the Correspondent Month 1972:1 - 1995:12)



Source: Figure based on ISTAT. Key words: IT NS (Italian tourist nights of stay); FOR NS (Foreign tourist nights of stay).

7.2.2 International Flows And Seasonality

At this point it is worth giving an account of the characteristics which are common to and/or differentiate the main source countries of tourism for Italy as a destination, that is: Belgium, France, Germany, Japan, Sweden, Switzerland, United Kingdom and United States. The number of arrivals and nights spent in Italian accommodation by foreign tourists are reported in Table 7.1. Note that the percentages are calculated over the whole period 1972-1995.

Table 7. 1 Number of Arrivals and Nights of Stay by Country of Residence: Average 1972 - 1995 (in percentage)

	Bel	Fra	Ger	Jap	Swe	Swi	UK	USA	Sum	Others	TOT
AR	2.57	10.08	29.85	2.70	1.30	5.36	6.64	10.61	69.10	30.90	100.00
NS	3.18	7.44	42.28	1.18	1.41	6.05	6.99	5.87	74.39	25.61	100.00

Source: Table based on ISTAT.

Two more exhaustive tables (Tables 7.2 and 7.3) are given in order to consider the evolution of the flows of tourism in Italy by countries of residence. As far as the countries under study are concerned, the number of arrivals as a whole (*Sum*) has a peak in 1972 with 72.8%, whereas the minimum level has been reached in 1975 with 65.3%. In terms of number of nights spent in registered accommodation, the maximum percentage is in 1989 with 83.2% and the minimum percentage in 1992 with 69.7%. Within the eight countries under analysis, the highest number of tourists that choose Italy as a destination are from Germany, followed by United States and France, and United Kingdom together with Switzerland (see also Table 7.1). The country with the lowest percentage among holiday-makers in Italy is Sweden with a downward trend over the period under consideration. Japan is included as it shows an upward trend along the three decades both in terms of arrivals and length of stay.

Note also that the other origin countries, as aggregated (*i.e. Others*), show the highest percentage of number of arrivals and nights spent in Italian accommodation in 1992.

Table 7. 2 Number of Arrivals of Tourists by Country of Residence: 1972-1995 (Percentages)

Years	Bel	Fra	Ger	Jap	Swe	Swi	UK	USA	Sum	Others	TOT
1972	2.8	11.4	23.9	1.7	1.5	4.6	8.2	18.7	72.8	27.2	100.0
1973	3.0	11.2	24.9	2.3	1.4	4.9	7.9	16.8	72.4	27.6	100.0
1974	3.1	9.8	26.6	2.1	1.4	5.4	6.8	14.5	69.7	30.3	100.0
1975	3.4	10.4	26.2	2.1	1.4	5.2	6.6	10.1	65.3	34.7	100.0
1976	3.2	12.3	26.9	2.7	1.3	5.8	6.5	13.6	72.4	27.6	100.0
1977	3.1	10.7	27.6	1.6	1.4	5.1	6.3	12.3	68.1	31.9	100.0
1978	3.2	10.5	29.5	1.4	1.4	5.4	6.7	10.9	69.0	31.0	100.0
1979	3.3	11.4	30.7	1.5	1.3	5.5	6.6	8.9	69.2	30.8	100.0
1980	3.1	11.4	30.9	1.3	1.4	5.3	7.5	8.6	69.6	30.4	100.0
1981	2.9	11.3	30.3	1.4	1.3	5.5	7.5	8.7	69.0	31.0	100.0
1982	2.7	11.7	31.1	1.4	1.3	5.6	7.2	9.6	70.5	29.5	100.0
1983	2.7	9.5	31.3	1.4	1.3	5.6	7.2	9.6	68.7	31.3	100.0
1984	2.1	10.1	30.0	1.5	1.3	5.8	6.4	15.1	72.2	27.8	100.0
1985	2.0	10.0	30.1	1.4	1.2	5.6	6.0	14.7	71.2	28.8	100.0
1986	2.3	10.8	33.6	1.7	1.5	6.1	7.0	7.1	70.1	29.9	100.0
1987	2.2	10.1	33.2	2.3	1.5	5.8	6.4	9.3	70.7	29.3	100.0
1988	2.3	9.7	33.3	2.7	1.5	5.9	6.2	9.0	70.4	29.6	100.0
1989	2.3	9.6	31.3	3.3	1.4	5.8	6.7	9.4	69.8	30.2	100.0
1990	2.3	9.5	28.4	3.7	1.3	5.3	6.6	10.2	67.3	32.7	100.0
1991	2.4	9.7	31.9	3.3	1.3	5.3	6.2	7.4	67.4	32.6	100.0
1992	2.3	8.8	29.6	3.5	1.2	5.0	6.3	9.4	66.3	33.7	100.0
1993	2.3	8.8	28.5	5.0	1.0	4.9	6.2	9.6	66.3	33.7	100.0
1994	2.2	8.6	29.5	5.2	1.0	4.7	6.2	9.6	67.0	33.0	100.0
1995	2.3	8.2	29.6	5.9	0.9	4.7	5.8	9.3	66.7	33.3	100.0

Source: Table based on ISTAT.

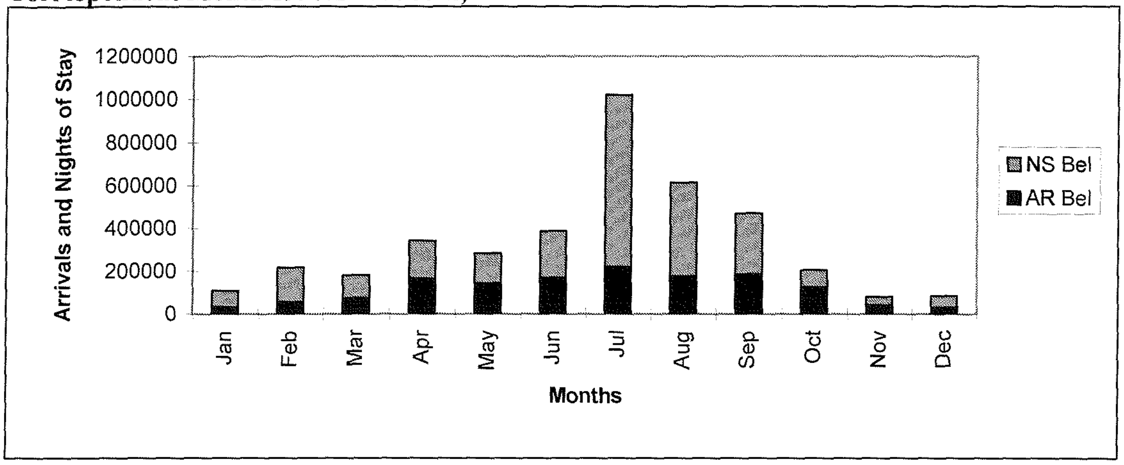
Table 7. 3 Number of Nights Spent by Tourists from Country of Residence: 1972-1995 (Percentages)

Years	Bel	Fra	Ger	Jap	Swe	Swi	UK	USA	Sum	Others	TOT
1972	3.3	8.7	38.7	0.7	1.9	5.5	7.8	9.6	76.2	23.8	100.0
1973	3.5	8.4	40.6	1.0	1.7	5.7	7.3	8.2	76.5	23.5	100.0
1974	3.6	7.4	42.2	0.9	1.6	6.1	6.2	7.3	75.3	24.7	100.0
1975	3.8	7.5	42.0	1.0	1.5	5.8	6.1	5.3	73.0	27.0	100.0
1976	4.2	9.0	41.2	1.2	1.6	6.4	6.4	6.9	76.9	23.1	100.0
1977	3.8	7.7	41.6	0.7	1.5	5.7	5.9	6.5	73.3	26.7	100.0
1978	4.0	7.4	43.3	0.6	1.6	5.8	6.3	5.3	74.2	25.8	100.0
1979	4.2	7.8	43.7	0.6	1.3	6.0	6.2	4.5	74.3	25.7	100.0
1980	3.8	7.8	43.9	0.5	1.4	6.9	7.2	4.3	75.9	24.1	100.0
1981	3.7	8.1	43.0	0.6	1.4	5.6	6.5	4.4	73.3	26.7	100.0
1982	3.4	8.3	43.5	0.6	1.4	6.0	7.1	4.9	75.2	24.8	100.0
1983	3.5	6.9	43.9	0.6	1.5	6.2	7.4	5.1	75.1	24.9	100.0
1984	2.5	7.4	43.5	0.7	1.3	6.6	6.5	7.7	76.2	23.8	100.0
1985	2.5	7.5	43.3	0.7	1.3	6.5	6.2	7.6	75.5	24.5	100.0
1986	2.7	7.6	45.1	0.7	1.4	6.7	7.2	4.1	75.5	24.5	100.0
1987	2.5	7.3	44.7	0.9	1.5	6.4	6.6	5.2	75.2	24.8	100.0
1988	2.7	7.0	44.9	1.1	1.5	6.6	6.2	5.0	74.9	25.1	100.0
1989	2.8	6.9	42.5	1.5	1.5	6.4	15.8	5.6	83.2	16.8	100.0
1990	2.8	7.3	38.7	1.9	1.4	6.0	7.1	6.3	71.5	28.5	100.0
1991	2.8	7.2	41.5	1.6	1.5	5.9	6.2	4.6	71.3	28.7	100.0
1992	2.9	6.7	39.7	2.3	1.3	5.7	5.2	5.9	69.7	30.3	100.0
1993	2.8	6.7	39.7	2.5	1.1	5.6	6.4	6.3	71.1	28.9	100.0
1994	2.7	6.4	40.1	2.6	1.0	5.3	6.7	6.4	71.2	28.8	100.0
1995	2.8	6.3	40.3	2.9	1.0	5.2	6.1	6.1	70.6	29.4	100.0

Source: Table based on ISTAT.

In Section 7.2.1, evidence has been given that domestic and foreign tourists show a different seasonal behaviour. Below, the seasonal pattern for number of arrivals and length of stay with respect to each origin country under study is compared. One can consider the seasonality that is calculated as an average for each month between January 1990 and December 1995. Starting with Belgium, one notices that tourists tend to arrive in Italy in a slightly larger number in July; however, the arrivals are roughly uniformly distributed between April and September. In terms of nights of stay, tourists from Belgium spend the longest holidays in July, August and September.

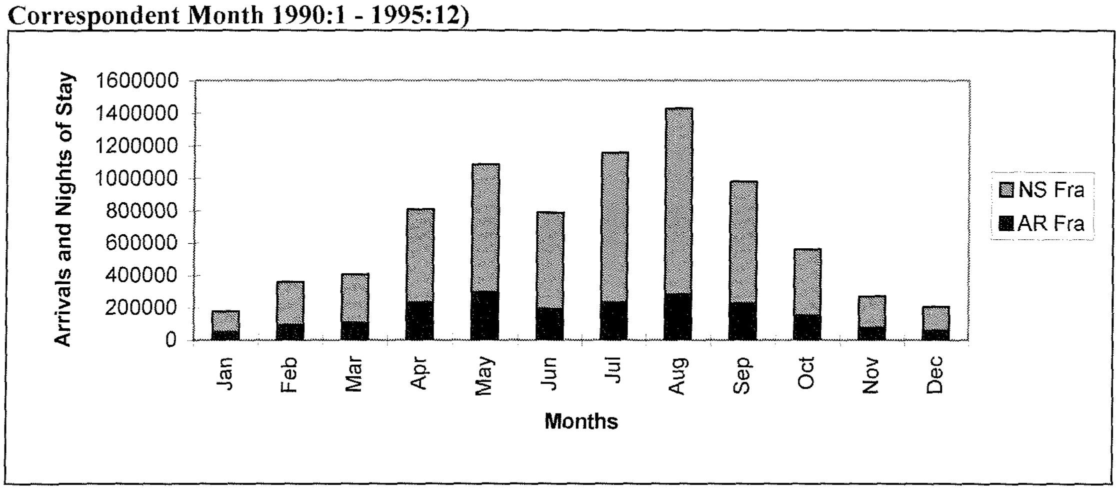
Figure 7.5 Arrivals and Nights of Stay for Belgium: Seasonality (Averages for Each of the Correspondent Month 1990:1 - 1995:12)



Source: Figure based on ISTAT. Key words: NS Bel (Nights of Stay tourists from Belgium); AR Bel (Tourists' arrivals from Belgium).

The peak month of arrivals for French tourists is May, followed by August. French, like Italians, spend the greatest length of time on their holidays in August and July. However, months during the low season, May and September, are also popular.

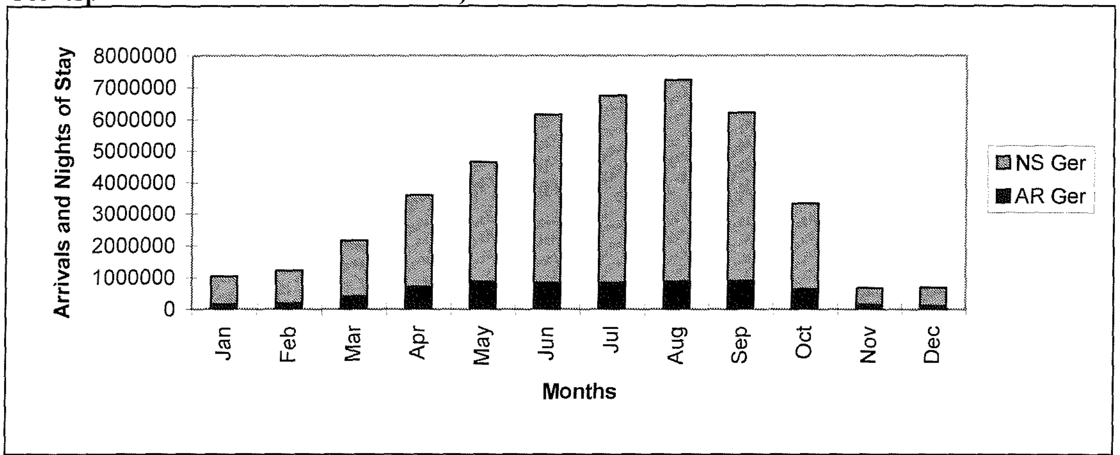
Figure 7.6 Arrivals vs Nights of Stay for France: Seasonality (Averages for Each of the Correspondent Month 1990:1 - 1995:12)



Source: Figure based on ISTAT. Key words: NS Fra (Nights of Stay tourists from France); AR Fra (Tourists' arrivals from France).

Germans arrivals present an uniform distribution between May and September, with some reduction in April and October. The peak month for the number of nights spent in Italian accommodation is August, followed by July, June and September.

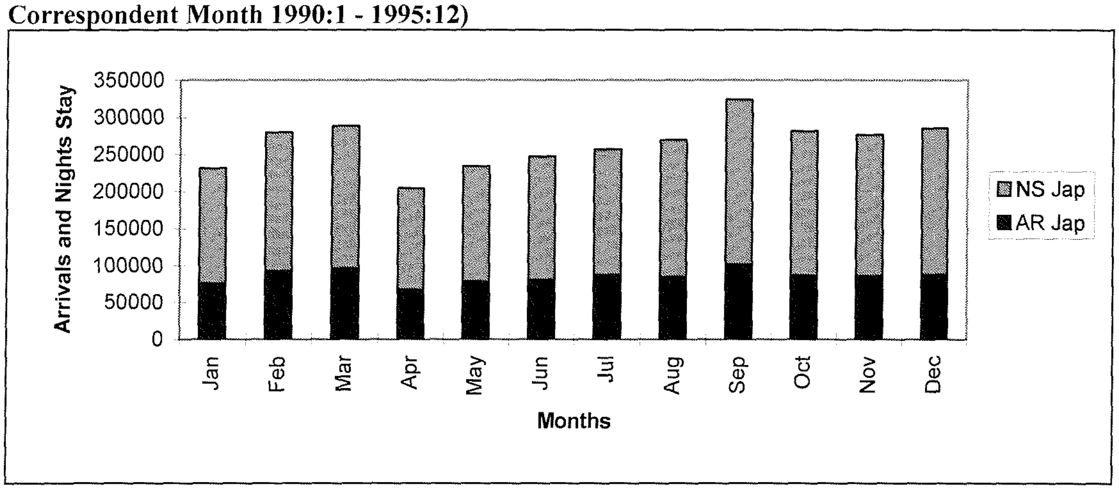
Figure 7.7 Arrivals and Nights of Stay for Germany: Seasonality (Averages for Each of the Correspondent Month 1990:1 - 1995:12)



Source: Figure based on ISTAT. Key words: NS Ger (Nights of Stay tourists from Germany); AR Ger (Tourists' arrivals from Germany).

More interesting is the seasonal pattern for Japanese tourists, which shows a nearly uniform distribution of arrivals throughout the year. September, March and February are the months with the highest number of arrivals. The same seasonal distribution can be noticed for the length of stay. Japanese more than other foreigners seem to prefer less crowed and low season months holidays, *i.e.* September, March and February.

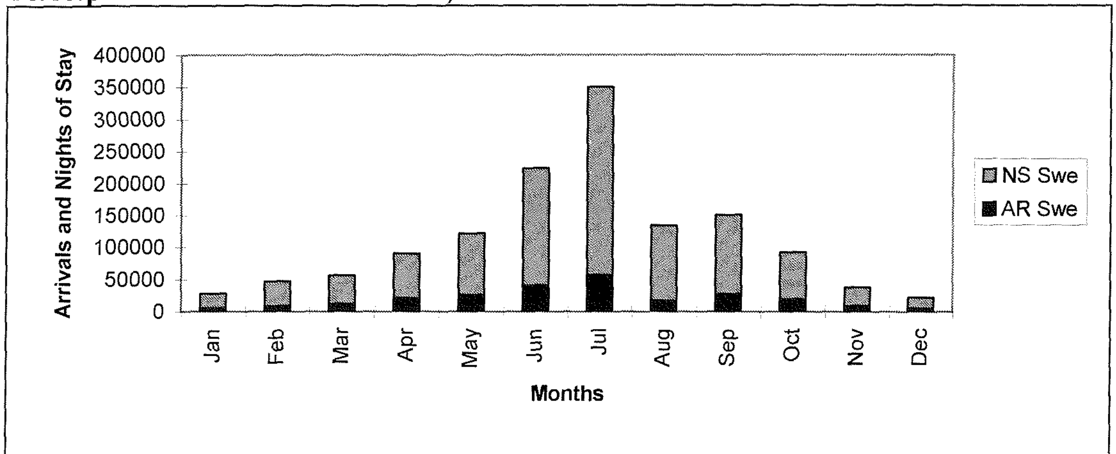
Figure 7.8 Arrivals and Nights of Stay for Japan: Seasonality (Averages for Each of the Correspondent Month 1990:1 - 1995:12)



Source: Figure based on ISTAT. Key words: NS Jap (Nights of Stay tourists from Japan); AR Jap (Tourists' arrivals from Japan).

The highest number of arrivals from Sweden occurs in July followed by June. The troughs occur in winter months. A similar seasonal distribution can be seen for length of stay.

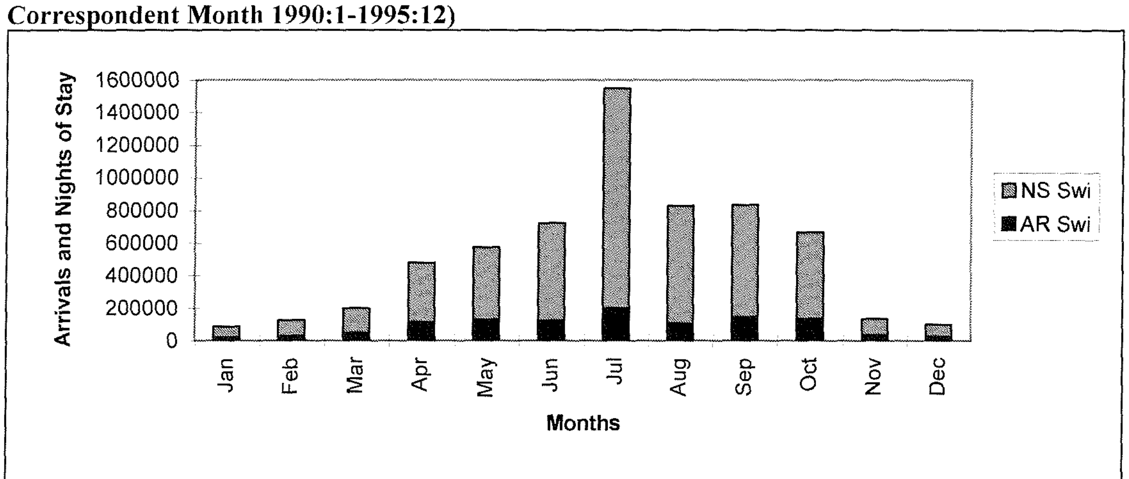
Figure 7.9 Arrivals and Nights of Stay for Sweden: Seasonality (Averages for Each of the Correspondent Month 1990:1-1995:12)



Source: Figure based on ISTAT. Key words: NS Swe (Nights of Stay tourists from Sweden); AR Swe (Tourists' arrivals from Sweden).

Arrivals from Switzerland are concentrated in July, where the longest period of stay also occurs. The months between April and October follow in terms of number of arrivals and August, September and June for nights of stay. The winter months represent the troughs.

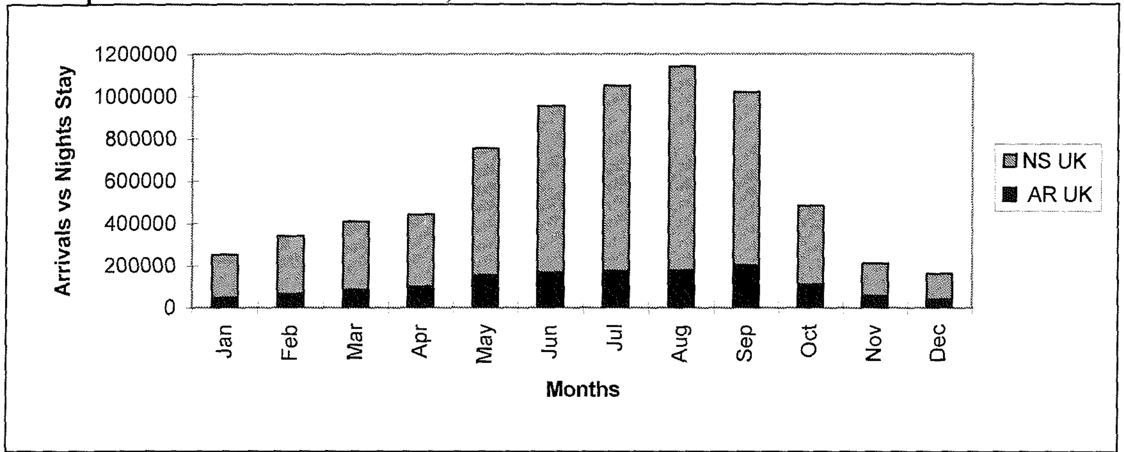
Figure 7.10 Arrivals and Nights of Stay for Switzerland: Seasonality (Averages for Each of the Correspondent Month 1990:1-1995:12)



Source: Figure based on ISTAT. Key words: NS Swi (Nights of Stay tourists from Switzerland); AR Swi (Tourists' arrivals from Swizterland).

As far as the United Kingdom is concerned, an almost equal distribution of arrivals of tourists can be noticed between June and September. The greatest number of nights spent in Italian accommodation is concentrated in August, followed by July, September and June.

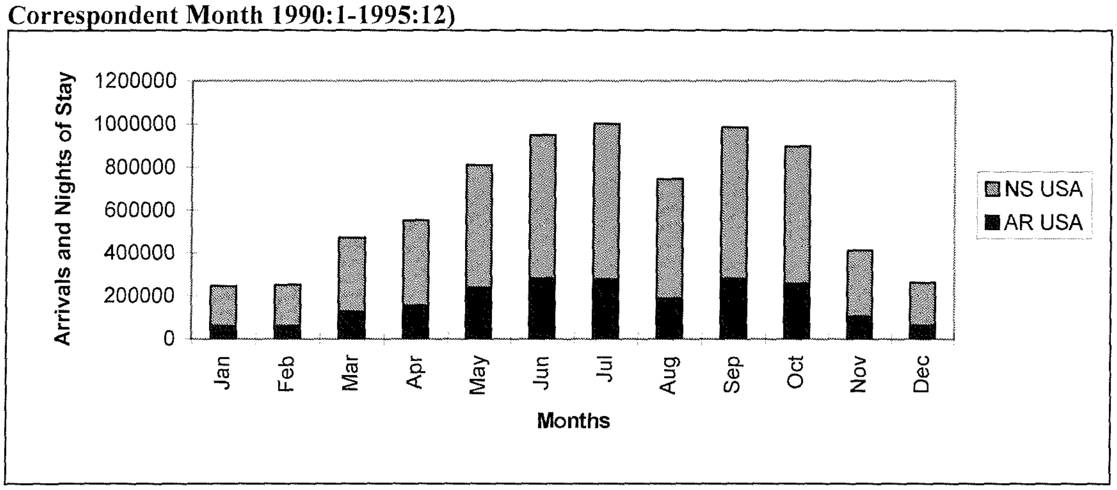
Figure 7.11 Arrivals and Nights of Stay for the UK: Seasonality (Averages for Each of the Correspondent Month 1990:1-1995:12)



Source: Figure based on ISTAT. Key words: NS UK (Nights of Stay tourists from the United Kingdom); AR UK (Tourists' arrivals from the United Kingdom).

June, September and July are months in which the greatest number of arrivals from the United States can be seen. A similar distribution can be noticed in terms of nights of stay period, with the highest concentration in July. Note also that August is less preferred than spring and autumn months.

Figure 7.12 Arrivals and Nights of Stay for USA: Seasonality (Averages for Each of the Correspondent Month 1990:1-1995:12)



Source: Figure based on ISTAT. Key words: NS USA (Nights of Stay tourists from the United States); AR USA (Tourists' arrivals from the United States).

7.3 NUMBERS VERSUS EXPENDITURE

The estimation of international tourism demand in Italy in this thesis is done by using tourist expenditure as the dependent variable. In the following subsections, some definitions and a comparison between numbers and expenditure are given.

7.3.1 Some Definitions

It is important to give a prior definition of tourism expenditure. According to the World Tourism Organisation (WTO) tourism expenditure is defined as "the total consumption expenditure made by a visitor or on behalf of a visitor for and during his/her trip and stay at destination" (Nordström, 1996, p.15). Total tourist expenditure measures a global quantity. In particular, tourist expenditure in any country can be expressed as the product of three factors: the number of tourists, average length of stay and average expenditure *per diem*. It may be very important to know which of the factors is responsible for a given change in expenditure. For example, a fall in expenditure may be accompanied by an increase in numbers, reflecting a decrease in the average length of stay and/or average expenditure *per diem*.

Tourist expenditure data are collected using three different methods, that is: bank records of foreign exchange transactions, surveys of tourists and surveys of tourism establishments. It appears that a good indicator for the real demand of tourism can be obtained by surveys that include information on private consumption behaviour for different kinds of goods and services such as accommodation, transportation, food and so on.

In Italy the main source of tourist expenditure data are bank records. In particular, tourist expenditure data are collected as bank records of foreign exchange transactions that can be considered as a proxy for tourism expenditure. The item "Foreign travel" in the balance of payments contains the expenditure of the "traveller". The "traveller", in this particular context, is defined as a person who spends a given period of time in another economy with a purpose different from working within that economy as an employee of the visited country and without becoming a citizen of that country. Travellers can be divided into two categories: a) excursionists, those who visit a foreign country for less than 24 hours; b) tourists, those who spend at least one night in the foreign country.

The item "Foreign travel" includes a consumption basket of goods and services such as: accommodation, refreshments, amusements, souvenirs and means of transportation within the visited country. However, it does not include expenditure for international travel. In more detail, the item "Foreign travel" includes the following components:

a) bank transfers on residents' and non-residents' account, for tourism, business, health, study and for other tourist services;

b) transactions with credit card issuers;

c) purchases/sales of petrol coupons;

d) forwarding/receiving of Italian coupons;

e) direct negotiation of bills, coins and other means of payments, denominated in foreign currencies or in lire, with residents and non-residents (traveller cheques, drawings from cash dispensers, bank cheques with value up to 20 million lire).

Ad hoc criteria have been used to avoid statistical problems. In the Seventies, for example, there was a realisation that the remittances of Italian banknotes by Swiss banks in Italy was due to capital investments rather than activities for tourism. Such remittances, originally exported to Italy illegally, were included into the item "movements of capital". However, from 1987 on, since this phenomenon was over the statistical computation was back to the *status quo ante*. Another under-estimation of tourist receipts and expenditures, in the Seventies, was given by the monetary restrictions in terms of the maximum amount of money that could be taken into another country. All the banknotes which circulated outside the bank system were not subject to any statistical computation. From the second half of the Eighties, given the increase of the maximum amount of money portable into another country, the quality of the data improved. In 1988, currency liberalisation extended the variety of means of payments that are used extensively in financing tourist expenditure (Banca D'Italia, 1995, pp. 7-10).

Ballatori and Vaccaro (1992) point out further limitations of tourism expenditure data. For example, they do not give any information about the motives for tourism (holiday, business, sport, health, etc); moreover, these data refer to the time in which the currency transaction occurred, whereas no information is given on the moment in which the expenditure takes place (pp. 206-216).

7.3.2 A Comparison Between Numbers And Expenditure

As previously stated, it would be of interest to better understand the relationship between numbers and expenditure. For this purpose international tourists' arrivals, nights of stay, nominal tourism receipts and real tourism receipts are compared. Note that the real tourism receipts are calculated as the ratio between nominal tourism receipts and the consumer price index in Italy (1990=100).

A graphical representation of the aforementioned series, expressed in logarithm, is given in Figure 7.13.

Figure 7.13 International Arrivals (*LAR*), Nights of Stay (*LNS*), Nominal Receipts (*LNT*-million lire) and Real Receipts (*LRT*-thousand lire). Figures in logarithm (1972-1995)

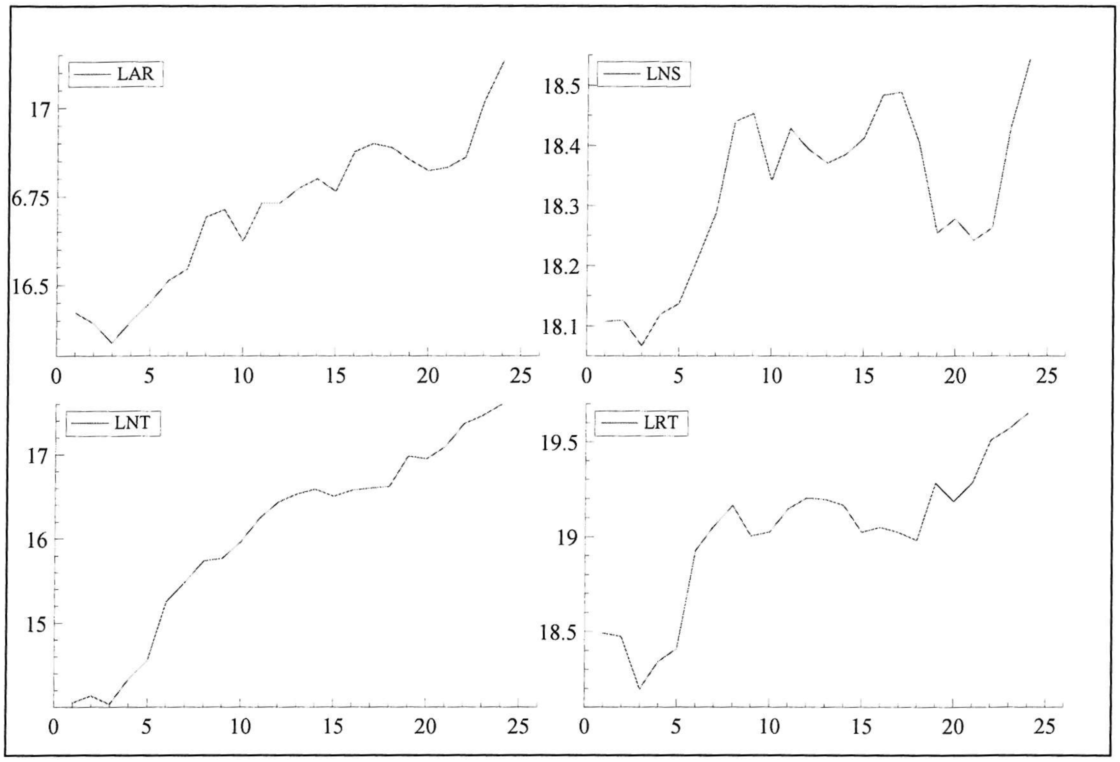


Table 7.4 gives a more exhaustive comparison amongst the series. Note that in this table one excludes 1990 as a year of comparison. Data after May 1990 are not comparable with the previous data, since new currency regulations were introduced due to the liberalisation of capital movements (see Bilancia Valutaria del Turismo, ISTAT, 1990). Accordingly, one omits the transition year, 1990, from Table 7.4. Later econometric estimations use data only up to May 1990.

Table 7.4 Foreign Arrivals, Nights of Stay and Receipts in Italy (1973-1995)

YEAR	Foreign Arrivals of Tourists	Annual Growth (%)	Foreign Tourists Nights of Stay	Annual Growth (%)	Nominal Tourism Receipts (million lire)	Annual Growth (%)	Real Tourism Receipts (lire)	Annual Growth (%)
1973	13.157.569	-2.9	73.264.129	0.2	1.377.000	8.7	105.303.531.071	-1.9
1974	12.441.657	-5.4	70.235.832	-4.1	1.244.600	-9.6	79.880.104.026	-24.1
1975	13.234.355	6.4	73.980.562	5.3	1.683.500	35.3	92.388.910.952	15.7
1976	13.929.798	5.3	75.298.862	1.8	2.101.200	24.8	98.773.484.395	6.9
1977	14.836.118	6.5	81.094.972	7.7	4.201.500	100.0	166.183.924.832	68.2
1978	15.321.451	3.3	87.552.283	8.0	5.334.100	27.0	188.889.253.040	13.7
1979	17.749.393	15.8	101.955.865	16.5	6.815.610	27.8	210.239.504.240	11.3
1980	18.121.622	2.1	103.282.488	1.3	7.034.200	3.2	178.997.268.604	-14.9
1981	16.579.848	-8.5	92.383.478	-10.6	8.585.200	22.0	182.781.171.954	2.1
1982	18.458.567	11.3	100.759.113	9.1	11.279.800	31.4	206.222.031.094	12.8
1983	18.478.878	0.1	97.297.512	-3.4	13.721.200	21.6	218.755.571.404	6.1
1984	19.265.301	4.3	95.162.370	-2.2	15.098.700	10.0	217.264.322.417	-0.7
1985	19.783.976	2.7	96.524.499	1.4	15.952.900	5.7	210.257.062.338	-3.2
1986	19.092.676	-3.5	99.286.309	2.9	14.691.000	-7.9	182.951.432.130	-13.0
1987	21.356.759	11.9	106.493.689	7.3	15.782.808	7.4	187.640.230.399	2.6
1988	21.851.403	2.3	107.030.118	0.5	16.138.880	2.3	182.635.382.874	-2.7
1989	21.607.711	-1.1	98.524.812	-7.9	16.444.000	1.9	175.125.686.271	-4.1
1991	20.241.217	-3.0	86.734.917	2.4	22.852.527	-3.4	215.082.607.059	-9.1
1992	20.424.982	0.9	83.642.567	-3.6	26.447.435	15.7	236.683.734.805	10.0
1993	21.025.353	2.9	85.430.773	2.1	34.625.046	30.9	296.658.969.013	25.3
1994	24.663.870	17.3	101.004.689	18.2	38.307.722	10.6	315.592.931.484	6.4
1995	27.581.077	11.8	113.000.571	11.9	43.717.611	14.1	341.833.147.846	8.3

Source: Calculations based on ISTAT, *Ufficio Italiano dei Cambi* and *Banca d'Italia*.

In the first period (1973-1989), it is interesting to note that in 8 out of 17 cases, arrivals, nights of stay, nominal and real receipts move in the same direction. In 1974, for example, Italian tourism declines both in terms of numbers and expenditure. The highest expansion in terms of receipts has been experienced in 1977 with a continuous growth in the following two years. The first half of the Eighties sees a decline in the growth of number of tourists. With the exception for the year 1982, where the number of foreign arrivals reach almost 18.5 million, the nights of stay in Italian registered accommodation were 100 million, the nominal receipts almost 11 thousand billion lire, and, finally, the real receipts two hundred billion lire. Note also that the decline of growth for Italian tourism in these years is picked up by the nominal receipts figures only in 1986, with a fall of almost 8% over the previous year. The year 1987 sees an increase in Italian tourism, followed by a fall in 1989.

In the second period (1991-1995), 3 out of the 5 times these series move in the same direction. After a decline in the growth of tourism, more evident in terms of numbers than receipts, there is a new upward trend which could be associated with the devaluation of the lira in 1992. Other factors have positively influenced Italian tourism such as the war in Yugoslavia which negatively affected the transit towards Turkey and Greece.

Overall, there are many instances in which the differences amongst the four series are of considerable magnitude and many others in which they move in a different direction. In terms of simple correlation analysis for the first period (1973-1989) the values are the following: $r(AR,NS)=0.92$, $r(NT,RT)=0.756$, $r(AR,NT)=0.95$, $r(AR,RT)=0.76$, $r(NS,NT)=0.83$ and, finally, $r(NS,RT)=0.84$. The highest positive correlation is given by the total number of foreign arrivals of tourists (AR) and the nominal tourism receipts (NT). Note also that the total number of nights spent in registered accommodation by foreigners (NS) shows a higher correlation with the real tourism receipts (RT) than the total number of arrivals with the latter series.

For the second period (1991-1995) the values are the following: $r(AR,NS)=0.99$, $r(NT,RT)=0.996$, $r(AR,NT)=0.91$, $r(AR,RT)=0.87$, $r(NS,NT)=0.85$ and, finally, $r(NS,RT)=0.81$. The highest positive correlation is given by the nominal tourism receipts (NT) and the real tourism receipts (RT), and by the total number of foreign arrivals of tourists (AR) and total number of nights of stay. The pair of total number of arrivals (*AR*) and nominal tourism receipts (*NT*) presents a correlation value equal to 0.91.

7.4 CONCLUSIONS

This chapter has been dedicated to a general discussion of the demand for tourism in Italy. A graphical basis has been provided for distinguishing domestic from international tourism demand. Differences have been identified both in terms of historic evolution and seasonality of demand. These findings encourage the author to distinguish the two components.

The characteristics for each of the main origin countries have been investigated. Germany, France, U.S.A., UK and Switzerland are the clients with the highest number of arrivals and nights of stay in Italy. These countries, except Switzerland, have also shown a more regular seasonal pattern, more visible in terms of nights of stay. The other source countries, that is Belgium, Sweden and Switzerland are characterised by an irregular seasonal distribution. Interestingly, Japan, that has shown an upward trend during the years between 1972 and 1995, is characterised by an almost uniform distribution in terms of arrivals and nights of stay.

Some definitions have been provided for tourism expenditure and a description of the method used by the Bank of Italy in collecting the tourism receipts data. The other aim of the chapter has been to make a comparison between numbers and expenditure. The sample period has been divided into two, as a discontinuity of the time series is due to different currency regulations introduced from June 1990. As far as the first sample period is concerned (1973-1989), a strong positive correlation has been found between number of foreign arrivals and nominal tourism receipts. The lowest correlation is obtained for the pair nominal and real receipts. Note also that the correlation between nights of stay and real tourism receipts has been found higher than the correlation for the pair arrivals and real tourism receipts. The latter finding seems to confirm the belief that the longer a tourist stays in a certain destination the more he/she is likely to spend (Sheldon, 1993).

CHAPTER 8.

ESTIMATING THE DEMAND FOR ITALIAN TOURISM

Aim of the Chapter:

To examine and model the international demand of tourism in Italy using monthly tourist expenditure.

8.1 INTRODUCTION

This chapter is dedicated to estimating Italian international tourism demand. As far as Italy is concerned, data on the actual amounts spent by tourists from the main origin countries do not exist for the period under study. Hence, tourist demand will be expressed in terms of tourist receipts defined in terms of foreign currency exchange transactions of value less than 20 millions of lire (see Chapter 7, Section 7.3.1). In the present study, monthly data will be employed for the period 1972:1-1990:5; as already pointed out in Chapter 7, a new currency regulation has been introduced from June 1990 on.

In this chapter, two distinct variables will be used as dependent variables: real tourist receipts and a weighted budget share for the main origin countries of tourism to Italy. In the majority of time-series empirical studies on tourism, real tourist receipts are employed as the dependent variable (*e.g.* Lee *et al.*, 1993; García-Ferrer and Queralt, 1997). On the other hand, the budget share variable is commonly used in panel data studies (Fujii *et al.*, 1985; Syriopoulos and Sinclair, 1993). The purpose of this chapter is to use the weighted average budget share of tourism in Italy in a time-series context. The aim is to understand which of the two variables best can be used to model the demand for tourism.

A "pre-modelling" analysis is carried out in order to identify the properties of the variables that one expects to influence tourism demand in Italy according to economic theory. Hence, once the integration and possibly cointegration status of such variables is established, one makes use of the LSE methodology. In this way, it will be possible to determine income and price elasticities that will be evaluated theoretically. It will also be possible to identify the explanatory power of other qualitative variables such as seasonal dummies and an "Easter" dummy included in the model.

At this point, a brief note has to be given on the supply constraint. One of the main assumptions in estimating the international demand for tourism in Italy is that one can assume the existence of no supply constraint. As Syriopoulos (1995) notes, "it is reasonable to accept that the supply of tourism does not impose any constraints on tourism demand" (p.321). This assumption is based on two arguments. The first argument is that hotels and tourist infrastructure are constructed to satisfy not only the current consumption but also the consumption in the future. Secondly, tourists make increasing use of accommodation other than registered accommodation, *e.g.* second houses, apartments and villas (there are many examples in Tuscany).

The chapter is divided in the following manner. Section 8.2 is dedicated to the use of a single equation rather than a system of equation modelling. In Section 8.3, and its subsections, definitions of the economic variables of interest are given and the integration and possible cointegration status of these variables is investigated. A linear and a non-linear model are estimated for the real tourism receipts. Finally, a further investigation is carried out treating the dependent variable as I(1). Section 8.4 and its subsections are articulated as follows. A trend analysis gives insight as to whether the seven origin countries under study (*i.e.* France, Germany, Japan, Sweden, Switzerland, UK and USA) constitute an appropriate aggregation in defining the weighted budget share. The choice of the weights will be discussed. Franses' seasonal unit roots test and the ADF test is carried out in order to establish the integration status of the economic series under study. A cointegration analysis amongst the I(1) variables is the objective of Section 8.4.5. In Sections 8.4.6 and 8.4.7, a linear and a non-linear model is run, respectively. Section 8.4.8 is dedicated to a discussion of the results obtained from the seven countries' aggregation for the budget share variable. Sections 8.9 and 8.10 provide a summary of the main findings and conclusions.

8.2 SINGLE EQUATION VERSUS SYSTEM OF EQUATIONS MODELS

In the majority of the studies of tourism demand, the single equation approach has been used. This approach has the advantage to allow the incorporation of variables in each equation that effect one particular country but not others. Moreover, this approach can also be used to estimate the short run as well as the long run dynamics. The main disadvantage of single equation models is that they do not link with microeconomic consumer behaviour theory. On the other hand, some authors suggest that a system of demand equations has the advantage of being able to provide a more rigorous link with economic theory. These models are able to establish the interdependencies, such as complementary or substitutability, amongst competitor countries (O'Hagan and Harrison, 1984; Syriopoulos and Sinclair, 1993; Papatheodorou, 1999). It is also possible to test different restrictions on a representative consumer's behaviour which are related to microeconomic theory. Negativity is the restriction which implies a negative relationship between demand and prices. Homogeneity asserts that a proportional change of a consumer expenditure and all prices does not affect the quantities purchased; symmetry asserts that the consumer choice is consistent; finally, adding-up for which the total expenditure equals the sum of individual expenditures. However, the main limit for these models is that they force the researcher to use the same explanatory variables in all equations of the system, though not important in explaining the demand for tourism in a particular country or countries under study.

As far as this study is concerned, another aspect to take into consideration is the availability of the data at a given frequency. The objective is to use monthly data that are not available from the official statistics (e.g. WTO). Moreover, the majority of the empirical studies on tourism expenditure, with very few exceptions (see Gonzàles and Moral, 1996; Seddighi and Shearing, 1997) employ annual data. On this basis, one will make use of the single equation approach.

The aim is to consider two possible specifications, which can be expressed in general terms as follows:

a) $$EXP = f(PR, RP, EX, SP, DV) \quad (8.2.1)$$

b) $$BS = f(PR, RP, EX, SP, DV) \quad (8.2.2)$$

where:

EXP = tourist receipts from foreigners.

