Department of Economics
University of Southampton
Southampton SO17 1BJ
UK

**Discussion Papers in
Economics and Econometrics**

MONTHLY, ANNUAL AND QUARTERLY FREQUENCIES: A COMPARISON OF MODELS FOR TOURISM IN SARDINIA

M Pulina
and
R J O'Brien

No. 0206

ISSN 0966-4246

DISCUSSION PAPERS
IN ECONOMICS AND ECONOMETRICS

MONTHLY, ANNUAL AND QUARTERLY FREQUENCIES:
A COMPARISON OF MODELS FOR TOURISM IN SARDINIA

M. Pulina
and
R.J. O'Brien

Department of Economics
University of Southampton

Discussion Paper No. 0206

August, 2002

Abstract

This paper constructs and estimates the demand for international tourism for the Italian Province of Sassari. The sample period under estimation is from 1972 to 1995. Three dynamic models are estimated at monthly, annual and quarterly data frequencies. Similarities and differences are explored amongst the three models, using recently developed econometric techniques. A "pre-modelling" data analysis is undertaken for the economic series of interest. By adopting the LSE "general-to-specific" methodology, dynamic estimations are run. A full range of diagnostic tests is provided. Short and long run income elasticities, negativity and substitutability are tested on the light of economic theory. On balance, evidence is found that the monthly and quarterly models present homogenous results in terms of seasonal and long run unit roots. Annual data show different and perhaps misleading results.

JEL classification: L83, C50
Keywords: Modelling tourism, general to specific modelling, Sassari, Sardinia.

ISSN 0966-4246

DISCUSSION PAPERS
IN ECONOMICS AND ECONOMETRICS

MONTHLY, ANNUAL AND QUARTERLY FREQUENCIES:A COMPARISON OF MODELS FOR TOURISM IN SARDINIA

M. Pulina and R.J. O'Brien

Department of Economics, University of Southampton

Discussion Paper No.

August, 2002

Abstract

This paper constructs and estimates the demand for international tourism for the Italian Province of Sassari. The sample period under estimation is from 1972 to 1995. Three dynamic models are estimated at monthly, annual and quarterly data frequencies. Similarities and differences are explored amongst the three models, using recently developed econometric techniques. A “pre-modelling” data analysis is undertaken for the economic series of interest. By adopting the LSE "general-to-specific" methodology, dynamic estimations are run. A full range of diagnostic tests is provided. Short and long run income elasticities, negativity and substitutability are tested on the light of economic theory. On balance, evidence is found that the monthly and quarterly models present homogenous results in terms of seasonal and long run unit roots. Annual data show different and perhaps misleading results.

JEL classification: L83, C50

I. Introduction

Studies of tourism demand have been undertaken in the Fifties; however, the dawn of a systematic economic analysis of tourism has been first seen with Gray (1966). In the Seventies, an increased number of empirical studies appeared in tourism literature. The determinants of international demand for tourism started to be analysed by applying economic concepts, econometric methodologies and forecasting tools (Artus, 1972; Archer, 1976). Crouch (1994a and 1994b) and Lim (1997) provide a comprehensive literature review for more than one hundred empirical studies over three decades on international tourism demand. In these surveys, a detailed account is provided of the type of data used, the methodologies adopted, and the dependent and explanatory variables employed. According to Lim (1997) and Sinclair (1998), extensive econometric effort still needs to be expended on the study of international tourism demand. Small sample sizes, lack of discussion of the appropriate functional forms, and failure to include the full range of diagnostic tests, are pointed out as some of the main deficiencies in empirical tourism demand studies. More advanced econometric approaches, including Hendry's methodology, seasonal and long run unit roots and cointegration analysis are still much neglected in 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).

In this paper, a model is formulated and estimated for the international demand for tourism in the Italian Province of Sassari (known also as north of Sardinia)[1]. The main proposition under investigation is the following: are there common findings by using different data frequencies (*i.e*. monthly, annual and quarterly data)? One of the suggestions given by Witt and Witt (1992) for further research is to estimate tourism models at different data frequencies. "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" (Witt and Witt 1992, p.171). Uysal and Roubi (1999) also point out the lack of research on the same ground. "The use of different data periods is one of the areas that would need further research in tourism demand and forecasting studies" (Uysal and Roubi 1999, p.116). The scope of this paper is to investigate this proposition using the sample period from 1972 to 1995.

The paper is structured in the following manner. In the first section, a brief discussion of the methodology adopted is given. In the second section, an investigation is carried out of the integration status of the variable of interest. Cointegration will be tested for the non-stationary variables. The third section is dedicated to a comparison of econometric models at three frequencies. A summary and a set of conclusions will be given in the last two sections.

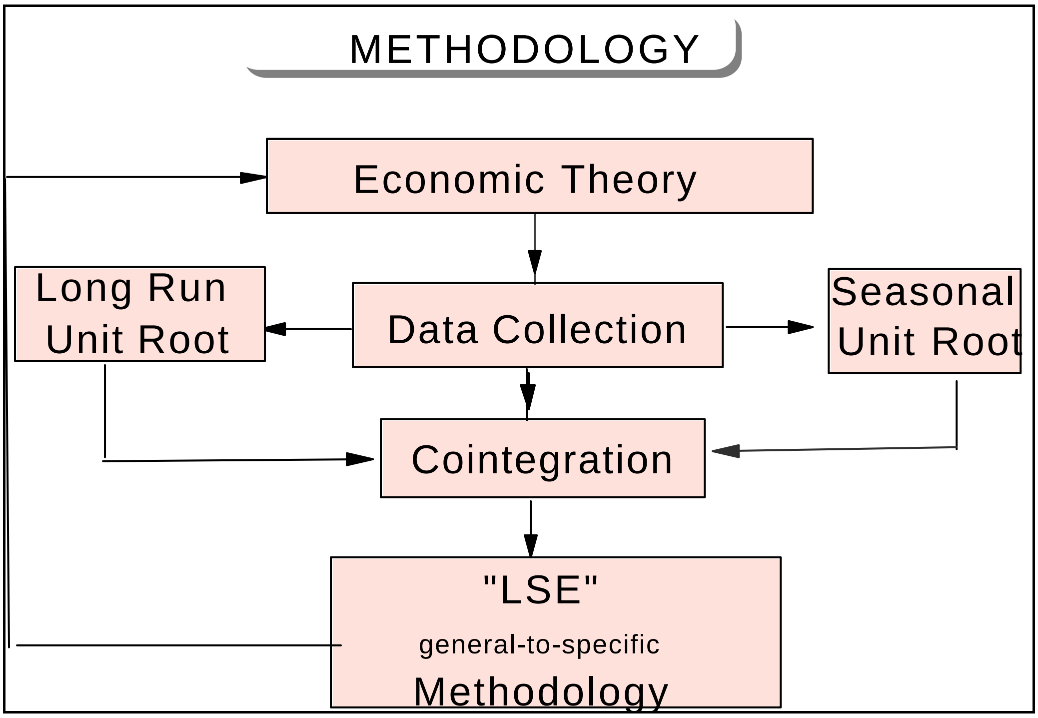
II. Methodology

The research steps followed in this paper are shown in Figure 1.

---

[1] Sassari Province, as reported by the Confcommercio (1994) sees the major quota of tourist flows in the island, equal to 54% relative to the other three Provinces, for the sample period under analysis. Moreover, few studies exist of the demand for tourism in the Province of Sassari (Solinas, 1992; D.E.I.S., 1995; Contu, 1997) and none of them makes use of the most recent econometric methodology.