BS = real weighted average budget share for the main source countries of tourists for Italy (*i.e.* France, Germany, Japan, Sweden, Switzerland, UK and USA). Note that countries such as Belgium have been not included, as the frequency of the data is not homogeneous. In particular, private consumption is available only with an annual frequency (see definition below). Further details will be given for the validity of countries' aggregation.

PR = income (as the weighted average industrial production index for the main origin countries).

RP = relative price (consumer price index for destination/weighted average consumer price index for origin countries).

EX = exchange rate as a weighted average for the main origin countries.

RSP = real substitute price (*i.e.* substitute price adjusted for the exchange rate).

DV = dummy variables.

Definitions of the above variables are given in Section 8.3 in more detail.

8.3 ITALIAN TOURIST RECEIPTS AS THE DEPENDENT VARIABLE

This section is dedicated to the estimation of tourist expenditure in Italy using the real tourist receipts (*LREXP*) as the dependent variable.

In Section 8.3.1, the definitions of the explanatory variables under study are provided. Section 8.3.2 is dedicated to the investigation of the integration status of the variables of interest. In Section 8.3.3 a cointegration analysis is carried out for the integrated I(1) variables. In Sections 8.3.4 and 8.3.6, the model is estimated.

8.3.1 Definition Of The Variables

In this section, a definition of the variables under study is provided on basis of the generic function (8.2.1). The dependent variable is constructed as follows:

A) Real Tourist Receipts (*REXP*).

$$REXP_t = EXPt / CPIit,t \qquad (8.3.1.1)$$

where:

EXP_t = Tourist receipts in current billions of lire, in month t (Source: Bank of Italy). As already stated, this is expressed in terms of foreign currency exchange transactions of value less than 20 million lire.

$CPI_{it,t}$ = Consumer price index (1990=100) in Italy, *it*, in month t (Source: ISTAT).

B) Income Proxy (*RPRa*).

The nominal weighted average income proxy with respect to the main origin countries, i, can be expressed by the following formula:

$$NPRt = \frac{\sum_{i=1}^{i=7} wi,t * PRi,t}{\sum_{i=1}^{i=7} wi,t} \qquad (8.3.1.2)$$

However, as dealing with a real dependent variable, the real weighted average income proxy is used and it is defined as follows:

$$RPRat = \frac{\sum_{i=1}^{i=7} wi,t * PRi,t}{\sum_{i=1}^{i=7} wi,t * Pi,t} \qquad (8.3.1.3)$$

where:

$PR_{i,t}$ = index of industrial production (1990=100), in country i in month t (Source: IFS *Datastream*).

$P_{i,t}$ = index of consumer price (1990=100), in country i in month t (Source: IFS IFS *Datastream*).

$w_{i,t}$ = This weight is formed taking into consideration the number of nights spent (say *NS*) by the tourists of each origin country i in all registered accommodation in year t (Source: ISTAT), and it is given by the following formula:

$$Wi,t = \frac{NSi,t}{\sum_{i=1}^{i=7} NSi,t} \qquad (8.3.1.4)$$

Use of substantial lags for the real income proxy implies that vacations are planned well in advance.

C) Relative Price (*RPa*).

The relative price represents the price of Italian tourism to the set of client countries i as previously listed. Such a variable can be expressed by the following formula:

$$RPt = \frac{CPIit,t}{CPIo,t} \qquad (8.3.1.5)$$

where:

$CPI_{it,t}$ = monthly consumer price index (1990=100) in Italy (Source: ISTAT).

$CPI_{o,t}$ = weighted average consumer price index, calculated as follows:

$$CPIo,t = \frac{\sum_{i=1}^{i=7} wi,t * CPIi,t}{\sum_{i=1}^{i=7} wi,t} \qquad (8.3.1.6)$$

where:

$CPI_{i,t}$ = monthly consumer price index (1990=100) in country i and month t (Source: IFS *Datastream*).

$w_{i,t}$ = the weights as defined in formula (8.3.1.4).

D) Exchange Rate (*EXa*).

The weighted average exchange rate with respect to the main origin countries, i, can be expressed by the following formula:

$$EXt = \frac{\sum_{i=1}^{i=7} wi,t * EXi,t}{\sum_{i=1}^{i=7} wi,t} \qquad (8.3.1.7)$$

where:

$EX_{i,t}$ = nominal exchange rate, in country i in month t (Source: *Banca d'Italia*).

$w_{i,t}$ = the weights as defined in formula (8.3.1.4).

E) Real Substitute Price (*RSP*).

In this case, the results achieved for the model of tourism in Sassari Province are followed. Evidence has been found that one could obtain a better specification by disaggregating the real substitute price for each of the pair destination/competitor country. Moreover, one could argue that the inclusion of a weighted average exchange rate for the competitors might create problems of multicollinearity, given the inclusion of the weighted average exchange rate for the source countries. As a reminder, the exchange rate for France is defined by the ratio (lira/dollar)/(franc/dollar), and so on for all the other competitors.

Hence, four different variables have been created which can be expressed as such:

$$RPS_{j,t} = \left(\frac{CPI_{i,t}}{CPI_{j,t}} \right) * \frac{1}{EX_{j,t}} \qquad (8.3.1.8)$$

where:

j = France, Greece, Portugal and Spain.

$CPI_{i,t}$ = monthly consumer price index in Italy (1990 =100) (Source: ISTAT).

$CPI_{j,t}$ = monthly consumer price index in country j (1990 =100) (Source: IFS *Datastream*).

$EX_{j,t}$ = monthly exchange rate, lira per unit of currency of country j (elaborated on IFS source).

The explanatory variables as defined above are represented in Figures 8.1 and 8.2 and expressed in logarithm.

Figure 8.1 Plots for (log) Real Tourist Receipts, Real Industrial Production Index (*LRPRa*), Relative Price (*LRPa*) and Exchange Rate (*LEXa*) (1972:1-1990:5)

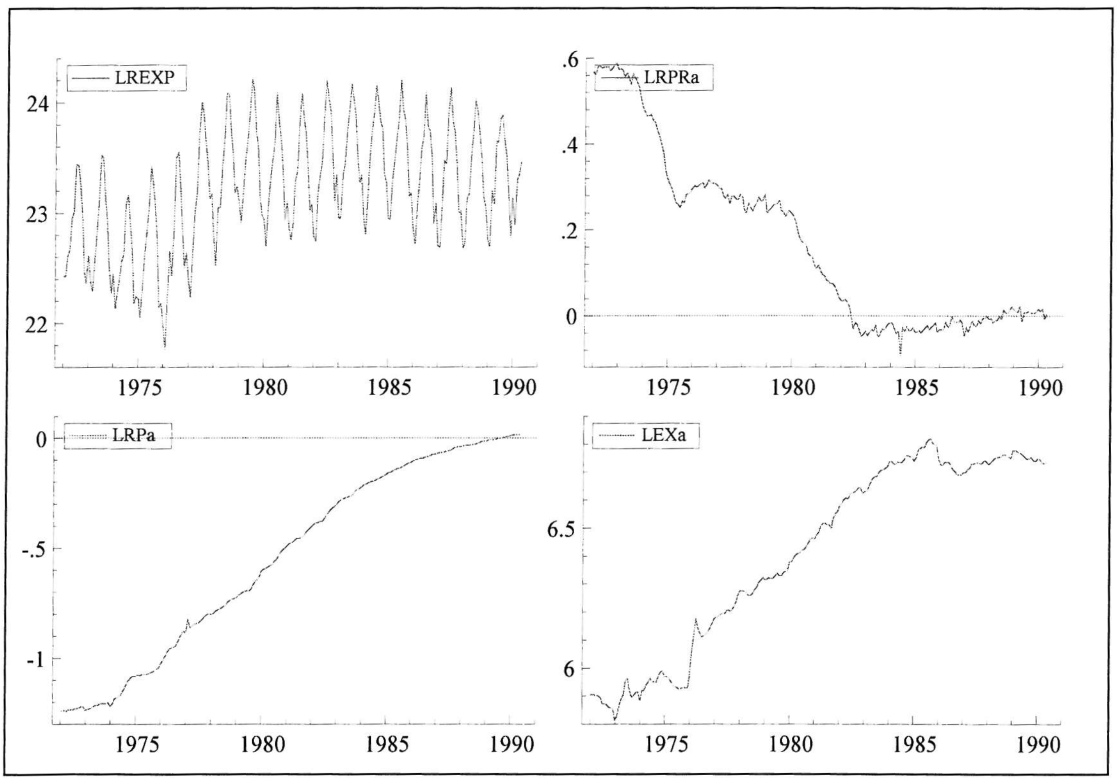
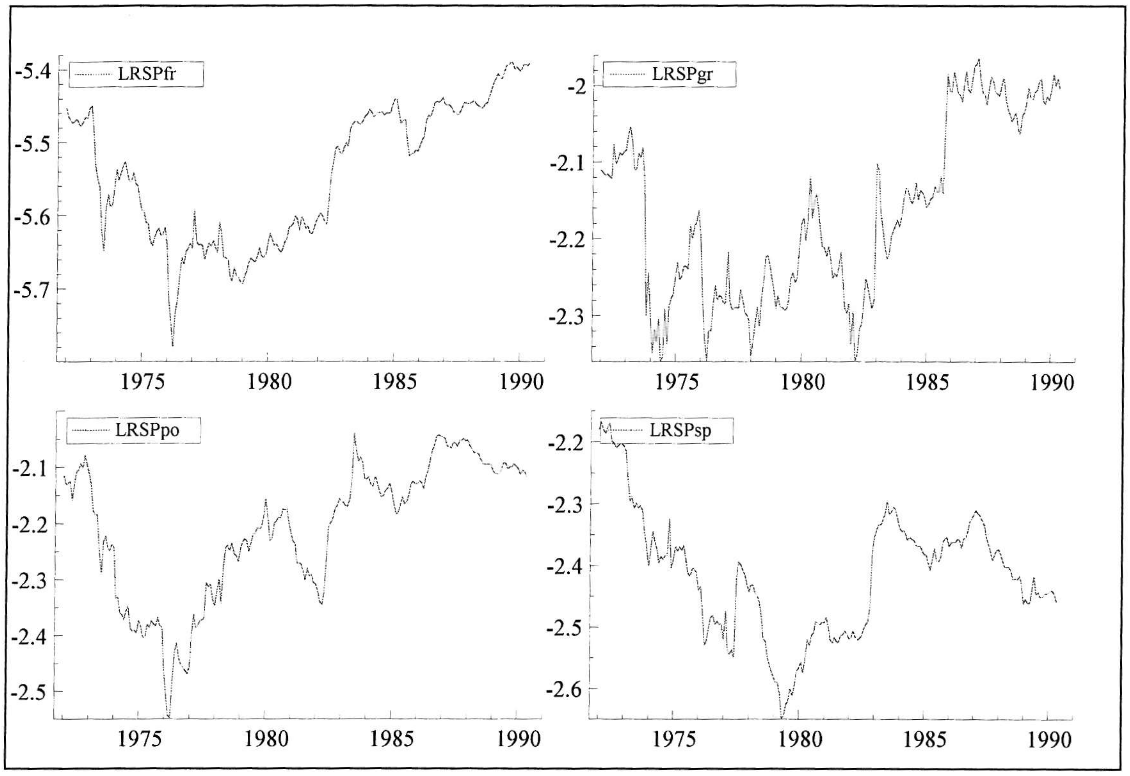


Figure 8.2 (Log) Real Substitute Price: France, Greece, Portugal and Spain (1972:1-1990:5)



As one can notice from Figure 8.1, *LRPRa* shows a downward trend which might make it a poor proxy for income.

8.3.2 Seasonal Unit Roots And Long Run Unit Roots

In this section, an account of the properties of each of the variables under study will be given. As already stated, all the series are expressed in logarithm and the analysis is be carried out for the period between 1972:1 and 1990:5.

The first step consists in testing for the existence of possible seasonal unit roots using Franses' (1991) test. Equation (2.6.2), where μ_t includes a constant, a trend and 11 seasonal dummies, is fitted by OLS to each of the series under study.

Table 8.1 Testing for Seasonal Unit Roots (1972:1-1990:5 - 221 Observations)

t-statistics	Variable							
	LREXP	***LRPRa***	***LRPa***	***LEXa***	***LRSPfr***	***LRSPgr***	***LRSPpo***	***LRSPsp***
π1	-1.031	-1.971	1.697	0.640	-3.420**	-2.685	-3.837****	-2.656
π2	2.231	3.267	2.797	2.752	4.017	4.094	2.656	4.357
π3	-2.722****	-0.890	-3.675****	-3.058****	-4.486****	-4.199****	-4.525****	-5.190****
π4	-4.511****	-5.926****	-7.596****	-6.028****	-3.770****	-4.834****	-4.482****	-4.515****
π5	-5.776****	-7.042****	-4.047****	-5.882****	-6.605****	-7.132****	-7.515****	-6.703****
π6	-5.861****	-7.134****	-5.708****	-6.658****	-6.980****	-6.349****	-7.151****	-6.623****
π7	-0.767****	-1.286****	-0.157	0.412	-2.089****	-1.531****	0.933	-1.639****
π8	-2.645	-1.836	-2.032	-2.627	-1.510	-2.357	-4.276****	-2.450
π9	-1.642	-4.684****	-2.235	-2.139	-6.758****	-5.843****	-7.090****	-5.379****
π10	-4.627****	-8.307****	-4.011****	-5.102****	-4.697****	-6.031****	-4.762****	-6.171****
π11	1.140	-3.048****	-0.727*	-1.877****	-4.854****	-4.098****	-4.224****	-4.477****
π12	-4.221****	-3.405**	-4.268****	-3.187*	-2.680	-3.025*	-3.297**	-2.447
F-statistics	***LREXP***	***LRPRa***	***LRPa***	***LEXa***	***LRSPfr***	***LRSPgr***	***LRSPpo***	***LRSPsp***
π3, π4	14.509****	18.006****	38.647****	24.396****	18.304****	21.941****	20.304****	25.871****
π5, π6	18.258****	27.440****	19.149****	22.221****	25.060****	25.900****	21.842****	24.085****
π7, π8	32.052****	22.979****	11.401****	12.984****	26.231****	35.247****	37.247****	40.535****
π9, π10	11.151****	34.731****	8.095***	13.174****	25.248****	24.478****	27.511****	23.112****
π11, π12	9.647****	24.481****	15.757****	13.940****	32.254****	28.652****	32.581****	26.929****
π3, ...,π12	40.101****	133.398****	34.102****	31.640****	117.868****	103.210****	124.799****	137.387****

Note: The four, three, two and one asterisks indicate that the seasonal unit root null hypothesis is rejected at the 1%, 5%, 10% and 20% level respectively.

As can be seen from Table 8.1, there is no evidence of seasonal unit roots. However, the null hypothesis of the existence of a long run unit root cannot be accepted for *LRSPfr* and *LRSPpo*. The latter result is in line with the ADF test. *LRSPpo*, in fact, appears to be I(0) about a trend; whereas, the long run unit root is accepted for *LRSPfr* by the ADF test (see Table 8.2). Note also that *LRSPgr* is found to be I(0) about a trend from the ADF test, whereas Franses' test suggests this variable to be non-stationary in the level. A divergence appears also for *LREXP* which by the ADF test is found to be stationary in the level (Table 8.2).

There follows the long run ADF unit roots test in order to establish the integration status of each of the variables (Table 8.2).

Table 8.2 Testing Long Run Unit Roots: 1972:1-1990:5

Series	*ADF*(1)	*LAG*(2)
LREXP(c)	- 3.98 **	5
LREXP(c,t)	- 3.57 *	6
LREXP(c,s)	- 1.24	12
DLREXP(c,s)	- 4.32 **	11
LREXP(c,t,s)	- 3.89 *	0
LRPRa(c)	- 2.89 *	3
LRPRa(c,t)	- 1.92	8
DLRPRa(c,t)	- 3.88 *	7
LRPRa(c,s)	- 2.89 *	3
LRPRa(c,t,s)	- 1.89	8
DLRPRa(c,t,s)	- 3.59 *	7
LRPa(c)	- 2.97 *	2
LRPa(c,t)	- 0.83	3
DLRPa(c,t)	- 6.58 **	2
LRPa(c,s)	- 3.03 *	2
LRPa(c,t,s)	- 0.82	3
DLRPa(c,t,s)	- 6.13 **	2
LEXa(c)	- 1.96	3
DLEXa(c)	- 8.40 **	2
LEXa(c,t)	- 0.39	3
DLEXa(c,s)	- 8.67 **	2
LEXa(c,s)	- 1.96	3
DLEXa(c,s)	- 8.08 **	2
LEXa(c,t,s)	- 0.33	3
DLEXa(c,t,s)	- 8.36 **	2
LRSPfr(c)	- 0.51	10
DLRSPfr(c)	- 5.56 **	9
LRSPfr(c,t)	- 3.06	10
DLRSPfr(c,t)	- 5.82 **	9
LRSPfr(c,s)	- 0.73	5
DLRSPfr(c,s)	- 6.63 **	4
LRSPfr(c,t,s)	- 2.95	5
DLRSPfr(c,t,s)	- 6.65 **	4
LRSPgr(c)	- 1.73	0
DLRSPgr(c)	-14.88 **	0
LRSPgr(c,t)	- 3.81 *	0
LRSPgr(c,s)	- 1.49	0
DLRSPgr(c,s)	-14.20 **	0
LRSPgr(c,t,s)	- 3.49 *	0
LRSPpo(c)	- 1.23	5
DLRSPpo(c)	- 7.54 **	4
LRSPpo(c,t)	- 3.63 *	5
LRSPpo(c,s)	- 1.19	2
DLRSPpo(c,s)	- 10.46 **	1
LRSPpo(c,t,s)	- 3.65 *	2
LRSPsp(c)	- 2.93 *	7
LRSPsp(c,t)	- 2.79	0
DLRSPsp(c,t)	-13.64 **	0
LRSPsp(c,s)	- 2.45	0
DLRSPsp(c,s)	-13.02 **	0
LRSPsp(c,t,s)	- 2.62	0
DLRSPsp(c,t,s)	-13.03 **	0

Notes for Table 8.2: (1) Augmented Dickey-Fuller (ADF) statistics with constant (c) critical values: 5%=-2.876 1%=-3.463; when c and t included c.v.: 5%=-3.433 1%=-4.005; c and s included c.v.: 5% = -2.876 1% = -3.463; c,t and s are included c.v.: 5% = -3.433 and 1% = -4.005;
(2) Number of lags set to the first statistically significant lag, testing downward and upon white residuals. Note that ADF(0) corresponds to the DF test.
(3) ** significant at 1%; * significant at 5%.

From the ADF test, it emerges that the dependent variable *LREXP* can be treated as stationary in the level. This finding seems to suggest that tourism in Italy is not increasing over time and it has reached maturity. The relative price (*LRPa*) appears to be I(0) when a constant, or constant and seasonals are included, otherwise to be stationary in the first difference. The real substitute price for Greece (*LRSPgr*) is found to be I(0) about a trend, as is the real substitute price for Portugal (*LRSPpo*). On the other hand, the exchange rate (*LEXa*), the real substitute price for France (*LRSPfr*) and for Spain (*LRSPsp*) are found to be I(1). Interestingly, the income proxy (*LRPRa*) appears to be non-stationary in the level. This finding is in line with the empirical results obtained in Hansen (1995). In this article, it is shown that the (log) U.S. real industrial production turns out to be a random walk.

8.3.3 Possible Cointegration Amongst I(1) Variables

Once the integration status of the variables of interest is established, the next step is to consider the possible existence of a cointegrating relationship amongst the I(1) variables.

A Johansen cointegration analysis is run by including the real substitute price for each of the main competitor countries for Italy *(i.e. LRSPfr*, *LRSPgr*, *LRSPpo* and *LRSPsp*). An unrestricted 13 lag system is estimated that indicates problems of non-normality in the residuals. Three impulse dummies have been created after inspecting the residuals for the equations for *LRSPgr* and *LRSPsp*, that is *i1977p1*, *i1982p12* and *i1983p1*. The inclusion of these dummies reduces the problems in the residuals but does not eliminate them. Hence, a Johansen cointegration analysis is run on an unrestricted 13 lag system including the aforementioned dummies and a time trend, treated unrestrictedly. The system can be reduced, according to the joint F-test, the SC and HQ criteria to 4 lags. One can conclude that the 4 variables do not appear to be cointegrated or stationary.

A second test is carried out on the pair *LRSPgr* and *LRSPpo*, that according to the ADF test have been found to be stationary in the level (see Table 8.2). Starting with an unrestricted 13 lag system, an impulse dummy (*i1983p1*) is included after inspecting the residuals for the *LRSPgr* equation. A time trend is also included. The system can be reduced parsimoniously to 11 lags, in accordance with the joint *F*-test. The results from the Johansen cointegration analysis suggest treating the two variables as stationary in agreement with the ADF test findings. The results are in Appendix F, Table F.1. Note also that the SC and HQ criteria suggest the estimation of a 1 lag system. Nevertheless, from the Johansen analysis, the conclusion is that the two variables are stationary.

The second investigation is carried out on the pair *LRSPfr* and *LRSPsp*. An unrestricted 13 lag system is run with the inclusion of an impulse dummy *i1977p7*, created after having inspected the residuals for *LRSPsp* equation. A constant, a time trend and seasonals are also included unrestrictedly which give the best results in terms of diagnostics. The 13 lag system can be reduced to a 7 lag system, according to the joint *F*-test. Also in this case, there is evidence for the two variables to be stationary. The complete Johansen analysis results are reported in Appendix F, Table F.2. Note, also, that the information criteria SC and HQ suggest running a one lag and two lag system, respectively. Nevertheless, the results indicate that *LRSPfr* and *LRSPsp* are stationary.

In conclusion, there is statistical evidence to believe that each of the real substitute prices can be treated as stationary in the level.

In accordance with economic theory, prices and exchange rate are expected to drift together in the long run. Thus, a cointegration analysis is done on the (log) relative price (*LRPa*) and (log) weighted average exchange rate (*LEXa*). An initial 13 lag system has been run, which includes a constant, trend and seasonals treated unrestrictedly. Impulse dummies created after inspecting the residuals, in order to avoid problems of non-normality, worsen the results in terms of diagnostics, so they are not included in the system. The system can be reduced further, up to 2 lags, as the restriction has been accepted by the joint *F*-test. Moreover, the 2 lag system is suggested also by the SC and HQ criteria (the complete results are reported in Appendix F, Table F.3). From the Johansen cointegration analysis, one infers that there is evidence for the presence of one cointegrating relationship between the two variables. From Table F.3, the equivalent cointegrating vector can be derived and it is given by the first row of β' matrix:

$$CI = LRPa - 0.77989\ LEXa \qquad (8.3.3.1)$$

The expectation is that the two cointegrated variables will present a long run coefficient of one. Thus, the following restriction: β=-1 is tested for the coefficient for the (log) weighted exchange rate; however, the null hypothesis has not been accepted at the 5% level from the χ^2 test[71]. This fact might be due to differences in the inflation rates in each of the countries under consideration.

A further ADF cointegration analysis is carried out on the above two variables. The static models, for the relative price and exchange rate, are estimated by OLS for the period 1972:1 to 1990:5, where a constant is included. The results are as follows:

$$LRPa = -8.8600 + 1.2963\ LEXa + \hat{u}_1 \qquad R^2 = 0.976 \quad DW = 0.08$$

and

$$LEXa = 6.8244 + 0.75271\ LRPa + \hat{u}_2 \qquad R^2 = 0.976 \quad DW = 0.08$$

The second step consists of estimating the static models where a constant and a time trend are included, which gives the following results:

$$LRPa = -5.1591 + 0.0035\ Trend + 0.6590\ LEXa + \hat{u}_1 \qquad R^2 = 0.99 \quad DW = 0.23$$

and

$$LEXa = 7.5963 - 0.00405\ Trend + 1.3437\ LRPa + \hat{u}_2 \qquad R^2 = 0.99 \quad DW = 0.23$$

The saved residuals $\hat{u}_1$ and $\hat{u}_2$, which can be interpreted as the deviations of the generic y_t from the long run path, are tested for a unit root under the null hypothesis of no-cointegration. The number of lags for the ADF test is set to the first statistically significant lag, testing downward and upon white residuals. The initial number of lags in the ADF test is set up to 13. The first significant lag is the third and the t-value for the corresponding coefficient is -1.82. MacKinnon's critical value is equal to -3.36[72] at the 5% that is greater, in absolute value, than -1.82. The null hypothesis cannot be rejected and, thus, there is no evidence for cointegration between the two variables of interest.

---

71 The results for the restriction test on the coefficient is: $\chi^2(1) = 4.9247$ [0.0265] *.

72 The estimated p =5% critical value for T=221 observations is the following: $C(p) = -3.3377 + (-5.967/221) + (-8.98/(221)^2)$.

The same finding has been obtained when, in the static model, a constant plus a trend are included. The critical value is -3.82[73] at the 5% level which is greater, in absolute value, than the t-value of the corresponding coefficient, -3.42, when the model is run with nine lags. Again, no evidence appears of the existence of cointegration between the relative price and weighted exchange rate.

The next step consists of regressing *LEXa* on *LRPa* with just the constant in the static model. Using the cointegration ADF test, based on the statistically significant lag approach, testing downwards, the lag length equals three and the correspondent t-value is -2.01. In this case, MacKinnon's critical value equals -3.36[74] at the 5% level. Thus, this critical value, in absolute term, is greater than the t-value for ρ. Therefore, there is statistical evidence for the existence of no-cointegration.

Including a constant and a trend in the static model, the results are the following. The critical value determined from MacKinnon's parameters, is -3.82[75]. The t-value for the correspondent coefficient for an ADF model of 9 lags equals -4.10 that, in absolute value, is greater than the critical value. Hence, there appears evidence for the existence of cointegration between the two variables. On balance, as argued before, one can consider the Johansen analysis to be more robust. It uses a simultaneous approach which involves the interdependencies of the variables under study.

A cointegration analysis has also been run for the three I(1) variables (*i.e.* real industrial production, relative price and exchange rate). Statistical evidence has been found for the existence of one cointegrating vector (see Appendix F, Table F.4); however, some difficulties appear in interpreting the results on an economic basis. An investigation follows in employing a non-linear transformation for the real industrial production index.

---

73 The estimated p = 5% critical value for T=221 observations is the following: C(p) = -3.7809 + (-9.421/221) + (-15.06/(221)^2).

74 The estimated p =5% critical value for T=221 observations is the following: C(p) = -3.3377 + (-5.967/221) + (-8.98/(221)^2).

75 The estimated p = 5% critical value for T=221 observations is the following: C(p) = -3.7809 + (-9.421/221) + (-15.06/(221)^2).

8.3.4 Estimation Using Real Tourist Receipts

In this section, monthly real tourist receipts are used in estimation for the period between 1972:1 and 1990:5. As already stated, the (log) real industrial production (*LRPRa*) has been found non-stationary in the level. Hence, such a variable needs an appropriate transformation in order to be included in the estimation of the real tourist receipts. As the dependent variable in use is a stationary variable, one can investigate the validity in adopting a logistic transformation. The initial formulation of the equation for the (log) tourist receipts (*LREXP*) can be expressed as:

LREXP = f(LRPRa, SLRPRa, DLRPa, DLEXa, CI, LRSPfr, LRSPgr, LRSPpo, LRSPsp,E, T, D)

Note that the log-linear specification has been tested against the linear form by adopting the Box and Cox test. Details are given in Appendix G.

The first step consists in running a VAR, by which it is possible to identify the lag size of the system. The first system for *LREXP* includes a constant, 11 seasonal dummies, the "Easter" dummy variable, two impulse dummies created in order to avoid problems of non-normality in the residuals (*i1976p5* and *i1990p1*), a time trend (note that all these variables are treated unrestrictedly), the first lag for the cointegrating vector *CI* (where the cointegrating vector is defined as: *CI = LRPa - 0.77989 LEXa*, from the results obtained using the Johansen cointegration analysis as reported in Section 8.3.3), 13 lags for each of the other explanatory variables and the dependent variable (treated as endogenous). Note that *SLRPRa* is the quadratic (log) real industrial production index that allows for non-linearities in the tourist expenditure.

A restricted 11 lag system, accepted by the joint *F*-test at the 1% level and in accordance with the HQ information criterion[76], is run. From Table 8.3, the diagnostic statistics show a good specification. The correlation of the actual and fitted values suggests that the equation explains almost 99.3% of the variance of the dependent variable. No problems appear in terms of diagnostic tests.

76

system	T	p		log-likelihood	SC	HQ	AIC
10	207	963	OLS	8662.4611	-58.887	-68.121	-74.695
11	207	1044	OLS	8807.2166	-58.198	***-68.210***	-75.094
12	207	1125	OLS	8914.8262	-57.151	-67.939	-76.134
13	207	1206	OLS	9041.7872	-56.291	-67.856	-76.360

System 12 --> System 11: F(81, 487) = 1.0499 [0.3713]
System 11 --> System 10: F(81, 545) = 1.6268 [0.0010] **

Table 8. 3 Statistical Tests of the Equation for the Real Tourism Expenditure (*LREXP*)

σ = 0.097250	RSS = 0.8606447747	
correlation of actual and fitted		
LREXP	0.99269	
LREXP	:Portmanteau 12 lags=	15.507
LREXP	:AR 1- 7 F(7, 84) =	1.747 [0.1091]
LREXP	:Normality Chi^2(2)=	0.125 [0.9393]
LREXP	:ARCH 7 F(7,77) =	1.207 [0.3091]

Hence, a further model with 11 lags is estimated for the *LREXP* equation; the final restricted parsimonious model is provided in Table 8.4.

Table 8. 4 Final Restricted Model for the Log Real Tourist Expenditure

EQ(2) Modelling LREXP by OLS (using spesa2.in7)
The present sample is: 1973 (1) to 1990 (5)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	9.3385	1.6945	5.511	0.0000	0.1547
LREXP_1	0.40047	0.053989	7.417	0.0000	0.2489
LREXP_5	0.28122	0.062791	4.479	0.0000	0.1078
LREXP_6	0.23182	0.065059	3.563	0.0005	0.0711
LREXP_7	-0.14220	0.062348	-2.281	0.0238	0.0304
LRSPfr	0.87387	0.36782	2.376	0.0187	0.0329
LRSPfr_2	-1.8858	0.40655	-4.638	0.0000	0.1147
LRSPfr_7	2.5869	0.41388	6.250	0.0000	0.1905
LRSPfr_9	-2.2403	0.64236	-3.488	0.0006	0.0683
LRSPfr_10	1.5368	0.54914	2.799	0.0057	0.0451
LRSPgr	-0.87381	0.15931	-5.485	0.0000	0.1534
LRSPgr_4	0.43685	0.16677	2.619	0.0096	0.0397
RLRSPpo	1.2086	0.27024	4.472	0.0000	0.1075
LRSPpo_5	2.8038	0.43033	6.516	0.0000	0.2037
LRSPpo_6	-1.1798	0.52378	-2.253	0.0256	0.0297
LRSPpo_7	-1.3036	0.37551	-3.471	0.0007	0.0677
RLRSPpo1	1.3108	0.38746	3.383	0.0009	0.0645
RLRSPsp	-0.77279	0.16994	-4.547	0.0000	0.1108
DLRPa_4	3.8478	1.0864	3.542	0.0005	0.0703
DLRPa_5	5.1379	1.1563	4.443	0.0000	0.1063
DLRPa_6	3.4381	1.1542	2.979	0.0033	0.0507
DLRPa_11	3.5258	1.0878	3.241	0.0014	0.0595
DLEXa	-1.3921	0.56667	-2.457	0.0151	0.0351
DLEXa_10	1.5171	0.51189	2.964	0.0035	0.0503
DLEXa_11	-1.2706	0.51145	-2.484	0.0140	0.0358
LRPRa_3	-0.56940	0.33881	-1.681	0.0947	0.0167
LRPRa_11	1.0250	0.38524	2.661	0.0086	0.0409
SLRPRa_3	1.7517	0.55768	3.141	0.0020	0.0561
SLRPRa_11	-2.8099	0.66886	-4.201	0.0000	0.0961
Easter	0.094853	0.037367	2.538	0.0121	0.0374
JAN	-0.20465	0.042243	-4.844	0.0000	0.1239
FEB	-0.082474	0.056715	-1.454	0.1478	0.0126
MAR	0.38007	0.060533	6.279	0.0000	0.1919
APR	0.54978	0.080613	6.820	0.0000	0.2189
MAY	0.91219	0.080619	11.315	0.0000	0.4354
JUN	0.99241	0.095081	10.437	0.0000	0.3962
JUL	1.3184	0.11317	11.649	0.0000	0.4498
AUG	0.92902	0.11832	7.852	0.0000	0.2708
SEP	0.62494	0.099795	6.262	0.0000	0.1911
OCT	0.41847	0.076043	5.503	0.0000	0.1543
NOV	0.017611	0.051495	0.342	0.7328	0.0007
i1976p5	-0.64869	0.10819	-5.996	0.0000	0.1780
i1990p1	0.48657	0.098351	4.947	0.0000	0.1285

R^2 = 0.976849 F(42,166) = 166.77 [0.0000] sigma = 0.091625
DW = 2.14 RSS = 1.393594366 for 43 variables and 209 observations

AR 1- 7 F(7,159) = 0.79369 [0.5936]
ARCH 7 F(7,152) = 0.86798 [0.5334]
Normality Chi^2(2)= 0.80387 [0.6690]
Xi^2 F(68, 97) = 0.91738 [0.6444]
RESET F(1,165) = 1.39350 [0.2395]

Note that restrictions on the lags of the non-linear coefficients (*LRPRa* and *SLRPRa*) are tested jointly by an *F*-test. The model does not present any particular problems in the residuals.

Few restrictions could be imposed, as one can see in Table 8.4: *RLRSPpo* is defined as the difference between the coefficients of the second and fourth lags of the substitute price for Portugal[77], *RLRSPpo1* is the difference between the coefficients of the ninth and tenth lags[78], *RLRSPsp* is the difference between the fifth and tenth lag coefficients[79]. No other restrictions have been accepted either by the joint *F*-test or by suggestion of the SC criterion.

Again, the final restricted model does not show any problems in the residuals and the model is overall well-specified.

The long run dynamics are reported in Table 8.5. The Wald test suggests that the long run coefficients are jointly significantly different from zero. An analysis of the short and long run dynamics is appropriate (see Tables 8.4 and 8.5). One might want to compare the long run dynamics with the short run dynamics as given by the previous table. The real substitute price for France (*LRSPfr*) shows a positive and statistically significant (*t*-statistic +2.04) long run coefficient. In Table 8.4, the short run coefficient ($LRSPfr_{t-2}$) presents a negative sign and a *t*-value equal to -4.64. The long run coefficient for *LRSPpo* is positive and statistically significant (*t*-statistic equals +2.22). The same positive sign appears in the short run, with a highly statistically significant coefficient (*t*-statistic equals +6.52 for the first significant lag, that is the fifth). As far as the real substitute price for Greece (*LRSPgr*) is concerned, in the long run, the coefficient shows the expected negative sign with a *t*-value equal to -3.22; whereas, in the short run the coefficient of the fourth lag is positive. The real substitute price for Spain presents a short run dynamic structure with the coefficient for the oscillation (*RLRSPsp*) presenting a negative sign. In conclusion, one cannot observe a simple pattern across the competing tourist destinations, which exhibit noticeable differences in both the short and long run.

---

[77] The restriction on the coefficient of the second and fourth lags, that present an opposite sign and similar magnitude, is accepted at the 5% level from the joint *F*-test (1,163) as the calculated value 0.000070 is smaller than the critical value 3.84; moreover, the SC criterion is minimised when the restriction is imposed: from -3.84058 to -3.86614.

[78] The restriction on the coefficient of the ninth and tenth lags is accepted at the 5% level from the joint *F*-test (1,164) as the calculated value 0.97 is smaller than the critical value 3.84; moreover, the SC criterion is minimised when such a restriction is imposed: from -3.86614 to -3.88583.

[79] The restriction on the coefficient of the fifth and tenth lags for the substitute price for Spain, is accepted at the 5% level. The joint *F*-test (1,165) value is 0.013 is smaller than the critical value 3.84; moreover, the SC criterion is minimised when such a restriction is imposed: from -3.88583 to -3.91131.

The relative price growth (*DLRPa*) shows a positive sign with a *t*-statistic equal to 3.11. Whereas, the first difference of the exchange rate (*DLEXa*) presents a negative long run coefficient that is not statistically significant (*t*-value -1.33). Note also that the cointegrating vector is not found to be statistically significant.

Table 8. 5 Solved Static Long Run Equation for *LREXP*

LREXP =	+40.83	+3.811 LRSPfr	-1.911 LRSPgr
(SE)	(8.615)	(1.871)	(0.5926)
	+1.401 LRSPpo	+69.74 DLRPa	-5.009 DLEXa
	(0.6305)	(22.45)	(3.764)
	+1.992 LRPRa	-4.627 SLRPRa	+0.4147 Easter
	(1.356)	(2.213)	(0.1918)
	-0.8948 JAN	-0.3606 FEB	+1.662 MAR
	(0.3081)	(0.2716)	(0.5182)
	+2.404 APR	+3.989 MAY	+4.339 JUN
	(0.786)	(1.184)	(1.339)
	+5.765 JUL	+4.062 AUG	+2.733 SEP
	(1.726)	(1.308)	(0.9399)
	+1.83 OCT	+0.077 NOV	-2.836 i1976p5
	(0.6499)	(0.2305)	(0.7632)
	+2.128 i1990p1	+5.284 RLRSPpo	+5.731 RLRSPpo1
	(0.6836)	(1.483)	(2.153)
	-3.379 RLRSPsp		
	(1.204)		

ECM = LREXP - 40.8329 - 3.81086*LRSPfr + 1.91062*LRSPgr - 1.40108*LRSPpo
- 69.7401*DLRPa + 5.00911*DLEXa - 1.99211*LRPRa + 4.62699*SLRPRa
- 0.414748*Easter + 0.894831*JAN + 0.36062*FEB - 1.66189*MAR - 2.40395*APR
- 3.98858*MAY - 4.33935*JUN - 5.76467*JUL - 4.06217*AUG - 2.73256*SEP
- 1.82978*OCT - 0.077003*NOV + 2.83643*i1976p5 - 2.12755*i1990p1
- 5.28449*RLRSPpo - 5.73148*RLRSPpo1 + 3.37904*RLRSPsp;

WALD test Chi^2(24) = 75.327 [0.0000] **

8.3.5 Estimating A Non-linear Model For The Real Tourist Receipts

As already mentioned, the restricted model in Table 8.4 can be called a non-linear model since it includes the square of the real industrial production. This quadratic term hypothesizes curvature in the graph of the response model relating the dependent variable (the total tourist expenditure) to the explanatory variable (the aggregated real industrial production). Hence, the points of maximum and/or minimum of each lag pair of the real industrial production index (*LRPRa* and *SLRPRa*) can be derived. For the third lag, given the generic equation:

$$y = a\,x^2 + b\,x + e$$

taking the partial derivative with respect to x (in this case *LRPRa*), and equating to zero, one finds that $x=0.16$; hence, as $a>0$ there is a minimum; however, x is greater than the smallest observation for *LRPRa* (*i.e.*-0.093202). For the eleventh lag, x is equal to 0.18; hence, as $a<0$ there is a maximum; in this case x is smaller than the largest observation for *LRPRa* (*i.e.* 0.590681). Given that a quadratic specification is used as a local approximation to a sigmoid curve, *e.g.* a logistic function, then the minimum should lie either below all the observations, or above all the observations if the sigmoid is reversed, going from high to low. So far, the values obtained suggest non-linearity, but not in the form that will accommodate a sigmoid transformation of a I(1) variable in an I(0) equation. However, these turning points have not been precisely estimated, and further analysis is needed.

Given the difficulties in using an I(1) variable to explain an I(0) variable, the use of a non-linear expression for the most significant lags of the (log) industrial production index will be investigated. The generic function for the real income proxy can be written as follows:

$$\theta\ (x, \mu, \sigma, \alpha) = \alpha\left(1 + e^{-(x-\mu)\ /\ \sigma}\right)^{-1}$$

where:

x = is the most significant lag for *LRPRa*, that is the eleventh and third lag, respectively;

μ = is the centre of the curvature;

σ = is the spread of the curvature;

α = is the impact parameter.

The mean and the standard error of *LRPRa* are used as starting values for the parameters μ (*mmu* which equals 0.12) and σ (*msig* which equals 0.21). The aim is to find the smallest RSS that corresponds to the maximum of the likelihood. In running the non-linear specification, the TSP package is used. The program is reported in Appendix G (Table G.3). The results obtained are provided in Table 8.6.

Table 8. 6 Non-linear Estimation for (log) Real Tourist Expenditure by TSP

Dependent variable: LREXP
Mean of dependent variable = 23.2305 Std. error of regression = .093521
Std. dev. of dependent var. = .537965 R-squared = .975883
Sum of squared residuals = 1.45187 Adjusted R-squared = .969781
Variance of residuals = .874620E-02 Durbin-Watson statistic = 2.13199

Log of Likelihood Function = 222.753
Number of Observations = 209

Parameter	Estimate	Standard Error	t-statistic
constant	8.70871	1.70631	5.10383
LREXP(-1)	.440847	.054119	8.14592
LREXP(-5)	.275400	.063671	4.32533
LREXP(-6)	.200937	.067075	2.99569
LREXP(-7)	-.184227	.063941	-2.88121
LRSPfr	.796134	.362670	2.19520
LRSPfr(-2)	-1.96245	.423037	-4.63894
LRSPfr(-7)	2.55358	.425734	5.99806
LRSPfr(-9)	-2.29417	.667382	-3.43757
LRSPfr(-10)	1.47403	.569316	2.58913
LRSPgr	- .819077	.163356	-5.01407
LRSPgr(-4)	.356904	.179291	1.99064
RLSPpo	1.15511	.276183	4.18241
LRSPpo(-5)	2.82234	.439880	6.41615
LRSPpo(-6)	-1.20271	.539467	-2.22945
LRSPpo(-7)	-1.22724	.388804	-3.15644
RLSPpo1	1.36442	.410582	3.32314
RLSPsp	- .694331	.172594	-4.02291
DLRPa(-4)	3.58945	1.16461	3.08210
DLRPa(-5)	4.61411	1.23251	3.74367
DLRPa(-6)	3.34473	1.21211	2.75942
DLRPa(-11)	3.43636	1.11207	3.09006
DLEXa	-1.49456	.583111	-2.56308
DLEXa(-10)	1.33121	.532032	2.50212
DLEXa(-11)	-1.35985	.530082	-2.56536
BETA*	- .151937	.054010	-2.81311
MMU	.323501	.011059	29.2536
MSIG	.646002E-02	.011653	.554344
ETA*	- .011287	.055862	- .202054
easter	.095587	.038413	2.48841
i1976p5	- .656908	.111077	-5.91399
i1990p1	.508124	.099663	5.09843
jan	- .183298	.043358	-4.22759
feb	- .048799	.058312	- .836857
mar	.408593	.062045	6.58539
apr	.539261	.082470	6.53889
may	.864576	.081137	10.6558
jun	.921625	.094336	9.76960
jul	1.22945	.112330	10.9450
aug	.822007	.117115	7.01882
sep	.533189	.098822	5.39546
oct	.353449	.075703	4.66887
nov	- .018303	.051721	- .353886

Note: * *beta* and *eta* are the coefficients for the logistic transformation of the income proxy.