Figure 1. Methodology



Appropriate economic theory is derived from the relevant literature. The next step consists in linking the theory with empirical practise. In this way, it is possible to derive an economic model for the international demand of tourism in the north of Sardinia as a function of certain quantitative and qualitative variables. In this study the generic demand function used is the following:

$$D = f\,(DIP, RP, EX, SP, DV)$$

where:
D = demand for tourism; DIP = disposable income; RP = relative price (destination/origin countries); EX = exchange rate; SP = substitute price (destination/competitor countries); DV = qualitative variables (such as seasonal dummies, impulse dummies, etc.).
The second step relates to data collection. From the raw data, approximations to the variables of interest are calculated on a monthly, quarterly and annual frequency.
The next phase involves a "pre-modelling" analysis for testing both possible long run and seasonal unit roots. The theory suggests that a series can be non-stationary in the level. In particular, a series whose first differences are stationary may be a random walk. To test this, one can use the so-called Augmented Dickey-Fuller (ADF) unit root test. In this paper, Dickey and Fuller's (1981) framework will be used. The ADF test consists in running equation (1):

$$?Y_t = ? + ?\,T + ???\,???\,?Y_{t-1}\;??\sum_{i?1}^{p}\;?_i\;?Y_{t-1}\;??\,?D\;?\;?_t \qquad (1)$$

where a constant, the first lag of the series, the lagged difference terms, a time trend (say T) and seasonal dummies (say D) are included. The augmentation is set to the first statistically significant lag testing downwards. Results for the ADF test will be given for each of the following possible combinations: equation (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 (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 ? =1. When failing to reject the null hypothesis, one treats the dependent variable as non-stationary[2].

The seasonal unit roots test, for quarterly and monthly series, is based on Hylleberg *et al*. (1990) and Franses (1991a and 1991b), respectively. These tests allow to study in a systematic manner characteristics and properties of international tourism seasonality in Sassari Province. Many recent 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. As Hyllerberg points out (see Hargreaves, 1994, pp.153-177), there are many different causes for seasonal variation. As far as tourism is concerned, a change in tourists' preferences or a 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.

For the quarterly time series, the HEGY seasonal unit roots procedure is adopted. 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). "There will be no seasonal unit roots if?? $_?$ and either ? $_?$ or ? $_?$ are different from zero, which therefore requires the rejection of both a test for ? $_?$ and a joint test for ? $_?$ and ? $_?$" (Hylleberg *et al*. 1990, p.223). The auxiliary equation is given by:

$$?^{?}(B)y_{4,t} = ??_t ? ??_1 ?y_{1,t-1} + ?_2 ?y_{2,t-1} ? ??_3 y_{3,t-2} ? ??_4 ?y_{3,t-1} ?? ??_t \qquad (2)$$

where $?^{?}(B)$ is a polynomial in the lag operator *B*, and $?_t$ represents the deterministic part, and in this particular study consists of a constant, a time trend and 3 seasonal dummies. The $y_{i,t}$ are linear combinations of lagged y_t values. Equation (2) is fitted by OLS[3].

One needs also to test for possible seasonal unit roots at a monthly frequency. 'Testing for unit roots in monthly time series is equivalent to testing for the significance of the parameters in the auxiliary regression' (Franses 1991a, p.202) estimated by Ordinary Least Squares (OLS):

$$?^{?}(B)y_{8,t} = ???_1 ?y_{1,t-1} + ?_2 ?y_{2,t-1} + ??_3 y_{3,t-1} + ??_4 ?y_{3,t-2} + ??_5 ?y_{4,t-1} + ?_6 ?y_{4,t-2} + ?_7 ?y_{5,t-1} + ??$$

$$? \;\; ??+?_8 y_{5,t-2} ?+ ??_9 ?y_{6,t-1} ?+ ??_{10} ?y_{6,t-2} ?+ ??_{11} y_{7,t-1} + ??_{12} ?y_{7,t-2} + ??_t + ??_t \qquad (3)$$

where, $?_t$, the deterministic part, consists of a constant, a time trend and 11 seasonal dummies. The null hypothesis of unit roots is tested both running a *t*-test of the separate ?'s, as well as the joint *F*-test of the pairs, and the ?'s in the interval $??_3 \ldots ?_{12}$[4]. If the null hypothesis is rejected one can treat the variable of interest as seasonally stationary. Both the monthly and quarterly seasonal unit roots tests are run in Microfit 4.0 (Pesaran e Pesaran, 1997).

'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' (Johansen 1995, p.34). Two or more variables are defined to be cointegrated when, though they are non-stationary in the levels, in the long run a linear combination of them converges towards a common equilibrium. Whenever necessary, the cointegration assumption will be tested amongst I(1) time series. Johansen's (1988) procedure is adopted since it is able to capture the interdependencies between time series[5].

---

2 All the results concerning with the ADF test are obtained using PcGive module of GiveWin 9.0 (Doornik and Hendry, 1996).

3 Critical values are provided in Hylleberg *et al*. (1990, pp.226-227).

4 Critical values for the seasonal unit roots test are given in Franses (1991b, pp.161-165).

5 Johansen and Juselius (1990), and Osterwald-Lenum (1992) critical values are employed.

The transformed variables will be employed for the estimation of a dynamic model where both the short and long run information are included. For this aim, the LSE "general-to-specific" methodology is used. In this paper, the main assumptions of the LSE econometric modelling are based on the comprehensive survey provided by Mizon (1996). The central concepts are congruency and encompassing. A model needs to be congruent with economic theory, past, present and future sample information, and the measurement system and encompass rival models. A model is robust when it is able to encompass a new set of information. The strategy of "general-to-specific" is argued 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 may also validate *a priori* economic theory. Statistical tests[6], information criteria[7] and a set of diagnostic tests[8] are used in order to evaluate these assumptions.

The final step of the methodology used is the feedback to economic theory (Figure 1). The results obtained from the congruent and encompassing model are compared with the theory. Income and price elasticities, and, in general, the capability of the independent variables to explain the dependent variable will be examined.

III. Economic Model and Definition of Variables

In the previous section, the generic demand function for tourism has been identified as derived from economic theory. A more precise definition, as used in the modelling phase, is given:

$$D = f\,(PR, RP, ER, SP, Easter, W, Trend, ID, Dummies)$$

where:

D = international demand for tourism. This variable is expressed as the total number of tourists' arrivals in all the registered accommodation in the North of Sardinia (source EPT Sassari (1972-1995). 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 (see Ballatori and Vaccaro, 1992). In the empirical analysis, this variable will be called *A*. A preliminary investigation, provided in Pulina (2002), has been carried out using the raw series of foreign arrivals. Such analysis encountered problems of non-normality and heteroscedasticity (at 1% level), which have been corrected with the adjustment of the dependent variable for the number of weekends in a month. As Baron (1989) points out "trading-day factors" might be important in the analysis of monthly data: these take into account the effects of four of five weekends 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. Given these assumptions, the dependent variable is defined as follows:

$$A = AR\,/\,N$$

where, *AR* is the total number of tourists' arrivals and *N* is the number of Saturdays in a month. For the quarterly and annual models, *N* is defined as the average number of Saturdays in a quarter and year, respectively. In Pulina (2002), a detailed analysis is provided.

---

6 In estimating an autoregressive distributed lag model the choice of the lag length is of extreme importance. In choosing the lag length the statistical joint *F*-test (or Wald test) is adopted. This test allows testing whether it is statistically significant to reduce the lag length by one.

7 The lag length of a model can be also chosen by making use of information criteria, that is Hannan-Quinn, denoted as HQ criterion; Schwartz, denoted as SC criterion; finally, Akaike, denoted as AIC criterion. The estimated information criteria are chosen so that they are minimised.

8 DW, Durbin-Watson statistic; AR, autocorrelation test; ARCH, conditional heteroscedasticity; NORM, normality test; HETER, heteroscedasticity test; RESET, functional form test; CHOW, prediction test; WALD, long run coefficients statistical significance test, excluding the constant.