The TSP package has encountered problems of maximisation; hence, GAUSS 3.2.32 has been used. This package uses double arithmetic precision to calculate the RSS from the regression. The analysis has proceeded by estimating a model where the first most significant lag is included, that is *LRPRa(-11)*. In this case, the RSS from

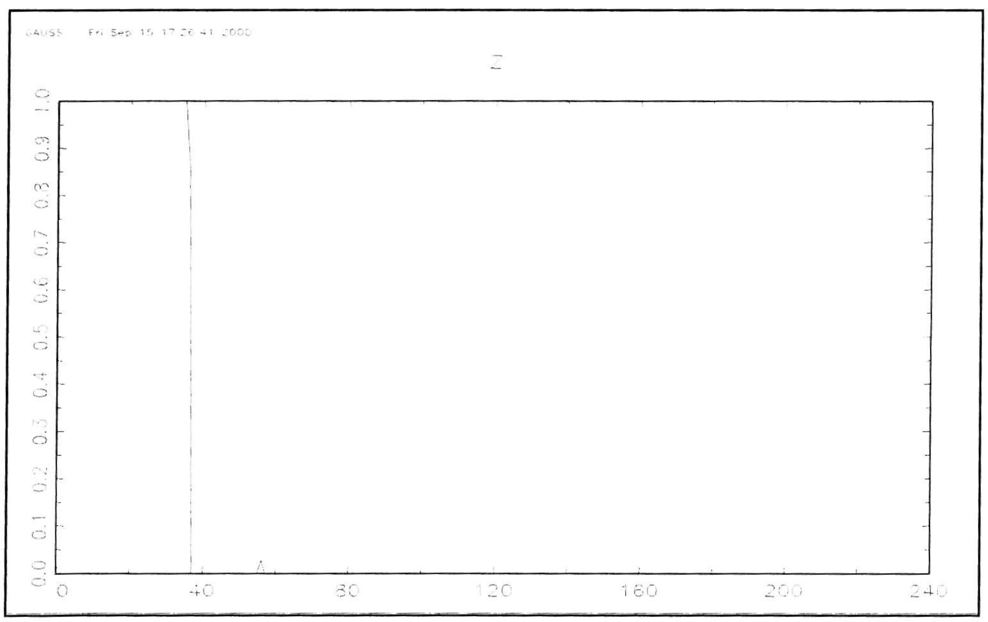
GAUSS (1.553668) is close to that for TSP (1.55372), at the starting points *mmu*=1.2 and *msig*=2.1. Fitting by non-linear least squares, Table 8.7 is obtained:

Table 8. 7 Results from the Non-linear Least Squares when including *LRPRa(-11)*

	RSS	MMU	MSIG
GAUSS a)	1.44297	0.316987	0.000148393
GAUSS b)	1.442524	0.318467	0.000580673
TSP	1.45277	0.324540	0.00556071

Gauss appears to find more than one minimum for the RSS; the GAUSS b) case is found with three different starting points for *mmu*. Arguably, one should stop at this point. The values of *msig* are so small that the logistic function has become very steep, that is the logistic transformation of *LRPRa* (say z_i[80]) is being turned into a shift dummy, and what is being represented is a change in the intercept, not an income effect. This can be seen in Figure 8.3.

Figure 8. 3 Logistic Transformation of *LRPRa*: plot of z_i



Next, inserting *LRPRa(-3)*, the RSS from GAUSS (1.5536275) is close to that for TSP (1.55368), at the starting points *mmu*=1.2 and *msig*=2.1. Fitting by non-linear least squares, Table 8.8 is obtained:

80 The generic equation for the logistic transformation of the variable x_i is the following: z_i = 1/1+EXP(-1*((x_i -μ)/σ)).

Table 8. 8 Results from the Non-linear Least Squares when including *LRPRa(-11)* and *LRPRa(-3)*

	RSS	MMU	MSIG
TSP	1.45187	0.323501	0.00646002
GAUSS a)	1.453473	0.322368	-1.469054
GAUSS b)	1.44282194	0.317041	0.00000425801
GAUSS c)	1.445763	0.309853	-0.000208184
GAUSS d)	1.43241964	0.6012872193	-0.0233751168

Cases a-c-d show negative values for *msig* that reverse the shape of the logistic transformation, which can be compensated for by reversing the sign of the coefficient on the logistic term. However, given the instability of parts of the calculation when *msig* is close to zero, it seems sensible to reparametrise, and optimise in terms of *msig** where *msig=exp(msig*)*. *Msig* will always be positive, with values close to zero corresponding to negative values of *msig**. The iterative process is better behaved, but gives three distinct minima, the two lowest being shown in Table 8.9.

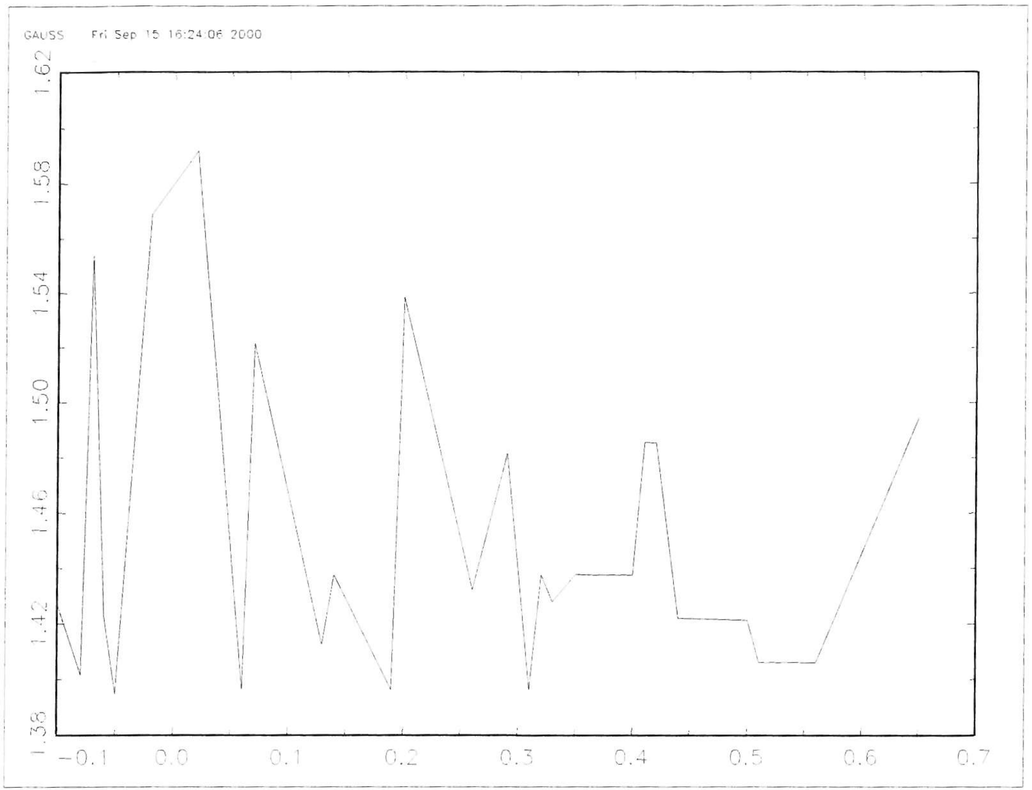
Table 8. 9 Reparametrisation for *msig**

	RSS	MMU	MSIG*
GAUSS e)	1.442372422	0.319385865	-9.597570598
GAUSS f)	1.43241964	0.6012872193	-2.38230481

As *exp*(-2.38230481)=0.09233751165, cases d) and f) are not identical. Thus, multiple minima appear again.

Finally, freeing the parameters of the two logit transformations, as *mmu1* and *msig1* for *LRPRa(-3)*, *mmu2* and *msig2* for *LRPRa(-11)*, the number of maxima seems to multiply even further. It seems sensible for *mmu* not to lie much outside the range of *LRPRa*, that is [-0.093202, 0.590681]. Accordingly, optimisations are started with *mmu1=mmu2* taking initial values over [-0.1, 0.7]. Deleting cases where convergence is not achieved within 300 iterations, one can graph the RSS against the starting point for 33 remaining cases, to obtain Figure 8.4.

Figure 8. 4 Multiple Minima for 33 cases: plot of RSS against the Starting Points [-0.1, 0.7]



There are a large number of distinct minima here, with the smallest being detailed in Table 8.10. Cases with *mmu1* and *mmu2* outside the range [-0.1, 0.7] have been eliminated.

Table 8. 10 Smallest Minima RSS with *mmu1* and *mmu2* inside the range [-0.1, 0.7]

RSS	MMU1	MSIG1*	MMU2	MSIG2*
1.394742183	0.02370718788	-7.194001639	0.319104224	-6.848891416
1.396104495	0.5635791969	-6.652379121	0.3184204	-8.276513355
1.396193344	0.5638461545	-9.861068655	0.319105256	-8.283935766
1.396193351	0.5638487143	-10.03780386	0.319093379	-8.254668349

It is clear that there may be considerable numerical and statistical objections to over-interpreting the results. It can be noted that the additional term in *LRPRa(-3)* reduces the RSS from 1.442524 (Table 8.7, case b) to 1.394742183, so an *F*-test gives ((1.442524-1.394742183)/3)/(1.394742183/(209-41))=1.918, which is smaller than the correspondent critical value at the 5% level (2.60). The $\chi^2(3)$ version of the test, 7.04, has a *p*-value of 7.1%. Thus, there is no evidence for the inclusion of a logistic term in *LRPRa(-3)* when testing at the 5% level.

Next, comparing a regression without logistic terms with the case just including *LRPRa(-11)*, one has a RSS equal to 1.6089936, the calculated $F(3,169)$=9.38, and the calculated $\chi^2(3)$=22.8, which are both significant at the 1% level. Thus, it seems that there is variation in the dependent variable that can be explained, even if the logistic form used seems not to give satisfactory modelling results. These experiments at non-linear modelling thus fail, in part, from numerical problems: there are multiple maxima in the likelihood. However, the inclusion of the *LRPRa(-11)*, results in a more important difficulty emerging: a sigmoid transfomation of a trending variable may be indistinguishable from a shift dummy, and the interpretation of its coefficient will either change or lose its meaning.

8.3.6 Estimating Italian Tourist Expenditure As I(1)

In this section, one will estimate a model considering the real tourist expenditure (*LREXP*) as non-stationary in the level, but stationary in the first difference. Some evidence for this variable to be a random walk has been found from Franses' unit roots test (see Table 8.1). Moreover, in the majority of the empirical studies, tourist expenditure is treated as integrated of order one (*e.g.* Lanza and Urga, 1995; Song *et al.*, 2000).

The aim is to test for the existence of a possible cointegrating vector with the other I(1) variables: real industrial production (*LRPRa*), relative price (*LRPa*) and exchange rate (*LEXa*). A Johansen analysis is run; a 13 lag system is estimated for the sample period 1972(1)-1990(5). The system also includes a constant, a trend and seasonals treated unrestrictedly, as it gives the best results in terms of diagnostic statistics. Non-normality problems cannot be avoided in each of the equations by including impulse dummies. The joint *F*-test indicates that the null hypothesis of a 12 lag system cannot be accepted at the 1% level. On the other hand, the information criteria suggest for a further parameter reduction; however, the diagnostic statistics have shown problems of heteroscedasticity and serial correlation. Hence, a 13 lag system has been estimated. From Table 8.11, the existence of one cointegrating vector is inferred.

Table 8. 11 Johansen Tests for the Number of Cointegrating Vectors

Ho	H_1	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
r=0	r=1	41.25**	30.93*	30.3	69.74**	52.30	54.6
r=1	r=2	17.51	13.13	23.8	28.49	21.37	34.6
r=2	r=3	9.09	6.82	16.9	10.99	8.24	18.2
r=3	r=4	1.90	1.42	3.7	1.90	1.42	3.7

Notes: (1) Adjusted by the degrees of freedom (see, Reimers, 1992).
(2) Critical values at a 5% level of confidence (see Osterward-Lenum, 1992).
and ** denotes rejection of the null (*i.e.* non-cointegration) at a 5% and 1% level, respectively.

Table 8. 12 The Eigenvalues $\hat{\lambda}$, Eigenvectors $\hat{\beta}$, and the Weights α

Eigenvalues $\hat{\lambda}$
(0.179874 0.0807209 0.0427554 0.00907855)

Standardized β' eigenvectors					Standardized $\hat{\alpha}$ coefficients				
LREXP	***LRPa***	***LEXa***	***LRPRa***		***LREXP***	-0.033	-0.592	-0.233	-0.0052
1.000	4.0310	-7.695	-5.604		***LRPa***	0.003	-0.042	0.008	-0.0004
0.026	1.000	-0.559	0.282		***LEXa***	0.014	0.145	-0.004	-0.0007
0.424	-2.384	1.000	-0.419		***LRPRa***	0.021	-0.064	-0.006	0.0039
-2.939	7.433	10.912	1.000						

The Johansen test shows that a cointegration equilibrium relationship exists between the real tourist expenditure, relative price, exchange rate and real income proxy. From Table 8.12, the cointegrating vector is the following:

$$CE = LREXP + 4.0310\, LRPa - 7.6952\, LEXa - 5.6041\, LRPRa \qquad (8.3.6.1)$$

According to the long run cointegrating vector, an increase in the exchange rate and real industrial production determines an increase in the foreign tourist receipts. Whereas, an increase in the relative price causes a decrease in the tourist demand as economic theory suggests.

Hence, a possible solution to the non-linearity problem in terms of the income proxy, for the *LREXP* equation, can be achieved by treating the dependent variable as I(1). The demand function for Italian tourism becomes:

$$DLREXP = f\,(DLRPRa,\ DLRPa,\ DLEXa,\ CE,\ LRSPfr,\ LRSPgr,\ LRSPpo,\ LRSPsp,\ E,\ T,\ D).$$

Thus, a model that includes current and lagged variables in the first difference is estimated. An initial unrestricted 13 lag system is run in order to identify the lag size. The equation for *DLREXP* includes a constant, 11 seasonal dummies, the "Easter" dummy variable, three impulse dummies created in order to avoid problems of non-normality in the residuals (*i.e. i1976p5*, *i1987p4* and *i1990p1*), a time trend (all these variables are treated unrestrictedly), CE_{t-1}, plus 13 lags for each of the other explanatory variables and the dependent variable (treated as endogeneous).

According to the joint F-test, a restricted system with 11 lags is accepted at the 5% level. Whereas, a 1 lag system, suggested by HQ and SC criteria, presents problems in the diagnostic statistics; on the other hand, a 12 lag system is suggested by the AIC criterion[81]. Hence, following the AIC criterion a 12 lag system is run. From Table 8.13, no problems appear in terms of diagnostic tests. The correlation of the actual and fitted values suggests that the equation explains 97.6% of the variance of the dependent variable.

Table 8. 13 Statistical Tests for *DLREXP* Equation

σ = 0.092256	RSS = 0.7915380845	
correlation of actual and fitted		
DLREXP	0.97607	
DLREXP	:Portmanteau 12 lags=	12.1180
DLREXP	:AR 1- 7 F(7, 86) =	0.6122 [0.7444]
DLREXP	:Normality Chi^2(2)=	1.8019 [0.4062]
DLREXP	:ARCH 7 F(7,79) =	0.2098 [0.9823]

The next step is to estimate a model with 12 lags for the *DLREXP* equation. The final restricted parsimonious model is shown in Table 8.14.

[81] The results for establish the lag size are the following:

system	T	p		log-likelihood	SC	HQ	AIC
1	207	208	OLS	6756.5017	**-59.922**	**-61.916**	-63.280
2	207	272	OLS	6808.7796	-58.778	-61.386	-63.785
3	207	336	OLS	6861.6150	-57.640	-60.862	-63.296
4	207	400	OLS	6927.1528	-56.624	-60.460	-63.929
5	207	464	OLS	6992.6103	-55.608	-60.057	-63.561
6	207	528	OLS	7061.3761	-54.624	-59.687	-63.226
7	207	592	OLS	7147.4385	-53.806	-59.483	-64.057
8	207	656	OLS	7211.1045	-52.773	-59.063	-63.673
9	207	720	OLS	7286.3875	-51.851	-58.756	-64.400
10	207	784	OLS	7357.6235	-50.891	-58.409	-64.088
11	207	848	OLS	7449.2635	-50.127	-58.259	-63.974
12	207	912	OLS	7539.9311	-49.355	-58.100	**-64.850**
13	207	976	OLS	7609.2319	-48.376	-57.735	-64.519

System 13 --> System 12: F(64, 456) = 0.87778 [0.7359]
System 12 --> System 11: F(64, 456) = 0.87778 [0.7359]
System 11 --> System 10: F(64, 548) = 1.4224 [0.0216] *

Table 8. 14 Final Short Run Model for *DLREXP*

EQ(2) Modelling DLREXP by OLS (using spesa4.in7)
The present sample is: 1973 (2) to 1990 (5)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-2.2647	0.83609	-2.709	0.0074	0.0411
DLREXP_1	-0.55601	0.053680	-10.358	0.0000	0.3855
DLREXP_2	-0.45400	0.061804	-7.346	0.0000	0.2399
DLREXP_3	-0.30625	0.053250	-5.751	0.0000	0.1621
DLREXP_4	-0.21975	0.049419	-4.447	0.0000	0.1036
DLREXP_10	-0.13092	0.044251	-2.958	0.0035	0.0487
DLREXP_12	0.23502	0.049190	4.778	0.0000	0.1178
DLRPa_4	3.0686	1.0407	2.949	0.0036	0.0484
DLRPa_5	3.9223	1.0133	3.871	0.0002	0.0806
DLRPa_6	3.3434	1.0641	3.142	0.0020	0.0546
DLRPa_11	2.0083	0.97640	2.057	0.0412	0.0241
DLEXa_10	2.2065	0.55229	3.995	0.0001	0.0854
DLEXa_11	-1.1881	0.47824	-2.484	0.0139	0.0348
LRSPfr_2	-1.1155	0.20778	-5.369	0.0000	0.1443
LRSPfr_7	1.7420	0.35443	4.915	0.0000	0.1238
LRSPfr_9	-1.0371	0.32712	-3.170	0.0018	0.0555
LRRSPfr	1.4493	0.55485	2.612	0.0098	0.0384
LRSPgr	-0.44830	0.13929	-3.218	0.0015	0.0571
LRSPgr_4	0.53687	0.14240	3.770	0.0002	0.0767
RLRSPgr	-0.57414	0.23844	-2.408	0.0171	0.0328
LRSPpo_2	0.77190	0.21735	3.551	0.0005	0.0687
LRRSPpo	-1.0802	0.21321	-5.067	0.0000	0.1305
LRSPsp_10	0.48570	0.13207	3.678	0.0003	0.0733
Trend	0.00097	0.00025	3.863	0.0002	0.0803
JAN	0.04182	0.045639	0.916	0.3608	0.0049
FEB	0.13573	0.040444	3.356	0.0010	0.0618
MAR	0.51489	0.048654	10.583	0.0000	0.3957
APR	0.72170	0.061218	11.789	0.0000	0.4484
MJ	0.93276	0.076903	12.129	0.0000	0.4625
JUL	1.0719	0.087624	12.232	0.0000	0.4667
AUG	0.73404	0.085864	8.549	0.0000	0.2994
SEP	0.44114	0.074975	5.884	0.0000	0.1684
OCT	0.28858	0.061020	4.729	0.0000	0.1157
NOV	-0.026224	0.048204	-0.544	0.5871	0.0017
i1976p5	-0.54056	0.097904	-5.521	0.0000	0.1513
i1990p1	0.53032	0.096638	5.488	0.0000	0.1497
i1987p4	0.23810	0.093376	2.550	0.0117	0.0366

R^2 = 0.920477 F(36,171) = 54.981 [0.0000] sigma = 0.0883626
DW = 2.21 RSS = 1.335160705 for 37 variables and 208 observations

AR 1- 7 F(7,164) =	1.7524	[0.1003]
ARCH 7 F(7,157) =	0.1767	[0.9897]
Normality Chi^2(2)=	3.2859	[0.1934]
Xi^2 F(59,111) =	0.7409	[0.8974]
RESET F(1,170) =	0.5598	[0.4554]

Few coefficient restrictions can be imposed applying the joint *F*-test as follows. *MJ* as the first restriction that involves the coefficients for the seasonal dummies May and June as they have a similar magnitude[82]. The second restriction *RLRSPgr*[83] is given by the difference between the sixth and seventh lag coefficients of the real substitute price for Greece. The further restriction is given by the difference between the coefficients for the fourth and fifth lag of the (log) real substitute price for Portugal (*LRRSPpo*)[84]. The final restriction involves the coefficients for the tenth and eleventh lag of the (log) real substitute price for France (*LRRSPfr*)[85].

The final model passes all diagnostic statistics and it is data admissible to its information set. From Tables 8.14 and 8.15, the growth in Italian tourist expenditure is positively related to the growth of the relative price with an elasticity equal to +3.07. The growth in the relative price shows a long run positive sign and it has a highly statistical significance (*t*-value +5.49). This finding suggests that foreign consumers do not respond quickly to short run changes in the Italian price with respect to home prices. As a reminder, from the Johansen long run analysis, it emerges that the negativity condition for the relative price does hold. The growth of the dependent variable is positively influenced by the growth in the exchange rate, both in the short run and in the long run. This finding confirms the Johansen long run elasticity (see 8.3.6.1). On the other hand, a permanent shift in the real substitute price for France (*LRSPfr*), which is stationary in the level, causes a permanent decrease in the growth of Italian expenditure in tourist goods and services. Moreover, a permanent shift in the real substitute price for Greece, Portugal and Spain causes a permanent increase in the growth of Italian expenditure. These findings seem to indicate that the competitor countries included in this study might not be the appropriate ones, with the only exception being France. The time trend plays a role in explaining the dependent

---

82 The restriction on the coefficient of the May and June dummy is accepted at the 5% level from the joint *F*-test (1,166) as the calculated value 0.15 is smaller than the critical value 3.84; moreover, the SC criterion is minimised when such a restriction is imposed: from -3.98751 to -4.01229.

83 The calculated *F*-test (1,167) equals 0.60 that is smaller than the critical value 3.84. Hence, the restriction on the coefficient of the sixth and seventh lag of the substitute price for Greece is accepted at the 5% level. Moreover, the SC criterion is minimised when such a restriction is imposed: from -4.01229 to -4.03435.

84 The calculated *F* (1,168) equals 0.77 and it is smaller than the critical value (3.84) at the 5% level. The SC criterion also suggests for this parameter reduction, it decreases from -4.03435 to -4.05541.

85 The calculated *F* (1,169) equals 0.33 and it is smaller than the critical value (3.84) at the 5% level. The SC criterion suggests for this parameter reduction, decreasing from -4.05541 to -4.07859.

variable, however, its impact is relative low. Finally, note that the growth of the real income proxy (*LRPRa*), the cointegrating vector (*CE*) and the "Easter" dummy do not appear in the final specification. Dropping the cointegrating vector means that long run dynamics disappear.

Table 8. 15 Static Long Run Equation for *DLREXP*

DLREXP =	-0.9313	+5.075 DLRPa	+0.4188 DLEXa
(SE)	(0.3481)	(0.9173)	(0.2487)
	-0.1688 LRSPfr	+0.03642 LRSPgr	+0.3174 LRSPpo
	(0.0822)	(0.05155)	(0.09088)
	+0.1997 LRSPsp	+0.0004007 Trend	+0.0172 JAN
	(0.05636)	(0.0001057)	(0.01866)
	+0.05581 FEB	+0.2117 MAR	+0.2968 APR
	(0.01605)	(0.01796)	(0.01782)
	+0.4407 JUL	+0.3018 AUG	+0.1814 SEP
	(0.01725)	(0.02256)	(0.02284)
	+0.1187 OCT	-0.01078 NOV	-0.2223 i1976p5
	(0.02)	(0.02004)	(0.04416)
	+0.2181 i1990p1	+0.0979 i1987p4	+0.3836 MJ
	(0.04355)	(0.03904)	(0.01639)
	-0.2361 RLRSPgr	-0.4442 LRRSPpo	+0.5959 LRRSPfr
	(0.1002)	(0.08952)	(0.2299)

ECM = DLREXP + 0.931252 - 5.07527*DLRPa - 0.418782*DLEXa + 0.168836*LRSPfr
- 0.036417*LRSPgr - 0.317404*LRSPpo - 0.19972*LRSPsp - 0.000400659*Trend
- 0.0171968*JAN - 0.0558117*FEB - 0.21172*MAR - 0.29676*APR - 0.440745*JUL
- 0.301837*AUG - 0.181395*SEP - 0.118663*OCT + 0.0107832*NOV + 0.222279*i1976p5
- 0.218066*i1990p1 - 0.097905*i1987p4 - 0.38355*MJ + 0.236084*RLRSPgr
+ 0.44418*LRRSPpo - 0.59593*LRRSPfr;
WALD test Chi^2(23) = 1341.4 [0.0000] **

8.4 BUDGET SHARE AS THE DEPENDENT VARIABLE

So far, real tourist receipts have been used as a proxy for the international demand for tourism to Italy. The current analysis involves an in-depth investigation of whether the real budget share (*LBSm*), defined in terms of a weighted average for the main origin countries, can be considered as a better proxy for the international demand for tourism.

8.4.1 Appropriate Aggregation

The aim is to investigate whether an aggregated definition is appropriate. One might consider individual budget shares for each of the source countries, that is France, Germany, Japan, Sweden, Switzerland, UK and USA, as follows:

$$BS_{i,t} = \lambda_{i,t}\, TEX_t \,/\, EX_{i,t}$$

where:

$$\lambda_{i,t} = \frac{NSi,t}{\sum_{i=1}^{i=7} NSi,t} = w_{i,t} \quad (8.4.1.1)$$

$\lambda_{i,t}$ is the weight defined in terms of the number of nights spent (*NS*) by the tourists of each origin country *i* in all registered accommodation in month *t* (Source: ISTAT). The budget share for each of the origin countries has been defined as the ratio between the total tourist expenditure in Italy (TEX_t) and the total private consumption ($EX_{i,t}$) in country *i*, both of them expressed in billions of lire. In order to be able to aggregate all the components one would expect common trends for each of the individual countries. A graphical representation is given below. Note that the (log) budget share for Sweden has been repeated in Figures 8.5 and 8.6.

Figure 8. 5 (Log) Budget Shares for: France, Germany, Japan and Sweden (1972:1-1990:5)

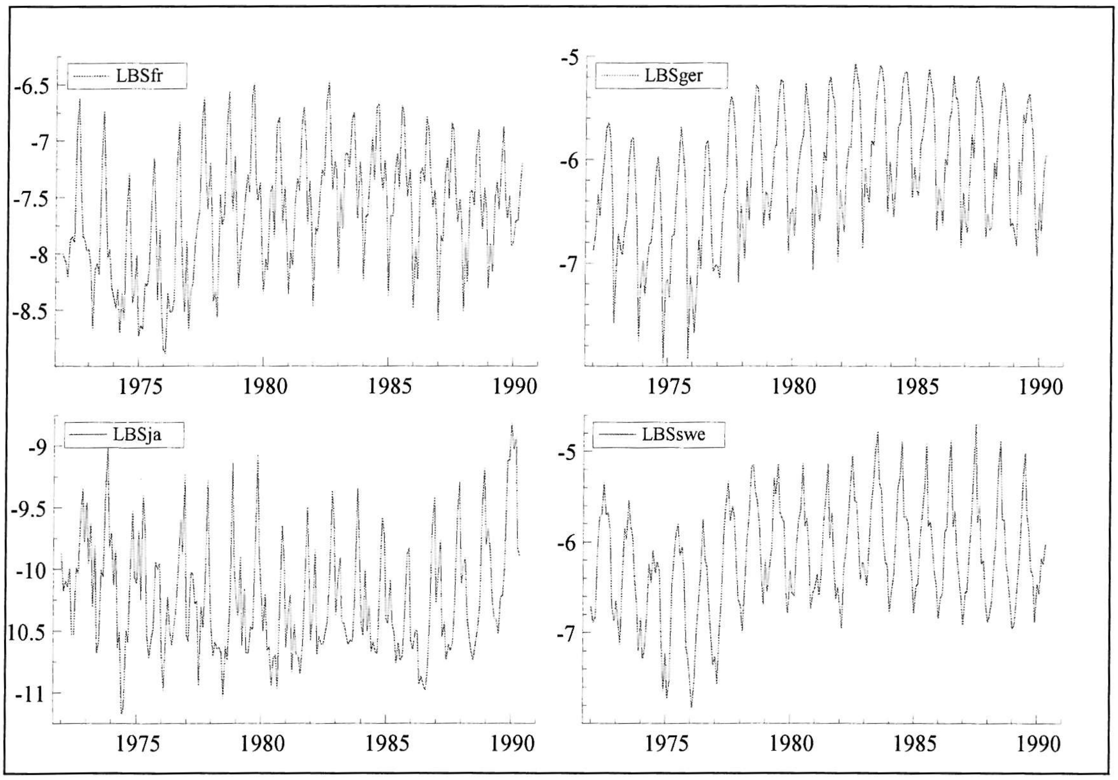
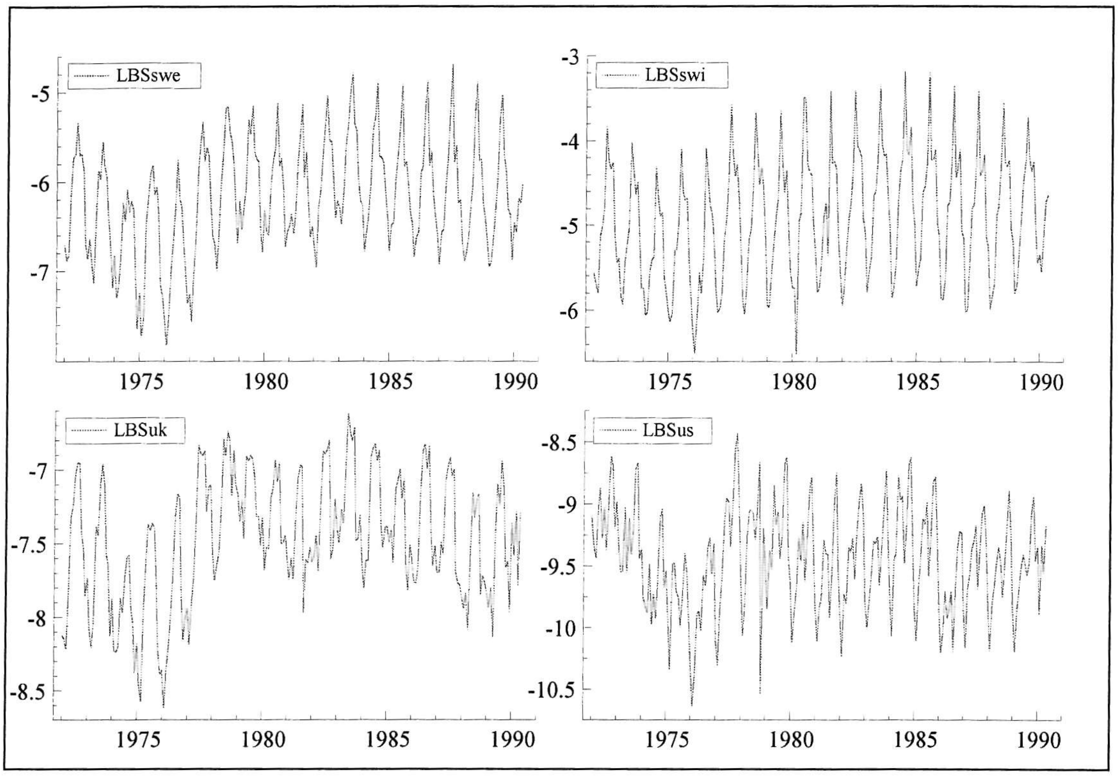


Figure 8. 6 (Log) Budget Shares for: Sweden, Switzerland, UK and USA (1972:1-1990:5)



Examining the graphs the main difference seems to be around 1975-1976. The other salient feature is the large seasonality (see Section 7.2.2, Chapter 7). In order to decide whether a common trend exists amongst the tourist origin countries, a regression is carried out for the budget share of each of the individual countries by including a constant, a time trend and 11 seasonal dummies. The aim is to consider the magnitude of the trend coefficient and to test the statistical significance of this coefficient. The complete results are given in Table H.1 (Appendix H). Similarities in terms of magnitude, sign and statistical significance of the trend coefficient exist amongst France, Germany, Sweden, Switzerland and the UK. On the other hand, Japan and the United States show significant differences from the other countries. This finding encourages an aggregation for only five countries. However, a seven countries aggregation is also included and similar results are expected.

There are still some doubts in terms of applying consumer theory. One of the main assumptions is that the series reflect constant preferences. The problem is that one is considering time-series (and not panel data) and it is possible that the consumers' preferences change over the time. In general, Deaton and Muellbauer (1980) consider cross-section studies. Problems of heteroscedasticity might also emerge in running the model for *LBSm*. It is also possible that 20 years ago people with average income did not want to spend their wages on foreign holidays; whereas, today there is more willingness to spend money on leisure abroad.

As already mentioned, budget shares are usually used with panel data and system of equations. There is explicit economic reasoning from micro foundations. A consumer allocates his/her income over a whole range of goods and services. Several stages can be identified within the budget allocation. The first stage consists of allocating income over a set of groups and services such as food, housing, holidays and so on. In the case of leisure, the second stage consists of allocating expenditure between destination subgroups (domestic tourism, holidays in Europe, in South America, in Australia, etc.). A further stage involves allocating expenditure within destinations of the same subgroup. Hence, on would model the budget share of Italy versus its close competitors.

A stronger assumption has been adopted in this chapter. The consumers, from the origin countries under study, are allocating their income over holidays in Italy and a whole range of goods and services. The main limitation of this assumption is that one does not underpin the modelling of the budget share with the main assumptions based on microeconomic theory of consumer demand. Moreover, the values of the dependent variable are very low. However, the procedure adopted in this chapter has several advantages. The first is that it allows a straightforward comparison between the results from the budget share and tourist receipts, as given earlier in the chapter. Furthermore, one of the main purposes of Chapter 8 is that of using time series at a monthly frequency. As already mentioned, there are grounds for believing that monthly series are the appropriate frequency for analysing and modelling the demand for tourism. Hence, one of the main problems is the lack of data at monthly frequency in the official statistics for the main competitors of Italy. Future work needs to be undertaken to expand the monthly database for other competitors.

8.4.2 Appropriate Weights

In defining the variables of interest, an investigation has been carried out on the appropriate weight to use. Some experiments have been done in understanding whether the (annual) weights calculated in terms of foreign nights of stay in registered accommodation ($w_{i,t}$) could approximate the values of the (annual) weights calculated by taking into consideration tourist expenditure expressed in terms of the currency of each source country ($x_{i,t}$). Note that in the comparison the nights spent by tourists are used instead of the number of tourists' arrivals. The former figures are more likely to be consistent with the fact that the average expenditure for a tourist (regardless of nationality) depends on the average length of stay in the destination country. The definitions for the two weights are:

$$wi,t = \frac{NSi,t}{\sum_{i=1}^{i=5} NSi,t} \approx xi,t = \frac{TEXi,t}{\sum_{i=1}^{i=5} TEXi,t}$$

The results are reported in Table 8.16.

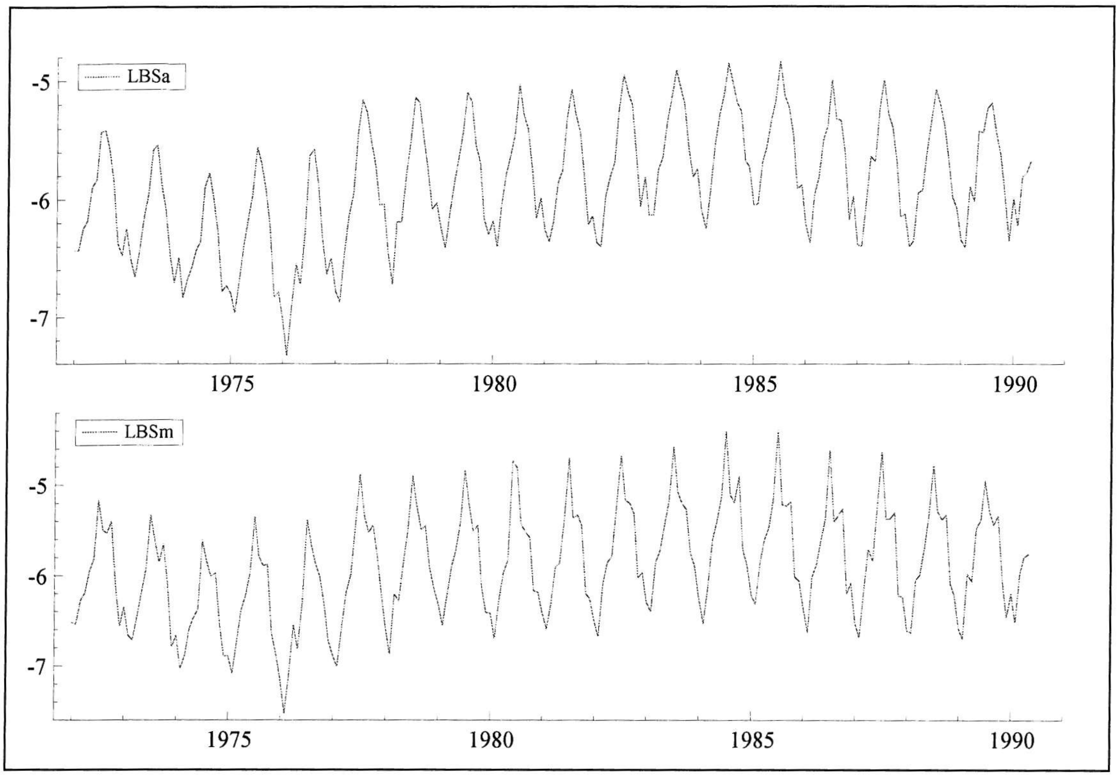
Table 8. 16 Comparison of Weights $w_{i,t}$ and $x_{i,t}$

$w_{i,t}$	70s	80s	90s	average	rank	
FRA	12.55	11.36	11.28	11.73	GER	66.24
GER	65.46	66.47	66.79	66.24	FRA	11.73
SWE	2.49	2.17	2.01	2.22	UK	10.36
SWI	9.22	9.70	9.41	9.44	SWI	9.44
UK	10.28	10.30	10.51	10.36	SWE	2.22
$x_{i,t}$	**70s**	**80s**	**90s**	**average**	**rank**	
FRA	14.54	17.85	14.99	15.79	GER	60.06
GER	59.26	56.31	64.61	60.06	FRA	15.79
SWE	1.47	0.90	0.59	0.98	SWI	13.95
SWI	14.23	15.02	12.59	13.95	UK	9.22
UK	10.51	9.92	7.23	9.22	SWE	0.98

One can observe that similarities between the two weights exist for most of the countries. However, it is argued that one of the limitations of the tourist balance of payments is that the data do not give any information about the origin of the tourist flows. Some currencies, such as dollars, are used more than others as a means of currency exchange (see Ballatori and Vaccaro, 1992). On balance, it seems that using $w_{i,t}$ as a weight should give a better definition of the variables of interest.

A second investigation has been done on the choice of using either annual or monthly weights in defining the real aggregated weighted average budget share (see definition provided in Section 8.4.3.1). Hence, two variables have been constructed as follows: *LBSm* using monthly weights and *LBSa* using the annual weight for the (log) real weighted average budget share. The graphs for the two variables are given in Figure 8.7.

Figure 8. 7 Annual versus Monthly Weights (1972:1-1990:5)



It appears that monthly weights tend to better highlight peaks, troughs and turning points of the seasonal distribution within the year, and in particular in the high season. Insofar, as some countries have different seasonal patterns (see Section 7.2.2, Chapter 7), one would wish to reflect this by using monthly weights. If not, *LBSm* and *LBSa* should be indistinguishable. The difference between them suggests using *LBSm*. This could also be inferred by inspection of the national seasonal patterns.

Note that the explanatory variables of interest have been defined in terms of annual weights. One could argue that holiday plans are made, in general, on an annual basis. Annual weights may be also thought to be more stable than monthly weights; more frequent observations might just reflect different seasonal patterns. A comparison can be done between Figure 8.8 and Figure H.1 (Appendix H). So, monthly weights will be used for *LBSm* and annual weights for the other variables.

8.4.3 Variables And Definitions

In this section, a definition of the variables under study is provided. The dependent variable is constructed as follows:

A) Real Aggregated Weighted Average Budget Share (*BSm*).

$$BS_t = \sum_{i=1}^{i=5} w^2{}_{i,t} * \left(\frac{TEXt}{EXPi,t} \right) \qquad (8.4.3.1)$$

where:

BS_t = Real aggregated weighted average budget share. Note that this variable has been deflated by dividing the total expenditure for tourism in Italy (TEX_t) by the total expenditure in the origin country i ($EXP_{i,t}$).

i = France, Germany, Sweden, Switzerland and United Kingdom, which represent an average 65% of the total countries originating tourism to Italy, for the period under study.

TEX_t = Total tourist receipts in current billions of lire, in month t (Source: Bank of Italy).

$EXP_{i,t}$ = Total expenditure in country i, in month t (Source: IFS *Datastream*, private consumption, average of quarterly data expressed in the currency of country i). Note that these figures have been calculated in billions of lire.

$w_{i,t}$ = This weight is formed by taking into consideration the number of nights spent (NS) by the tourists of each origin country i in all registered accommodation in month t (Source: ISTAT), and it is given by the following formula:

$$w_{i,t} = \frac{NSi,t}{\sum_{i=1}^{i=5} NSi,t} \qquad (8.4.3.2)$$

At this point a note is due. Ideally, the budget share should be defined as follows:

$$BS_t = \sum_{i=1}^{i=5} \alpha_{i,t} * \left(\frac{TEXi,t}{EXPi,t} \right) \qquad (8.4.3.3)$$

where:

$$\sum_{i=1}^{5} \alpha_{i,t} = \frac{w_{i,t}}{\sum_{i=0}^{5} w_{i,t}} \qquad (8.4.3.4)$$

and where $w_{i,t}$ is defined as in (8.4.3.2). However, tourism receipts disaggregated by country ($TEX_{i,t}$) are available only at an annual frequency; hence, the monthly total receipts have been used to calculate this variable, as mentioned before. One can define: $TEX_{i,t} = \lambda_{i,t}\, TEX_t$. Hence, definition (8.4.3.3) can be written as:

$$BS_t = \sum_{i=1}^{i=5} \alpha_{i,t} * \lambda_{i,t} \left(\frac{TEXt}{EXPi,t} \right) \qquad (8.4.3.5)$$

One method to approximate $\lambda_{i,t}$ is to use the monthly share of nights spent by tourists in all registered accommodation, as follows:

$$\lambda_{i,t} = \frac{NSi,t}{\sum_{i=1}^{i=5} NSi,t} = w_{i,t} \quad (8.4.3.6)$$

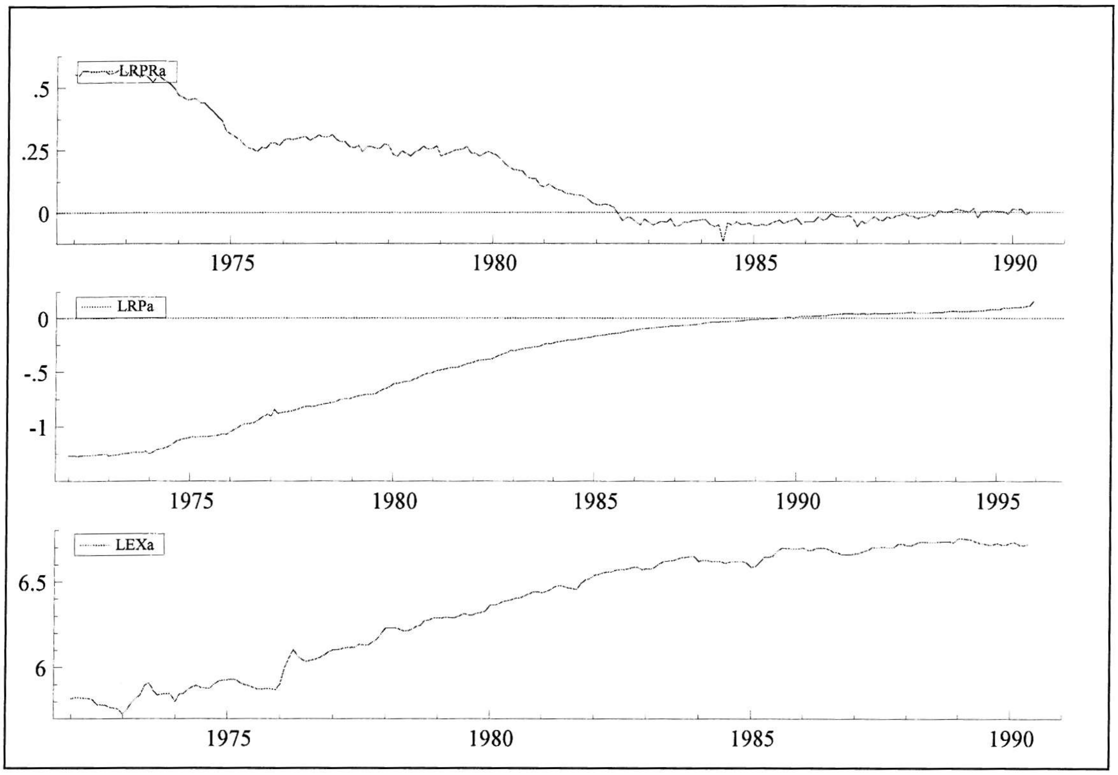
By substituting $\lambda_{i,t}$ by $w_{i,t}$ and re-defining $\alpha_{i,t}$ in terms of $w_{i,t}$ in definition (8.4.3.5), the final formula of the budget share (8.4.3.1) is obtained.