PR = weighted average of the industrial production index (1990=100). This variable has been used as proxy of the income index for which monthly data are not available (Gonzàles and Moral, 1995; García-Ferrer and Queralt, 1997). *PR* is defined as follows:

$$PR_t = \frac{\sum_{i=1}^{i=7} w_{i,t} * PR_{i,t}}{\sum_{i=1}^{i=7} w_{i,t}}$$

where:

i = Belgium, France, Germany, Sweden, Switzerland, United Kingdom and United States. These origin countries represent the highest quota for Sassari Province.

$PR_{i,t}$ = industrial production index (1990=100) seasonally adjusted, in country i in month t (Source: IFS-datastream).

$w_{i,t}$ = take into account the amount of tourists coming from the origin country i in year t (Source: ISTAT), and it is given by the following formula:

$$w_{i,t} = \frac{AR_{i,t}}{\sum_{i=1}^{i=7} AR_{i,t}} \qquad (4)$$

Note that the weights vary over time, to reflect the changing importance of different constituents of the average being calculated. The weights are allowed to change annually rather than monthly. Annual weights may be thought to be more stable than the monthly weights. One could argue that more frequent changes might just reflect different seasonal patterns.

RP = relative price. It represents the price of north of Sardinia tourism relative to the set of clients countries (i) as above listed. It can be thought as a proxy of the competitive price between origin and destination country (Martin and Witt, 1987). Such a variable is expressed by the following formula:

$$RP_t = \frac{CPI_{ss,t}}{CPI_{o,t}}$$

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:

$$CPI_{o,t} = \frac{\sum_{i=1}^{i=7} w_{i,t} * CPI_{i,t}}{\sum_{i=1}^{i=7} w_{i,t}}$$

where:

$CPI_{i,t}$ = monthly consumer price index (1990=100) in country i and month t (Source: IFS-datastream).

Note that the weights ($w_{i,t}$) are defined as in (4).

ER = weighted average exchange rate. The weighted exchange rate with respect to the main origin countries (i) can be expressed by the following formula:

$$ER_t = \frac{\sum_{i=1}^{i=7} w_{i,t} * ER_{i,t}}{\sum_{i=1}^{i=7} w_{i,t}}$$

where:

$ER_{i,t}$ = nominal exchange rate, in country i in month t (Source: *Banca d'Italia*).

$w_{i,t}$ = as in formula (4).

SP = substitute price. It represents the price of north of Sardinia tourism relative to the set of competitor countries in the Mediterranean area. This variable is defined by the following formula:

$$SP_t = \frac{CPI_{ss,t}}{CPI_{c,t}}$$

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:

$$CPI_{c,t} = \frac{\sum_{i=1}^{4} ?_{i,t} * CPI_{i,t}}{\sum_{i=1}^{4} ?_{i,t}}$$

i = France, Greece, Portugal and Spain.

$CPI_{i,t}$ = monthly consumer price index (1990=100) in country i and month t (Source: IFS-datastream).

$?_{i,t}$ = weights are defined as follows:

$$?_{i,t} = \frac{AR_{i,t}}{\sum_{i=1}^{4} AR_{i,t}}$$

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.

Easter = "Easter" *dummy*. This variable is included 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 falls between the 26th March and the 22nd April. The dummy variable "Easter" has, therefore, 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 (see Gonzàles and Moral, 1996).

W = climate variable (Source: University of Agriculture of Sassari). Such a variable is expressed as the average temperatures for the Province of Sassari. One is interested in considering if weather conditions, that can be also regarded as a component of the tourism supply, have an impact in explaining international demand for tourism in Sassari Province (McIntosh *et al.*, 1995).

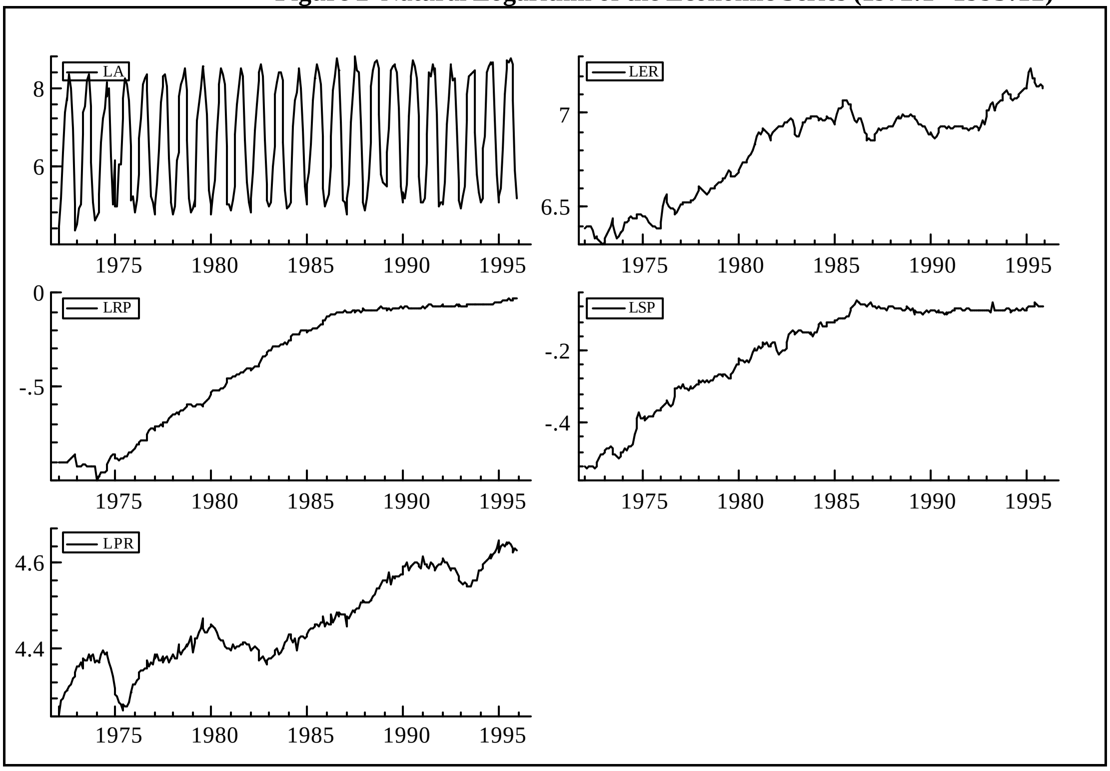
T = trend. Two main streams of thought are in the literature. First, a time trend is included in the model in order to pick up possible changes in consumers' tastes for a specific destination over time. Second, the time trend variable is recognised as causing problems of multicollinearity with other explanatory variables such as income (see Crouch, 1994b). In this study, a time trend is included in the final restricted model as having a statistically significant coefficient.

ID = impulse dummies. These qualitative variables are constructed in order to avoid non-normality problems in the residuals. Using monthly data, such dummies are not always easy to interpret. Possible factors for outliers could be related to particular events, such as strikes for boats or planes, or particular discounts for holiday packages 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.

Dummies = seasonal dummies. Such variables, with either a monthly or quarterly frequency, have been included to evaluate seasonal factors and cyclical holidays effecting the international demand for tourism to Sassari Province.

Graphs of each economic series are provided in Figure 2 at a monthly frequency. In each case the natural logarithm of the variables is used.

Figure 2 Natural Logarithm of the Economic Series (1972:1 - 1995:12)



IV. Unit Roots and Cointegration: a Monthly, Quarterly and Annual Frequency Comparison

In this section, results for seasonal and long run unit roots, and cointegration analyses are reported for the economic variables with a monthly, quarterly and annual frequency, respectively.

Equation (3) is fitted by OLS for each of the five time series mentioned above, for the sample period from 1972:1 to 1995:12. The results are reported in Table 1.