B) Income Proxy (*RPRa*), Relative Price (*RPa*) and Exchange rate (*EXa*) are defined as reported in Section 8.3.1 for five countries' aggregation only.

C) Real Substitute Price (*RSPj*) is defined in a disaggregated manner for the main competitors in the Mediterranean area, that is France, Greece, Portugal and Spain (see Section 8.3.1).

The plot of *LBSm* is given in Figure 8.7 (Section 8.4.1). The graphs of *LRPRa*, *LRPa* and *LEXa* are represented in Figure 8.8, and *LRSPfr*, *LRSPgr*, *LRSPpo* and *LRSPsp* in Figure 8.2 (Section 8.3.1). All the variables mentioned are expressed in logarithm.

Figure 8. 8 Real Industrial Production Index (*LRPRa*), Relative Price (*LRPa*) and Exchange Rate (*LEXa*) (1972:1-1990:5)



8.4.4 Seasonal Unit Roots And Long Run Unit Roots

In this section, an account will be given of the properties of each of the variables under study. As already stated, all the series will be expressed in logarithm and the analysis will be carried out for the period between 1972:1 and 1990:5. It is worth noting that from both the Franses and ADF tests similar results are expected for the five and seven countries aggregations.

The first step consists of testing for the existence of possible seasonal unit roots using Franses' (1991) test. Equation (2.6.2), where μ_t includes a constant, a trend and 11 seasonal dummies, is fitted by OLS to each of the series under study.

Table 8. 17 Testing Seasonal Unit Roots (1972:1-1990:5; 221 Obs. - 5 Countries Aggregation)

t-statistics	Variable			
	LBSm	***LRPRa***	***LRPa***	***LEXa***
π1	0.114	-1.960	-1.971	0.259
π2	-4.050****	-3.532****	-3.724****	-3.572****
π3	-1.286	-0.686	-3.253****	-5.227****
π4	-5.313****	-5.773****	-5.321****	-1.911
π5	-4.834****	-5.865****	-4.294****	-6.792****
π6	-4.374****	-5.874****	-5.023****	-6.377****
π7	0.407	-0.348***	-1.981****	-1.298
π8	-2.801	-2.720	-0.806	-1.217
π9	-2.145	-4.512****	-4.429****	-5.868****
π10	-4.363****	-6.363****	-4.905****	-4.230****
π11	0.844	-2.592****	-3.337****	-4.670****
π12	-5.366****	-3.507***	-3.586***	-0.692
F-statistics	***LBSm***	***LRPRa***	***LRPa***	***LEXa***
π3, π4	15.233****	17.025****	17.831****	16.156****
π5, π6	11.842****	18.724****	12.626****	24.040****
π7, π8	15.102****	25.846****	16.047****	13.217****
π9, π10	9.521****	21.898****	15.159****	19.183****
π11, π12	17.723****	22.419****	30.109****	17.473****
π3, ...,π12	24.362****	132.813****	179.160****	81.051****

Note: The four, three, two and one asterisks indicate that the seasonal unit root null hypothesis is rejected at the 1%, 5%, 10% and 20% level respectively.

As emerges from Table 8.17, there is no evidence of seasonal unit roots. However, the null hypothesis of the existence of a long run unit root cannot be rejected for all the variables under consideration. As a reminder, the explanatory variables *LRSPfr*, *LRSPgr*, *LRSPpo* and *LRSPsp* (real substitute price for each of the competitors) have not presented unit roots in the seasonals (see Table 8.1).

Table 8.18 includes the long run ADF unit root test in order to establish the integration status of each of the variables.

Table 8. 18 Testing Long Run Unit Roots: 1972:1-1990:5 (5 Countries Aggregation)

Series	*ADF*(1)	*LAG*(2)
LBSm(c)	- 3.71 **	6
LBSm(c,t)	- 5.23 **	6
LBSm(c,s)	- 4.17 **	0
LBSm(c,t,s)	- 3.95 *	1
LRPRa(c)	- 3.30 *	1
LRPRa(c,t)	- 1.78	8
DLRPRa(c,t)	- 4.03 **	7
LRPRa(c,s)	- 3.33 *	1
LRPRa(c,t,s)	- 1.84	8
DLRPRa(c,t,s)	- 3.71 *	7
LRPa(c)	- 3.14 *	2
LRPa(c,t)	- 1.67	1
DLRPa(c,t)	- 6.97 **	2
LRPa(c,s)	- 3.23 *	2
LRPa(c,t,s)	- 0.91	3
DLRPa(c,t,s)	- 6.56 **	2
LEXa(c)	- 2.20	5
DLEXa(c)	- 7.56 **	4
LEXa(c,t)	- 0.37	5
DLEXa(c,s)	- 7.93 **	4
LEXa(c,s)	- 1.96	1
DLEXa(c,s)	- 9.73 **	0
LEXa(c,t,s)	- 1.29	1
DLEXa(c,t,s)	- 9.87 **	0

Notes:
(1) Augmented Dickey-Fuller (ADF) statistics with constant (*c*) critical values: 5%=-2.876 1%=-3.463; constant and Trend (*c,t*) included c.v.: 5%=-3.433 1%=-4.005; Constant and Seasonals (*c,s*) included c.v.: 5% = -2.876 1% = -3.463; Constant and Trend and Seasonals (*c,t,s*) included c.v.: 5% = -3.433 and 1% = -4.005;
(2) Number of lags set to the first statistically significant lag, testing downward and upon white residuals. Note that ADF(0) corresponds to the DF test.
(3) ** significant at the 1% ; * significant at the 5% level.

The ADF test suggests the dependent variable (*LBSm*) is stationary in the level. The income proxy (*LRPRa*) is non-stationary in the level. The relative price (*LRPa*) is I(0) when a constant, or constant and seasonals are included otherwise it is stationary in the first difference. The (log) exchange rate (*LEXa*) is found to be I(1). All these results confirm those obtained for the seven countries' case. As a reminder from Table 8.2, the real substitute price for Greece (*LRSPgr*) is I(0) about a trend as well as the real substitute price for Portugal (*LRSPpo*). On the other hand, the real substitute price for France (*LRSPfr*) and for Spain (*LRSPsp*) are found to be I(1).

8.4.5 Possible Cointegration Amongst I(1) Variables

The next step of the analysis consists in considering the possible existence of a cointegrating relationship between the relative price (*LRPa*) and (*LEXa*). The expectations are twofold: these two variables should drift together in the long run, in accordance with economic theory, and similar results are expected to the seven countries aggregation case.

An initial 13 lag system has been run, which includes a constant, a trend and 11 seasonals treated unrestrictedly. Impulse dummies, created after inspecting the residuals in order to avoid problems of non-normality, worsen the results in terms of diagnostics, so they are not included in the system. The system can be reduced further to 2 lags, as suggested by the SC and HQ criteria (the complete results are reported in Appendix H, Table H.2). From a Johansen cointegration analysis, there is evidence for the presence of one cointegrating relationship between the two variables. From Table H.2, the cointegrating vector can be derived and it is given by the first row of the β' matrix:

$$CI = LRPa \ - 0.99691\ LEXa \qquad (8.4.5.1)$$

The expectation is that the two cointegrated variables will present a long run coefficient of one. Thus, the restriction: β=-1 is tested for the coefficient for the (log) exchange rate and the null hypothesis fails to be rejected at the 5% level from the χ^2 test[86]. This finding is consistent with economic theory for which, in the long run, there appear no significant differences in the inflation rates for the countries under consideration. As a reminder from Section 8.3.3, the long run coefficient restriction could not be accepted when including Japan and USA as source countries of tourism to Italy.

As for the seven countries aggregation case, a cointegration analysis has also been run for the three I(1) variables (*i.e.* real industrial production, relative price and exchange rate). Statistical evidence has been found for the existence of one cointegrating vector (see Appendix H, Table H.3); however, there are difficulties in interpreting the results on an economic basis. Hence, an investigation follows in employing a non-linear transformation for the real industrial production index.

86 The results for the restriction test on the coefficient is: $\chi^2(1) = 3.7093$ [0.0541] .

The cointegration analysis results for each of the real substitute prices have been given in Section 8.3.3. Statistical evidence has been found for each of real substitute prices to be treated as stationary in the level.

8.4.6 Estimating The Weighted Aggregated Budget Share

The first aim of this analysis is to estimate a model using the weighted budget share for a five countries aggregation as the dependent variable.

As stated in the previous sections, the real income proxy (*LRPRa*) has been found to be non-stationary in the level. Thus, this variable needs an appropriate transformation in order to be included in the estimation of the real budget share. In this section, an account of the use of a logistic transformation for *LRPRa* is given.

The Box and Cox procedure has been applied in order to give a statistical foundation for the choice of the logarithmic functional form. The complete results for this analysis are reported in Appendix I.

The initial formulation of the equation for the aggregated budget share (*LBSm*) is the following:

$$LBSm = f\,(LRPRa,\ SLRPRa,\ DLRPa,\ DLEXa,\ CI,\ LRSPfr,\ LRSPgr,\ LRSPpo,\ LRSPsp,\ Easter,\ T,\ SEAS,\ D) \qquad (8.4.6.1)$$

By using a VAR, it is possible to identify the lag size of the system. The system for *LBSm* includes: a constant, 11 seasonal dummies (*SEAS*), the "Easter" dummy variable (*Easter*), three impulse dummies created in order to correct for problems of non-normality in the residuals (*D* that is *i1980p6, i1989p5* and *i1989p12*), a time trend (*T*) (all these variables are treated unrestrictedly), the first lag for the cointegrating vector (*CI*), plus 13 lags for each of the other explanatory variables and the dependent variable (treated as endogenous). Note that *SLRPRa* in (8.4.6.1) is the quadratic (log) real production index which allows for non-linearities in the budget share. The period under study is from 1972:1 to 1990:5.

According to the joint F-test a restricted system with 11 lags is accepted at the 1% level. Whereas, at least a 13 lag system is suggested by the AIC criterion and a 1 lag system is suggested by the SC and HQ criteria[87]. Note, however, that the latter system has presented problems of serial correlation, heteroscedasticity and non-normality in the residuals. Hence, an initial very unrestricted 13 lag system is chosen in accordance with the AIC criterion.

From Table 8.19, the diagnostic statistics show a good specification. The correlation of the actual and fitted values suggests that the equation explains 99.4% of the variance of the dependent variable. No problems appear in terms of diagnostic tests.

Table 8. 19 Statistical Tests of the Equation for the Weighted Average Budget Share (*LBSm*)

σ = 0.110776 RSS = 0.8835287202	
correlation of actual and fitted	
LBSM	0.99423
LBSM	:Portmanteau 12 lags= 10.75
LBSM	:AR 1- 7 F(7, 65) = 0.468 [0.8544]
LBSM	:Normality Chi^2(2)= 0.054 [0.9733]
LBSM	:ARCH 7 F(7,58) = 0.433 [0.8776]

A model with 13 lags is estimated for the equation of *LBSm* and the parsimonious model obtained is reported in Table 8.20.

87

system	T	p		log-likelihood	SC	HQ	AIC
1	207	243	OLS	7767.2399	-68.786	-71.116	-73.046
2	207	324	OLS	7859.9880	-67.595	-70.702	-72.942
..	....	.....		.....	....	....	....
10	207	981	OLS	8552.8216	-57.364	-66.771	-73.636
11	207	1053	OLS	8680.4680	-56.742	-66.840	-73.869
12	207	1134	OLS	8803.4244	-55.843	-66.718	-75.057
13	207	1215	OLS	8952.3705	-55.196	-66.847	-75.496

```
System 11 --> System 10: F(72, 506) =    1.5808 [0.0029] **
System 12 --> System 11: F(81, 480) =    1.1977 [0.1305]
System 13 --> System 12: F(81, 422) =    1.3012 [0.0529]
```

Table 8. 20 Final Model for the Aggregated Budget Share (*LBSm*)

EQ(1) Modelling LBSm by OLS (using LBSq5.in7)
The present sample is: 1973 (3) to 1990 (5)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	0.18732	3.2028	0.058	0.9534	0.0000
LBSm_3	0.21205	0.053976	3.929	0.0001	0.0939
LBSm_11	0.15896	0.054602	2.911	0.0042	0.0538
LBSm_12	0.32590	0.054234	6.009	0.0000	0.1951
LRPRa_1	-1.6292	0.63039	-2.584	0.0107	0.0429
LRPRa_2	1.7767	0.70564	2.518	0.0129	0.0408
LRPRa_4	-1.4645	0.67118	-2.182	0.0307	0.0310
LRPRa_12	1.5817	0.45190	3.500	0.0006	0.0760
SLRPRa_1	1.2384	1.4958	0.828	0.4090	0.0046
SLRPRa_2	3.2580	1.6518	1.972	0.0504	0.0254
SLRPRa_4	0.85284	1.4502	0.588	0.5574	0.0023
SLRPRa_12	-1.0602	0.80590	-1.316	0.1903	0.0115
DLRPa_1	-2.9373	1.2111	-2.425	0.0165	0.0380
DLRPa_2	-4.2135	1.3007	-3.239	0.0015	0.0658
DLRPa_3	-2.8408	1.1300	-2.514	0.0130	0.0407
DLRPa_5	5.6023	1.2946	4.327	0.0000	0.1117
DLRPa_6	4.0221	1.3539	2.971	0.0035	0.0559
DLRPa_7	3.8810	1.1837	3.279	0.0013	0.0673
DLRPa_9	3.9333	1.1743	3.350	0.0010	0.0700
DLRPa_10	2.9606	1.2603	2.349	0.0201	0.0357
DLRPa_11	5.7830	1.3197	4.382	0.0000	0.1142
DLRPa_12	6.9637	1.3496	5.160	0.0000	0.1516
DLRPa_13	7.9210	1.2197	6.494	0.0000	0.2206
DLEXa_2	1.6060	0.65508	2.452	0.0154	0.0388
DLEXa_3	2.2301	0.65738	3.392	0.0009	0.0717
DLEXa_6	1.6125	0.68758	2.345	0.0203	0.0356
DLEXa_10	1.4869	0.65011	2.287	0.0236	0.0339
DLEXa_13	-2.1964	0.59113	-3.716	0.0003	0.0848
RLRSPfr	-3.2253	0.65614	-4.916	0.0000	0.1395
LRSPfr_7	-2.1331	0.44652	-4.777	0.0000	0.1328
RLRSPfr1	1.9834	0.71368	2.779	0.0062	0.0493
RLRSPpo	2.0757	0.34304	6.051	0.0000	0.1973
LRSPfr_12	3.3464	0.69374	4.824	0.0000	0.1351
LRSPfr_13	-2.8206	0.60223	-4.684	0.0000	0.1283
LRSPgr	-1.2939	0.18846	-6.865	0.0000	0.2403
RLRSPgr	1.4710	0.29993	4.904	0.0000	0.1390
LRSPgr_6	-0.49106	0.19567	-2.510	0.0132	0.0406
LRSPgr_12	-0.96124	0.19415	-4.951	0.0000	0.1413
LRSPpo_6	1.5918	0.34127	4.664	0.0000	0.1274
LRSPpo_8	-1.5518	0.32906	-4.716	0.0000	0.1299
RLRSPsp	-0.71504	0.35157	-2.034	0.0437	0.0270
LRSPsp_10	0.91210	0.23567	3.870	0.0002	0.0913
CI_1	2.3457	0.47884	4.899	0.0000	0.1387
i1980p6	0.69481	0.11264	6.168	0.0000	0.2034
I1989P5	0.45579	0.10697	4.261	0.0000	0.1086
i1989p12	-0.38827	0.10792	-3.598	0.0004	0.0799
Jan	-0.14139	0.041999	-3.367	0.0010	0.0707
Feb	-0.18752	0.062258	-3.012	0.0030	0.0574
Mar	0.086108	0.070476	1.222	0.2237	0.0099
Apr	0.28023	0.084015	3.335	0.0011	0.0695
May	0.33459	0.10467	3.197	0.0017	0.0642
Jun	0.41264	0.11653	3.541	0.0005	0.0776
Jul	0.92968	0.10728	8.666	0.0000	0.3351
Aug	0.48025	0.086942	5.524	0.0000	0.1700
Sep	0.23460	0.080298	2.922	0.0040	0.0542
Oct	0.35421	0.069229	5.117	0.0000	0.1494
Nov	0.052624	0.040947	1.285	0.2007	0.0110
Trend	0.0076918	0.0010649	7.223	0.0000	0.2593

R^2 = 0.981018 F(57,149) = 135.1 [0.0000] sigma = 0.0989364
DW = 1.74 RSS = 1.458471906 for 58 variables and 207 observations

```
AR 1- 7 F( 7,142) =       1.0636 [0.3901]
ARCH 7  F( 7,135) =       0.5032 [0.8308]
Normality Chi^2(2)=       2.5215 [0.2834]
Xi^2    F(100, 48) =      0.5480 [0.9940]
RESET   F( 1,148) =       0.4902 [0.4849]
```

Five restrictions could be imposed, as one can see in Table 8.22: *RLRSPfr* is defined by the difference between the fifth and sixth lags[88], *RLRSPfr1* is given by the difference between the eighth and ninth lags[89], *RLRSPsp* is defined by the difference between the fifth and seventh lags[90], *RLRSPpo* is given by the difference between the tenth and twelfth lags[91] and, finally, *RLRSPgr* is defined as the difference between the the first and second lags of the corresponding variables[92]. No other restrictions have been accepted either by the joint *F*-test or by suggestion of the SC criterion. As one can notice, no problems appear in the residuals. The long run dynamics are provided in Table 8.21.

[88] The restriction on the coefficient of the fifth and sixth lags, that present an opposite sign and similar magnitude, is accepted at the 5% level from the joint *F*-test (1,139) as the calculated value 0.017 is smaller than the critical value 3.84; moreover, the SC criterion is minimised when the restriction is imposed.

[89] The restriction on the coefficient of the eighth and ninth lags is accepted at the 5% level from the joint *F*-test (1,141). The calculated value equals 0.11 and it is smaller than the critical value 3.84; moreover, the SC criterion is minimised when such a restriction is imposed.

[90] The restriction on the coefficient of the fifth and seventh lags, which present an opposite sign and similar magnitude, is accepted at the 5% level from the joint *F*-test (1,140) as the calculated value 0.27 is smaller than the critical value 3.84; moreover, the SC criterion is minimised when such a restriction is imposed.

[91] The restriction on the coefficient of the tenth and twelfth lags, which present an opposite sign and similar magnitude, is accepted at the 5% level from the joint *F*-test (1,142) as the calculated value 2.77 is smaller than the critical value 3.84; moreover, the SC criterion is minimised when this restriction is imposed.

[92] The restriction on the coefficient of the first and second lags, which present an opposite sign and similar magnitude, is accepted at the 5% level from the joint *F*-test (1,143) as the calculated value 2.55 is smaller than the critical value 3.84; note also that this restriction is suggested by the SC criterion.

Table 8. 21 Long run Equation for *LBSm*

LBSm =	+0.618		+0.8732	LRPRa	+14.15	SLRPRa
(SE)	(10.6)		(1.193)		(3.69)	
	+102.5	DLRPa	+15.64	DLEXa	-5.303	LRSPfr
	(34.67)		(7.064)		(1.78)	
	-9.06	LRSPgr	+0.132	LRSPpo	+3.009	LRSPsp
	(2.119)		(0.7597)		(1.193)	
	+7.739	CI	+2.292	i1980p6	+1.504	I1989P5
	(2.429)		(0.6901)		(0.5195)	
	-1.281	i1989p12	-0.4665	Jan	-0.6187	Feb
	(0.467)		(0.1538)		(0.203)	
	+0.2841	Mar	+0.9246	Apr	+1.104	May
	(0.2549)		(0.3736)		(0.4452)	
	+1.361	Jun	+3.067	Jul	+1.584	Aug
	(0.4414)		(0.6978)		(0.377)	
	+0.774	Sep	+1.169	Oct	+0.1736	Nov
	(0.22)		(0.2)		(0.1276)	
	+0.02538	Trend	-10.64	RLRSPfr	-2.359	RLRSPsp
	(0.00737)		(3.309)		(1.323)	
	+6.544	RLRSPfr1	+6.848	RLRSPpo	+4.853	RLRSPgr
	(2.975)		(2.136)		(1.543)	

ECM = LBSm - 0.61803 - 0.873197*LRPRa - 14.1507*SLRPRa - 102.525*DLRPa
- 15.6358*DLEXa + 5.30265*LRSPfr + 9.0604*LRSPgr - 0.131956*LRSPpo
- 3.00927*LRSPsp - 7.73917*CI - 2.29235*i1980p6 - 1.50378*I1989P5
+ 1.28101*i1989p12 + 0.466497*Jan + 0.618695*Feb - 0.284093*Mar
- 0.924561*Apr - 1.1039*May - 1.36142*Jun - 3.06727*Jul - 1.58446*Aug
- 0.77401*Sep - 1.16864*Oct - 0.17362*Nov - 0.0253773*Trend + 10.6412*RLRSPfr
+ 2.35912*RLRSPsp - 6.54382*RLRSPfr1 - 6.84834*RLRSPpo - 4.85322*RLRSPgr;

WALD test Chi^2(29) = 164.11 [0.0000] **

8.4.7 Weighted Aggregated Budget Share: A Non-linear Estimation

Again, from Table 8.20, there appear difficulties in identifying the points of maximum or minimum for each lag pair of the real industrial production index (*LRPRa* and *SLRPRa*). Taking the partial derivative with respect to the most significant lag, the twelfth, from the generic equation:

$$y = a\,x^2 + b\,x + e$$

and equating to zero, it is found that x=0.75; hence, as $a<0$ there is a maximum; moreover, x is greater than the largest observation for *LRPRa* (*i.e.*0.57664), hence the existence of a maximum is reasonable. For the second lag, x=-0.27; hence, as $a>0$ there is a minimum; moreover, in this case x is smaller than the smallest observation for *LRPRa* (*i.e.* -0.11784), hence the presence of a minimum seems reasonable. Analysing the first lag, x equals 0.66; as $a>0$ there is a minimum; however, in this case x is greater than the smallest observation for *LRPRa* (*i.e.* -0.11784), hence the presence of a minimum does not seem reasonable. Finally, taking the partial derivative with respect to the fourth lag, x=0.86; hence, as $a>0$ there is a minimum; however, x is greater than the smallest observation for *LRPRa* (*i.e.* -0.11784), hence the presence of a minimum does not seem reasonable. As already mentioned, given that a quadratic specification is used as a local approximation to a sigmoid curve, *e.g.* a logistic function, then the minimum should lie below all the observations. Therefore, the values obtained suggest non-linearity, but not in the form that will accommodate an I(1) variable in an I(0) equation. These turning points have not been precisely estimated for each of the lags. Hence, as in the tourist receipts case, there are difficulties in using an I(1) variable to explain an I(0) variable.

These findings lead to the use of a non-linear expression for the most significant lags of the (log) real industrial production index. The generic logistic function for the real income proxy is used, as follow:

$$\theta\ (x, \mu, \sigma, \alpha) = \alpha\left(1 + e^{-(x-\mu)\ /\ \sigma}\right)^{-1}$$

where:

x = is the most significant lag for *LRPRa*; in this case the twelfth, second, first and fourth lag, respectively;

μ = is the centre of the curvature;

σ = is the spread of the curvature;

α = is the impact parameter.

The mean and the standard error of *LRPRa* have used as starting values of the parameters μ (*mu* equals 0.16) and σ (*sig* equals 0.20). The aim is to find the smallest RSS that corresponds to the maximum of the likelihood. In running the non-linear specification, the TSP package is used. The results obtained are provided in Table 8.22.

Table 8. 22 Non-linear Estimation for *LBSm* Equation

Dependent variable: LBSM
Mean of dependent variable = -5.87951
Std. dev. of dependent var. = .610731
Sum of squared residuals = 2.13953
Variance of residuals = .014169

Std. error of regression = .119034
R-squared = .972631
Adjusted R-squared = .962662
Durbin-Watson statistic = 1.23277

Log of Likelihood Function = 179.523
Number of Observations = 207

Parameter	Estimate	Standard Error	t-statistic
Costant	-.052981	3.93037	-.013480
LBSm(-3)	.323801	.063664	5.08607
LBSm(-11)	.117858	.065122	1.80979
LBSm(-12)	.334415	.064777	5.16252
DLRPa(-1)	-2.39930	1.45434	-1.64975
DLRPa(-2)	-3.37960	1.54909	-2.18167
DLRPa(-3)	-2.21154	1.36078	-1.62520
DLRPa(-5)	5.19525	1.53715	3.37978
DLRPa(-6)	4.61145	1.56033	2.95543
DLRPa(-7)	3.28867	1.39160	2.36323
DLRPa(-9)	2.34484	1.40462	1.66937
DLRPa(-10)	1.69740	1.49997	1.13162
DLRPa(-11)	3.94681	1.59810	2.46969
DLRPa(-12)	5.31729	1.61000	3.30267
DLRPa(-13)	7.09624	1.44528	4.90995
DLEXa(-2)	.213693	.782453	.273107
DLEXa(-3)	.996495	.797460	1.24959
DLEXa(-6)	1.26046	.826786	1.52453
DLEXa(-10)	1.53816	.775091	1.98449
DLEXa(-13)	-1.51996	.696093	-2.18355
BETA*	.708976	5.94785	.119199
MU	.821939	7.27207	.113027
SIG	.616720	3.56528	.172980
ZETA*	6.95274	57.1332	.121693
ETA*	-3.25872	26.5956	-.122529
GAMMA*	-1.66176	13.6719	-.121546
RLRSPfr	-1.55714	.744263	-2.09219
LRSPfr(-7)	-.704626	.489475	-1.43956
LRSPfr1	1.61527	.861444	1.87508
LRSPfr(-12)	2.65582	.826836	3.21203
LRSPfr(-13)	-2.18844	.707142	-3.09477
LRSPgr	-1.00105	.223302	-4.48296
RLRSPgr	1.10429	.357929	3.08522
LRSPgr(-6)	-.231731	.231579	-1.00065
LRSPgr(-12)	-.488882	.230606	-2.11999
RLRSPpo	1.71293	.410506	4.17273
LRSPpo(-6)	1.41928	.410658	3.45612
LRSPpo(-8)	-1.37680	.394902	-3.48643
RLRSPsp	-.500353	.421400	-1.18736
LRSPsp(-10)	.660525	.277783	2.37784
CI(-1)	.917245	.584938	1.56811
trend	.453333E-02	.124337E-02	3.64600
jan	-.159901	.050195	-3.18562
feb	-.132227	.074963	-1.76388
mar	.179928	.084500	2.12934
apr	.397849	.100523	3.95778

may	.503651	.125258	4.02091
jun	.565623	.139747	4.04747
jul	1.00498	.128531	7.81892
aug	.545199	.104557	5.21437
set	.292521	.096618	3.02762
oct	.291204	.082648	3.52342
nov	.049449	.049137	1.00636
i1980p6	.662178	.135470	4.88799
i1989p5	.399103	.128192	3.11331
i1989p12	-.326182	.129194	-2.52476

Notes: *beta*, *zeta*, *eta* and *gamma* are the coefficients for the logistic transformation for the twelfth, second, first and fourth lag of the income proxy (*LRPRa*)

A number of tests has been undertaken in order to achieve the best specification. Four distinct cases have been considered starting with the inclusion of the first most significant lag (*i.e.* $LRPRa_{t-12}$ where RSS equals 2.34725), and then introducing one by one the other statistically significant lags with the following order: $LRPRa_{t-2}$ (where RSS=2.10702), $LRPRa_{t-1}$ (where RSS=2.13862) and, finally, $LRPRa_{t-4}$ (where RSS=2.13953). Hence, a test is run on the joint statistical significance of the coefficients for the second, first and fourth lag, that is H0: *beta*, *zeta*, *gamma=0*. From the F-test (3,152), the calculated value equals 4.92 that is greater than the critical value, 2.60, at the 5% level. Therefore, the unrestricted model holds. The null hypothesis has been tested also by a LR(3) test. The calculated value is 19.2 greater than the critical value (7.81) at the 5% level. Again, the null hypothesis cannot be accepted.

From Table 8.22, the parameters *mu* and *sig* do not look significant. Hence, the next stage of the test involves testing these parameters. In this case, if *mu* lies within the observations, and *sig* is large, the logistic transformation of a variable will be arbitrarily highly correlated with the level of the variable, and the linear model is approximately nested within the logistic model. Hence, an approximately likelihood ratio test can be carried out. The calculated LR(2) equals 7.36[93] that is greater than the critical value (5.99) at the 5% level of significance. Hence, the linearity specification is rejected.

So far, the same parameters (*mu* and *sig*) have been considered for all the lags of *LRPRa*. The next step of the investigation involves freeing the parameters for each of the logistic transformation of *LRPRa*, as follows:

93 The log-likelihood for the linear version equals (175.844) and the log-likelihood for the logistic version equals (179.523).

A) *beta*, *mu* and *sig* for $LRPRa_{t-12}$;

B) *zeta*, *mu1* and *sig1* for $LRPRa_{t-2}$;

C) *eta*, *mu2* and *sig2* for $LRPRa_{t-1}$;

D) *gamma*, *mu3* and *sig3* for $LRPRa_{t-4}$;

Hence, the aim is to test the following hypotheses:

H0: $mu1 = mu2 = mu3 = mu$ and $sig1 = sig2 = sig3 = sig$

H1: $mu1 \neq mu2 \neq mu3 \neq mu$ and $sig1 \neq sig2 \neq sig3 \neq sig$

by running a LR(6) test. In this case, the log-likelihood for the restricted (*Lres*) model equals 179.523 and the log-likelihood for the unrestricted (*Lures*) model equals 179.444. The TSP package fails to find the maximum as the log-likelihood in the unrestricted model is worse. A further experiment has been done by freeing off fewer parameters at a time, in *mu* and *sig* pairs. The following tests are more illustrative than conclusive. The first parameters (*i.e. mu1* and *sig1*) are for the second lag of the income proxy. The *F*-test (2,150) equals 9.90 greater than the critical value (3.00) at the 5% level[94]. The LR(2) equals 26.14 greater than the critical value (5.99) at the 5% level[95]. Hence, in both the cases the restriction cannot be accepted. The same tests are run for the second set of parameters (*mu2* and *sig2*). The *F*-test (2,148) presents a negative sign (-26.53) as the RSS for the unrestricted model equals 5.89203, that is greater than the RSS in the restricted model. On the other hand, the LR(2) is equal to 5.46 smaller than the critical value (5.99) at the 5% level. Freeing the third set of parameters (*mu3* and *sig3*) leads to the following results. The calculated *F*-test (2,146) equals 328.47 greater than 3.00, hence the unrestricted model has to be run. However, the LR(2) presents a negative sign (-31.76). It is worth noting that similar findings have been reached when using different starting values for *mu* and *sig*[96]. In conclusion, TSP does not seem to be accurate enough in handling the maximisation and it does not improve the likelihood of the restricted model but even makes it worse.

Hence, the analysis is continued considering the four lags for the logistic transformation of the income proxy where only *mu* and *sig* are freely estimated (see

---

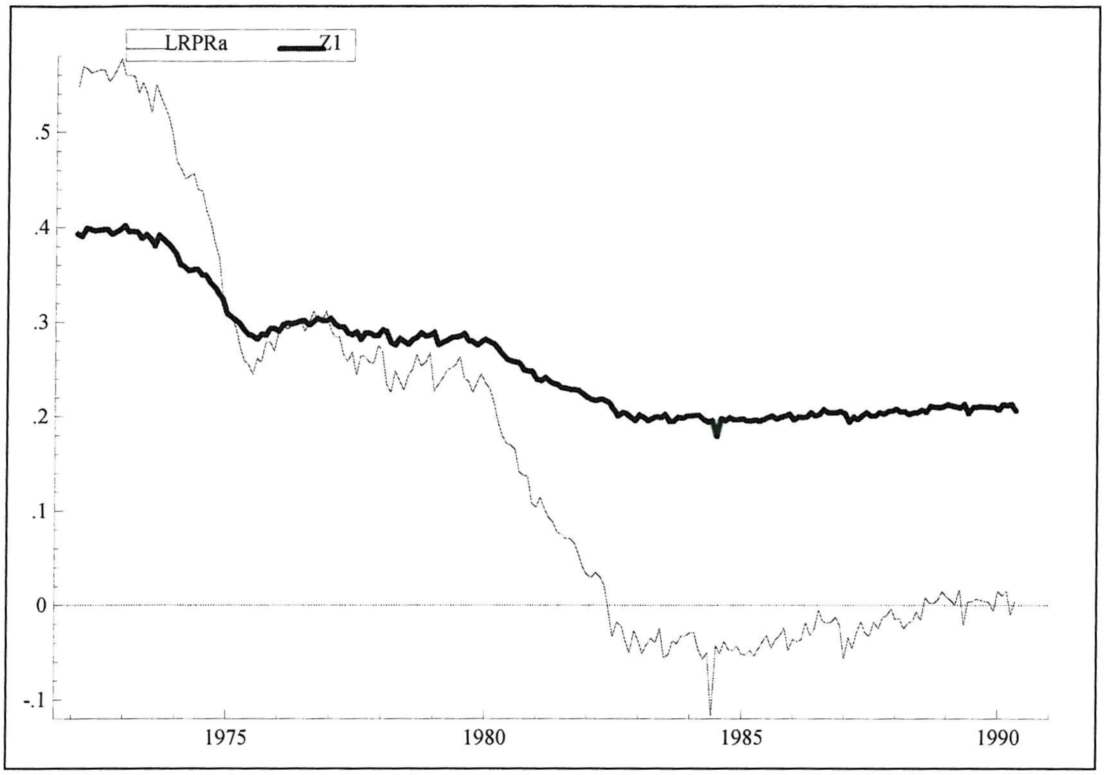
94 The RSS for the restricted model equals 2.13953 and the RSS for the unrestricted model is 1.88996.

95 The log-likelihood for the restricted model is 179.523 and for the unrestricted model equals 192.592.

96 One solution has been pursued starting the unrestricted model from the restricted maximum, that is setting $mu = 0.821939$ and $sig = 0.616720$ (see Table 8.22). Nevertheless, the TSP package has shown similar problems in achieving the maximisation.

Table 8.22). The plot of $Z_{i,t}$ (*i.e.* the logistic transformation of $LRPRa_{i,t}$[97]) on *LRPRa* is given in Figure 8.9.

Figure 8. 9 *LBSm*: Plot of $LRPRa_{i,t}$ on $Z_{i,t}$



From Figure 8.9, the logistic transformation of the income variable seems to be acting as a dummy variable. However, to have a better understanding of this variable ($Z_{i,t}$), a normalisation can been done for *LRPRa* and z_i in terms of their own mean and standard deviation in the following manner:

$$NLRPRa_{i,t} = [(LRPRa_{i,t} - \mu) / \sigma]$$

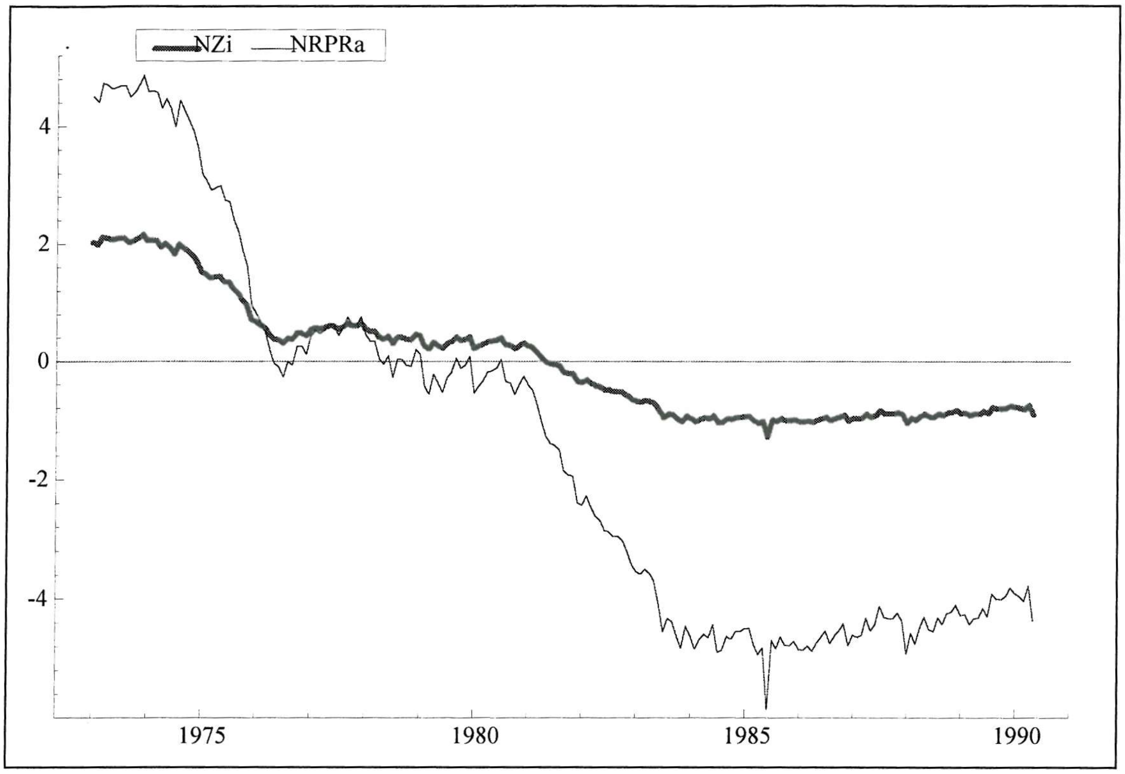
where μ is the mean of $LRPRa_{i,t}$ equal to 0.16 and σ is its standard deviation equal to 0.20. The logistic transformation of the income proxy ($Z_{i,t}$), normalised for its own mean and standard deviation, is given by the following formula:

$$NZ_{i,t} = [(Z_{i,t} - mu)/ sig]$$

where *mu* is the mean of the transformed variable equal to 0.26 and *sig*, its standard deviation, equals 0.064. The graphical representation of the two variables is given in Figure 8.10.

[97] To note that the logistic transformation of the generic variable x_i is the following: $z_i = 1/1+EXP(-1*((x_i - \mu)/\sigma))$.

Figure 8. 10 *LBSm*: Plot of $NLRPRa_{i,t}$ on $NZ_{i,t}$



Visual inspection of Figure 8.10 suggests either an intercept shift or a bounded downward stochastic trend. The transformed income proxy decreases in two stages with a small rise in between. Given that a logistic shape is expected, in this particular case, a lot of the observations lie between floor and ceiling. Nevertheless, the interpretation of such a variable does appear to be ambiguous. Hence, the economic theory has largely vanished.

Some conclusions with respect to the budget share model seem to be due. The model reported in Table 8.22 has been run in PcGive in order to have a more immediate and straightforward comparison with the other models estimated so far. The tests provided by PcGive should be considered as asympotically valid. In Table 8.23, the long run dynamics are reported. From the Wald test, the null hypothesis that the long run coefficients, excluding the constant term, are jointly equal to zero has to be rejected.

Table 8. 23 *LBSm*: Long Run Dynamics

```
Solved Static Long Run equation
       LBSm =     +7.543                   +151.4 DLRPa           +14.96 DLEXa
(SE)          (    18.03)          (     81.59)           (     12.75)
                   +3.675 Zi                +44.44 Z2               -18.6 Z1
              (    6.247)          (     23.94)           (     14.47)
                   -8.543 Z4                -8.875 RLRSPfr         -1.071 LRSPfr
              (    12.73)          (     5.523)           (     2.163)
                   +8.873 RLRSPfr1          -10.55 LRSPgr          +6.012 RLRSPgr
              (     6.16)          (     4.414)           (     3.214)
                   +9.112 RLRSPpo          +0.3045 LRSPpo           -2.18 RLRSPsp
              (    4.796)          (      1.36)           (     2.411)
                   +4.032 LRSPsp            +6.199 CI            +0.02719 Trend
              (    2.484)          (     3.932)           (   0.01425)
                  -0.7948 Jan              -0.6832 Feb             +0.858 Mar
              (    0.368)          (    0.3843)           (    0.6129)
                   +1.972 Apr               +2.486 May             +2.713 Jun
              (    1.028)          (     1.249)           (     1.193)
                   +5.081 Jul               +2.724 Aug              +1.35 Sep
              (    2.039)          (      1.07)           (    0.4962)
                   +1.437 Oct              +0.2168 Nov             +3.488 i1980p6
              (   0.4421)          (    0.2371)           (     1.707)
                   +2.095 I1989P5           -1.807 i1989p12
              (    1.149)          (         1)

ECM = LBSm - 7.54337 - 151.448*DLRPa - 14.9617*DLEXa - 3.67459*Zi
     - 44.4379*Z2 + 18.5984*Z1 + 8.5434*Z4 + 8.8748*RLRSPfr + 1.07068*LRSPfr
     - 8.8734*RLRSPfr1 + 10.5464*LRSPgr - 6.0118*RLRSPgr - 9.11176*RLRSPpo
     - 0.304501*LRSPpo + 2.18036*RLRSPsp - 4.03166*LRSPsp - 6.19918*CI
     - 0.0271857*Trend + 0.794784*Jan + 0.683205*Feb - 0.857955*Mar
     - 1.97241*Apr - 2.48557*May - 2.71293*Jun - 5.08065*Jul - 2.7236*Aug
     - 1.35036*Sep - 1.43737*Oct - 0.216777*Nov - 3.48752*i1980p6 -
       + 2.09453*I1989P5 + 1.80659*i1989p12;

WALD test Chi^2(31) = 60.431 [0.0012] **
```

From Table 8.22, the Italian budget share of tourism is influenced by its own history. This is also consistent with the adjustment of the dependent variable to changes in the right hand side variables. The dependent variable shows a strong dependence on the past relative price growth; the long run coefficient has a positive sign and a *t*-value of +1.86. Hence, there is evidence to believe that foreigners do not show a prompt response to changes in the Italian price with respect to home prices. On the other side, the budget share is negatively influenced by the exchange rate growth in both the short and long run. Note also that the cointegrating vector (CI_{t-1}) presents a positive sign denoting that *LRPa* (the relative price) and *LEXa* (the weighted exchange rate) have an opposite effect on the real budget share. The substitute price for France (*LRSPfr*) shows the expected negative elasticity both in the short and long run, though these are not statistically significant. A negative elasticity occurs also for the substitute price for Greece (*LRSPgr*). In this case, the long run coefficient is statistically significant with a *t*-value equal to -2.39. On the other hand, a positive elasticity of substitution is determined for Portugal and Spain. In this case, short run coefficients are statistically significant, whereas the long run coefficient is not statistically significant for Portugal (*t*-value +0.22) and it is just statistically significant for Spain (*t*-value +1.62). Amongst the other variables, the time trend shows an upwards trend in popularity for Italian tourism and the seasonal dummies are, in general, statistically significant at the 5% level, with the highest coefficient for the month of July.

8.4.8 A Seven Countries Aggregation For The Budget Share (*LBS7m)*

In this section, an account is given of the main findings obtained by using the weighted budget share (*LBS7m*) with an aggregation for seven origin countries. The expectation is to obtain results similar to the five countries aggregation case.

The dependent variable is defined as in Section 8.4.3 but for a seven countries' aggregation; the explanatory variables of interest (*i.e.* income proxy *LRPRa*, relative price *LRPa*, weighted exchange rate *LEXa* and substitute price *LRSPj* for the four competitor countries) are defined as in Section 8.3.1. The first part of the analysis concerns the integration status of *LBS7m*. Running Franses' test, no evidence is found for the existence of seasonal unit roots (Table 8.24). From the ADF test, it emerges that *LBS7m* can be treated as an I(0) variable (Table 8.25).

Table 8. 24 Seasonal Unit Roots for *LBS7m*

t-statistics	Variable ***LBS7m***	*t*-statistics	Variable ***LBS7m***	*F*-statistics	***LBS7m***
$\pi 1$	-1.165	$\pi 7$	0.709	$\pi 3$, $\pi 4$	13.134****
$\pi 2$	-3.843****	$\pi 8$	-3.202**	$\pi 5$, $\pi 6$	12.856****
$\pi 3$	-0.137	$\pi 9$	-1.749	$\pi 7$, $\pi 8$	17.902****
$\pi 4$	-5.123****	$\pi 10$	-5.037****	$\pi 9$, $\pi 10$	13.201****
$\pi 5$	-4.905****	$\pi 11$	1.119	$\pi 11$, $\pi 12$	11.561****
$\pi 6$	-4.844****	$\pi 12$	-4.549****	$\pi 3$, ...,$\pi 12$	21.895****

Note: The two and four asteristics indicate that the unit root null hypthesis is rejected at the 10% and the 1% level, respectively.