No evidence of seasonal unit roots emerges from Table 1. Thus the seasonal pattern can be treated as deterministic. For the long run unit roots a comparison can be made between Tables 1 and 2. The dependent variable, *LA*, can be treated as stationary in the level (or I(0)) from both tables. The income proxy (*LPR*), exchange rate (*LER*), relative price (*LRP*) and substitute price (*LSP*) are all non-stationary in the level (Table 1, top row). Again, from the ADF test (Table 2), *LRP* and *LER* appear I(1); however, *LPR* and *LSP*[9] can be treated as stationary in the level (or I(0)).

9 Experiments have shown as the substitute price can be considered stationary in the level when a constant is included, whereas such a variable is I(1) when a constant and a time trend are included (Table 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 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 1 Testing for Seasonal Unit Roots (1972:1-1995:12)					
t-statistics	Variable				
	LA	***LPR***	***LER***	***LRP***	***LSP***
??	-3.502 ***	-3.081	-1.659	0.071	-1.152
??	-4.715 ***	-4.627 ***	-4.914 ***	-4.262 ***	-4.509 ***
??	1.966	-1.467	-6.812 ***	-6.056 ***	-5.623 ***
??	-6.907 ***	-6.133 ***	-3.357 *	-3.537 **	-3.843 ***
??	-6.679 ***	-6.944 ***	-7.508 ***	-7.550 ***	-6.581 ***
??	-7.387 ***	-6.790 ***	-6.332 ***	-7.841 ***	-6.584 ***
??	2.516	-2.757 ***	-1.776 ***	-2.977 ***	-2.544 ***
??	-4.627 ***	-1.197	-1.221	-0.356	-0.671
??	-2.117	-4.428 ***	-6.675 ***	-5.479 ***	-5.705 ***
???	-6.017 ***	-7.998 ***	-2.716	-5.969 ***	-6.266 ***
???	1.606	-3.364 ***	-4.932 ***	-3.498 ***	-4.016 ***
???	-4.919 ***	-3.215 *	-1.650	-3.114	-2.712
F-statistics	***LA***	***LPR***	***LER***	***LRP***	***LSP***
??????	26.594 ***	20.207 ***	27.867 ***	26.654 ***	19.847 ***
??????	27.506 ***	25.696 ***	26.822 ***	32.300 ***	25.126 ***
??????	16.902 ***	33.898 ***	24.417 ***	22.591 ***	22.531 ***
???????	18.848 ***	32.151 ***	26.385 ***	22.831 ***	21.111 ***
????????	12.723 ***	24.846 ***	30.397 ***	25.052 ***	36.727 ***
??????????????	24.184 ***	208.198 ***	94.600 ***	186.018 ***	150.043 ***

Notes: ***, ** and * indicate that the seasonal unit root null hypothesis is rejected at the 1%, 5% and 10% level, respectively.

Table 2 Augmented Dickey-Fuller Unit Root Test (1972:1 - 1995:12)					
Time Series	*ADF*(1)	*LAG*(2)	***Time Series***	*ADF*(1)	*LAG*(2)
			LPR(c)	- 0.46	3
LA(c)	- 3.87 **	9	DLPR(c)	- 11.77 **	2
LA(c,t)	- 4.36 **	10	LPR(c,t)	- 3.84 *	8
LA(c,s)	- 4.06 **	2	LPR(c,s)	- 0.45	3
LA(c,s,t)	- 6.33 **	2	DLPR(c,s)	- 7.73 **	2
			LPR(c,s,t)	- 3.77 *	8
LRP(c)	- 2.16	12	**LER**(c)	- 1.68	1
DLRP(c)	- 3.11 *	11	DLER(c)	- 10.70 **	1
LRP(c,t)	- 0.64	12	LER(c,t)	- 2.20	1
DLRP(c,t)	- 3.78 *	11	DLER(c,t)	- 10.72 **	1
LRP(c,s)	- 2.89 *	0	LER(c,s)	- 1.61	1
LRP(c,s,t)	- 0.47	12	DLER(c,s)	- 10.51 **	1
DLRP(c,s,t)	- 3.76 *	11	LER(c,s,t)	- 2.11	1
			DLER(c,s,t)	- 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,s,t)	- 1.18	0			
DLSP(c,s,t)	- 15.13**	0			

Notes: * and ** 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.