Table 8. 25 ADF Test for *LBS7m* Defined for Seven Countries Aggregation

Series	*ADF*(1)	*LAG*(2)
LBS7m(c)	- 3.68 **	6
LBS7m(c,t)	- 5.16 **	6
LBS7m(c,s)	- 4.30 **	0
LBS7m(c,t,s)	- 3.60 *	1

Notes:
(1) Augmented Dickey-Fuller (ADF) statistics with constant (*c*) critical values: 5%=-2.876 1%=-3.463; constant and trend (*c,t*) included c.v.: 5%=-3.433 1%=-4.005; constant and seasonals (*c,s*) included c.v.: 5% = -2.876 1% = -3.463; constant and trend and seasonals (*c,t,s*) included c.v.: 5% = -3.433 and 1% = -4.005.
(2) Number of lags set to the first statistically significant lag, testing downward and upon white residuals. Note that ADF(0) corresponds to the DF test.
(3) ** significant at the 1% ; * significant at the 5% level.

The integration and cointegration status of the explanatory variables is the same as that reported in Section 8.3.1 (see Tables 8.1 and 8.2) and Section 8.3.2.

The initial formulation of the equation for the aggregated budget share (*LBS7m*) is the following:

LBS7m = f (LRPRa, SLRPRa, DLRPa, DLEXa, CI, LRSPfr, LRSPgr, LRSPpo, LRSPsp, Easter, T, SEAS, D) (8.4.8.1)

It is worth noting that the logarithmic specification has been tested against a linear specification. From the Box-Cox test, it emerges that the log-linear form is much better than the linear specification[98] .

By using a VAR, it is possible to identify the lag size of the system. This system for *LBS7m* includes: a constant, 11 seasonal dummies, the "Easter" dummy variable, four impulse dummies created in order to avoid problems of non-normality in the residuals (*1976p5*, *1978p11*, *1980p6* and *i1989p11*), a time trend (all these variables are treated unrestrictedly), the first lag for the cointegrating vector *CI* (see expression (8.3.3.1)), plus 13 lags for each of the other explanatory variables and the dependent variable (treated as endogenous). Note that *SLRPRa* in (8.4.8.1) is the quadratic (log) real income proxy which allows for non-linearities in the budget share. The period under study is from January 1972 to May 1990.

98 The calculated χ^2 equals 101.55 that is greater than the tabulated critical value, 3.84, at the 5% level; hence, the null hypothesis cannot be accepted, that is the two models are empirically different. Moreover, the SSELL (1.780189128) is smaller than SSEL/($\overline{\text{BSm}}$ G)^2 (4.748716) value, hence, the log-linear specification is adopted.

According to the joint F-test, a restricted system with 11 lags could be accepted at the 1% level. Whereas, at least a 13 lag system is suggested by the AIC criterion[99]. Following the latter suggestion, an initial very unrestricted 13 lag system is run.

From Table 8.26, the diagnostic statistics show a good specification. The correlation of the actual and fitted values suggests that the equation explains 99.3% of the variance of dependent variable. No problems appear in terms of diagnostics.

Table 8. 26 Statistical Tests of the Equation for *LBS7m*

σ = 0.137549 RSS = 1.343303926	
correlation of actual and fitted	
LBS7M	0.99345
LBS7M	:Portmanteau 12 lags= 7.7223
LBS7M	:AR 1- 7 F(7, 66) = 0.39959 [0.8991]
LBS7M	:Normality Chi^2(2)= 0.44394 [0.8009]
LBS7M	:ARCH 7 F(7,59) = 0.16044 [0.9918]

Hence, a model with 13 lags is estimated for the equation of *LBS7m*, and the final parsimonious restricted model obtained is shown in Table 8.27.

99

system	T	p		log-likelihood	SC	HQ	AIC
10	207	981	OLS	8621.9644	-58.032	-67.439	-74.304
11	207	1062	OLS	8761.4051	-57.292	***-67.476***	-74.651
12	207	1143	OLS	8889.2594	-56.441	-67.401	-74.887
13	207	1224	OLS	9012.1735	-55.542	-67.279	***-76.074***

System 11 --> System 10: F(81, 532) = 1.5236 [0.0038] **
System 12 --> System 11: F(81, 474) = 1.2333 [0.0963]
System 13 --> System 12: F(81, 416) = 1.0363 [0.4025]

Note that from the joint F-test a VAR(11) has to be estimated; the same conclusion is reached using the HQ criterion. The SC criterion suggests further coefficient reductions.

Table 8. 27 Final Model for *LBS7m*

EQ(1) Modelling LBS7m by OLS (using LBSq100.in7)					
The present sample is: 1973 (3) to 1990 (5)					
Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-1.3392	1.7196	-0.779	0.4372	0.0035
LBS7m_1	0.21972	0.053626	4.097	0.0001	0.0889
LBS7m_3	0.13780	0.055282	2.493	0.0136	0.0349
LBS7m_5	0.21008	0.054958	3.823	0.0002	0.0783
LBS7m_12	0.17067	0.056294	3.032	0.0028	0.0507
LRPRa_4	-0.92236	0.29463	-3.131	0.0021	0.0539
SLRPRa_4	2.7168	0.66809	4.067	0.0001	0.0877
LRSPfr_2	-1.5206	0.34169	-4.450	0.0000	0.1033
LRSPfr_10	0.74464	0.27917	2.667	0.0084	0.0397
LRSPgr	-0.57944	0.20099	-2.883	0.0044	0.0461
LRSPgr_6	-0.62961	0.20750	-3.034	0.0028	0.0508
LRSPpo_5	0.64127	0.28794	2.227	0.0272	0.0280
LRSPpo_10	-1.5322	0.47461	-3.228	0.0015	0.0571
LRSPpo_11	2.1501	0.64715	3.322	0.0011	0.0603
LRSPpo_12	-1.5408	0.43811	-3.517	0.0006	0.0671
DLRPa_6	3.5186	1.3700	2.568	0.0111	0.0369
DLRPa_13	5.7159	1.4544	3.930	0.0001	0.0824
DLEXa	-1.6325	0.63119	-2.586	0.0105	0.0374
CI_1	1.5539	0.32159	4.832	0.0000	0.1195
Easter	0.15403	0.052206	2.950	0.0036	0.0482
Jan	0.037682	0.049705	0.758	0.4494	0.0033
Feb	0.12269	0.074910	1.638	0.1033	0.0154
Mar	0.51975	0.078355	6.633	0.0000	0.2037
Apr	0.82233	0.13589	6.051	0.0000	0.1755
May	1.0137	0.13634	7.435	0.0000	0.2432
Jun	1.2814	0.14750	8.688	0.0000	0.3050
Jul	1.8175	0.18127	10.026	0.0000	0.3689
Aug	1.1048	0.16594	6.658	0.0000	0.2049
Sep	0.92276	0.13048	7.072	0.0000	0.2253
Oct	0.69480	0.11956	5.811	0.0000	0.1641
Nov	-0.11864	0.078911	-1.503	0.1346	0.0130
I1976P5	-0.61493	0.13782	-4.462	0.0000	0.1037
I1978P11	0.52987	0.12881	4.114	0.0001	0.0896
i1980p6	0.73096	0.12756	5.730	0.0000	0.1603
I1989P5	0.47028	0.12653	3.717	0.0003	0.0743
R^2 = 0.975812 F(34,172) = 204.08 [0.0000]sigma = 0.120283					
DW = 2.06 RSS = 2.488510454 for 35 variables and 207 obs.					
AR 1- 7 F(7,165) =	1.1512 [0.3339]				
ARCH 7 F(7,158) =	0.63698 [0.7248]				
Normality Chi^2(2)=	1.2654 [0.5312]				
Xi^2 F(51,120) =	0.95924 [0.5569]				
RESET F(1,171) =	0.33849 [0.5615]				
Tests of parameter constancy over: 1989 (6) to 1990 (5)					
Forecast Chi^2(12)=	18.303 [0.1068]				
Chow F(12,160) =	1.3682 [0.1862]				

Restrictions on the lags of the non-linear coefficients (*LRPRa* and *SLRPRa*) have been jointly tested by an *F*-test. As already stated, the equation for *LBS7m* allows for non-linearities in the coefficients of the real income proxy. To identify the point of either a maximum or a minimum, the generic equation for the fourth lag needs to be considered:

$$y = a x^2 + b x + e$$

Taking the partial derivative with respect to x (in this case *LRPRa*), and equating to zero, one finds that $x=0.17$; hence, as $a>0$ there is a minimum. However, x is greater than the smallest observation for *LRPRa* (*i.e.*-0.093202); whereas, the condition for the estimates to be consistent with an underlying sigmoid shape, requires x to lie outside the range of observations for *LRPRa*.

As for the previous cases, a non-linear transformation of the real income proxy is attempted. The generic function for *LRPRa* is as follows:

$$\theta\ (x, \mu, \sigma, \alpha) = \alpha\left(1 + e^{-(x-\mu)\ /\ \sigma}\right)^{-1}$$

where:

x = fourth lag for *LRPRa*;

μ = centre of the curvature, say *mu* 0.12 (as the mean of *LRPRa* for the period 1972:1-1990:5);

σ = spread of the curvature, say *sig* 0.21 (as standard error of *LRPRa*);

α = is the impact parameter, say *beta*.

The non-linear equation is run in TSP with the results reported in Table 8.28.

Table 8. 28 Non-linear Estimation for *LBS7m* Equation

Dependent variable: LBS7M

Mean of dependent variable = -6.18545	Std. error of regression = .124803
Std. dev. of dependent var. = 0.70670	R-squared = .974165
Sum of squared residuals = 2.66347	Adjusted R-squared = .968877
Variance of residuals = 0.01558	Durbin-Watson statistic = 1.99946
Log of Likelihood Function = 156.825	
Number of Observations = 207	

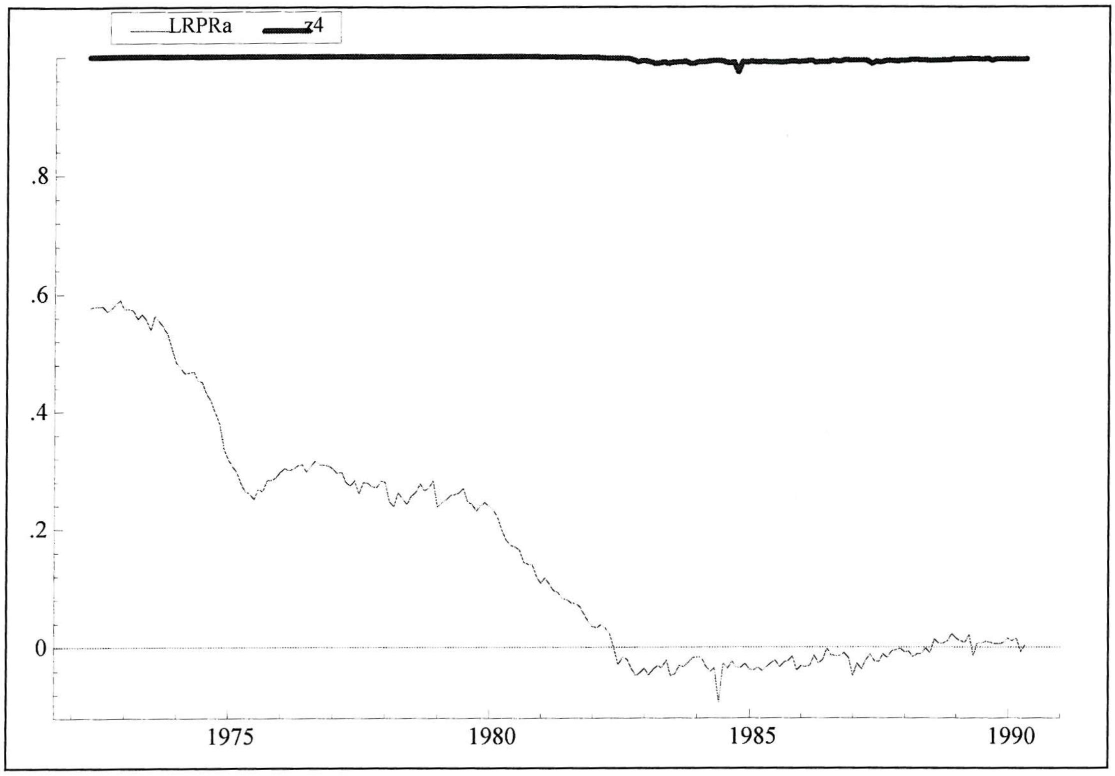
Parameter	Estimate	Standard Error	t-statistic
constant	4.29509	1194.82	.359477E-02
LBS7m(t-1)	.243160	.055221	4.40338
LBS7m(t-3)	.123033	.056980	2.15925
LBS7m(t-5)	.175006	.055715	3.14111
LBS7m(t-12)	.17478	.058105	3.00797
BETA*	-6.1168	1195.53	-.511634E-02
MU	-.31295	12.8891	-.024280
SIG	.060822	.209158	.290793
LRSPfr(t-2)	-.811386	.298226	-2.72071
LRSPfr(t-10)	.545988	.286960	1.90266
LRSPgr	-.42514	.206361	-2.06018
LRSPgr(t-6)	-.315319	.195136	-1.61589
LRSPpo(t-5)	.63696	.306426	2.0787
LRSPpo(t-10)	-1.26780	.48855	-2.59503
LRSPpo(t-11)	1.98150	.66941	2.96006
LRSPpo(t-12)	-1.31743	.45986	-2.86483
DLRPa(t-6)	2.90154	1.42217	2.04022
DLRPa(t-13)	4.68791	1.45619	3.21930
DLEXa	-1.84372	.652961	-2.82363
CI(t-1)	.655587	.194003	3.37925
EASTER	.182302	.053613	3.40037
jan	.012723	.050747	.250719
feb	.070911	.074719	.949044
mar	.476386	.079785	5.97090
apr	.701942	.135603	5.17644
may	.904316	.137324	6.58527
jun	1.16928	.149659	7.81296
jul	1.67327	.183945	9.09656
aug	.970004	.168216	5.76642
sep	.832133	.133087	6.25254
oct	.630806	.123488	5.10823
nov	-.167889	.080591	-2.08322
i1976p5	-.696141	.140752	-4.94589
i1978p11	.533306	.133571	3.99266
i1980p6	.707795	.132252	5.35188
i1989p5	.455248	.131619	3.45884

Note: * *beta* is the coefficient for the logistic transformation for the fourth lag of the income proxy (*LRPRa*).

The plot of $LRPRa_{i,t}$ on $Z4_{i,t}$ (*i.e.* the logistic transformation of $LRPRa_{i,t}$[100]) is given in Figure 8.11.

100 The logistic transformation of the fourth lag of the income proxy is given by the following: $z4_t = 1/1+\exp(-1*((LRPRa_{t-4} - \mu)/\sigma))$.

Figure 8. 11 *LBS7m*: Plot of $LRPRa_{i,t}$ on $Z4_{i,t}$



The transformed variable ($Z4_i$) seems to show a flat transformation at the beginning and later on it presents some movements. However, the scale of the transformed variable is so small that there are some difficulties in understanding the change in the series. Hence, a normalisation of the series is done as follows:

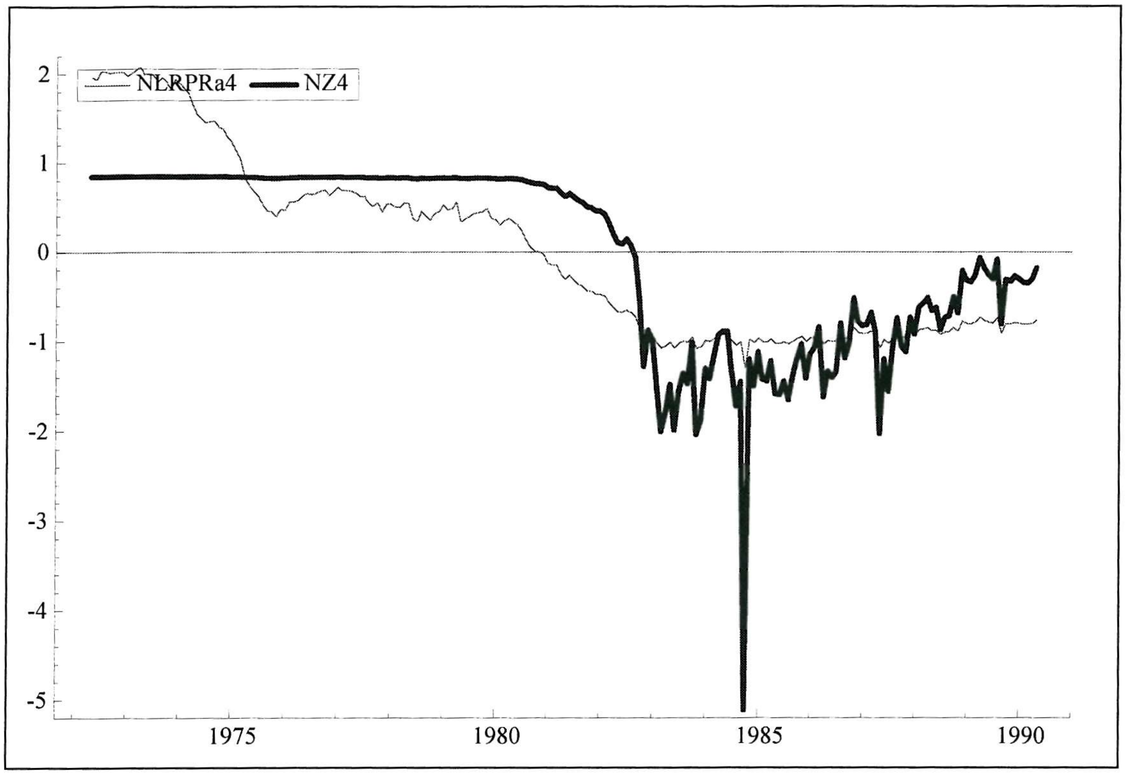
$$NLRPRa_{i,t} = (LRPRa_{i,t} - \mu) / \sigma$$

where μ, the mean of the income proxy, is equal to 0.12 and σ, the standard deviation, equals 0.21. The normalised logistic transformation of the income proxy, $NZ4_t$, is equal to:

$$NZ4_{i,t} = [(Z4_{i,t} - mu)/ sig]$$

where *mu* is the mean of *Z4* equal to 0.996 and *sig*, its standard deviation, is equal to 0.0044. The plot of the two normalised variables is given in Figure 8.12.

Figure 8. 12 *LBS7m*: Plot of $NLRPRa_{i,t}$ on $NZ4_{i,t}$



From Figure 8.12, the transformed logistic income proxy does not appear to be a single period impulse dummy. A better understanding can be obtained by calculating the proportion of the sum of the squares of this impulse dummy contained in its smallest observations (*i.e.* 1984:10=-5.11055 and 1983:11=-2.04211)[101]. In more detail, the smallest observation, 1984:10, counts for 12.09% of the total sum of the squares, and the observation for 1983:11 counts for 1.93%. Hence, it can be concluded that the logistic transformation of the income variable has been turned into an approximation of an impulse dummy variable. Again, as for the five aggregation case, estimates do not support an economic interpretation.

From Tables 8.28 and 8.29, some considerations on the other explanatory variables included in the final restricted model have to be reported. Note that the long run dynamics for the *LBS7m* equation have been obtained from PcGive, as the tests provided by this package can be considered as asympotically valid.

101 The sum of the squares for $NZ4_{i,t}$ equals 215.997 and is obtained by applying the following formula: $SS = \sum (x_i - \overline{x})^2$

Table 8. 29 ***LBS7m*****: Long Run Dynamics**

```
LBS7m =             +25.22              -32.72 z4             -1.102 LRSPfr
(SE)          (      9.941)       (      12.25)         (      0.842)
                    -2.493 LRSPgr       +0.1838 LRSPpo         +25.04 DLRPa
              (    0.6904)        (    0.7708)          (      9.291)
                    -6.213 DLEXa        +2.195 CI              +0.6159 Easter
              (     2.557)        (    0.4368)          (     0.2427)
                   +0.0318 Jan          +0.2137 Feb            +1.592 Mar
              (    0.1686)        (    0.2728)          (     0.5503)
                    +2.344 Apr          +3.025 May             +3.929 Jun
              (    0.8084)        (    0.9502)          (       1.11)
                    +5.639 Jul          +3.272 Aug             +2.807 Sep
              (      1.45)        (    0.8816)          (     0.7226)
                    +2.133 Oct          -0.572 Nov             -2.366 I1976P5
              (     0.561)        (    0.3177)          (     0.7187)
                    +1.811 I1978P11     +2.394 i1980p6         +1.557 I1989P5
              (    0.6303)        (    0.7586)          (     0.5856)

ECM = LBSm - 25.2225 + 32.7161*z4 + 1.10235*LRSPfr + 2.49332*LRSPgr -
0.183805*LRSPpo - 25.0435*DLRPa + 6.2125*DLEXa - 2.19533*CI -
0.615929*Easter - 0.0317971*Jan - 0.213657*Feb - 1.59163*Mar +
- 2.34392*Apr -3.02541*May - 3.92862*Jun -5.63939*Jul - 3.27245*Aug +
- 2.80724*Sep - 2.1326*Oct + 0.571954*Nov + 2.36629*I1976P5 +
- 1.81055*I1978P11 - 2.39428*i1980p6 - 1.55655*I1989P5;

WALD test Chi^2(23) = 129.17 [0.0000] **
```

As derived from Table 8.28, the Italian budget share of tourism involves a rather strong dependence on its own history. The substitute price for France presents a negative coefficient both in the short and long run, though, in the latter case, the *t*-value equals -1.31. A negative substitution elasticity holds in the Greek case and the long run coefficient is statistically significant at the 5% level. The coefficient for the substitution price for Portugal shows a positive sign. However, the long run coefficient does not turn out to be statistically significant. The dependent variable is positively influenced by the relative price growth; the long run coefficient has a positive sign and a *t*-value of +2.69. The exchange rate growth enters the equation with a negative sign both in the short and long run, with a statistically significant coefficient at the 5% level. Note also that the cointegrating vector (CI_{t-1}) presents a positive sign denoting that *LRPa* (the relative price) and *LEXa* (the weighted exchange rate) are having an opposite effect on the real budget share. Interestingly, the Easter dummy has a contribution in explaining the foreign demand of tourism in Italy expressed in terms of expenditure. The seasonal dummy coefficients are, in general, statistically significant at the 5% level with the highest increase of the budget share in the months of July and June, respectively.

8.5 SUMMARY

In this section, the main economic findings for the Italian tourism demand are summarised. In Table 8.30, the economic results from the tourist receipts model are reported. As a reminder, no conclusive findings have been found for the model when treating the dependent variable (*LREXP*) as stationary in the level. Hence, the international tourism expenditure are treated as integrated of order one (*DLREXP*). Income and price elasticities are reported considering both the short and long run behaviour. Note that the long run income and price elasticities are derived from the Johansen cointegration analysis.

Table 8. 30 *DLREXP*: Short and Long Run Elasticities for Italian Tourism Demand

Elasticities	Monthly Model *DLREXP* (221 obs. 1972:1-1990:5) (Tables 8.14-8.15)
INCOME (long run) (2)	5.60
INCOME growth (long run)	-
INCOME growth (short run)	-
REL.PRICE (long run) (2)	-4.03
REL.PRICE growth (long run)	5.07 (5.53)
REL.PRICE growth (short run)	3.07 (2.95)
EX.RATE (long run) (2)	7.69
EX.RATE growth (long run)	0.42 (1.67)
EX.RATE growth (short run)	2.11 (3.99)
SUB.PRICEfr(long run) (4)	-0.17 (-2.05)
SUB.PRICEfr(short run)	-1.11 (-2.02)
SUB.PRICEgr(long run) (4)	0.04 (0.71)
SUB.PRICEgr(short run)	0.54 (3.77)
SUB.PRICEpo(long run) (4)	0.32 (3.49)
SUB.PRICEpo(short run)	0.77 (3.55)
SUB.PRICEsp(long run) (4)	0.20 (3.54)
SUB.PRICEsp(short run)	0.49 (3.67)

Notes: (1) t -values are given in parenthesis.
(2) Note that the long run elasticities for the income proxy, relative price and exchange rate are from the Johansen cointegration analysis (see 8.7.1.4 and 8.8.2.1).
(3) Note that the short run elasticity corresponds to the first significant lag in the model (see Pindyck and Rubinfeld, p. 377, 1991).
(4) Permanent shifts in the I(0) variables and long run changes in the growth of tourist expenditure in Italy.

The equation for *DLREXP* presents correct signs with respect to economic theory for the growth exchange rate and substitution price for France. A positive substitution elasticity appears for the other competitors showing disagreement with economic theory. The results from the Johansen cointegration analysis have been included as they show coherence with economic theory. It is interesting to note that the absolute long run elasticities are relatively high. This finding is in line with the empirical results achieved by Kulendran and Witt (2001). In this study, the estimated elasticities from the Johansen cointegration analysis result are higher than the elasticities obtained by running least squares regression models.

In Table 8.31, the main results for the budget share equation, for both a five and seven aggregation countries, are presented.

Table 8. 31 *LBSm - LBS7m*: Short and Long Run Elasticities for Italian Tourism Demand

Elasticities	**LBSm (221 obs.) (Tables 8.22 - 8.23)**	**LBS7m (221 obs.) (Tables 8.28 - 8.29)**
INCOME VARIABLE	by logistic transformation: a shift of intercept	by logistic transformation: approximation impulse dummy
REL.PRICE growth (long run)	151.4 (1.86)	25.04 (2.69)
REL.PRICE (short run)	- 2.40 (-1.65)	2.90 (2.04)
EX.RATE growth (long run)	14.96 (1.17)	-6.21(-2.43)
EX.RATE growth(short run)	0.21 (0.27)	-
CI (long run)	6.20 (1.58)	2.19 (5.02)
CI (short run)	0.92 (1.57)	0.66 (3.38)
SUB.PRICE fr (long run)	-1.07 (-0.49)	-1.10 (-1.31)
SUB.PRICE fr (short run)	- 0.70 (-1.44)	- 0.81 (-2.72)
SUB.PRICE gr (long run)	-10.55 (-2.31)	-2.49 (-3.61)
SUB.PRICE gr (short run)	- 0.23 (-1.00)	-0.31 (-1.62)
SUB.PRICE po (long run)	0.30 (0.22)	0.18 (0.24)
SUB.PRICE po (short run)	1.42 (3.46)	0.64 (2.08)
SUB.PRICE sp (long run)	4.03 (1.62)	-
SUB.PRICE sp (short run)	0.66 (2.38)	-

Notes: (1) t -values are given in parenthesis.
(3) Note that the short run elasticity corresponds to the first significant lag in the model (see Pindyck and Rubinfeld, p. 377, 1991).

Overall, the model for the *LBS7m* equation seems more congruent with economic theory in terms of magnitude, sign and statistical significance of the coefficients.

Some comparisons with other empirical studies for Italy might be interesting. To this end, Syriopoulos' (1995) study is considered where annual tourist expenditure data are used in estimating a disaggregated dynamic model of demand for Italian tourism for the main source countries (France, Germany, Sweden, UK and USA). The sample period covers the years between 1960 and 1987 and the dependent variable is expressed in the first difference. In terms of long run elasticities, a positive income elasticity emerges for Italy. The highest value is for the UK (2.40) and the lowest value is for Germany (1.00). Negative price elasticities are found with the highest value for

Germany (1.61) and the lowest value for the USA (0.38). A negative substitution price elasticities in terms of competitors is found with the highest value for Sweden (3.30) and the lowest value for the USA (0.32). Note also that the income elasticity in the short run has turned out to be not statistically significant for the UK and the USA.

Song *et al.* (2000) estimate the UK demand for outbound tourism, expressed in terms of expenditure, to twelve destinations, amongst which is Italy. The sample period is from 1965 up to 1994. Table 8.32 summarises the results for Italy.

Table 8. 32 Song *et al.* (2000): Results for Italy

Elasticities for Italy	Short and Long Run (30 obs. 1965 - 1994) (Tables 8.7- 8.11-8.12)
INCOME (long run)	1.74 (5.93)
INCOME (short run)	1.74 (2.46)
REL.PRICE (long run)	- 1.03 (-4.17)
REL.PRICE (short run)	- 0.33 (-1.16)

Note: This table reproduces results from page 617.

In conclusion, differences emerge amongst empirical studies due to various causes such as sample periods, data frequency and heterogeneity in the data aggregation. In the present study, where monthly data have been employed, the use of a different data aggregation for the main origin countries of tourism to Italy has given rather notable differences.

8.6 CONCLUSIONS

In this chapter, a dynamic model for Italian tourism has been estimated. Monthly data have been used for the sample period between January 1972 and May 1990. A logarithmic specification has been adopted as suggested by the Box and Cox (1964) test. One of the aims of this chapter has been to identify which dependent variable best approximates Italian international tourism demand. For this purpose, two distinctive dependent variables have been chosen. The first variable is the tourist receipts collected from the Italian balance of payments. The second variable is a weighted average budget share for the main countries originating tourists' flows to Italy.

The first model has involved the estimation of the real tourist receipts. Seasonal and long run unit roots have been tested. From the ADF test, evidence has been found that the real tourist receipts could be treated as stationary in the level. On the other hand, the income proxy (*LRPRa*) has been found to be non-stationary in the level but in the first difference. Hence, a non-linear specification in the income proxy variable has been used in estimating Italian tourism. However, no clear conclusion has been reached by using a logistic transformation. Several minima for the residual sum of the squares have been found. Following Franses' test results and other empirical studies (see Song *et al.* 2000), an experiment has been done using expenditure in first differences. By applying a Johansen cointegration analysis, statistical evidence has been found that the real tourist receipts, the real income proxy, the relative price and exchange rate drift together in the long run. Interestingly, the long run relationship amongst these variables has given the expected sign in accordance with economic theory. In this way, it has been possible to estimate a dynamic model that includes both the short and long run information. Some unexpected results have been found in terms of substitution elasticity.

A further aim of this chapter has been to estimate Italian tourism demand by using a weighted budget share as the dependent variable. An initial investigation has involved an understanding of the appropriate level of aggregation. Seven main origin countries of tourism to Italy have been considered: France, Germany, Japan, Sweden, Switzerland, the UK and the USA. From graphical inspection and statistical analysis, the weighted budget share should have included all of these countries, except Japan and the USA that show differences in the magnitude, sign and statistical significance of the trend coefficient. As a next step, a twofold analysis has been carried out for five and seven countries aggregation, respectively. Interestingly, though using a different level of aggregation, similar results have been achieved in terms of the integration and cointegration status of the main economic variables of interest. In particular, the real industrial production index has been found stationary in the first difference. Hence, a non-linear specification has been undertaken in both of these cases. The results have shown that the logistic transformation of the real income proxy has given an ambiguous interpretation in the five countries' model, whereas the income proxy has been transformed into an approximation to an impulse dummy in the seven countries' model. Hence, in both cases the economic interpretation has vanished. Overall, the seven countries aggregation has shown a better specification. Considering the adjusted R-squared, the first model has been able to explain 96% of the variance of the dependent variable, whereas the second model has been able to explain almost 97% of the variance of the budget share. As already reported in Section 8.9, the seven countries' aggregation model has shown better results in terms of economic theory for the explanatory variables included in the final restricted model.

Finally, it is worth noting that the "Easter" dummy included in the three models has turned out statistically significant only in the budget share equation for the seven countries aggregation. This finding contrasts with the other two specifications and with Gonzàles and Moral (1995) study. In the latter, the coefficient of the Easter dummy has been found statistically significant only when using tourists' arrivals as the dependent variable.

CHAPTER 9.

GENERAL DISCUSSION

Aim of the Chapter:

To discuss the contributions of this thesis to the tourism literature, bearing in mind the initial propositions on which this work is based.

9.1 INTRODUCTION

This chapter contains a general discussion of the main contributions of this thesis to the existing tourism literature. The findings obtained from the empirical work are structured to reflect the main propositions and aims of this thesis as given in Chapter 1. The first proposition is concerned with the capability of more advanced econometric approaches to give insight into modelling and estimating the demand for tourism. As a second proposition, one gives grounds, both in terms of evolution of tourists' flows, seasonality and econometric findings, for separating domestic from international tourism. The third proposition investigates whether estimates of tourism demand at different time frequencies can be reconciled. Finally, the fourth and last investigation examines the capability of the estimated models to satisfy economic theory, and whether any conflict between theory and econometric findings emerges from this analysis.

9.2 ADVANCED ECONOMETRIC TOOLS AND TOURISM DEMAND

One of the questions which has been answered in this empirical work is: can advanced econometric approaches give more insight into modelling and understanding tourism demand? More advanced econometric tools, largely used in other applied econometric studies, have transferred to the analysis of tourism demand the possibility to investigate the characteristics and properties of the economic series under study. First, the integration status of a variable as well as any possible cointegration relationship between variables can be detected. Next, seasonal unit roots tests and mis-specification tests have led to the investigation of seasonal patterns; evidence has been found for the existence of seasonal structural breaks which have been carefully dummied out. From the Johansen cointegration analysis, evidence has been found for the existence of long run relationships amongst I(1) variables. For example, the relative price (Sassari/main origin countries, *LRP*) and the weighted exchange rate (*LER*) for the main source countries have been found to be cointegrated. The same conclusion has been reached for the relative price Italy/main origin countries (*LRPa*) and the exchange rate (*LEXa*), considering both seven and five countries' aggregations. These findings have led to the inclusion in the models both short and long run information. Moreover, the use of dynamic modelling has given more insight into the differences between short and long run income and price elasticities as discussed below.

Hence, by means of more advanced econometric approaches, it has been possible to discover properties and relationships between economic series which are still much neglected in the tourism literature. A more rigorous testing procedure, by making use of the LSE methodology, has revealed mis-specification problems, and other problems in the residuals (e.g. heteroscedasticity) that have hardly been considered in the tourism literature. The finding of this thesis shows how more advanced econometric approaches give the researcher help in modelling and estimating the demand for tourism.

9.3 DOMESTIC AND INTERNATIONAL DEMAND FOR TOURISM

In testing the second proposition, evidence for distinguishing domestic from international tourism has been given. The analysis holds for tourism in the north of Sardinia. Several differences can be pointed out from a graphical inspection alone as shown in Chapter 3. The first substantial difference is evident in the historic evolution of the flows of tourists, over the period under study. Moreover, a graphical analysis has shown different seasonal distributions for the two components. The seasonal distribution for foreign tourists shows overall smaller variations for the months between June and September, with the highest value in July. On the other hand, the seasonality of the domestic flows of tourists presents a more irregular distribution, with the highest peak in August. On this basis, the two components can be considered as different.

The next step is to find econometric evidence for modelling international and domestic demand for tourism separately. A list of the main differences that have been discovered follows.

1) Trading-day factors -

The first element that indicates the need to estimate two separate models relates to the "trading-day" factors. The dependent variable, that is the total number of foreign tourists in all registered accommodation, has to be adjusted in order to take into account the number of Saturdays in each month for the period under consideration. A comparison between the models both with and without the normalisation has been made. The model with the raw series as the dependent variable has encountered problems of non-normality and heteroscedasticity (at the 1% level), which have been corrected after imposing the normalisation.

A different normalisation is required for the number of domestic arrivals. Three different models have been estimated. The first model includes the raw data; the second includes the adjustment for the number of Saturdays and the third the normalisation for the number of Sundays in a month. From a first round estimation, the first and third models can be considered as superior. The residual sum of the squares have presented smaller values than in the case where the dependent variable has been corrected for the number of Saturdays in a month. However, from a second round estimation, the model that includes the raw data as the dependent variable has been chosen; it presents, in fact, the best specification in terms of diagnostic tests. Problems of mis-specification (*RESET* test at the 1% level) appear in the model with the normalisation for the number of Sundays. This finding suggests that domestic tourists are likely to arrive in Sassari Province any day of the week by either boat or plane. Foreign tourists, on the other hand, are much more constrained by the arrival day, as they are more likely to use charter flights that occur mainly at the weekends as far as the period of study is concerned[102].

2) Structural Breaks -

Relevant characteristics have been discovered in modelling the domestic

102 It is worth noting that in recent years there are more international flights available to the north of Sardinia during the week. It would be interesting in further work to find out if this circumstance affects the decisions of holiday-makers.

demand for tourism in the north of Sardinia. The existence of seasonal unit roots at some frequencies are detected by applying Franses' (1991) test. This could be thought a symptom of non-stationarity that might be due to structural breaks. A preliminary Chow test (1967) carried out on all the coefficients of the variables of interest, seasonal dummies, and using Andrews' (1993) critical values indicate that a structural break is evident in the seasonal pattern. Analogous results have been found both in the monthly and quarterly data models. Two structural changes occur. The first, between the first and second half of the Eighties, and the second occurs between the Eighties and Nineties. Changes in the tastes of domestic consumers, who seem to prefer to spend their holidays in the peak months (*i.e.* July, August and September), are observed. On the other hand, a general decrease of tourism demand has been detected in the off-peak months, *i.e.* April and October.

The previous results suggest once more the validity of modelling domestic and international demand for tourism separately. Note that no structural breaks have been detected in estimating the international demand for tourism in the north of Sardinia.

9.4 MONTHLY, QUARTERLY AND ANNUAL DATA

In this section, an account is given of the main similarities and differences in employing data points at different time frequencies. The first part is dedicated to the international demand for tourism in Sassari Province and the second subsection is dedicated to the domestic demand.

9.4.1 International Demand For Tourism At Different Time Frequencies

Another purpose of this thesis has been to estimate models at different time frequencies. This section gives an account of the main similarities and differences from the estimation of the international demand for tourism in Sassari Province. In order to have homogeneity in the results, the same economic series have been used in running the three models.

1) SIMILARITIES

a) Long run and seasonal unit roots.

The first relevant analogy appears in the characteristics and properties of the economic series under study when using monthly and quarterly data. The ADF test suggests that the adjusted series of foreign arrivals (*LA*), the nominal weighted average industrial production index (*LPR*) and the substitute price (*LSP*) are stationary in the level. On the other hand, the relative price (*LRP*) and weighted average exchange rate (*LER*) are I(1). No seasonal unit roots have been detected by applying Franses' test.

b) Cointegration analysis.

The I(1) variables, *LRP* and *LER*, have been used to test for possible cointegration adopting the Johansen analysis. Homogenous results have been achieved from the cointegration analysis for the three data frequencies. The restriction on the long run coefficient (say β), *i.e.* β=-1, has been accepted in all cases. This result suggests that there are no major differences in the inflation rates in the countries under consideration.

c) Long run elasticities.

The three models have given the same positive sign for the income proxy (*LPR*) which is in line with economic theory. An increase in the income causes a rise in tourism demand in the north of Sardinia. Homogeneous results have been obtained in terms of sign of the coefficient for the substitute price (*LSP*). However, it presents a positive sign that contrasts with economic theory. Note that the magnitude of the substitute price coefficient is slightly smaller in the monthly model than in the other two models. Common results have been achieved for the coefficient sign and magnitude of the cointegrating vector (*CI*). In particular, an increase in *CI*, determined either by an increase in the relative price (*LRP*) or a decrease in the exchange rate (*LER*), decreases the international demand for tourism in the long run. Notably, the exchange rate growth (*DLER*) has been found not statistically significant in any of the three models.

d) Short run elasticities.

No common results have been reached in terms of short run dynamics. It is worth noting that a static final model has been achieved in both the annual and quarterly case. The last finding suggests that the models converge rapidly to their long run equilibria.

- DIFFERENCES

a) Long run unit roots.

The first relevant difference appears when applying the ADF test to annual data. The integration status of the economic series under study differs considerably from that observed for the other time frequencies. For example, the dependent variable (*LA*) and the income proxy (*LPR*) have been found to be non-stationary in the level and difficulties have appeared in establishing the integration status of the exchange rate (*LER*). As already pointed out in Chapter 4, the use of a small number of observations (24 in total) causes a lack of power of the ADF test, leading to a relatively frequent failure to reject the null hypothesis of an unit root.

b) Cointegration analysis.

Given the differences encountered in the ADF test, the cointegration analysis for *LRP* and *LER*, when using annual data, has been based on the ADF test results obtained in the monthly and quarterly cases.

c) Long run elasticities.

Though the sign of the income proxy (*LPR*) has been found to be positive, its elasticity varies across the three models. The annual model shows a quite high coefficient (+2.34), which is in line with some empirical results for the Italian case (Clauser, 1991; Witt and Witt, 1992). The monthly data model presents an income elasticity above unity (+1.06) which denotes a not very strong preference of foreign tourists for Sardinian tourism as in the annual case. A coefficient less than one (+0.79) has been obtained in the quarterly model case; this result seems to suggest that the marketing policy in Sardinia needs improvement in order to attract a higher number of foreign tourists. On the other hand, the quarterly model gives a statistically negative price elasticity in terms of first difference of the relative price (*DLRP*). An increase in the growth of relative price is associated with a decrease in the foreign demand for tourism.

d) Short run elasticities.

The monthly model is the only one that gives insight into the differences between short and long run dynamics. It is interesting to note that this model shows a high income elasticity (+2.56) in the short run. Moreover, the coefficient for the substitute price presents a very high positive elasticity in the short run, in contrast again with current economic theory.

9.4.2 Domestic Demand For Tourism At Different Time Frequencies

This section is dedicated to the similarities and differences encountered in the estimation of the domestic demand for tourism in Sassari Province.

- SIMILARITIES

a) Long run and seasonal unit roots.

As for the international case, homogeneous results have been obtained from the ADF test when using monthly and quarterly economic series. The raw series of domestic arrivals (*LAR*), the Italian industrial production index (*LPR*) and the substitute price (*LSP*) are I(0). The relative price (*LREP)* is I(1). Moreover, no seasonal unit roots have been detected by applying Franses' test as shown in Chapter 5.

b) Long run elasticities.

The three data models have given the same results in terms of long run income elasticity. The income proxy (*LPR*) coefficient has presented the expected positive sign, in line with economic theory, and an elasticity less than one, as Malacarni (1991) found for the Italian case (that is +0.92). However, the magnitude of the income proxy coefficient is slightly smaller in the annual model than in the other two models. In each of the three models, homogeneous results have been obtained in terms of the sign of the coefficient for the substitute price (*LSP*). Though highly statistically significant, it presents an unusual positive sign. On the other hand, the relative price growth elasticity does not turn out to be statistically significant in the long run.

c) Short run elasticities.

The short run income elasticity is found to be positive and less than one. These findings are in line with economic theory. The short run elasticity is less than the long run elasticity which confirms other empirical studies (Syriopoulos, 1995; Song *et al.* 2000). This suggests that if income increases then Italian tourists adjust relatively slowly in the short run and substantially in the long run. Again, the monthly, quarterly and annual models show a positive sign for the substitute price coefficient. Moreover, the relative price does not turn out to be statistically significant in the short run.

- DIFFERENCES

a) Long run unit roots.

Using 24 data points, the annual series of number of domestic arrivals and the index of industrial production are found to be I(1) by applying the ADF test. These results diverge from the ADF test findings using the other time frequencies. Again, the lack of power of the ADF test is confirmed by the use of a small number of observations.

On balance, the models estimated with monthly and quarterly data have given the most homogenous results. Similar findings, in fact, have been achieved in terms of characteristics and properties of the series under study, in terms of short and long run income and price elasticities as well as in terms of magnitude of the coefficients.

9.5 ARE THE ECONOMIC PROPOSITIONS ALWAYS SATISFIED?

The next proposition explored in this thesis relates to the capability of the estimated models to satisfy economic theory. Are income elasticity, negativity and substitutability always satisfied in the estimated dynamic models?

Starting with the income elasticity, overall these results confirm those obtained in other empirical studies. In the majority of the present estimations, the long run income elasticity has been found to be positive which suggests tourism to be a normal good. In estimating the annual international tourism for the Province of Sassari, the income elasticity has been found to be greater than one. This result implies Sardinian tourism is viewed as a luxury good; this elasticity value confirms the results achieved by Malacarni (1991) for Italy. However, the income elasticity findings for the monthly and quarterly cases cannot be compared with other empirical studies; there are no other studies available either for Sardinia or for Italy. In the monthly data case, in fact, the long run income elasticity has been found to be just above unity (+1.06) and less than unity (+0.79) in the quarterly data case. This fact suggests that foreign tourists consider Sardinian tourism as a necessity good. The results in estimating domestic tourism confirm that Italians view Sassari tourism as a necessity good. This finding, that holds for each of the time frequencies used in estimation, is in line with the results obtained by Malacarni (1991) for the Italian case.

In general, the negativity hypothesis has been satisfied in the quarterly model used for the international demand for tourism. The relative price growth has presented a negative and statistically significant coefficient.

The major conflict between econometric results and *a priori* economic theory has been found in the positive sign for the nominal substitute price coefficient. This finding has encouraged a further investigation (see Chapter 6). An unequivocal answer has not been achieved by including the real substitution price in an aggregated manner. Further investigation of the individual components (*i.e.* the real substitute price for each of the main Sardinian competitors) has led to heterogeneous results both in the short and long run dynamics. On balance, the empirical findings seem to suggest that France and Portugal can be viewed as the main competitor countries for tourism in the north of Sardinia.

In estimating the Italian tourism demand in terms of expenditure, four real substitution prices have been included for the main Italian competitors in the Mediterannean area (France, Greece, Portugal and Spain). Again, heterogeneous results have been achieved in the three equations (that is tourist receipts equation and budget share equations for five and seven-origin countries' aggregations). On balance, from the empirical findings, France and Greece can be considered as the main competitors for Italian tourism. These findings seem to confirm Syriopoulos' (1995) study in which "the performance of the "substitute effective price" variable (effective price in a destination relative to competitor destinations) was not statistically satisfactory in all cases" (p.331).

9.6 ECONOMIC THEORY AND ECONOMETRIC ANALYSIS

The last proposition explored in this thesis relates to the existence of any conflict between economic theory and econometric results.

In Chapter 8, evidence has been found that the (log) real weighted industrial production (*LRPRa*) is stationary in the first difference. This finding suggests the real industrial production is drifting increasingly as the time goes on and it confirms Hansen's (1995) results when employing U.S. annual macroeconomic time series. In Hansen's article, there is evidence for the (log) real per capita GNP to be I(0) and the (log) real industrial production to be I(1) when the first difference of the unlogged unemployment rate is used as a covariate.

Given the previous result, a non-linear specification in the income proxy has been pursued in estimating the international demand of tourism to Italy, expressed both in terms of tourist receipts and budget share. In both cases, the logistic transformation of the income proxy has turned into a dummy variable. This is clear for the seven countries aggregation, where the transformed proxy looks like an impulse dummy. For the five countries aggregation visual inspection suggests either an intercept shift, or a bounded downward (stochastic) trend. Transformation into a dummy makes the economic interpretation vanish. The expectation is, in fact, that the logistic transformation of the income proxy is able to describe the economic relationship existing between the international tourist expenditure for Italy and the real income proxy for the main source countries. Economic theory suggests that below a certain income level the tourist expenditure is expected to be zero. As income starts rising, Italian tourist expenditure should increase until the point in which a maximum is reached. Hence, at a certain level of income tourist expenditure reaches its maximum. For further increases in the income, the level of tourist expenditure is expected to become stable around a certain value.

However, any conflict with economic theory arising from the current empirical work, far from invalidating the theory, can be thought of as the result of different causes. For example, proxy variables may not completely express economic reality; one is constrained by the availability of the data. One, for instance, is obliged to employ the industrial production index as a measure of individuals' income instead of using disposable income when dealing with models at a monthly frequency.

Arguably, in Chapters 4, 5 and 6 the use of the nominal industrial production index as the income proxy has, in general, given the positive expected sign in the long run. However, exceptions can be detected in the negative sign resulting in the coefficients of the income proxy whenever oscillations are involved. Furthermore, some discrepancies have been obtained in estimating the international demand for Sassari Province both in terms of long and short run dynamics (see Chapter 6). In this case, a negative income elasticity has been obtained. The last findings denote that the industrial production seems possibly to be detecting "over-time" effects overstated by the use of such a proxy. That is, while the trend in industrial production differs from that in disposable income, it may reflect short run movements in income: for example, more overtime working in periods of prosperity. In the short run, this may reduce the demand for leisure in general, and tourism in particular. It does seem that for the period in question industrial production is a reasonable proxy, not ideal, but satisfactory for disposable income.

Interestingly, some economic ground has been found in the equation for the Italian tourist receipts growth. As a reminder, Franses' unit roots test has suggested the Italian real tourist receipts is non-stationary in the level. This finding is also in line with other empirical studies existing in the tourism literature (e.g. Lanza and Urga, 1995). By applying the Johansen cointegration analysis, the possible existence of a long run relationship has been tested for the four I(1) variables (*i.e.* real tourist receipts (*LREXP*), real industrial production (*LRPRa*), relative price (*LRPa*) and weighted average exchange rate (*LEXa*)). From this test, evidence has been found that the previous variables drift together in the long run. This finding is consistent with the international demand for Italian tourism being linked to variables identified by economic theory in the long run. Note also that the long run coefficients show the expected sign: a positive income elasticity for the real industrial production; negativity holds for the relative price and a positive long run coefficient turns out correctly for the weighted exchange rate.

9.7 CONCLUSIONS

This chapter has been dedicated to the contributions of this thesis to the tourism literature. An account has been given in terms of the initial propositions of the investigation.

In Section 9.2, the contribution from analysing tourism demand with more advanced econometric techniques has been discussed. Franses (1991) and Hylleberg *et al.* (1990) seasonal unit root tests, ADF test, Johansen cointegration analysis, and a series of diagnostic tests following the LSE methodology have given new knowledge and understanding in estimating tourism demand for the Province of Sassari and for Italy. In Section 9.3, a discussion for separating international from domestic demand for tourism has been given. Differences and similarities in modelling and estimating tourism demand at different time frequencies are reported in Section 9.4. If economic theory is relevant, one would expect income elasticity, negativity and substitutability to be satisfied; Section 9.5 has been dedicated to the validation of these *a priori* propositions as derived from this empirical work. Some incongruities between economics and econometrics emerging from this thesis have been the objectives of the last section.

CHAPTER 10.

CONCLUSIONS

Sinclair and Stabler (1997), and Sinclair (1998), have emphasised that there is still a requirement to develop research into the demand for tourism. They also suggest that there are still many relevant aspects not explicitly taken into consideration by empirical studies in the tourism literature. Amongst others, there is a lack of discussion of functional forms. The inclusion of diagnostic tests in addition to the usual *t*-test, *F*-test, R-squared adjusted and D-W statistics is also desirable. Problems of heteroscedasticity are hardly considered. Short and long run elasticities are still much neglected. At this point, Sinclair and Stabler (1997) write "with few exceptions, notably Syriopoulos (1995), the majority of studies have assumed that demand depends on current income but not on past or expected future income" (p. 39). This thesis makes a new contribution in analysing and modelling the demand of tourism. The aim has been estimating the demand of tourism in Italy as a whole and a particular emphasis has been given to tourism in the Province of Sassari (Sardinia). Several contributions to knowledge have been made by this empirical work.

- Economic theory has been tested by employing more advanced econometric modelling, such as: seasonal and long run unit root testing, the Johansen cointegration analysis and the LSE general-to-specific methodology.
- Greater sample sizes have been used with a minimum of 24 observations, when using annual data, to a maximum of 288 observations when employing monthly data. This has established similarities and differences both in the pre-modelling and estimation phase.
- For each of the estimated models, statistical evidence has been given for adopting the logarithmic functional form. A full range of diagnostic tests has been included in the estimation procedure.
- Elasticities have been estimated. The derivation of short and long run dynamics has been possible by including actual and lagged dependent and explanatory variables.

In Chapter 2, a description of the main methodological steps adopted in the thesis has been given.

Chapters 3, 4, 5 and 6 have been dedicated to analysing, modelling and estimating tourism demand in the Province of Sassari. New information has been given on both international and domestic tourist flows. The analysis in Chapter 3 has highlighted major differences between domestic and international demand for tourism in the north of Sardinia. In terms of the historic evolution of tourist flows, domestic arrivals in all registered accommodation have shown an upward trend in the years between 1972 and 1995, with only a few exceptions. On the other hand, the trend for international demand appears to be more influenced by economic events during the three decades, such as higher standards of living, monetary restrictions and exchange rate depreciation. From an initial graphical inspection, major differences have also emerged in the seasonality of the two components. This finding has been confirmed by the pre-modelling and estimation analysis in Chapters 4 and 5. While the international seasonal pattern has turned out to be deterministic, a varying seasonal pattern has been found for domestic tourism demand. Evidence for the latter result has emerged from Franses' seasonal unit roots test as well as from problems of mis-specification in the models, which does not include seasonal parameter changes. A further distinction between the two components is in the so-called "trading-day" factor. From the econometric analysis, the need to normalise the total international tourists' arrivals for the number of Saturdays in a month has emerged. In Chapter 5, the domestic demand model has given the best specification without any correction for the dependent variable. Hence, this study has discovered characteristics and properties of the economic series which can be used as a basis for further forecasting exercises employing time-series and econometric models in this thesis. However, it is worth noting that forecasting has not been the objective of this thesis.

In Chapter 3, an account has been given of the main explanatory variables that could have an impact in explaining the demand of tourism in accordance with economic theory. The scope of Chapters 4, 5, 6 and 8 has been to test economic theory by adopting the Johansen cointegration analysis and the LSE methodology. Firstly, the integration status of the economic series of interest has been established by the ADF test. Evidence has been found for the relative price (*LRP*) and exchange rate (*LER*) to be random walks and thus I(1). Furthermore, a Johansen cointegration analysis has suggested that *LRP* and *LER* converge to a common equilibrium path in the long run, which satisfies the *a priori* theory. The same conclusion has been reached for the relative price and exchange rate within the Italian model. A log-linear form, suggested by the Box and Cox test, has been chosen for all the models estimated in this thesis. Short and long run income and price elasticities have been obtained via a dynamic analysis and inclusion of I(0) and I(1) variables. The main discrepancy between economic theory and econometric results is that of a positive sign for the coefficient of the nominal substitution price, that is Sassari/main competitors in the Mediterranean area (France, Greece, Portugal and Spain). Economic theory states that a decrease in tourism demand is expected when the price in a certain destination increases with respect to other competitors, *ceteris paribus*. The unexpected result has led to a further investigation in Chapter 6, where the monthly tourism demand is modelled with the inclusion of the exchange rate for the competitor countries. However, the results have confirmed a positive sign for the coefficient of the real substitution price defined in an aggregated manner. Therefore, an investigation has been undertaken in analysing the properties of the individual components of the substitution price, which has led to the inclusion of four distinctive real substitute prices for each of the competitors. The ADF test and cointegration tests have given higher homogeneity for the real substitute price of each of the individual competitors that has been included both in the foreign and domestic demand models. However, the final models obtained have not given a definitive answer in terms of short and long run substitute price elasticities. Statistical evidence has shown France and Portugal appear to be the main competitors for tourism in the north of Sardinia. Heterogeneous results have also been achieved in estimating international tourism in Italy (Chapter 8). On balance, France and Greece have appeared to be the main competitors for Italian tourism.

This empirical work has investigated whether homogenous results can be obtained by employing series at different frequencies. On balance, from Chapters 4 and 5, monthly and quarterly models have given more homogenous and robust results both in the pre-modelling and final estimated values when compared with annual data alone. However, it has also under-pinned the validity of employing monthly as well as annual time series. By using monthly variables, it has been possible to include in the model an "Easter" dummy which has entered with a satisfactory t-value. This finding has confirmed that the seasonal dummies are not able to capture the Easter holiday effect due to its mobility between March and May. On the other hand, the use of annual data has given the possibility to test the validity of using the Italian industrial production index (*LPR*) as a proxy for the personal disposable income (*LPDIN*). However, some caution should be sounded as just 10 observations have been used (that is 1983-1992) in doing such an analysis. The use of annual data has also given the possibility to include in the model two supply variables, that is: arrivals of boats and flights in the north of Sardinia. The estimation results, for the international demand of tourism, have shown a satisfactory determination in terms of statistical significance and signs. Both of the two components present a positive sign indicating that the higher the mean of transportation supplied the higher the number of foreign arrivals. These results have suggested a more careful analysis. The problem of simultaneity has been investigated via the Durbin-Wu-Hausman's test. From the analysis, it has appeared that the number of international flights to the north of Sardinia is to be considered as endogenous. However, it has not been in the scope of the study to fit a model for such a variable. Moreover, the total number of arrivals of boats has been found to be predetermined. One inference is that the total number of boats is picking up other determinants that affect tourism demand. In the domestic demand model (Chapter 5), the coefficient for the national arrivals of boats to the north of Sardinia has been found to be statistically significant. The issue of possible simultaneity has also been investigated by the Durbin-Wu-Hausman's test. The null hypothesis of no simultaneity has been rejected and the supply variable has to be treated as endogenous. One can conclude that a combination of different time frequency models is able to give more insight in understanding tourism demand as well as its components.

Chapter 8 has involved the study of Italian tourism as a whole. Two separate dependent variables have been constructed: the real tourism expenditure (*LREXP*), expressed in terms of real tourism receipts in Italy and an aggregated budget share for the main origin countries (*LBSm*). The sample period under study is from 1972:1 up to 1990:5. From June 1990 on, in fact, the data are collected with new currency regulations which are not comparable with the previous data. The new contribution is the use of monthly data, pre-modelling analysis and the use of the LSE methodology. The aim has been to establish which variable better represents tourism demand.

As far as *LREXP* is concerned, a model has been estimated by using a quadratic and a non-linear equation, respectively. Note, in fact, that the weighted average real industrial production for the main source countries (*LRPRa*), has been found to be I(1) from both the ADF test and Franses' unit roots test. A quadratic transformation was not found to be an adequate approximation to a logistic transformation. Hence, a non-linear model has been estimated. From this analysis, multiple maxima in the likelihood have been detected giving evidence that the logistic form used does not turn out to be satisfactory.

A further experiment has been carried out in estimating the tourist receipts growth (*DLREXP*) as the dependent variable, in accordance to Franses' test. From a Johansen cointegration analysis, the I(1) variables have been found to have a long run relationship as suggested by economic theory.

Several investigations have been undertaken for analysing and estimating the weighted budget share (*LSBm*). The first investigation has involved a deeper understanding on the appropriate degree of aggregation. From a graphical inspection and statistical testing, evidence has been found for a five countries aggregation (*i.e.* France, Germany, Sweden, Switzerland and UK, with the exclusion of Japan and USA). However, for either the five or seven countries aggregations, similar results have been obtained in terms of integration and cointegration status of the variables under study. In particular, the income proxy (*LRPRa*) has been found I(1) and a quadratic and non-linear model, respectively, have been run in both the cases. Notably, the results for the income proxy coefficient have failed to satisfy economic theory. In each of the three models, in fact, the logistic transformation of the income proxy has turned out to be interpretable as an approximation to a dummy variable. Overall, the models for the tourist receipts growth has given results more in line with economic theory.

From the empirical findings of the thesis, some important implications emerge for both private and public operators.

- The analysis of the "trading-day" factor has highlighted different choices in the timing of holiday trips by foreign and domestic tourists. International tourists seem to choose Saturdays as the departure day. Italians, on the other hand, seem to prefer Sundays rather than Saturdays as the departure day and, in fact, any day of the week. This empirical finding can assist the private sector in terms of using price discrimination for consumers choosing Sassari Province as the destination for their holidays.

- A graphical inspection of the possible existence of a capacity constraint in accommodation supplied in Sassari Province gives useful information for both private and public operators. There is no evidence that supply is a constraint on demand. Moreover, the peak month, August, shows an average rate of utilisation of almost 71%. June, July and September present an average rate of utilisation of around 40%. These findings suggest that the objective of the private and public sector does not seem to be that of increasing the accommodation capacity but using the existing capacity in off-peak months.
- Seasonality is one of the main features of tourism activity. Hence, the understanding of the seasonality is a necessity for both private and public operators. In this thesis, the use of monthly time series has given a deeper insight into the characteristics of the seasonal pattern for both the international and domestic demand components. There are reasons for believing that the public administration, at a regional level, should adopt promotion policies to encourage a de-seasonalisation process, in particular for the domestic demand. The objective of the local authorities should also be that of promoting Sassari tourism in off-peak months for Northern European clients. It is worth recalling that some econometric evidence has been found for the existence of a positive correlation between international demand and temperatures in Sassari Province.
- Econometric and graphical analysis has demonstrated the existence of structural changes in the domestic seasonal pattern. In both the 1980s and 1990s, evidence has been found of a further extension of the high season by Italians. This finding represents useful information for the public sector. The supply of public infrastructure and natural resources need to be examined in order to avoid possible negative externalities. In Sardinia, for example, one of the main problems is the lack of adequate water resources. Increasing consumption of water, in peak months (the driest months of the year) by the increased number of users, can create a negative impact for locals not only in the tourist season but also, more and more, in off-season months.

- The existence of an "Easter" effect has suggested the importance of "second holidays" for both foreigners and Italians. Hence, the "Easter" factor constitutes new information for the operators in tourism activity. The private sector can adopt price discrimination for tourist consumers, together with higher standard of quality of the goods and services supplied during "second holidays" periods. The local authorities should be aware of these effects in order to improve the quality of public goods and services supplied in off-peak months.
- Econometric evidence has also indicated that France and Portugal are substitute destinations for Sassari Province, and France and Greece for Italy as a whole. This finding is useful for both private and public operators, who shoud consider promoting Sassari and Italian tourism on a competitive base with these substitute countries.

APPENDIX A

In this appendix, the analysis for the series of the foreign arrivals of tourists without adjustment for the number of weekends per month is given.

The degree of augmentation appropriate when testing for unit roots is a subject for debate (see Hansen, 1995; Caporale and Pittis, 1997). Hansen (1995) suggests a Covariate Augmented Dickey Fuller (CADF) testing procedure for unit roots which produces more precise estimates than a conventional ADF test. Caporale and Pittis (1997) find that a number of macroeconomic time series are non-stationary using the ADF, but are indeed stationary when using a CADF test. They regard this as supporting the case of CADF tests, on the grounds that one is controlling the size of the test, and working for improvements in power as rejection of the null corresponds to acceptance of stationarity.

On the basis of these assumptions a previous seasonal unit roots test is performed on the unadjusted series *LAR*. The first notable change with respect to the adjusted series (*i.e. LA*) is the need for at least three impulse dummies (*i.e. i1994p4, i1992p7, i1995p9*[103]). Evidence for using such dummies is detected from the existence of serial correlation and from an inspection of the residuals. The equation (2.6.2) (Chapter 2) is fitted by OLS for the (log) foreign arrivals. Note that in this case μ_t also includes the three impulse dummies in order to correct for the presence of serial correlation. For *LAR* the presence of seasonal unit roots cannot be accepted at a general 1% level of significance[104] (Table A.1), both performing the *t*-tests of the separate π's (with exception of π_3, π_7 and π_{11}) and the *F*-test of the pairs of π's, as well as the joint *F*-test of $\pi_3 = .. = \pi_{12} = 0$.

103 *i1994p4*, *i1992p7* and *i1995p9* take the value 1 in *1994(4)*, in *1992(7)* and *1995(9)* respectively, and 0 elsewhere.

104 The critical values for the seasonal unit roots test are provided in Franses (1991) pp.161-165.

Table A. 1 Testing for Seasonal Unit Roots

t-statistics	Variable *LAR*	*t*-statistics	Variable *LAR*	*F*-statistics	*LAR*
π1	-3.456 *	π7	2.074	π3, π4	21.450 ***
π2	-4.660 ***	π8	-4.281 ***	π5, π6	29.441 ***
π3	0.661	π9	-2.818 *	π7, π8	16.504 ***
π4	-6.504 ***	π10	-7.050 ***	π9, π10	25.358 ***
π5	-6.785 ***	π11	1.543	π11, π12	12.359 ***
π6	-7.665 ***	π12	-4.834 ***	π3, ...,π12	25.349 ***

Notes: The one and the three asteristics indicate that the unit root null hypthesis is rejected at the 10% and the 1% level, respectively.

The main conclusion is that the arrivals of foreign tourists, *LAR*, can be considered as including a deterministic seasonal pattern and, furthermore, as the null hypothesis of a long run unit root is not accepted (*i.e.* H_0: $\pi_1 = 0$) this variable can be modelled as a stationary process, *i.e.* I(0). The latter result has also been confirmed by running the Augmented Dickey-Fuller (ADF) test where the null hypothesis of the presence of a unit root cannot be accepted at a 1% level of significance[105].

The explanatory variables reported in Section 4.2.1 remain unchanged for an unrestricted system with 13 lags[106] for the period from 1972:1 to 1995:12 that was run. The system includes a constant, 11 seasonal dummies, a trend, four impulse dummies (*i1974p12*, *i1985p3*, *i1991p11* and *i1995p3*) in order to avoid non-normality problems in the residuals. A final dummy has been constructed in order to take into account the "Easter holiday" effect.

The analysis is concentrated on the demand for tourism and the results for the unrestricted model using the unadjusted series are reported below.

---

105 When including the only constant, the value for the ADF statistic is -3.64 (for 9 lags). When including the constant and trend, the calculated value is -4.08 (for 10 lags). When including the constant and seasonals, the calculated value is -4.22 (for 2 lags). Finally, when including the constant, trend and seasonals, the calculated value is -7.99 (for 1 lag). Hence, the null hypothesis fails to be accepted in all the four cases as the calculated values are greater than the correspondent non standard critical values at the 5% level, as provided by PcGive 9.0.

106 The tests of reduction of the system order, provided by PcFiml 9.0, is the following:

system	T	p		log-likelihood	SC	HQ	AIC
12	274	400	OLS	5952.2022	-35.252	-38.410	-41.447
13	274	425	OLS	6015.5330	-35.203	-38.557	-40.909
System 13 lags > System 12 lags: F(25, 688) = 3.6507 [0.0000] **							

As one can notice the restriction for twelve lags cannot be accepted at the 1% level; note also that the HQ criterion suggests for at least a 13 lag system.

Table A. 2 Results from the Unrestricted Model for the Foreign Demand of Tourism

EQ(1) Modelling LAR by OLS (using For.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-2.0130	3.0043	-0.670	0.5037	0.0024
LAR_1	0.33170	0.060414	5.491	0.0000	0.1395
LAR_2	0.15265	0.056290	2.712	0.0073	0.0380
LAR_3	0.021708	0.057073	0.380	0.7041	0.0008
LAR_4	-0.011691	0.056543	-0.207	0.8364	0.0002
LAR_5	-0.010077	0.055629	-0.181	0.8565	0.0002
LAR_6	0.069763	0.055305	1.261	0.2087	0.0085
LAR_7	-0.075899	0.055620	-1.365	0.1740	0.0099
LAR_8	0.038961	0.055597	0.701	0.4843	0.0026
LAR_9	-0.12750	0.055885	-2.281	0.0237	0.0272
LAR_10	0.0031767	0.057085	0.056	0.9557	0.0000
LAR_11	0.16427	0.056877	2.888	0.0043	0.0429
LAR_12	-0.040836	0.057287	-0.713	0.4768	0.0027
LAR_13	0.045468	0.056111	0.810	0.4188	0.0035
LPR	0.14790	1.3449	0.110	0.9126	0.0001
LPR_1	0.73795	1.5535	0.475	0.6353	0.0012
LPR_2	0.014950	1.6305	0.009	0.9927	0.0000
LPR_3	1.2220	1.6067	0.761	0.4479	0.0031
LPR_4	0.97956	1.5558	0.630	0.5297	0.0021
LPR_5	-2.9090	1.5101	-1.926	0.0556	0.0196
LPR_6	1.6123	1.5123	1.066	0.2878	0.0061
LPR_7	-1.5896	1.4805	-1.074	0.2843	0.0062
LPR_8	-0.16279	1.4993	-0.109	0.9137	0.0001
LPR_9	-1.0700	1.5751	-0.679	0.4978	0.0025
LPR_10	1.7597	1.5415	1.142	0.2551	0.0070
LPR_11	-1.9868	1.5280	-1.300	0.1951	0.0090
LPR_12	1.0390	1.4840	0.700	0.4847	0.0026
LPR_13	0.78701	1.2095	0.651	0.5160	0.0023
LSP	3.0073	2.8053	1.072	0.2851	0.0061
LSP_1	-7.3839	3.7848	-1.951	0.0526	0.0201
LSP_2	6.0701	3.3247	1.826	0.0695	0.0176
LSP_3	3.3728	3.0946	1.090	0.2772	0.0063
LSP_4	-9.0145	3.1422	-2.869	0.0046	0.0424
LSP_5	5.2909	3.1321	1.689	0.0928	0.0151
LSP_6	-2.5112	3.1243	-0.804	0.4226	0.0035
LSP_7	0.28581	3.1390	0.091	0.9276	0.0000
LSP_8	-3.1074	3.1088	-1.000	0.3188	0.0053
LSP_9	3.3053	3.0432	1.086	0.2788	0.0063
LSP_10	1.2973	3.0692	0.423	0.6730	0.0010
LSP_11	6.7969	3.1930	2.129	0.0346	0.0238
LSP_12	-19.417	3.2830	-5.915	0.0000	0.1583
LSP_13	12.420	2.5893	4.797	0.0000	0.1101
DLRP	-4.8092	2.3228	-2.070	0.0398	0.0225
DLRP_1	0.21883	2.0720	0.106	0.9160	0.0001
DLRP_2	0.38943	1.6488	0.236	0.8135	0.0003
DLRP_3	-3.4200	1.6503	-2.072	0.0396	0.0226
DLRP_4	2.9342	1.6723	1.755	0.0810	0.0163
DLRP_5	1.0237	1.6899	0.606	0.5454	0.0020
DLRP_6	2.4033	1.6940	1.419	0.1577	0.0107
DLRP_7	-1.5430	1.6930	-0.911	0.3633	0.0044
DLRP_8	0.13397	1.6828	0.080	0.9366	0.0000
DLRP_9	4.3895	1.6401	2.676	0.0081	0.0371
DLRP_10	-0.61011	1.6641	-0.367	0.7143	0.0007
DLRP_11	-3.5089	2.1334	-1.645	0.1017	0.0143
DLRP_12	5.1743	2.0492	2.525	0.0124	0.0331
DLRP_13	-0.087405	1.4982	-0.058	0.9535	0.0000
DLER	-0.83466	0.73495	-1.136	0.2576	0.0069
DLER_1	1.2727	0.78261	1.626	0.1056	0.0140
DLER_2	-1.4080	0.79222	-1.777	0.0772	0.0167

DLER_3	-1.1212	0.77933	-1.439	0.1519	0.0110
DLER_4	0.55321	0.78382	0.706	0.4812	0.0027
DLER_5	-0.017506	0.80350	-0.022	0.9826	0.0000
DLER_6	-1.4858	0.80956	-1.835	0.0681	0.0178
DLER_7	0.60890	0.79795	0.763	0.4464	0.0031
DLER_8	-0.019835	0.82074	-0.024	0.9807	0.0000
DLER_9	0.75799	0.81628	0.929	0.3543	0.0046
DLER_10	-0.93636	0.82720	-1.132	0.2591	0.0068
DLER_11	0.060552	0.84866	0.071	0.9432	0.0000
DLER_12	-0.53218	0.83599	-0.637	0.5252	0.0022
DLER_13	0.48566	0.79089	0.614	0.5399	0.0020
CI_1	-0.28081	0.20939	-1.341	0.1815	0.0096
Trend	-6.6839e-005	0.00078538	-0.085	0.9323	0.0000
easter	0.52977	0.068423	7.743	0.0000	0.2437
i1974p12	1.3191	0.26408	4.995	0.0000	0.1183
i1985p3	0.76916	0.19102	4.027	0.0001	0.0802
i1991p11	-0.59135	0.18176	-3.254	0.0014	0.0538
i1995p3	0.61112	0.19294	3.167	0.0018	0.0512
JA	0.33984	0.13137	2.587	0.0104	0.0347
FE	0.60280	0.21298	2.830	0.0052	0.0413
MAR	0.96498	0.27712	3.482	0.0006	0.0612
AP	1.4988	0.33285	4.503	0.0000	0.0983
MAY	2.3952	0.37595	6.371	0.0000	0.1791
JUN	2.0778	0.41124	5.052	0.0000	0.1207
JUL	1.9575	0.40170	4.873	0.0000	0.1132
AU	1.5323	0.37038	4.137	0.0001	0.0843
SE	1.4875	0.31386	4.739	0.0000	0.1077
OT	0.50369	0.25399	1.983	0.0488	0.0207
NO	-0.49905	0.15060	-3.314	0.0011	0.0557

R^2 = 0.990628 F(87,186) = 225.99 [0.0000] sigma = 0.16343
DW = 2.01 RSS = 4.96794508 for 88 variables and 274 observations

AR 1- 7 F(7,179) =	1.8525 [0.0800]
ARCH 7 F(7,172) =	0.8725 [0.5295]
Normality Chi^2(2)=	4.9620 [0.0837]
Xi^2 F(158, 27) =	0.2605 [1.0000]
RESET F(1,185) =	0.1918 [0.6619]

After a general-to-specific simplification, one obtains a parsimonious model as reported in Table A.3.

Table A. 3 Results from the Parsimonious Model for the Foreign Demand of Tourism

EQ(2) Modelling LAR by OLS (using For.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	2.6891	0.60149	4.471	0.0000	0.0746
LAR_1	0.31795	0.046389	6.854	0.0000	0.1593
LAR_2	0.18469	0.045643	4.046	0.0001	0.0619
LAR_9	-0.099788	0.045825	-2.178	0.0304	0.0188
LAR_11	0.15070	0.046565	3.236	0.0014	0.0405
RLPR	1.8605	0.78308	2.376	0.0183	0.0223
LSP_8	-3.4268	1.5586	-2.199	0.0288	0.0191
LSP_9	5.9937	1.8141	3.304	0.0011	0.0422
LSP_12	-10.118	1.8409	-5.496	0.0000	0.1086
LSP_13	8.1853	1.5670	5.223	0.0000	0.0991
easter	0.53934	0.060266	8.949	0.0000	0.2441
i1974p12	1.5228	0.17205	8.851	0.0000	0.2401
i1985p3	0.84225	0.17260	4.880	0.0000	0.0876
i1991p11	-0.49827	0.16838	-2.959	0.0034	0.0341
i1995p3	0.55709	0.16959	3.285	0.0012	0.0417
JA	0.37596	0.094733	3.969	0.0001	0.0597
FE	0.57799	0.12728	4.541	0.0000	0.0768
MAR	0.94282	0.15373	6.133	0.0000	0.1317
AP	1.4639	0.19123	7.655	0.0000	0.1911
MAY	2.2977	0.19712	11.656	0.0000	0.3539
JUN	2.0427	0.21327	9.578	0.0000	0.2700
JUL	1.9601	0.20737	9.452	0.0000	0.2648
AU	1.5729	0.21579	7.289	0.0000	0.1764
SE	1.5116	0.17907	8.441	0.0000	0.2232
OT	0.55079	0.14361	3.835	0.0002	0.0560
NO	-0.43545	0.096889	-4.494	0.0000	0.0753

R^2 = 0.987511 F(25,248) = 784.39 [0.0000] sigma = 0.163387
DW = 2.09 RSS = 6.620467844 for 26 variables and 274 observations

AR 1- 7 F(7,241) = 1.0143 [0.4217]
ARCH 7 F(7,234) = 1.0594 [0.3906]
Normality Chi^2(2)= 9.9227 [0.0070] **
Xi^2 F(34,213) = 1.7956 [0.0070] **
RESET F(1,247) = 2.7585 [0.0980]

A couple of coefficients restrictions are attempted. *RLPR* is given by the difference between the coefficients of the third and fifth lag[107], respectively, of the index of industrial production. Such a restriction has been accepted at the 5% level from the joint *F*-test, where the $F(1,247)=1.66$ is smaller than the conventional critical value (3.84); also the SC criterion suggests this result (in the unrestricted model it is equal to -3.17653 and in the restricted model is -3.19033).

As one can notice the *May*, *Jun* and *Jul* dummies present coefficients almost of the same size and same sign. A restriction on these dummies has been attempted. The appropriate *F* statistic for the restriction is 5.53 with q=2 degrees of freedom in the numerator and N-K=248 in the denominator. This value is greater than the critical

[107] Note that the coefficient for the third lag equals +2.01 with a *t*-value of 2.54, whereas the fifth lag equals -1.77 with a *t*-value of -1.87. Hence, in the short run the income elasticity is positive and greater than one.

value of the F distribution at a 5% level, thus failing to accept the null hypothesis the restriction does not hold. Hence, the model reported in Table A.3 has been kept.

The results from Table A.3 are quite different to the ones obtained using *LA* (*i.e.* the adjusted series for the foreign arrivals) as the dependent variable (see Table 4.7). In terms of diagnostic statistics, the model where the unadjusted series of arrivals of tourists are considered (Table A.3) presents problems of non-normality and heteroscedasticity at the 1% level. However, the value of the R^2 denotes a good fit. Moreover, as the relevant F-statistic indicates, the overall significance of the regression is satisfactory.

In terms of significance of parameter coefficients, the cointegrating vector does not influence the foreign demand of tourism, as well as the first difference of the relative price (*DLRP*) and of the weighted exchange rate (*DLER*). The lags of foreign arrivals, as explanatory variables, present an average positive sign. The statistical significance of the first, second, ninth and eleventh lag of the dependent variable confirms that adjustment is not rapid. The income proxy enters the final equation with its difference (*RLPR*) which has a positive sign. The substitute price shows a negative elasticity in the short run and a positive elasticity in the long run (see Table A.4), which is in conflict with economic theory.

As assumed from the analysis of the unrestricted model, the "Easter" dummy (*EASTER*) has a particular importance in explaining the pattern of international tourism. Such a variable has been constructed giving the value one in the Easter month and zero otherwise[108]. The time trend does not have any particular power in explaining the demand for tourism.

The long run dynamics are reported in Table A.4. The long run multipliers and the standard errors are, in general, well-specified. The long run responses of the foreign demand for tourism to changes in the expanatory variables are also statistically significant.

One can conclude that this model presents mis-specification in the residuals with heteroscedasticity and non-normality problems.

---

108 For a more detailed description of the construction of the Easter dummy see Chapter 3.

Table A. 4 Solved Static Long Run Equation

```
LAR =            +6.023              +1.421 LSP             +1.208 easter
(SE)         (    0.312)        (    0.1793)           (    0.2583)
                 +3.411 i1974p12     +1.887 i1985p3         -1.116 i1991p11
             (   0.7249)        (    0.4916)           (    0.4291)
                 +1.248 i1995p3      +0.8421 JA             +1.295 FE
             (   0.4347)        (     0.274)           (     0.342)
                 +2.112 MAR          +3.279 AP              +5.147 MAY
             (    0.425)        (    0.5326)           (    0.7099)
                 +4.575 JUN           +4.39 JUL             +3.523 AU
             (     0.52)        (    0.4879)           (    0.4696)
                 +3.386 SE           +1.234 OT             -0.9754 NO
             (    0.465)        (     0.286)           (    0.3296)
                 +4.167 RLPR
             (    1.923)

ECM = LAR - 6.0233 - 1.42136*LSP - 1.20805*easter - 3.41083*i1974p12
     - 1.88655*i1985p3 + 1.11607*i1991p11 - 1.24782*i1995p3 - 0.842098*JA
     - 1.29464*FE - 2.1118*MAR - 3.27899*AP - 5.14658*MAY - 4.57549*JUN
     - 4.39043*JUL - 3.52307*AU - 3.38575*SE - 1.23372*OT + 0.975364*NO
     - 4.16739*RLPR;

WALD test Chi^2(18) = 267.3 [0.0000] **
```

In this appendix, the main purpose is to give a better explanation for the choice of the adjustment for the dependent variable in terms of the number of weekends in each month. So far, evidence has been given that the short run model for the adjusted data (see Table 4.6) produces better results in terms of diagnostic tests than the model without such adjustment (Table A.3). Given that *LN* is the (log) number of weekends (*i.e.* Saturdays) in each month, a further investigation has been done in order to check if using *LA=LAR-LN*, *i.e.* forcing the coefficient on *LN=1*, can be considered as an over-adjustment for the number of weekends. A preliminary investigation has involved the comparison between the restricted model as given in Table A.3 (say Model A) and an unrestricted model in which *LN* and its lags (*i.e. LN(-1), LN(-2), LN(-9), LN(-11)*) were included (say Model A^*), in order to test for their joint significance. The RSS from Model A shows a value of 6.620467844 for 26 variables and 274 observations, whereas the RSS for Model A^* shows a value of 6.45003987 for 31 variables and 274 observations. The appropriate F statistic is 1.13 with q=5 degrees of freedom in the numerator and N-K=243 in the denominator. This value is smaller than the critical value of the F distribution at a 5% level (*i.e.* 2.21), thus failing to reject the null hypothesis the restriction holds.

A second investigation has involved the comparison between the final parsimonious model as given in Table 4.6 (Model B) and an unrestricted model in which *LN* and its lags *i.e. LN(-1), LN(-2), LN(-3)* and *LN(-11)* are included (say Model

B*), in order to test for their joint significance. The RSS from Model B is 9.639300629 for 25 variables and 274 observations, whereas the RSS from the unrestricted model (Model B*) is 7.300511685 for 30 variables and 274 observations. The joint significance of the (log) number of weekends is tested by performing an F-test. The value of the F statistic equals 15.62 with q=5 degrees of freedom in the numerator and N-K=244 in the denominator, thus the restricted model cannot be accepted at the 5% level.

Hence, the main conclusion from this analysis is that there is no statistical evidence for the adjustment for the number of weekends. In fact, from the previous analysis one infers that Model A*, where the (log) number of weekends are included, does not reject Model A; the opposite conclusion is reached from the comparison between Model B and Model B*, where the (log) number of weekends are included. However, one chooses Model B in which the adjustment for the number of weekends has been involved. Firstly, Model B can be considered better than Model A in that it does not present signs of non-normality as well as heteroscedasticity (at least at the 1% level). On the other hand, Model B* is difficult to interpret. Therefore, Model B, as reported in Tables 4.6 and 4.7 with the White correction for heteroscedasticity at the 5% level, has been chosen. It is important to assess that it might be possible to obtain a better model specification by running a model in which the motivation for tourism is taken into account. It might be the case that a distinction between the number of "business holiday trips" and "holiday trips" could determine better results in terms of both statistical and economic performance. However, as already pointed out in Chapter 3, such a distinction is not available from the official statistics as far as the period under study is concerned.

APPENDIX B

This Appendix, provides a detailed description of the formulas for each of the explanatory variables used.

A) Industrial Production Index (*PR*).

This variable has been used as a proxy of the income index for which monthly data are not available. Thus, the *PR* variable is a weighted average of the industrial production index (1990=100) for the origin countries, that is:

$$PR_t = \frac{\sum_{i=1}^{i=7} w_{i,t} * PR_{i,t}}{\sum_{i=1}^{i=7} w_{i,t}} \qquad \text{(B.1)}$$

where:

i = Belgium, France, Germany, Sweden, Switzerland, United Kingdom and United States.

$PR_{i,t}$ = industrial production index (1990=100) seasonally adjusted, in country i in month t (Source: IFS).

$w_{i,t}$ = takes into account the number of tourists coming from the origin country i in year t (Source: ISTAT), and is given by the following formula:

$$w_{i,t} = \frac{AR_{i,t}}{\sum_{i=1}^{i=7} AR_{i,t}} \qquad \text{(B.2)}$$

Note that the weights vary over time, to reflect the changing importance of different constituents of the average being calculated. Moreover, the weights are allowed to change annually rather than monthly. Annual weights may be thought to be more stable than the monthly weights. Firstly, one could argue that holiday plans are made on an annual basis and, secondly, more frequent observations might just reflect different seasonal patterns.

B) Relative Price (*RP*).

The relative price represents the price of north of Sardinia tourism to the set of clients countries (i) as listed above. Such a variable can be expressed by the following formula:

$$RP_t = \frac{CPI_{ss,t}}{CPI_{o,t}} \qquad \text{(B.3)}$$

where:

$CPI_{ss,t}$ = monthly consumer price index (1990=100) in Sassari (Source: ISTAT)

$CPI_{o,t}$ = weighted average consumer price index, calculated as follows:

$$CPI_{o,t} = \frac{\sum_{i=1}^{i=7} w_{i,t} * CPI_{i,t}}{\sum_{i=1}^{i=7} w_{i,t}} \qquad \text{(B.4)}$$

where

$CPI_{i,t}$ = monthly consumer price index (1990=100) in country i and month t (Source: IFS).

Note that the weights ($w_{i,t}$) are defined as in (B.2).

C) Exchange Rate (*ER*).

The weighted exchange rate with respect to the main origin countries, i, can be expressed by the following formula:

$$ER_t = \frac{\sum_{i=1}^{i=7} w_{i,t} * ER_{i,t}}{\sum_{i=1}^{i=7} w_{i,t}} \qquad \text{(B.5)}$$

where:

$ER_{i,t}$ = nominal exchange rate, in country i in month t (Source: *Banca d'Italia*).

$w_{i,t}$ = as in formula (B.2.).

D) Substitute Price (*SP*).

The substitute price represents the price of north of Sardinia tourism to the set of competitor countries in the Mediterranean area. This variable can be expressed by the following formula:

$$SP_t = \frac{CPI_{ss,t}}{CPI_{c,t}} \qquad \text{(B.6)}$$

where:

$CPI_{ss,t}$ = monthly consumer price index (1990=100) in Sassari (Source: ISTAT).

$CPI_{c,t}$ = weighted average consumer price index for the competitor countries, calculated as follows:

$$CPI_{c,t} = \frac{\sum_{i=1}^{i=4} \alpha_{i,t} * CPI_{i,t}}{\sum_{i=1}^{i=4} \alpha_{i,t}} \qquad (B.7)$$

where:

i = France, Greece, Portugal and Spain.

$CPI_{i,t}$ = monthly consumer price index (1990=100) in country i and month t (Source: IFS).

$\alpha_{i,t}$ = weights are defined as

$$\alpha_{i,t} = \frac{AR_{i,t}}{\sum_{i=1}^{i=4} AR_{i,t}} \qquad (B.8)$$

where $AR_{i,t}$ are the number of tourists' arrivals in the each of the competitor country, i, from the following origin countries: Belgium, Germany, Sweden, Switzerland, United Kingdom and United States (Source: OECD - Tourism Policy and International Tourism in OECD Member Countries; World Tourism Organisation). These weights are allowed to vary annually.

APPENDIX C

As pointed out in Section 2.7.1 (Chapter 2), one of the approaches used for testing cointegration in single equations is the (Augmented) Dickey-Fuller test.

In this specific case, one tests the null hypothesis of no-cointegration for the (log) relative price (*LRP*) and the (log) exchange rate (*LER*) which are found to be I(1). The first step consists in estimating the static model (2.7.1.1), where a constant is included, by OLS that gives the following results:

$$LRP = -8.8112 + 1.2403\, LER + \hat{u}_1 \qquad R^2 = 0.89 \qquad DW = 0.05$$

and

$$LER = 7.0718 + 0.72048\, LRP + \hat{u}_2 \qquad R^2 = 0.89 \qquad DW = 0.05$$

Using (2.7.1.2) in Chapter 2, where neither the constant nor the time trend are included, the saved residuals $\hat{u}_1$ and $\hat{u}_2$, that can be interpreted as the deviations of the generic y_t from the long run path, are tested for a unit root under the null hypothesis of no-cointegration.

The number of lags for the ADF test is set to the first statistically significant lag, testing downward and upon white residuals. Starting with 13 lags in the ADF equation (2.7.1.2), since monthly data are employed, a nine lag model is chosen. The *t*-value for the coefficient ρ, as from equation (2.7.1.2) equals -1.94. The MacKinnon's critical value equals -3.36[109] at the 5% that is greater, in absolute value, than the calculated value (-1.94). As a conclusion the null hypothesis cannot be rejected.

The same finding has been obtained, when in the static model a constant and a trend are included. In this case, the critical value is -3.81[110] greater, in absolute value, than the *t*-value for ρ in equation (2.7.1.2), that is greater than -1.63. Once more, no evidence appears of the existence of cointegration between the relative price and weighted exchange rate.

A similar conclusion is reached when regressing *LER* on *LRP*. When including only the constant, a final nine lag model is run. The MacKinnon's critical value is

---

[109] The estimated p =5% critical value for T=288 observations is the following: C(p) = -3.3377 + (-5.967/288) - (8.98/(288)^2).

[110] The estimated p =5% critical value for T=288 observations is the following: C(p) = -3.7809 + (-9.421/288)-(15.06/(288)^2).

equal to -3.36[111] at the 5% level. Thus, this critical value, in absolute value, is greater than the t-value for ρ, that is -2.13. Once more, there is no evidence for *LER* and *LRP* to be cointegrated.

Including a constant and a trend in the static model the results are the following. Starting with 13 lags, the ADF model can be reduced to a nine lag model. The critical value determined from MacKinnon's parameters equals -3.81[112] greater, in absolute value, than the t-value for ρ, -2.32. Thus, one fails to reject the null hypothesis of no-cointegration.

111 The estimated p =5% critical value for T=288 observations is the following: C(p) = -3.3377 + (-5.967/288) - (8.98/(288)^2).

112 The estimated p =5% critical value for T=288 observations is the following: C(p) = -3.7809 + (-9.421/288)-(15.06/(288)^2).

APPENDIX D

Table D. 1 Results from the Unrestricted System for the Foreign Demand of Tourism

EQ(1) Modelling LA by OLS (using For.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-6.2097	3.5340	-1.757	0.0805	0.0163
LA_1	0.14588	0.063487	2.298	0.0227	0.0276
LA_2	0.17748	0.058208	3.049	0.0026	0.0476
LA_3	0.14979	0.058788	2.548	0.0116	0.0337
LA_4	-0.072701	0.059743	-1.217	0.2252	0.0079
LA_5	-0.080373	0.057826	-1.390	0.1662	0.0103
LA_6	0.15955	0.058750	2.716	0.0072	0.0381
LA_7	-0.068016	0.058674	-1.159	0.2478	0.0072
LA_8	-0.020914	0.058474	-0.358	0.7210	0.0007
LA_9	-0.085676	0.058126	-1.474	0.1422	0.0115
LA_10	-0.0091134	0.059838	-0.152	0.8791	0.0001
LA_11	0.14456	0.059799	2.417	0.0166	0.0305
LA_12	0.048631	0.059715	0.814	0.4165	0.0036
LA_13	-0.021793	0.059147	-0.368	0.7130	0.0007
LPR	2.4572	1.5909	1.545	0.1242	0.0127
LPR_1	-2.5097	1.8339	-1.369	0.1728	0.0100
LPR_2	0.18422	1.9262	0.096	0.9239	0.0000
LPR_3	4.2027	1.8991	2.213	0.0281	0.0257
LPR_4	-0.40972	1.8547	-0.221	0.8254	0.0003
LPR_5	-4.1738	1.7586	-2.373	0.0186	0.0294
LPR_6	3.7797	1.7611	2.146	0.0332	0.0242
LPR_7	-2.9978	1.7124	-1.751	0.0817	0.0162
LPR_8	-0.70805	1.7504	-0.405	0.6863	0.0009
LPR_9	-0.25425	1.8377	-0.138	0.8901	0.0001
LPR_10	-0.23661	1.8009	-0.131	0.8956	0.0001
LPR_11	-0.33207	1.7830	-0.186	0.8525	0.0002
LPR_12	2.9046	1.7026	1.706	0.0897	0.0154
LPR_13	-0.85059	1.4072	-0.604	0.5463	0.0020
LSP	5.0420	3.2934	1.531	0.1275	0.0124
LSP_1	-12.463	4.4673	-2.790	0.0058	0.0402
LSP_2	7.9403	3.9317	2.020	0.0449	0.0215
LSP_3	6.1001	3.6334	1.679	0.0949	0.0149
LSP_4	-8.8141	3.6653	-2.405	0.0172	0.0302
LSP_5	3.0815	3.6208	0.851	0.3958	0.0039
LSP_6	-2.0694	3.6052	-0.574	0.5667	0.0018
LSP_7	0.59679	3.6274	0.165	0.8695	0.0001
LSP_8	-2.5955	3.5889	-0.723	0.4705	0.0028
LSP_9	2.4527	3.5093	0.699	0.4855	0.0026
LSP_10	0.053402	3.5428	0.015	0.9880	0.0000
LSP_11	7.9198	3.6812	2.151	0.0327	0.0243
LSP_12	-17.081	3.8163	-4.476	0.0000	0.0972
LSP_13	10.516	3.0079	3.496	0.0006	0.0617
DLRP	-5.5295	2.7221	-2.031	0.0436	0.0217
DLRP_1	0.94715	2.4145	0.392	0.6953	0.0008
DLRP_2	0.47026	1.9413	0.242	0.8089	0.0003
DLRP_3	-2.7296	1.9199	-1.422	0.1568	0.0107
DLRP_4	0.87495	1.9449	0.450	0.6533	0.0011
DLRP_5	3.1008	1.9737	1.571	0.1179	0.0131
DLRP_6	2.6501	1.9766	1.341	0.1817	0.0096
DLRP_7	-0.11063	1.9707	-0.056	0.9553	0.0000
DLRP_8	-1.1741	1.9644	-0.598	0.5508	0.0019
DLRP_9	4.1033	1.9116	2.147	0.0331	0.0242
DLRP_10	2.6200	1.9398	1.351	0.1784	0.0097

DLRP_11	-4.1293	2.4721	-1.670	0.0965	0.0148
DLRP_12	3.1050	2.4195	1.283	0.2010	0.0088
DLRP_13	-0.16339	1.7594	-0.093	0.9261	0.0000
DLER	-0.44192	0.83027	-0.532	0.5952	0.0015
DLER_1	1.5721	0.90972	1.728	0.0856	0.0158
DLER_2	-1.0007	0.92474	-1.082	0.2806	0.0063
DLER_3	-1.5677	0.90996	-1.723	0.0866	0.0157
DLER_4	-0.39249	0.90969	-0.431	0.6666	0.0010
DLER_5	1.0067	0.93875	1.072	0.2850	0.0061
DLER_6	-1.9645	0.94703	-2.074	0.0394	0.0226
DLER_7	0.29166	0.93741	0.311	0.7560	0.0005
DLER_8	0.35391	0.96449	0.367	0.7141	0.0007
DLER_9	0.00013191	0.95847	0.000	0.9999	0.0000
DLER_10	0.23640	0.97028	0.244	0.8078	0.0003
DLER_11	-1.6063	0.98935	-1.624	0.1062	0.0140
DLER_12	0.19634	0.98173	0.200	0.8417	0.0002
DLER_13	-1.0005	0.92131	-1.086	0.2789	0.0063
CI_1	-0.53921	0.24319	-2.217	0.0278	0.0258
Trend	-0.00044102	0.00092104	-0.479	0.6326	0.0012
easter	0.45700	0.078908	5.792	0.0000	0.1528
JA	0.46404	0.14856	3.124	0.0021	0.0498
FE	0.86825	0.24345	3.566	0.0005	0.0640
MAR	1.1219	0.32163	3.488	0.0006	0.0614
AP	1.7367	0.38185	4.548	0.0000	0.1001
MAY	2.8395	0.43154	6.580	0.0000	0.1888
JUN	2.6435	0.47455	5.571	0.0000	0.1430
JUL	2.3952	0.46691	5.130	0.0000	0.1239
AU	2.0559	0.42734	4.811	0.0000	0.1107
SE	2.0174	0.36207	5.572	0.0000	0.1430
OT	1.0048	0.28968	3.469	0.0006	0.0608
NO	-0.14077	0.16837	-0.836	0.4042	0.0037
i1974p12	1.6152	0.31044	5.203	0.0000	0.1270
i1979p3	-0.71270	0.22204	-3.210	0.0016	0.0525
i1985p3	0.75918	0.22515	3.372	0.0009	0.0576
i1991p11	-0.67610	0.21253	-3.181	0.0017	0.0516

```
R^2 = 0.987197  F(87,186) = 164.85 [0.0000]  sigma = 0.190905
DW = 1.90 RSS = 6.778747518 for 88 variables and 274 observations

AR 1- 7 F( 7,179) =     0.6192 [0.7397]
ARCH 7  F( 7,172) =     0.0872 [0.9989]
Normality Chi^2(2)=     5.4792 [0.0646]
Xi^2     F(158, 27) =   0.2156 [1.0000]
RESET    F( 1,185) =    0.0644 [0.8000]
```

Table D. 2 Results from the Parsimonious Model for the Foreign Demand of Tourism

EQ(2) Modelling LA by OLS (using For.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	-2.9643	1.6942	-1.750	0.0814	0.0123
LA_1	0.12242	0.052346	2.339	0.0202	0.0218
LA_2	0.11169	0.050861	2.196	0.0290	0.0192
LA_3	0.14233	0.050973	2.792	0.0056	0.0307
LA_11	0.10552	0.051057	2.067	0.0398	0.0171
LPR_3	2.7919	0.65650	4.253	0.0000	0.0685
LPR_7	-2.1930	0.63541	-3.451	0.0007	0.0462
LSP_1	-4.4956	1.8955	-2.372	0.0185	0.0224
LSP_2	5.4563	1.9429	2.808	0.0054	0.0311
LSP_11	4.6694	1.9418	2.405	0.0169	0.0230
LSP_12	-4.9055	1.9021	-2.579	0.0105	0.0263
CI_1	-0.33050	0.15803	-2.091	0.0375	0.0175
easter	0.42555	0.072204	5.894	0.0000	0.1237
i1974p12	1.4990	0.21065	7.116	0.0000	0.1707
i1979p3	-0.57881	0.20537	-2.818	0.0052	0.0313
i1985p3	0.67640	0.20816	3.249	0.0013	0.0412
i1991p11	-0.60355	0.20454	-2.951	0.0035	0.0342
JA	0.27667	0.11173	2.476	0.0140	0.0243
FE	0.66665	0.17009	3.919	0.0001	0.0588
MAR	1.0998	0.20344	5.406	0.0000	0.1062
AP	1.7745	0.24598	7.214	0.0000	0.1746
MAY	2.8104	0.25330	11.095	0.0000	0.3335
JUN	2.7937	0.26841	10.408	0.0000	0.3057
JUL	2.7677	0.25952	10.664	0.0000	0.3162
AU	2.5117	0.25219	9.960	0.0000	0.2874
SE	2.3538	0.20971	11.224	0.0000	0.3387
OT	1.1862	0.17007	6.975	0.0000	0.1651
NO	0.010627	0.11189	0.095	0.9244	0.0000

R^2 = 0.981956 F(27,246) = 495.84 [0.0000] sigma = 0.197067
DW = 1.91 RSS = 9.553523239 for 28 variables and 274 observations

AR 1- 7	F(7,239) =	1.0187	[0.4186]
ARCH 7	F(7,232) =	0.41841	[0.8903]
Normality	Chi^2(2)=	3.7361	[0.1544]
Xi^2	F(38,207) =	1.3836	[0.0803]
RESET	F(1,245) =	1.6856	[0.1954]

APPENDIX E

Table E. 1 *LAR* - Program for Performing Chow Structural Break Test

```
1 options crt;
2 freq n;
3 smpl 1 288;
4 ?1972.1-1995.12;
4 load (file='ital2.csv')lar lam las lpr dlp lsp e lw jan feb mar apr
  may jun jul aug sep oct nov dec i1974p11 i1987p3 i1992p3 i1993p3 dlrep
  lrep trend;
5 ?1973.3-1995.12;
5 smpl 15 288;
6 OLSQ lar c lar(-1)-lar(-13) lpr(-1)-lpr(-13) dlrep(-1)-dlrep(-13)
  lsp(-1)-lsp(-13) e lw i1974p11 i1992p3 i1993p3 jan feb mar apr may jun jul
  aug sep oct nov;
7 RRSS=@SSR;
8 ?1973=year3,1978=year7;
8 do j=7 to 23;
9 set i= 12*j;
10 smpl 15 i;
11 GENR C2=0;
12 GENR JAN2=0;
13 GENR FEB2=0;
14 GENR MAR2=0;
15 GENR APR2=0;
16 GENR MAY2=0;
17 GENR JUN2=0;
18 GENR JUL2=0;
19 GENR AUG2=0;
20 GENR SEP2=0;
21 GENR OCT2=0;
22 GENR NOV2=0;
23 GENR LAR1=0;
24 GENR LAR2=0;
25 GENR LAR3=0;
26 GENR LAR4=0;
27 GENR LAR5=0;
28 GENR LAR6=0;
29 GENR LAR7=0;
30 GENR LAR8=0;
31 GENR LAR9=0;
32 GENR LAR10=0;
33 GENR LAR11=0;
34 GENR LAR12=0;
35 GENR LAR13=0;
36 GENR LPR1=0;
37 GENR LPR2=0;
38 GENR LPR3=0;
39 GENR LPR4=0;
40 GENR LPR5=0;
41 GENR LPR6=0;
42 GENR LPR7=0;
43 GENR LPR8=0;
44 GENR LPR9=0;
45 GENR LPR10=0;
46 GENR LPR11=0;
```

```
GENR LPR12=0;
GENR LPR13=0;
GENR LSP1=0;
GENR LSP2=0;
GENR LSP3=0;
GENR LSP4=0;
GENR LSP5=0;
GENR LSP6=0;
GENR LSP7=0;
GENR LSP8=0;
GENR LSP9=0;
GENR LSP10=0;
GENR LSP11=0;
GENR LSP12=0;
GENR LSP13=0;
GENR DLREP1=0;
GENR DLREP2=0;
GENR DLREP3=0;
GENR DLREP4=0;
GENR DLREP5=0;
GENR DLREP6=0;
GENR DLREP7=0;
GENR DLREP8=0;
GENR DLREP9=0;
GENR DLREP10=0;
GENR DLREP11=0;
GENR DLREP12=0;
GENR DLREP13=0;
SET h=i+1;
smpl h 288;
GENR c2=c;
GENR jan2=jan;
GENR feb2=feb;
GENR mar2=mar;
GENR apr2=apr;
GENR may2=may;
GENR jun2=jun;
GENR jul2=jul;
GENR aug2=aug;
GENR sep2=sep;
GENR oct2=oct;
GENR nov2=nov;
GENR LAR1=lar(-1);
GENR LAR2=lar(-2);
GENR LAR3=lar(-3);
GENR LAR4=lar(-4);
GENR LAR5=lar(-5);
GENR LAR6=lar(-6);
GENR LAR7=lar(-7);
GENR LAR8=lar(-8);
GENR LAR9=lar(-9);
GENR LAR10=lar(-10);
GENR LAR11=lar(-11);
GENR LAR12=lar(-12);
GENR LAR13=lar(-13);
GENR LPR1=lpr(-1);
GENR LPR2=lpr(-2);
GENR LPR3=lpr(-3);
GENR LPR4=lpr(-4);
```

```
106 GENR LPR5=lpr(-5);
107 GENR LPR6=lpr(-6);
108 GENR LPR7=lpr(-7);
109 GENR LPR8=lpr(-8);
110 GENR LPR9=lpr(-9);
111 GENR LPR10=lpr(-10);
112 GENR LPR11=lpr(-11);
113 GENR LPR12=lpr(-12);
114 GENR LPR13=lpr(-13);
115 GENR LSP1=lsp(-1);
116 GENR LSP2=lsp(-2);
117 GENR LSP3=lsp(-3);
118 GENR LSP4=lsp(-4);
119 GENR LSP5=lsp(-5);
120 GENR LSP6=lsp(-6);
121 GENR LSP7=lsp(-7);
122 GENR LSP8=lsp(-8);
123 GENR LSP9=lsp(-9);
124 GENR LSP10=lsp(-10);
125 GENR LSP11=lsp(-11);
126 GENR LSP12=lsp(-12);
127 GENR LSP13=lsp(-13);
128 GENR DLREP1=dlrep(-1);
129 GENR DLREP2=dlrep(-2);
130 GENR DLREP3=dlrep(-3);
131 GENR DLREP4=dlrep(-4);
132 GENR DLREP5=dlrep(-5);
133 GENR DLREP6=dlrep(-6);
134 GENR DLREP7=dlrep(-7);
135 GENR DLREP8=dlrep(-8);
136 GENR DLREP9=dlrep(-9);
137 GENR DLREP10=dlrep(-10);
138 GENR DLREP11=dlrep(-11);
139 GENR DLREP12=dlrep(-12);
140 GENR DLREP13=dlrep(-13);
141 smpl 15 288;
142 OLSQ lar c lar(-1)-lar(-13) lpr(-1)-lpr(-13) dlrep(-1)-dlrep(-13)
    lsp(-1)-lsp(-13) e lw i1974p11 i1992p3 i1993p3 jan feb mar apr may jun
    jul aug sep oct nov c2 jan2 feb2 mar2 apr2 may2 jun2 jul2 aug2 sep2 oct2
    nov2 lar1 lar2 lar3 lar4 lar5 lar6 lar7 lar8 lar9 lar10 lar11 lar12
    lar13 lpr1 lpr2 lpr3 lpr4 lpr5 lpr6 lpr7 lpr8 lpr9 lpr10 lpr11 lpr12
    lpr13 lsp1 lsp2 lsp3 lsp4 lsp5 lsp6 lsp7 lsp8 lsp9 lsp10 lsp11 lsp12
    lsp13 dlrep1
142 dlrep2 dlrep3 dlrep4 dlrep5 dlrep6 dlrep7 dlrep8 dlrep9 dlrep10
    dlrep11 dlrep12 dlrep13;
143 print i @SSR;
144 F=(((RRSS-@SSR)/64)/(@SSR/(@NOB-133)));
145 smpl i i;
146 print F;
147 enddo;
148 end;
```

Table E. 2 *LAR* - Program for Checking for Seasonal Parameters Changes

```
options crt;
freq n;
smpl 1 288;
?1972.1-1995.12;
load (file='ital2.csv')lar lam las lpr dlp lsp e lw jan feb mar apr
  may jun jul aug sep oct nov dec i1974p11 i1987p3 i1992p3 i1993p3 dlrep;
?1973.2-1995.12;
smpl 15 288;
?OLSQ lar c lar(-1)-lar(-13) lpr(-1)-lpr(-13) dlrep(-1)-dlrep(-13)
  lsp(-1)-lsp(-13) e lw i1974p11 i1992p3 i1993p3 jan feb mar apr may jun jul
  aug sep oct nov;
?RRSS=@SSR;
?1973=year2,1984=year13;
do j=13 to 19;
set i= 12*j;
smpl 15 i;
GENR C2=0;
GENR JAN2=0;
GENR FEB2=0;
GENR MAR2=0;
GENR APR2=0;
GENR MAY2=0;
GENR JUN2=0;
GENR JUL2=0;
GENR AUG2=0;
GENR SEP2=0;
GENR OCT2=0;
GENR NOV2=0;
SET h=i+1;
smpl h 288;
GENR c2=c;
GENR jan2=jan;
GENR feb2=feb;
GENR mar2=mar;
GENR apr2=apr;
GENR may2=may;
GENR jun2=jun;
GENR jul2=jul;
GENR aug2=aug;
GENR sep2=sep;
GENR oct2=oct;
GENR nov2=nov;
smpl 15 288;
supres @logl @coef @ses;
OLSQ (silent) lar c lar(-1)-lar(-13) lpr(-1)-lpr(-13)
  dlrep(-1)-dlrep(-13) lsp(-1)-lsp(-13) e lw i1974p11 i1992p3 i1993p3 jan
  feb mar apr may jun jul aug sep oct nov c2 jan2 feb2 mar2 apr2 may2 jun2
  jul2 aug2 sep2 oct2 nov2;
URSS=@SSR;
smpl i i;
print i URSS;
smpl 15 288;
supres @logl @coef @ses;
OLSQ (silent) lar c lar(-1)-lar(-13) lpr(-1)-lpr(-13)
  dlrep(-1)-dlrep(-13) lsp(-1)-lsp(-13) e lw i1974p11 i1992p3 i1993p3 jan
  feb mar apr may jun jul aug sep oct nov apr2 jun2 jul2 aug2;
RRSS=@SSR;
smpl i i;
print i RRSS;
```

```
smpl 15 288;
F=(((RRSS-URSS)/8)/(URSS/(@NOB-81)));
smpl i i;
print F;
enddo;
?1990=year19,1994=year23;
do k=19 to 23;
set n=12*k;
smpl 15 n;
GENR C3=0;
GENR JAN3=0;
GENR FEB3=0;
GENR MAR3=0;
GENR APR3=0;
GENR MAY3=0;
GENR JUN3=0;
GENR JUL3=0;
GENR AUG3=0;
GENR SEP3=0;
GENR OCT3=0;
GENR NOV3=0;
SET l=n;
smpl l 288;
GENR c3=c;
GENR jan3=jan;
GENR feb3=feb;
GENR mar3=mar;
GENR apr3=apr;
GENR may3=may;
GENR jun3=jun;
GENR jul3=jul;
GENR aug3=aug;
GENR sep3=sep;
GENR oct3=oct;
GENR nov3=nov;
smpl 15 288;
supres @logl @coef @ses;
OLSQ (silent) lar c lar(-1)-lar(-13) lpr(-1)-lpr(-13)
  dlrep(-1)-dlrep(-13) lsp(-1)-lsp(-13) e lw i1974p11 i1992p3 i1993p3 jan
  feb mar apr may jun jul aug sep oct nov c3 jan3 feb3 mar3 apr3 may3 jun3
  jul3 aug3 sep3 oct3 nov3;
USS=@SSR;
smpl l l;
print l USS;
smpl 15 288;
supres @logl @coef @ses;
OLSQ (silent) lar c lar(-1)-lar(-13) lpr(-1)-lpr(-13)
  dlrep(-1)-dlrep(-13) lsp(-1)-lsp(-13) e lw i1974p11 i1992p3 i1993p3 jan
  feb mar apr may jun jul aug sep oct nov apr3 jun3 aug3 sep3 oct3;
RSS=@SSR;
smpl l l;
print l RSS;
smpl 15 288;
FF=(((RSS-USS)/7)/(USS/(@NOB-81)));
smpl l l;
print FF;
enddo; end;
```

Table E. 3 Results from the Unrestricted System for the Domestic Demand of Tourism

EQ(1) Modelling LAR by OLS (using vdomlar1.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	1.7478	0.81396	2.147	0.0330	0.0237
LAR_1	0.38074	0.066066	5.763	0.0000	0.1488
LAR_2	0.048142	0.069510	0.693	0.4894	0.0025
LAR_3	0.0030978	0.066150	0.047	0.9627	0.0000
LAR_4	0.10788	0.065040	1.659	0.0988	0.0143
LAR_5	-0.15857	0.063814	-2.485	0.0138	0.0315
LAR_6	0.015455	0.066789	0.231	0.8172	0.0003
LAR_7	-0.046713	0.065977	-0.708	0.4798	0.0026
LAR_8	-0.028143	0.069883	-0.403	0.6876	0.0009
LAR_9	0.065871	0.066148	0.996	0.3206	0.0052
LAR_10	-0.010144	0.066661	-0.152	0.8792	0.0001
LAR_11	0.16078	0.069354	2.318	0.0215	0.0275
LAR_12	0.14191	0.072039	1.970	0.0503	0.0200
LAR_13	-0.0016780	0.063473	-0.026	0.9789	0.0000
LPR	-0.085595	0.25083	-0.341	0.7333	0.0006
LPR_1	-0.0077824	0.27178	-0.029	0.9772	0.0000
LPR_2	-0.68728	0.27959	-2.458	0.0149	0.0308
LPR_3	0.74267	0.29163	2.547	0.0117	0.0330
LPR_4	0.16924	0.29156	0.580	0.5623	0.0018
LPR_5	0.028723	0.29499	0.097	0.9225	0.0000
LPR_6	0.27744	0.29688	0.935	0.3512	0.0046
LPR_7	-0.26316	0.29648	-0.888	0.3759	0.0041
LPR_8	0.17956	0.30431	0.590	0.5559	0.0018
LPR_9	-0.56994	0.29166	-1.954	0.0522	0.0197
LPR_10	-0.19479	0.29109	-0.669	0.5042	0.0024
LPR_11	0.51833	0.28882	1.795	0.0743	0.0167
LPR_12	0.29354	0.27091	1.084	0.2799	0.0061
LPR_13	-0.13497	0.24506	-0.551	0.5825	0.0016
LSP	-0.64898	0.97676	-0.664	0.5072	0.0023
LSP_1	1.9312	1.3185	1.465	0.1447	0.0112
LSP_2	-2.6493	1.2677	-2.090	0.0380	0.0225
LSP_3	1.9388	1.2923	1.500	0.1352	0.0117
LSP_4	-1.6154	1.2704	-1.272	0.2051	0.0084
LSP_5	0.90200	1.2740	0.708	0.4798	0.0026
LSP_6	1.5719	1.2625	1.245	0.2147	0.0081
LSP_7	-2.1950	1.2530	-1.752	0.0814	0.0159
LSP_8	1.3004	1.2408	1.048	0.2959	0.0057
LSP_9	1.4233	1.2168	1.170	0.2436	0.0071
LSP_10	-1.3774	1.1979	-1.150	0.2517	0.0069
LSP_11	-0.27902	1.1973	-0.233	0.8160	0.0003
LSP_12	-1.0631	1.2743	-0.834	0.4052	0.0036
LSP_13	1.0841	0.99087	1.094	0.2753	0.0063
DLREP	2.0523	1.1637	1.764	0.0794	0.0161
DLREP_1	-0.51820	1.2143	-0.427	0.6701	0.0010
DLREP_2	2.3245	1.2154	1.913	0.0573	0.0189
DLREP_3	-0.24840	1.2253	-0.203	0.8396	0.0002
DLREP_4	-0.17517	1.2466	-0.141	0.8884	0.0001
DLREP_5	-0.38739	1.2713	-0.305	0.7609	0.0005
DLREP_6	-0.10116	1.2611	-0.080	0.9362	0.0000
DLREP_7	0.086550	1.3086	0.066	0.9473	0.0000
DLREP_8	-0.84412	1.2514	-0.675	0.5008	0.0024
DLREP_9	-0.61847	1.2237	-0.505	0.6138	0.0013
DLREP_10	-0.24303	1.1906	-0.204	0.8385	0.0002
DLREP_11	0.18527	1.1736	0.158	0.8747	0.0001
DLREP_12	0.040284	1.1686	0.034	0.9725	0.0000
DLREP_13	0.94620	0.95181	0.994	0.3214	0.0052
LW	0.057102	0.060803	0.939	0.3489	0.0046
E	0.16347	0.036018	4.538	0.0000	0.0978
i1974p11	0.29702	0.10678	2.782	0.0060	0.0391
i1992p3	0.38033	0.094417	4.028	0.0001	0.0787

i1993p3	-0.18133	0.097432	-1.861	0.0643	0.0179
jan	0.12267	0.073012	1.680	0.0946	0.0146
jan2	0.073022	0.081555	0.895	0.3717	0.0042
feb	0.23525	0.10321	2.279	0.0238	0.0266
mar	0.18798	0.10856	1.732	0.0850	0.0155
apr	0.45177	0.11566	3.906	0.0001	0.0743
apr2	0.58764	0.12917	4.549	0.0000	0.0982
apr3	0.38338	0.13197	2.905	0.0041	0.0425
may	0.27499	0.13125	2.095	0.0375	0.0226
may3	0.22747	0.14478	1.571	0.1178	0.0128
jun	0.26362	0.14023	1.880	0.0616	0.0183
jun2	0.42852	0.16037	2.672	0.0082	0.0362
jun3	0.50522	0.16031	3.152	0.0019	0.0497
jul	0.38669	0.13892	2.784	0.0059	0.0392
jul2	0.50424	0.15928	3.166	0.0018	0.0501
jul3	0.54595	0.16619	3.285	0.0012	0.0537
aug	0.53047	0.13621	3.895	0.0001	0.0739
aug2	0.69222	0.15700	4.409	0.0000	0.0928
aug3	0.82937	0.16517	5.021	0.0000	0.1172
sep	0.48937	0.13501	3.625	0.0004	0.0647
sep3	0.33492	0.14929	2.243	0.0260	0.0258
oct	-0.24626	0.11004	-2.238	0.0264	0.0257
oct3	-0.45147	0.11629	-3.882	0.0001	0.0735
nov	-0.14206	0.075034	-1.893	0.0598	0.0185

```
R^2 = 0.991966  F(83,190) = 282.66 [0.0000] sigma = 0.0834002
DW = 1.89 RSS = 1.321562962 for 84 variables and 274 observations

AR 1- 7 F( 7,183) =      1.669 [0.1191]
ARCH 7  F( 7,176) =      0.782 [0.6036]
Normality Chi^2(2)=      2.806 [0.2459]
Xi^2    F(139, 50) =     0.463 [0.9998]
RESET   F( 1,189) =      0.935 [0.3347]
```

Table E. 4 Restricted Model for the Domestic Demand for Tourism (*LAR)*

EQ(2) Modelling LAR by OLS (using vdomlar1.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	2.0170	0.59384	3.397	0.0008	0.0464
LAR_1	0.38879	0.042620	9.122	0.0000	0.2599
LAR_5	-0.082737	0.033508	-2.469	0.0142	0.0251
LAR_11	0.11983	0.046145	2.597	0.0100	0.0277
LAR_12	0.21298	0.053721	3.965	0.0001	0.0622
LPR_2	-0.63561	0.18668	-3.405	0.0008	0.0466
LPR_3	0.69440	0.19749	3.516	0.0005	0.0496
LPR_11	0.27851	0.10747	2.592	0.0101	0.0276
LSP_6	1.7557	0.68195	2.575	0.0106	0.0272
LSP_7	-1.3783	0.67999	-2.027	0.0438	0.0170
E	0.15053	0.030347	4.960	0.0000	0.0940
i1974p11	0.26246	0.085739	3.061	0.0025	0.0380
i1992p3	0.33178	0.085295	3.890	0.0001	0.0600
i1993p3	-0.20539	0.088583	-2.319	0.0213	0.0222
jan	0.064519	0.029784	2.166	0.0313	0.0194
jan2	0.017218	0.041616	0.414	0.6794	0.0007
feb	0.053268	0.030584	1.742	0.0829	0.0126
mar	0.067664	0.055816	1.212	0.2266	0.0062
apr	0.30116	0.071988	4.183	0.0000	0.0688
apr2	0.44638	0.079546	5.612	0.0000	0.1173
apr3	0.25874	0.081703	3.167	0.0017	0.0406
may	0.26545	0.076452	3.472	0.0006	0.0484
may3	0.24405	0.086992	2.805	0.0054	0.0321
jun	0.30782	0.078279	3.932	0.0001	0.0613
jun2	0.47421	0.094396	5.024	0.0000	0.0962
jun3	0.52908	0.097862	5.406	0.0000	0.1098
jul	0.44712	0.084166	5.312	0.0000	0.1064
jul2	0.53435	0.10295	5.190	0.0000	0.1021
jul3	0.58188	0.10894	5.341	0.0000	0.1074
aug	0.58053	0.084387	6.879	0.0000	0.1664
aug2	0.75038	0.10204	7.353	0.0000	0.1858
aug3	0.84420	0.10831	7.794	0.0000	0.2040
sep	0.47808	0.082347	5.806	0.0000	0.1245
sep3	0.34407	0.10339	3.328	0.0010	0.0446
oct	-0.17967	0.064626	-2.780	0.0059	0.0316
oct3	-0.36181	0.075497	-4.792	0.0000	0.0883
nov	-0.092644	0.031171	-2.972	0.0033	0.0359

R^2 = 0.990374 F(36,237) = 677.33 [0.0000] sigma = 0.0817401
DW = 2.04 RSS = 1.583503899 for 37 variables and 274 observations

AR 1- 7 F(7,230) =	0.66927	[0.6980]
ARCH 7 F(7,223) =	0.53684	[0.8061]
Normality Chi^2(2)=	0.80300	[0.6693]
Xi^2 F(45,191) =	0.95665	[0.5554]
RESET F(1,236) =	2.74260	[0.0990]

Table E. 5 Unrestricted Model for the Adjusted Series of Domestic Arrivals of Tourism (*LAS*)

EQ(1) Modelling LAS by OLS (using vdomlas.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	1.3484	0.99936	1.349	0.1788	0.0092
LAS_1	-0.067165	0.062013	-1.083	0.2801	0.0060
LAS_2	0.24507	0.056578	4.332	0.0000	0.0878
LAS_3	0.17885	0.057876	3.090	0.0023	0.0467
LAS_4	-0.12697	0.053567	-2.370	0.0188	0.0280
LAS_5	-0.072638	0.054815	-1.325	0.1867	0.0089
LAS_6	0.098669	0.053169	1.856	0.0650	0.0174
LAS_7	-0.22282	0.052336	-4.257	0.0000	0.0850
LAS_8	-0.055809	0.053711	-1.039	0.3001	0.0055
LAS_9	0.22804	0.055502	4.109	0.0001	0.0797
LAS_10	0.11075	0.057790	1.916	0.0568	0.0185
LAS_11	0.12802	0.055180	2.320	0.0214	0.0269
LAS_12	0.19155	0.060956	3.142	0.0019	0.0482
LAS_13	-0.065012	0.059596	-1.091	0.2767	0.0061
LPR	-0.12353	0.34123	-0.362	0.7177	0.0007
LPR_1	-0.55750	0.38161	-1.461	0.1456	0.0108
LPR_2	-0.11972	0.39133	-0.306	0.7600	0.0005
LPR_3	1.0590	0.39550	2.678	0.0080	0.0355
LPR_4	-0.45925	0.40430	-1.136	0.2574	0.0066
LPR_5	0.67004	0.40598	1.650	0.1005	0.0138
LPR_6	-0.11037	0.41742	-0.264	0.7918	0.0004
LPR_7	-0.32972	0.40621	-0.812	0.4180	0.0034
LPR_8	0.70465	0.41804	1.686	0.0935	0.0144
LPR_9	-0.40769	0.41544	-0.981	0.3276	0.0049
LPR_10	-0.60094	0.39886	-1.507	0.1335	0.0115
LPR_11	1.0253	0.39778	2.577	0.0107	0.0329
LPR_12	0.75499	0.37719	2.002	0.0467	0.0201
LPR_13	-0.93834	0.33940	-2.765	0.0062	0.0377
DLREP	0.35673	1.7534	0.203	0.8390	0.0002
DLREP_1	-0.38541	1.8954	-0.203	0.8391	0.0002
DLREP_2	-0.36148	1.8944	-0.191	0.8489	0.0002
DLREP_3	1.7248	1.9204	0.898	0.3702	0.0041
DLREP_4	0.42767	1.9382	0.221	0.8256	0.0002
DLREP_5	-3.4200	1.9508	-1.753	0.0812	0.0155
DLREP_6	0.10136	1.9476	0.052	0.9585	0.0000
DLREP_7	3.4152	1.9201	1.779	0.0769	0.0160
DLREP_8	1.7821	1.8449	0.966	0.3352	0.0048
DLREP_9	1.3535	1.8194	0.744	0.4578	0.0028
DLREP_10	0.65137	1.8113	0.360	0.7195	0.0007
DLREP_11	0.47043	1.7289	0.272	0.7858	0.0004
DLREP_12	2.9078	1.7004	1.710	0.0888	0.0148
DLREP_13	-0.80787	1.3203	-0.612	0.5413	0.0019
LSP	-2.6388	1.6165	-1.632	0.1042	0.0135
LSP_1	2.4848	2.2673	1.096	0.2745	0.0061
LSP_2	2.6155	2.2487	1.163	0.2462	0.0069
LSP_3	-2.2205	2.2731	-0.977	0.3298	0.0049
LSP_4	-2.2738	2.2844	-0.995	0.3208	0.0051
LSP_5	3.6506	2.3407	1.560	0.1205	0.0123
LSP_6	-1.2773	2.3514	-0.543	0.5876	0.0015
LSP_7	-0.74013	2.2935	-0.323	0.7473	0.0005
LSP_8	-0.65523	2.3134	-0.283	0.7773	0.0004
LSP_9	2.0002	2.2889	0.874	0.3833	0.0039
LSP_10	-1.5110	2.2653	-0.667	0.5056	0.0023
LSP_11	2.9199	2.2785	1.282	0.2015	0.0084
LSP_12	-3.8829	2.3012	-1.687	0.0931	0.0144
LSP_13	1.7563	1.7033	1.031	0.3038	0.0054
LW	-0.18247	0.085228	-2.141	0.0335	0.0230
E	0.14147	0.047178	2.999	0.0031	0.0441
i1987P3	-0.36229	0.12758	-2.840	0.0050	0.0397
i1992p3	0.38884	0.12956	3.001	0.0030	0.0442

```
jan              0.018517     0.083392     0.222  0.8245  0.0003
feb              0.066374      0.11915     0.557  0.5781  0.0016
mar               0.12682      0.13102     0.968  0.3343  0.0048
apr               0.43581      0.13994     3.114  0.0021  0.0474
apr2              0.57830      0.15207     3.803  0.0002  0.0690
may               0.45269      0.15397     2.940  0.0037  0.0424
may3              0.18186      0.16950     1.073  0.2846  0.0059
jun               0.39849      0.17072     2.334  0.0206  0.0272
jul               0.45094      0.16523     2.729  0.0069  0.0368
jul2              0.75933      0.18723     4.056  0.0001  0.0778
jul3              0.99358      0.20288     4.897  0.0000  0.1095
aug               0.99762      0.17497     5.702  0.0000  0.1429
aug2               1.3508      0.19543     6.912  0.0000  0.1968
aug3               1.5347      0.21561     7.118  0.0000  0.2062
sep                1.1334      0.17565     6.452  0.0000  0.1759
sep2               1.2666      0.19183     6.603  0.0000  0.1827
oct               0.23524      0.13928     1.689  0.0928  0.0144
oct3            -0.063136      0.14261    -0.443  0.6585  0.0010
nov              -0.12432     0.083929    -1.481  0.1401  0.0111

R^2 = 0.983939  F(78,195) = 153.16 [0.0000]  sigma = 0.116731
DW = 1.91 RSS = 2.65708676 for 79 variables and 274 observations

AR 1- 7 F( 7,188) =     1.1357 [0.3425]
ARCH 7  F( 7,181) =    0.56329 [0.7850]
Normality Chi^2(2)=     3.4725 [0.1762]
Xi^2    F(134, 60) =   0.41231 [1.0000]
RESET   F( 1,194) =     2.9156 [0.0893]
```

Table E. 6 Parsimonious Model for the Adjusted Series (*LAS*)

EQ(2) Modelling LAS by OLS (using vdomlas.in7)
The present sample is: 1973 (3) to 1995 (12)

Variable	Coefficient	Std.Error	t-value	t-prob	PartR^2
Constant	1.7304	0.80147	2.159	0.0318	0.0191
LAS_2	0.17458	0.047447	3.679	0.0003	0.0534
LAS_3	0.14409	0.045036	3.199	0.0016	0.0409
LAS_4	-0.15358	0.045497	-3.376	0.0009	0.0453
LAS_7	-0.21891	0.043479	-5.035	0.0000	0.0955
LAS_9	0.24211	0.049738	4.868	0.0000	0.0899
LAS_11	0.10330	0.047872	2.158	0.0319	0.0190
LAS_12	0.23857	0.047943	4.976	0.0000	0.0935
LPR_1	-0.75765	0.24853	-3.048	0.0026	0.0373
LPR_3	0.90705	0.27354	3.316	0.0011	0.0438
LPR_11	0.33429	0.15290	2.186	0.0298	0.0195
LSP_5	0.38693	0.15094	2.563	0.0110	0.0266
E	0.11282	0.043077	2.619	0.0094	0.0278
i1987P3	-0.33437	0.12288	-2.721	0.0070	0.0299
i1992p3	0.32108	0.12304	2.610	0.0096	0.0276
jan	0.026145	0.059603	0.439	0.6613	0.0008
feb	0.086802	0.084992	1.021	0.3081	0.0043
mar	0.15369	0.10337	1.487	0.1384	0.0091
apr	0.35897	0.11808	3.040	0.0026	0.0371
apr2	0.46276	0.12906	3.586	0.0004	0.0508
may	0.21548	0.11580	1.861	0.0640	0.0142
may3	-0.029320	0.13631	-0.215	0.8299	0.0002
jun	0.25561	0.12331	2.073	0.0392	0.0176
jul	0.37851	0.11169	3.389	0.0008	0.0457
jul2	0.56739	0.12857	4.413	0.0000	0.0751
jul3	0.76246	0.14440	5.280	0.0000	0.1041
aug	0.73407	0.11308	6.492	0.0000	0.1494
aug2	1.0166	0.13604	7.473	0.0000	0.1888
aug3	1.1498	0.15128	7.600	0.0000	0.1940
sep	0.75703	0.10708	7.070	0.0000	0.1724
sep2	0.83828	0.11735	7.143	0.0000	0.1753
oct	0.055854	0.084659	0.660	0.5101	0.0018
oct3	-0.21173	0.10288	-2.058	0.0407	0.0173
nov	-0.13259	0.064957	-2.041	0.0423	0.0171

R^2 = 0.979974 F(33,240) = 355.89 [0.0000] sigma = 0.117493
DW = 2.05 RSS = 3.313086534 for 34 variables and 274 observations

AR 1- 7 F(7,233) = 1.0985 [0.3648]
ARCH 7 F(7,226) = 0.51629 [0.8218]
Normality Chi^2(2)= 4.7013 [0.0953]
Xi^2 F(44,195) = 0.65444 [0.9517]
RESET F(1,239) = 7.4809 [0.0067] **

Table E. 7 VAR(1) for Testing Industrial Production as Proxy for Personal Disposable Income

```
EQ(1)Estimating the unrestricted reduced form by OLS(using
annoprox1.in7)
The present sample is:  2 to 10

URF Equation 1 for LPR
Variable     Coefficient    Std.Error  t-value  t-prob
LPR_1            0.95766      0.36274    2.640  0.0385
LPDIN_1        -0.047736      0.11680   -0.409  0.6970
Constant         0.66395      0.68319    0.972  0.3687
sigma = 0.0269842   RSS = 0.004368878634

URF Equation 2 for LPDIN
Variable     Coefficient    Std.Error  t-value  t-prob
LPR_1            0.40468      0.12476    3.244  0.0176
LPDIN_1          0.84260     0.040172   20.975  0.0000
Constant        -0.24166      0.23497   -1.028  0.3434
sigma = 0.00928059   RSS = 0.0005167758224

correlation of URF residuals
                        LPR       LPDIN
 LPR                 1.0000
 LPDIN             -0.43554      1.0000

standard deviations of URF residuals
          LPR        LPDIN
     0.026984    0.0092806

loglik = 79.22673  log|\Omega| = -17.6059  |Omega| = 2.25859e-008
T = 9
log|Y'Y/T| = -9.82394
R^2(LR) = 0.999583  R^2(LM) = 0.729638
F-test on all regressors except unrestricted,
F(4,10) = 119.9 [0.0000] **
variables entered unrestricted:  Constant
F-tests on retained regressors, F(2, 5)
     LPR_1     12.8304 [0.0107] *     LPDIN_1    222.471 [0.0000] **

correlation of actual and fitted
          LPR        LPDIN
      0.94580      0.99948

LPR     :Portmanteau  2 lags=  1.9672
LPDIN   :Portmanteau  2 lags=  5.6089
LPR     :AR 1- 1 F( 1,  5) =   2.6958 [0.1615]
LPDIN   :AR 1- 1 F( 1,  5) =   5.369e-006 [0.9982]
LPR     :Normality Chi^2(2)=   2.222     [0.3292]
LPDIN   :Normality Chi^2(2)=   1.7685    [0.4130]
LPR     :ARCH 1  F( 1,  4) =   0.11804   [0.7485]
LPDIN   :ARCH 1  F( 1,  4) =   1.0317    [0.3672]
LPR     :Xi^2    F( 4,  1) =   0.080196  [0.9758]
LPDIN   :Xi^2    F( 4,  1) =   0.23397   [0.8925]
Vector portmanteau  2 lags=    9.7464
Vector AR 1-1  F( 4,  6) =     0.36975   [0.8226]
Vector normality Chi^2( 4)=    6.6848    [0.1535]
Vector  Xi^2  Chi^2( 12) =     9.8957    [0.6251]
Vector  Xi*Xj Chi^2( 15) =    12.694     [0.6259]
```

APPENDIX F

Table F. 1 Cointegration Analysis for the Real Substitute Price for Greece and Portugal

system	T	p		log-likelihood	SC	HQ	AIC
1	208	10	COINT	1558.2764	***-14.727***	-14.822	-14.983
2	208	14	COINT	1564.9739	-14.689	***-14.822***	-15.048
3	208	18	COINT	1567.7096	-14.612	-14.784	-15.074
4	208	22	COINT	1569.8325	-14.530	-14.740	-15.095
5	208	26	COINT	1572.2129	-14.450	-14.699	-15.117
6	208	30	COINT	1577.5609	-14.399	-14.686	-15.169
7	208	34	COINT	1580.9661	-14.329	-14.654	-15.202
8	208	38	COINT	1584.3920	-14.259	-14.623	-15.235
9	208	42	COINT	1586.8838	-14.181	-14.582	-15.258
11	208	50	COINT	1594.3170	-14.047	-14.525	-15.330
12	208	54	COINT	1598.0031	-13.980	-14.496	-15.365
13	208	58	COINT	1601.2977	-13.909	-14.463	***-15.397***

System 13 --> System 12: F(4, 356) = 1.4209 [0.2265]
System 12 --> System 11: F(4, 360) = 1.6092 [0.1714]
System 11 --> System 10: F(4, 364) = 2.4026 [0.0495] *

eigenvalue	loglik for	rank
	1581.49	0
0.0998829	1592.44	1
0.01793	1594.32	2

Ho	H_1	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
r=0	r=1	21.89**	19.57*	16.9	25.65**	22.94*	18.2
r=1	r=2	3.76*	3.37	3.7	3.76*	3.37	3.7

standardized beta' eigenvectors

LRSPgr	LRSPpo
1.0000	0.87546
-1.0467	1.0000

standardized alpha coefficients

LRSPgr	-0.021506	0.050780
LRSPpo	-0.053549	-0.004728

long-run matrix Po= alpha* beta', rank 2

	LRSPgr	LRSPpo
LRSPgr	-0.074658	0.031952
LRSPpo	-0.048601	-0.051608

Number of lags used in the analysis: 11
Variables entered unrestricted: Constant Trend i1983p1

Table F. 2 Cointegration Analysis for the Real Substitute Price for France and Spain

system	T	p		log-likelihood	SC	HQ	AIC
1	208	32	COINT	1741.3195	***-15.922***	-16.228	-16.743
2	208	36	COINT	1751.0308	-15.913	***-16.257***	-16.837
3	208	40	COINT	1753.1919	-15.831	-16.213	-16.858
4	208	44	COINT	1754.8687	-15.745	-16.165	-16.874
5	208	48	COINT	1756.7471	-15.660	-16.119	-16.892
6	208	52	COINT	1758.8359	-15.577	-16.075	-16.912
7	208	56	COINT	1764.5024	-15.529	-16.065	-16.966
8	208	60	COINT	1767.4955	-15.455	-16.029	-16.995
9	208	64	COINT	1772.9011	-15.405	-16.017	-17.047
10	208	68	COINT	1774.8735	-15.321	-15.971	-17.066
11	208	72	COINT	1776.9216	-15.238	-15.926	-17.086
12	208	76	COINT	1777.3798	-15.140	-15.866	-17.090
13	208	80	COINT	1780.0307	-15.063	-15.827	***-17.116***

System 13 --> System 12: $F(24, 334)$ = 1.0787 [0.3662]
System 12 --> System 11: $F(20, 338)$ = 1.0793 [0.3695]
System 11 --> System 10: $F(16, 342)$ = 1.3151 [0.1850]
System 10 --> System 9: $F(12, 346)$ = 1.4741 [0.1318]
System 9 --> System 8: $F(8, 350)$ = 1.8027 [0.0754]
System 8 --> System 7: $F(4, 354)$ = 1.2827 [0.2764]
System 7 --> System 6: $F(4, 358)$ = 2.4718 [0.0443] *

eigenvalue	loglik for rank	
	1753.32	0
0.0623249	1760.01	1
0.0422638	1764.50	2

Ho	H_1	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
r=0	r=1	13.39	12.48	16.9	22.37*	20.86*	18.2
r=1	r=2	8.98**	8.38**	3.7	8.98**	8.38**	3.7

standardized beta' eigenvectors

LRSPfr	LRSPsp
1.0000	-0.27873
-0.99564	1.0000

standardized alpha coefficients

LRSPfr	-0.082477	-0.014271
LRSPsp	-0.012539	-0.078992

long-run matrix Po= alpha* beta', rank 2

	LRSPfr	LRSPsp
LRSPfr	-0.068268	0.0087179
LRSPsp	0.066109	-0.075497

Number of lags used in the analysis: 7

Variables entered unrestricted: Constant i1977p7 Trend Seasonal Seasonal_1 Seasonal_2 Seasonal_3 Seasonal_4 Seasonal_5 Seasonal_6 Seasonal_7 Seasonal_8 Seasonal_9 Seasonal_10

Table F. 3 Cointegration Analysis for the Relative Price and Exchange Rate

system	T	p		log-likelihood	SC	HQ	AIC
1 lag	208	30	COINT	1943.3406	-17.916	-18.203	-18.686
2 lags	208	34	COINT	1966.3230	***-18.034***	***-18.359***	-18.907
3 lags	208	38	COINT	1968.2954	-17.951	-18.314	-18.926
....	....	....		....	......	......	....
10 lags	208	66	COINT	1987.1322	-17.413	-18.044	-19.107
11 lags	208	70	COINT	1989.9080	-17.337	-18.006	-19.134
12 lags	208	74	COINT	1994.5701	-17.280	-17.987	-19.179
13 lags	208	78	COINT	1998.5028	-17.215	-17.960	***-19.216***

System 13 lags --> System 12 lags: F(4, 336) = 1.6033 [0.1731]
System 12 lags --> System 11 lags: F(4, 340) = 1.9267 [0.1056]
... .
System 4 lags --> System 3 lags: F(4, 372) = 2.2196 [0.0664]
System 3 lags --> System 2 lags: F(4, 376) = 0.8956 [0.4666]
System 2 lags --> System 1 lag : F(4, 380) = 11.0990 [0.0000] **

eigenvalue	loglik for	rank
	1953.06	0
0.107397	1964.88	1
0.0138206	1966.32	2

Ho	H_1	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
r=0	r=1	23.63**	23.18**	16.9	26.53**	26.02**	18.2
r=1	r=2	2.895	2.839	3.7	2.895	2.839	3.7

standardized beta' eigenvectors

LRPa	LEXa
1.0000	-0.77989
-4.9114	1.0000

standardized alpha coefficients

LRPa	-0.035660	-0.0028824
LEXa	0.15540	-0.0034362

long-run matrix Po=alpha*beta', rank 2

	LRPa	LEXa
LRPa	-0.021503	0.024929
LEXa	0.17227	-0.12463

Number of lags used in the analysis: 2
Variables entered unrestricted:Trend Seasonal Seasonal_1 Seasonal_2 Seasonal_3 Seasonal_4 Seasonal_5 Seasonal_6 Seasonal_7 Seasonal_8 Seasonal_9 Seasonal_10 Constant

Table F. 4 Cointegration Analysis for Real Industrial Production (*LRPRa*), Relative Price (*LRPa*) and Exchange Rate (*LEXa*)

eigenvalue	loglik for rank	
	2880.72	0
0.115695	2893.51	1
0.0787522	2902.04	2
0.000459869	2902.09	3

Ho:rank=p	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
p == 0	25.57*	22.25	23.8	42.73**	37.18*	34.6
p <= 1	17.06*	14.85	16.9	17.16	14.93	18.2
p <= 2	0.096	0.083	3.7	0.10	0.08	3.7

standardized beta' eigenvectors

LRPRa	LRPa	LEXa
1.0000	-0.20819	0.61705
0.16833	1.0000	-0.64085
3.3898	10.116	1.0000

standardized alpha coefficients

LRPRa	-0.080158	-0.031148	-0.00016882
LRPa	-0.014504	-0.065039	0.00020729
LEXa	-0.038281	0.16525	0.00029502

long-run matrix Po= alpha* beta', rank 3

	LRPRa	LRPa	LEXa
LRPRa	-0.085973	-0.016167	-0.029669
LRPa	-0.024749	-0.059922	0.032938
LEXa	-0.0094650	0.17620	-0.12922

Number of lags used in the analysis: 9
Variables entered unrestricted:
Constant Trend

APPENDIX G

In this appendix, an account is given of the investigation done in assessing whether the logarithmic specification is better than the linear form, in estimating the international demand for tourism in Italy.

An initial Box and Cox (1964) procedure has been applied in order to give a statistical foundation on the choice of the log-linear functional form. First of all, one tests for the integration status of the variables expressed in a linear specification. By applying an ADF: *REXP* (real tourist receipts) is found to be I(0), together with *SPpo* (substitute price, Italy/Portugal) and *SPgr* (substitute price, Italy/Greece); whereas, *RPRa* (real income proxy), *RPa* (relative price, Italy/origin countries), *EXa* (weighted average exchange rate for origin countries), *SPfr* (substitute price, Italy/France) and *SPsp* (substitute price, Italy/Spain) are I(1).

Hence, the Johansen procedure is adopted for testing the existence of possible cointegration between the relative price (*RPa*) and exchange rate (*EXa*). An initial unrestricted k=13 VAR, which includes a constant, monthly seasonals and a trend unrestrictedly, can be reduced to a two lag system in accordance with the joint F-test, SC and HQ criteria. The resulting cointegrating vector for the linear specification is the following:

$$ECL = RPa - 0.00071743\ EXa$$

The next step consists of running a Johansen cointegration analysis for the substitute price, as it has already been done for the logarithmic specification. The first analysis is for *SPpo* and *SPgr* which are stationary in the level. One would expect these two series to be stationary also from the Johansen testing. An initial 13 lag system, including a constant and time trend unrestrictedly, can be reduced to a 11 lag system. The finding is that the two economic series are stationary.

Table G. 1 Johansen Tests for the Number of Cointegrating Vectors

Ho	H_1	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
r=0	r=1	21.94**	19.62*	16.9	26.94**	24.09**	18.2
r=1	r=2	5.00*	4.47*	3.7	5.00*	4.47*	3.7

Notes: (1) Adjusted by the degrees of freedom (see, Reimers, 1992)
(2) Critical values at a 5% level of confidence (see Osterward-Lenum, 1992).
(3) * and ** denotes rejection of the null (*i.e.* non-cointegration) at a 5% and 1% level, respectively.

Hence, one runs a Johansen cointegration analysis for *SPfr* and *SPsp* which are stationary in the first difference. An initial 13 lag system, including a constant and time trend unrestrictedly, can be reduced to a 5 lag system. The finding is that the two economic series are stationary as inferred from Table G.2.

Table G. 2 Johansen Tests for the Number of Cointegrating Vectors

Ho	H_1	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
r=0	r=1	16.20**	15.42	16.9	30.07**	28.63**	18.2
r=1	r=2	13.87**	13.21**	3.7	13.87**	13.21**	3.7

Notes:
(1) Adjusted by the degrees of freedom (see, Reimers, 1992)
(2) Critical values at a 5% level of confidence (see Osterward-Lenum, 1992).
* and ** denotes rejection of the null (*i.e.* non-cointegration) at a 5% and 1% level, respectively.

On the basis of these findings, a Box and Cox test is carried out. One runs an unrestricted 13 lag tourism demand equation expressed both in logarithm and linear form. The independent variables are defined as before. *SEAS* are the dummies included in the model.

1) logarithmic form

$$LREXP_t = a_1 + a_2\, LREXP_{t-1\ldots} + a_3\, LRPRa_{t\ldots} + a_4\, SLRPRa_t\ldots + a_5\, DLEXa_t\ldots + \\ + a_6\, DLRPa_{t\ldots} + a_7\, CI_{t-1} + a_8\, LRSPfr + a_9\, LRSPgr + a_{10}\, LRSPpo + \\ + a_{11}\, LRSPsp + a_{12}\, E + a_{13}\, T + a_{14}\, Seas + e_t$$

and

2) linear form

$$REXP_t = a_1 + a_2\, REXP_{t-1\ldots} + a_3\, RPRa_{t\ldots} + a_4\, SRPRa_t\ldots + a_5\, DEXa_t\ldots + a_6\, DRPa_{t\ldots} + \\ + a_7\, ECL_{t-1} + a_8\, RSPfr + a_9\, RSPgr + a_{10}\, RSPpo + a_{11}\, RSPsp + \\ + a_{12}\, E + a_{13}\, T + a_{14}\, Seas + e_t$$

The sum of the squared errors from the logarithmic form (SSE_{LL}) is equal to 0.8391724493, whereas the sum of the squared errors for the linear form (SSE_L) equals 1.56E+20. The null hypothesis that the two models are empirically equivalent is tested and then which of the two models fits the data better. One needs to calculate the sum of the squared errors for the linear model with ($REXP/\overline{\mathrm{REXP}}_G$) as the dependent variable. $\overline{\mathrm{REXP}}_G$ is the geometric mean and is defined as follows:

$$\overline{\mathrm{REXP}}_G = exp\left\{\frac{1}{T}\sum_{t=1}^{T} ln REXP_t\right\}$$

For the latter model, the sum of the squared errors (that is $SSE_L/(\overline{\mathrm{REXP}}_G)^2$) equals 1.081542. The calculated χ^2 is equal to 26.26 that is greater than the tabulated critical value, 3.84, at the 5% level; hence, the null hypothesis cannot be accepted, that is the two models are empirically different. Moreover, the log-linear specification is "much better" than the linear specification as the SSE_{LL} is smaller than $SSE_{L}\,/(\overline{\mathrm{REXP}}_{G})^2$ value.

Table G. 3 Non-Linear Model for the (log) Real Tourism Expenditure

```
1  options crt;
2  freq n;
3  smpl 1 288;
4  ?1972.1-1995.12;
4  load (file='c:\prova\exaggr\LREXPAG.csv')lrexp lrspfr lrspgr lrsppo
   lrspsp dlrp dlex ci lrpr slrpr e jan feb mar apr may jun jul aug sep oct
   nov dec trend i1976p5 i1990p1 RLSPpo RLSPpo1 RLSPsp;
5  ?1973.1-1990.5;
5  smpl 13 221;
6  ?OLSQ lrexp c lrexp(-1) lrexp(-5) lrexp(-6) lrexp(-7) lrspfr
   lrspfr(-2) lrspfr(-7) lrspfr(-9) lrspfr(-10) lrspgr lrspgr(-4) rlsppo
   lrsppo(-5) lrsppo(-6) lrsppo(-7) rlsppo1 rlspsp dlrp(-4) dlrp(-5) dlrp(-6)
   dlrp(-11) dlex dlex(-10) dlex(-11) lrpr(-3) lrpr(-11) slrpr(-3) slrpr(-11)
   e i1976p5 i1990p1 jan feb mar apr may jun jul aug sep oct nov;
6  FRML LRESPQ lrexp = a + b*lrexp(-1) +cii*lrexp(-5) +d*lrexp(-6)
   +ee*lrexp(-7) + f*lrspfr +i*lrspfr(-2) + j*lrspfr(-7) +k*lrspfr(-9)
   +l*lrspfr(-10) +m*lrspgr +n*lrspgr(-4) +o*rlsppo +p*lrsppo(-5)
   +q*lrsppo(-6) +r*lrsppo(-7) +s*rlsppo1 +t*rlspsp +u*dlrp(-4) +v*dlrp(-5)
   +w*dlrp(-6) +y*dlrp(-11) + z*dlex +alfa*dlex(-10) +delta*dlex(-11)
6  + beta*(1/(1+EXP(-1*((LRPR(-11)-MMU)/MSIG)))) +
   eta*(1/(1+EXP(-1*((LRPR(-3)-MMU)/MSIG))))+ g*e + h*i1976p5 + ni*i1990p1
   +elle*jan + emme*feb + enne*mar +zeta*apr +acca*may +esse*jun +gi*jul
   +effe*aug +bi*sep +ciii*oct +di*nov;
7  PARAM a b cii d ee f i j k l m n o p q r s t u w y v z alfa delta
   beta eta ni elle g h emme enne zeta acca esse gi effe bi ciii di;
8  CONST MMU .12 MSIG .21 MMU .12 MSIG .21;
9  LSQ LRESPQ;
10 PARAM MMU MSIG;
11 LSQ LRESPQ;
12 end;
```

APPENDIX H

Table H. 1 Common Trend Analysis

TREND	coeff.	t-value
FRA	0.0027	9.10
GER	0.0038	13.57
JAP	-0.00008	- 0.22
SWE	0.0025	8.02
SWI	0.0025	9.55
UK	0.0016	5.20
USA	-0.0004	- 0.39

Figure H. 1 Real Industrial Production Index (*LRPRm*), Relative Price (*LRPm*) and Exchange Rate (*LEXm*) (1972:1-1990:5 - 5 countries aggregation)

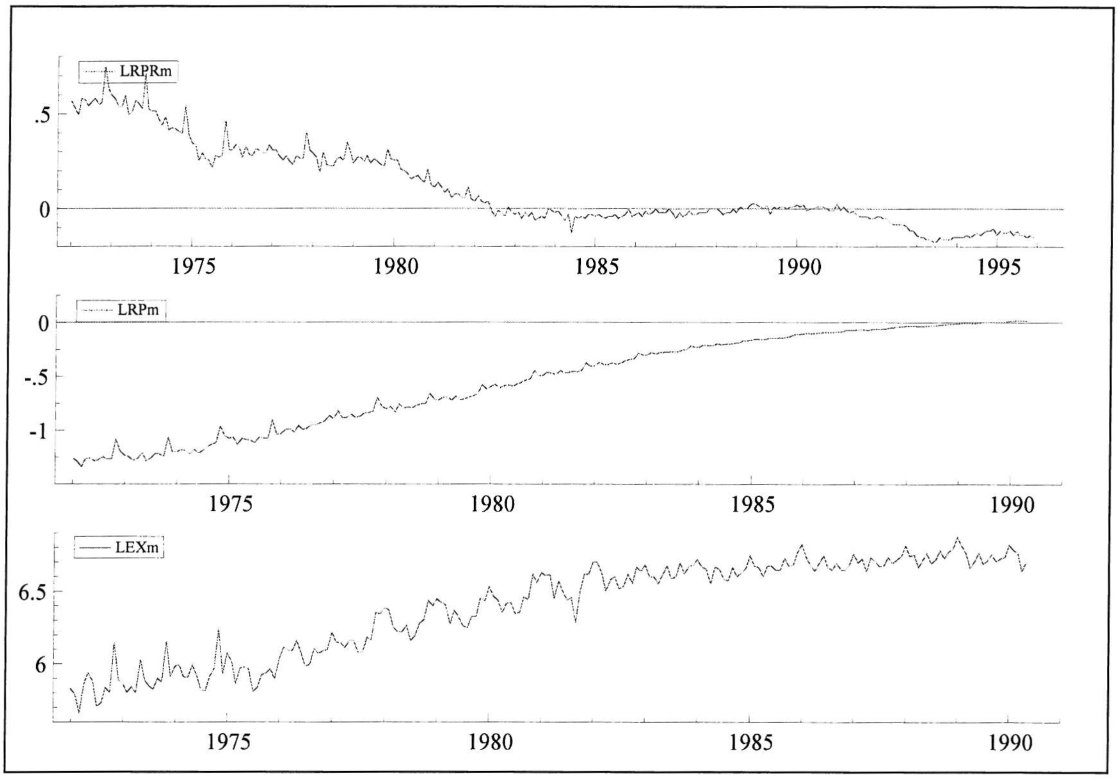


Table H. 2 Cointegration Analysis for the Relative Price and Exchange Rate

system	T	p		log-likelihood	SC	HQ	AIC
1	208	30	COINT	1905.3568	-17.551	-17.838	-18.321
2	208	34	COINT	1931.4005	***-17.699***	***-18.024***	-18.571
....		....		....	....	.....	....
9	208	62	COINT	1947.0492	-17.131	-17.723	-18.722
10	208	66	COINT	1949.9641	-17.056	-17.687	-18.750
11	208	70	COINT	1952.9712	-16.982	-17.651	-18.779
12	208	74	COINT	1960.7161	-16.954	-17.661	-18.853
13	208	78	COINT	1963.3398	-16.877	-17.622	***-18.878***

```
System  2 --> System  1: F( 4, 380) =   12.672 [0.0000] **
System  3 --> System  2: F( 4, 376) =   0.5883 [0.6713]
System  4 --> System  3: F( 4, 372) =   1.6918 [0.1512]
.....          ........                  .....
System 10 --> System  9: F( 4, 348) =   1.2278 [0.2987]
System 11 --> System 10: F( 4, 344) =   1.2524 [0.2886]
System 12 --> System 11: F( 4, 340) =   3.2246 [0.0129] *
System 13 --> System 12: F( 4, 336) =   1.0663 [0.3731]
```

eigenvalue	loglik for rank	
	1916.27	0
0.126145	1930.29	1
0.0105796	1931.40	2

Ho	H_1	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
r=0	r=1	28.05**	27.51**	16.9	30.26**	29.68**	18.2
r=1	r=2	2.21	2.17	3.7	2.21	2.17	3.7

standardized beta' eigenvectors

LRPa	LEXa
1.0000	-0.99691
-0.3304	1.0000

standardized alpha coefficients

LEXa	-0.10969	0.021587
LRPa	0.043375	0.013886

long-run matrix Po=alpha*beta', rank 2

	LEXa	LRPa
LEXa	-0.11682	0.13093
LRPa	0.038786	-0.029354

Number of lags used in the analysis: 2

Variables entered unrestricted:

Trend Seasonal Seasonal_1 Seasonal_2 Seasonal_3 Seasonal_4 Seasonal_5 Seasonal_6 Seasonal_7 Seasonal_8 Seasonal_9 Seasonal_10 Constant

Table H. 3 Cointegration Analysis for Real Industrial Production (*LRPRa*), Relative Price (*LRPa*) and Exchange Rate (*LEXa*)

eigenvalue	loglik for	rank
	2862.37	0
0.121612	2875.85	1
0.0575214	2882.01	2
0.00241357	2882.26	3

Ho:rank=p	λ_{max}	$\lambda_{max}(1)$	C.V.(2)	λ_{trace}	$\lambda_{trace}(1)$	C.V.(2)
p == 0	26.97*	22.3	23.8	39.8*	32.91	34.6
p <= 1	12.32	10.19	16.9	12.82	10.61	18.2
p <= 2	0.50	0.41	3.7	0.50	0.42	3.7

standardized beta' eigenvectors

LRPRa	LRPa	LEXa
1.0000	-0.66035	1.4944
0.045892	1.0000	-1.1082
0.10451	0.22367	1.0000

standardized alpha coefficients

LRPRa	-0.10597	-0.065215	0.0013233
LEXa	0.029159	-0.035512	0.0089947
LRPa	-0.041019	0.044879	0.0026422

long-run matrix Po= alpha* beta', rank 3

	LRPRa	LRPa	LEXa
LRPRa	-0.10883	0.0050617	-0.084778
LEXa	0.028469	-0.052755	0.091927
LRPa	-0.038683	0.072557	-0.10839

Number of lags used in the analysis: 12

Variables entered unrestricted:

Constant Trend

APPENDIX I

In this appendix, a detailed account of the specification form adopted in estimating the real aggregated budget share (*LBSm*) is given. A non-linear transformation for the income proxy is considered.

The initial formulation of the equation for the aggregated budget share (*LBSm*) is as follows:

$$LBSm = f\,(LRPRa,\ SLRPRa,\ DLRPa,\ DLEXa,\ CI,\ LRSPfr,\ LRSPgr,\ LRSPpo,\ LRSPsp,\ E,\ T,\ D) \qquad \text{(I.1)}$$

A Box and Cox (1964) procedure has been applied in order to give a statistical foundation on the choice of the log-linear functional form. Firstly, one tests for the integration status of the variables expressed in a linear specification, that is: *BSm* (aggregated real budget share), *RPRa* (real income proxy), *RPa* (relative price, Italy/origin countries), *EXa* (weighted average exchange rate for origin countries), *SPfr* (substitute price, Italy/France), *SPgr* (substitute price, Italy/Greece), *SPpo* (substitute price, Italy/Portugal) and *SPsp* (substitute price, Italy/Spain). From applying the ADF, *BSm*, *SPpo* and *SPgr* have been found to be stationary in the level; whereas, *RPRa* *RPa*, *EXa*, *SPfr* and *SPsp* are I(1).

Hence, the Johansen procedure is adopted for testing the existence of possible cointegration between the relative price (*RPa*) and exchange rate (*EXa*). An initial unrestricted k=13 VAR, which includes a constant, monthly seasonals and a trend unrestrictedly, can be reduced to a two lag system in accordance with the joint F-test, SC and HQ criteria. The resulting cointegrating vector for the linear specification is the following:

$$LCI = RPa - 0.0012363\ EXa$$

The next step consists in running a Johansen cointegration analysis for the substitute price, as has already been done for the logarithmic specification. The first analysis is for *SPpo* and *SPgr* which are stationary in the level. One would expect these two series to be stationary also from the Johansen testing. An initial 13 lag system, including a constant and time trend unrestrictedly, can be reduced to an 11 lag system. The finding shows that the two economic series are stationary.

Table I. 1 Johansen Tests for the Number of Cointegrating Vectors

Ho	H_1	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
r=0	r=1	21.94**	19.62*	16.9	26.94**	24.09**	18.2
r=1	r=2	5.00*	4.47*	3.7	5.00*	4.47*	3.7

Notes:
(1) Adjusted by the degrees of freedom (see, Reimers, 1992).
(2) Critical values at a 5% level of confidence (see Osterward-Lenum, 1992).
* and ** denotes rejection of the null (*i.e.* non-cointegration) at a 5% and 1% level, respectively.

Hence, one runs a Johansen cointegration analysis for *SPfr* and *SPpo* which are stationary in the first difference. An initial 13 lag system, including a constant and time trend unrestrictedly, can be reduced to a 5 lag system. The finding is that the two economic series are stationary as inferred from Table I.2.

Table I. 2 Johansen Tests for the Number of Cointegrating Vectors

Ho	H_1	λ_{max}	λ_{max} (1)	C.V.(2)	λ_{trace}	λ_{trace}(1)	C.V.(2)
r=0	r=1	16.20**	15.42	16.9	30.07**	28.63**	18.2
r=1	r=2	13.87**	13.21**	3.7	13.87**	13.21**	3.7

Notes:
(1) Adjusted by the degrees of freedom (see, Reimers, 1992).
(2) Critical values at a 5% level of confidence (see Osterward-Lenum, 1992).
* and ** denotes rejection of the null (*i.e.* non-cointegration) at a 5% and 1% level, respectively.

On the basis of these findings, a Box and Cox test is carried out. One runs an unrestricted 13 lag tourism demand equation expressed both in a logarithmic and linear form. The independent variables are defined as before.

1) logarithmic form

$$LBSm_t = a_1 + a_2\, LBSm_{t-1\ldots} + a_3\, LRPRa_{t\ldots} + a_4\, SLRPRa_{t}\ldots + a_5\, DLEXa_{t}\ldots +$$
$$+ a_6\, DLRPa_{t\ldots} + a_7\, CI_{t-1} + a_8\, LRSPfr + a_9\, LRSPgr + a_{10}\, LRSPpo +$$
$$+ a_{11}\, LRSPsp + a_{12}\, E + a_{13}\, T + a_{14}\, Seas + e_t$$

and

2) linear form

$$BSm_t = a_1 + a_2\, BSm_{t-1\ldots} + a_3\, RPRa_{t\ldots} + a_4\, SRPRa_{t}\ldots + a_5\, DEXa_{t}\ldots + a_6\, DRPa_{t\ldots} +$$
$$+ a_7\, LCI_{t-1} + a_8\, RSPfr + a_9\, RSPgr + a_{10}\, RSPpo + a_{11}\, RSPsp +$$
$$+ a_{12}\, E + a_{13}\, T + a_{14}\, Seas + e_t$$

The sum of the squared errors from the logarithmic form (SSE_{LL}) is equal to 1.223519728, whereas the sum of the squared errors for the linear form (SSE_L) equals 0.0000291621223. One wants to test the null hypothesis that the two models are empirically equivalent and find out which of the two models fits the data better. The sum of the squared errors for the linear model with ($BSm/\overline{BSm}_G$) as the dependent variable is calculated. Note that $\overline{\mathrm{BSm}}_{G}$ is the geometric mean defined as follows:

$$\overline{\mathrm{BSm}}_{G} = exp\left\{ \frac{1}{T} \sum_{t=1}^{T} ln\, BSm_t \right\}$$

For the latter model, the sum of the squared errors (that is $SSE_L/(\overline{\mathrm{BSm}}_{G})^2$) equals 3.806690317. The calculated χ^2 is equal to 117.47 that is greater than the tabulated critical value, 3.84, at the 5% level; hence, the null hypothesis cannot be accepted, that is the two models are empirically different. Moreover, one infers that the logarithmic specification is "much better" than the linear specification as the SSE_{LL} is smaller than $SSE_L/(\overline{\mathrm{BSm}}_{G})^2$ value. Hence, the logarithmic functional form is adopted for the demand function as given in (I.1).

REFERENCES

Andrews, D.W.K. (1993). "Tests for Parameter Instability and Strutural Change With Unknown Change Point". *Econometrica*, Vol.61: 821-856.

Arbel, A., and S.A. Ravid (1985). "On Recreation Demand: A Time-Series Approach". *Applied Economics*, no.17: 979-990.

Archer, B.H. (1976). *Demand Forecasting in Tourism*. Occasional Papers in Economics, no.9, Bangor: University of Wales Press.

Artus, J.R. (1972). "An Econometric Analysis of International Travel". *IMF Staff Papers*, no.19: 579-614.

Ballatori, E., and G. Vaccaro (1992). *"Elementi di Statistica per il Turismo"*. FrancoAngeli: Milano.

Baloglu, S., and K.W. McCleary (1999). "U.S. International Pleasure Travelers' Images of Four Mediterranean Destinations: A Comparison of Visitors and Nonvisitors". *Journal of Travel Research,*38: 144-152.

Banca D'Italia (1995). *Tematiche Istituzionali. Manuale della Bilancia dei Pagamenti in Italia*. Fonti Statistiche e Metodi di Elaborazione: 7-10.

Banerjee, A., J.J. Dolado, J.W. Galbraith, and D.F. Hendry (1993). *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data.* Oxford Univerity Press: New York.

Baron, R. (1989). *Travel and Tourism Data.* Euromonitor Publication Limited: Wiltshire.

Beaulieu, J.J., and J.A. Miron (1993). "Seasonal Unit Roots in Aggregate U.S. Data". *Journal of Econometrics*, no.55: 305-328.

Blackwood & Partners (1994). *Interviste Turisti Stranieri*. ESIT, Cagliari.

Bond, M.E. (1979). "The World Trade Model: Invisibles". *International Monetary Fund Staff Papers*, 26: 257-333.

Box, G.E.P., and D.R. Cox (1964). "An Analysis of Transformations". *Journal of the Royal Statistical Society, Series B,* 26(2): 211- 43.

Caporale, G.M., and N. Pittis (1997). "Unit Root Testing Using Covariates: Some Theory and Evidence." *Centre for Economic Forecasting, London Business School*, D.P. no. 20-97: 1-31.

Carraro, C., and M. Manente (1996). "The TRIP and STREP Forecasting Models of Tourism Flows: from World-Wide to Regional Analysis in the Italian Case". *Fondazione Eni Enrico Mattei*, November: 1-19.

Chow, G.C. (1960). "Tests of Equality Between Sets of Coefficients in Two Regressions". *Econometrica,* 28: 591-605.

Chow, G.C. (1967). "Technological Change and the Demand for Computers". *The American Economic Review*, 57: 1117-1130.

Clauser, O. (1991). "Ruolo della Spesa Promozionale nelle Esportazioni Turistiche Italiane". *Politica del Turismo,* 2: 19-29.

Confcommercio (1994). *Il Sistema Turistico nel Nord Sardegna*, Sassari.

Contu, G. (1997). *Domanda ed Offerta nel Nord Sardegna e Ruolo dei Prezzi e Qualità*. Tesi di Laurea in Scienze Politiche, Università degli Studi di Sassari.

Cooley, T.F., and S.F. LeRoy (1985). "Atheoretical Macroeconometrics. A Critique". *Journal of Monetary Economics*, 16: 283-308.

CRENoS (2000). *Settimo Rapporto di Previsione sull'Economia della Sardegna.* Centro Ricerche Economiche Nord Sud. Università degli Studi di Cagliari.

Crouch, G.I. (1994). "The Study of International Tourism Demand: A Survey of Practice". *Journal of Travel Research*, 32(4): 41-55.

Deaton, A., and J. Muellbauer (1980). "An Almost Ideal Demand System". *The American Economic Review*, 70(3): 312-326.

D.E.I.S. (1995). *Il Settore Turistico in Corsica e Sardegna. Un'analisi comparata.* Università degli Studi di Sassari.

Delbono, F., and G. Fiorentini (1987). *Economia del Turismo.* La Nuova Italia Scientifica: Roma.

Di Matteo, L., and R. Di Matteo (1993). "The Determinants of Expenditures by Canadian Visitors to the United States". *Journal Of Travel Research*, 31(4): 34-42.

Dickey, D.A., and W.A. Fuller (1981). "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root". *Econometrica,* 49(4): 1061-1072.

Doornik, J.A., and D.F. Hendry (1996). *PC FIML 9.0 - Interactive Econometric Modelling of Dynamic Systems.* Institute of Economics and Statistics: University of Oxford.

Doornik, J.A., and D.F. Hendry (1996). *PC GIVE 9.0 - Empirical Econometric Modelling Using.* International Thomson Business Press: London.

Dwyer, L., P. Forsyth, and P. Rao (2000). "The Price Competitiveness of Travel and Tourism: a Comparison of 19 Destinations". *Tourism Management*, 21: 9-22.

Enders, W., and C.W.J. Granger (1998). "Unit-Root Tests and Asymmetric Adjustment With an Example Using the Term Structure of Interest Rates". *American Statistical Association. Journal of Business and Economic Statistics*, 16(3): 304-311.

Engle, R.E., and C.W.J. Granger (1987). "Co-Integration and Error Correction: Representation, Estimation, and Testing". *Econometrica*, 55(2): 251-276.

Engle, R.E., D.F. Hendry, and J.F. Richard (1983). "Exogeneity". *Econometrica*, 51: 277-304.

Formica, S., and M. Uysal (1996). "The Revitalization of Italy as a Tourist Destination". *Tourism Management,* 17(5): 323-331.

Franses, P.H. (1991a). "Model Selection and Seasonality in Time Series". *Institute Research Series*, Tinbergen, 10: 94-165.

Franses, P.H. (1991b). "Seasonality, Non-Stationarity and the Forecasting of Monthly Time Series". *International Journal of Forecasting*, 7: 199-208.

Fujii, E.T, M. Khaled, and J. Mak, (1985). "An Almost Ideal Demand System For Visitor Expenditures". *Journal of Transport and Policy,* 19: 161-171.

Garcìa-Ferrer, A., and R.A. Queralt (1997). "A Note on Forecasting International Tourism Demand in Spain". *Departamento de Anàlisis Econòmico, Madrid.*

Gardini, A. (1984). "La Domanda Turistica, Un'Analisi Econometrica". *Politica del Turismo,* 2(1): 179-195.

Garratt, A., K. Lee, M.H. Peresan, and Y. Shin, (1998). "A Long Run Structural Macroeconometric Model of the UK." Unpublished paper : 1-28.

Gonzàles, P., and P. Moral (1995). "An Analysis of the International Tourism Demand in Spain". *International Journal of Forecasting*, 11: 233-251.

Gonzàles, P., and P. Moral (1996). "Analysis of Tourism Trends in Spain". *Annals of Tourism Research*, 23(4): 739-754.

Granger, C.W.J., and J. Hallman (1991). "Nonlinear Transformations of Integrated Time Series". *Journal of Time Series Analysis*, 12(3): 207-224.

Granger, C.W.J., and T.H. Lee (1989). "Investigation of Production, Sales, and Inventory Relationships Using Multicointegration and Non-Symmetric Error-Correction Models". *Journal of Applied Econometrics*, 4: S145-S159.

Grasselli, P. (1982). *Economia e Politica del Turismo.* FrancoAngeli: Milano.

Gray, H.P. (1966). "The Demand for International Travel by the United States and Canada". *International Economic Review,* 7(1): 83-92.

Griffiths, W., R.C. Hill, and G. Judge (1993). *Learning and Practising Econometrics* . John Wiley and Sons: WEA.

Hamilton, J.D. (1994). *Time Series Analysis.* Princeton: Princeton University Press.

Hansen, B.E. (1995). "Rethinking the Univariate Approach to Unit Root Testing - Using Covariates to Increase Power". *Econometric Theory,* 11: 1148-1171.

Hargreaves, C.P. (1994). *Nonstationary Time Series Analysis and Cointegration.* New York: Oxford University Press.

Harris, R. (1995). *Cointegration Analysis in Econometric Modelling.* Prentice Hall and Harvester Wheatsheaf: London.

Hartmann, R. (1986). "Tourism, Seasonality and Social Change". *Leisure Studies,* 5: 25-33.

Hendry, D.F. (1993). "The Roles of Economic Theory and Econometrics in the Analysis of Economic Time Series". Unpublished paper, Institute of Economics and Statistic, Oxford. Presented to the European Meeting of the Econometric Society, Uppsala, September.

Hendry, D.F. (1995). *Dynamic Econometric.* New York: Oxford University Press.

Hendry, D.F., and G.E. Mizon (1990). "Procrustean Econometrics: or Stretching and Squeezing Data". Chapter 7 in Granger (ed.) (1990a), *Modelling Economic Series: Readings in Econometric Methodology.* Oxford: Oxford University Press.

Hendry, D.F., and H-M. Krolzig (1999a). "Computer Automation of General-to-Specific Model Selection Procedures". Discussion Paper, Institute of Economics and Statistics and Nuffield College, Oxford: 1-26.

Hendry, D.F., and H-M. Krolzig (1999b). "Improving on "Data Mining Reconsidered"" by K.D. Hoover and S.J. Perez". *Royal Economic Society*, 2: 202-219.

Hoover, K.D., and S.J. Perez (1999). "Data Mining Reconsidered: Encompassing and the General-to-Specific Approach to Specification Search". *Econometrics Journal*, 2: 1-25.

Hylleberg, S., R.F. Engle, C.W.J. Granger, and B.S. Yoo (1990). "Seasonal Integration and Cointegration". *Journal of Econometrics,* 44: 215-238.

Hylleberg, S., and M. Paldam (1991). *New Approaches to Empirical Macroeconomics.* Cambridge: Blackwell Publishers.

Johansen, S. (1988) "Statistical Analysis of Cointegration Vectors". *Journal of Economic Dynamics and Control,* 12: 231-254.

Johansen, S. (1995). *Likelihood Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

Johansen, S., and K. Juselius (1990). "Maximum Likelihood Estimation and Inference on Cointegration - With Applications to the Demand for Money". *Oxford Bulletin of Economics and Statistics,* 52(2): 169-210.

Kulendran, N. (1995). "Forecasting International Quarterly Tourist Arrivals to Australia with Time-Series Models". *Working Papers*, Victoria University Melbourne, No.3/95: 1-44.

Kulendran, N., and S.F. Witt (2001). "Cointegration Versus Least Squares Regression". *Annals of Tourism Research*, 28(2): 291-311.

Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag.

Lanza, A., and G. Urga (1995). "Tourism Specialisation and Economic Growth: A Cross Country Study Using Multivariate Cointegration". *Discussion Paper*, LBS, 10:1-17.

Leamer, E.E. (1985). "Vector Autoregressions for Causal Inference", in K. Brunner and A. Meltzer (eds.), *Understanding Monetary Regimes* (supplement to *Journal of Monetary Economics*), 22: 255-303.

Lee, C., T.Var, and T.W. Blaine (1996). "Determinants of Inbound Tourist Expenditures". *Annals of Tourism Research*, 23 (3): 527-541.

Lickorish, L.J. (1997). "Travel Statistics - the Slow Move Forward". *Tourist Management*, 18 (8): 491-497.

Lim, C. (1997). "Review of International Tourism Demand Models". *Annals of Tourism Research*, 24 (4): 835-849.

Lim, C. (1999). "A Meta-Analytic Review of International Tourism Demand". *Journal of Travel Research* 37: 273-284.

Lim, C. and M. McAleer (2000). "A Seasonal Analysis of Asian Tourist Arrivals to Australia". *Applied Economics*, 32: 499-509.

Lim, C. and M. McAleer (2001). "Monthly Seasonal Variations. Asian Tourism to Australia". *Annals of Tourism Research*, 28(1): 68-82.

Maddala, G.S. (1992). *Introduction to Econometrics*. New York: MacMillan Publishing Company 2nd ed..

Makridakis, S., S.C. Wheelwright, and V.E. McGee (1989). *Forecasting: Methods and Applications*. New York: Wiley.

Malacarni, L. (1991). "Analisi Econometrica della Domanda Turistica in Italia". *Politica del Turismo,* 2: 19-29.

Martin, C.A., and S.F. Witt (1987). "Tourism Demand Forecasting Models: Choice of Appropriate Variable to Represent Tourists Cost of Living". *Tourism Management,* 8: 223-245.

Martin, C.A., and S.F. Witt (1988). "Substitute Prices in Models of Tourism Demand". *Annals of Tourism Research,* 15(2): 255-268.

Martin, C.A., and S.F. Witt (1989). "Forecasting Tourism Demand: a Comparison of the Accuracy of Several Quantitative Methods". *International Journal of Forecasting,* 5: 7-19.

Masberg, B.A. (1998). "Defining the Tourism: Is it Possible? A View from the Convention and Visitors Bureau". *Journal of Travel Research,* 37: 67-70.

McIntosh, R.W., C.R. Goeldner, and J.R.B. Ritchie (1995). *Tourism. Principles, Practices, Philosophies*. New York: Wiley.

Ministero del Turismo e dello Spettacolo (1995). *Sesto Rapporto Sul Turismo Italiano*, Mercury, Firenze.

Mizon, G.E. (1996). "Progressive Modelling of Macroeconomic Time Series - The LSE Methodology", Ch. 4 in Hoover, K.D. (ed) *Macroeconometrics.* California: Kluwer Academic Publishers.

Nordström, J. (1996). "Tourism Satellite Account for Sweden 1992-93". *Tourism Economics,* 2(1): 13-42.

O'Hagan, J.W., and M.J. Harrison (1984). "Market Shares of US Tourism Expenditure in Europe: an Econometric Analysis". *Applied Economics,* 16: 919-931.

Osterwald-Lenum, M. (1992). "A Note With Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics". *Oxford Bulletin of Economics and Statistics,* 54: 461-472.

Papatheodorou, A. (1999). "The Demand for International Tourism in the Mediterranean Region". *Applied Economics,* 31: 619-630.

Pesaran, M.H., and B. Pesaran (1997). *Working with Microfit 4.0.* O.U.P.

Pindyck, R.S., and D.L. Rubinfeld (1991). *Econometric Models and Economic Forecasts*. Singapore: International Editions.

Qiu, H., and J. Zhang (1995). "Determinants of Tourism and Expenditures in Canada". *Journal of Travel Research,* 34(2): 43-49.

Quayson, J., and T. Var (1982). "A Tourism Demand Function for the Okanagan, BC". *Tourism Management,* 3: 108-115.

Raeside, R., J. Adams, R. Rush, and K. Haldane (1997). "Modelling and Forecasting Domestic Tourism in Scotland". *Paper presented at the 17th International Symposium on Forecasting, Barbados*: 6-11.

Reimers, H.E. (1992). "Comparison of Tests for Multivariate Cointegration". *Statistical Papers,* 33: 335-359.

Rugg, D.D. (1973). "The Choice of Journey Destination: A Theoretical and Empirical Analysis". *Review of Economics and Statistics,* 55(1): 64-71.

Seddighi, H.R., and D.F. Shearing (1997). "The Demand for Tourism in North East England with Special Reference to Northumbria: An Empirical Analysis". *Tourism Management,* 18(8): 499-511.

Sheldon, P.J. (1993). "Forecasting Tourism: Expenditure Versus Arrivals". *Journal of Travel Research*, 32: 13-19.

Sims, C.A. (1980). "Macroeconomics and Reality". *Econometrica*, 48 (1): 1-48; reprinted as Chapter 8 in Granger (ed.) (1990a).

Sinclair, M.T. (1998). "Tourism and Economic Development. A Survey". *Journal of Development Studies*, 34(5): 1-51.

Sinclair, M.T., and M. Stabler (1997). *The Economics of Tourism.* Routledge: London and New York.

Smeral, E. (1988). "Tourism Demand, Economic theory, and Econometrics: An Integrated Approach". *Journal of Travel Research,* 26(4): 38-42.

Smith, S.L.J. (1995). *Tourism Analysis.* Essex: Longman.

Solinas, C. (1992). *Il Turismo nel Nord Sardegna. Una Valutazione dell'Impatto Economico nella Spesa Turistica*, no. 99: CIRIEC.

Song, H., P. Romilly, and X. Liu (2000). "An Empirical Study of Outbound Tourism Demand in the UK". *Applied Economics*, 32: 611-624.

Summary, R. (1987). "Estimation of Tourism Demand by Multivariable Regression Analysis: Evidence from Kenya". *Tourism Management,* 8(4): 317-322.

Syriopoulos, T.C. (1995). "A Dynamic Model of Demand for Mediterranean Tourism". *International Review of Applied Economics,* 9(3): 318-336.

Syriopoulos, T.C., and M.T. Sinclair (1993). "An Econometric Study of Tourism Demand: the AIDS Model of US and European Tourism in Mediterranean Countries". *Applied Economics,* 25: 1541-1552.

Tremblay, R. (1989). "Pooling International Tourism in Western Europe". *Annals of Tourism Research,* 16: 477-491.

Uysal, M., and M.S. El Roubi (1999). "Artificial Neural Networks Versus Multiple Regression in Tourism Demand Analysis". *Journal of Travel Research*, 38:111-118.

Var, T., G. Mohammad, and O. Icoz (1990). "Factors Affecting International Tourism Demand for Turkey". *Annals of Tourism Research*, 17 (4): 606-611.

Vogt, G.M., and C. Wittayakorn (1998). "Determinants of the Demand for Thailand's Exports of Tourism". *Applied Economics,* 30: 711-715.

Webb, R.H. (1995). "Forecasts of Inflation from VAR Models". *Journal of Forecasting,* Vol. 14: 267-285.

Witt, S., and C.A. Witt (1992). *Modelling and Forecasting Demand in Tourism.* London: Academic Press.