Johansen's maximum likelihood procedure is used to establish whether a cointegrating relationship exists between the I(1) variables (*LRP* and *LER*). 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). 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:

$$X_t \ ?= ?? \ D_t ?? \ ? \ _1 X_{t-1} \ ???????? \ ? \ _{13} X_{t-13} \ ? ???_t \qquad (t = 1,.....T)$$

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 is possibly picking up the first oil shock. A VAR (13), as above specified, has been re-estimated with *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 and information criteria SC and HQ it is possible to reduce the system to a VAR (3). This lag gives a satisfactory portmanteau test statistic for serial correlation; however, non-normality as well as heteroscedasticity problems have not been eliminated.

To test the cointegration hypothesis one makes use of the procedure presented in Johansen (1988). The results of the eigenvalue and eigenvector calculations are given in Table 3[10].

Table 3 Eigenvalues $\hat{?}$, Eigenvectors $\hat{?}$ Weights $\hat{?}$				
Eigenvalues $\hat{?}$				
(0.0737	0.0063)			
Standardized $\hat{?}$??eigenvectors		Standardized $\hat{?}$ coefficients		
LRP	***LER***	***LRP***	-0.02	0.019
1.00	-1.09	***LER***	0.01	-0.017
-0.43	1.00			

Table 4 reports the results of the tests for reduced rank. The test statistics are the maximal eigenvalue ??? $_{max}$) and the trace statistics ?? $_{trace}$).

Table 4 Johansen Tests for the Number of Cointegrating Vectors							
Ho	H_1	$?_{max}$	$?_{max}$???)	C.V.(2)	$?_{trace}$	$?_{trace}$??)	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, therefore, one can reject the hypothesis that $r=0$, at least at the 1% level, concluding that there is one cointegrating relationship.

The coefficient estimates are found in Table 4 as the first row of ? ' matrix and the cointegrating relationship is defined as follows:

$$CI = LRP - 1.0893LER$$

[10] All the results concerning with the cointegration testing are obtained using PcFiml module of GiveWin 9.0 (Doornik and Hendry, 1996).

The coefficient for the (log) weighted exchange rate has been tested for the following restriction: *? =-1* ?and such a restriction has been accepted at the 1% level from the ?^{2} test[11]. In this way one can model the following cointegrating vector:

$$CI = LRP - 1 * LER \qquad (5)$$

Hence, in the long run, there are no main price differentials amongst the origin countries under analysis. This finding validates the theoretical assumption that in the short run price and exchange rates should be used separately, whereas a real exchange rate should be used in the long run in accordance to the Purchasing Power Parity (PPP) theory.

The next step of the investion consists in establishing whether analogies can be found between quarterly and monthly time series within the "pre-modelling" phase. Following the Hylleberg *et al.* (1990) methodology the results reported in Table 5 are obtained.

Table 5 Testing for Seasonal Unit Roots (1972:1 - 1995:4)

t-statistics	Variable				
	LA	***LPR***	***LSP***	***LER***	***LRP***
??	-3.93 **	-3.23 *	-1.04	-1.70	0.05
??	-3.32 ***	-5.26 ****	-6.05 ****	-7.13 ***	-5.54 ****
??	-6.45 ****	-4.17 ****	-2.82	-3.32 *	-4.18 ****
??	-2.19 **	-7.41 ****	-5.87 ****	-5.04 ****	-6.34 ****
F-statistics	***LA***	***LPR***	***LSP***	***LER***	***LRP***
??????	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

.

A comparison between Tables 1 and 5 shows no difference in terms of seasonal unit roots. In both the cases, each economic variable presents a deterministic seasonal pattern.

In Table 6, the results from the ADF test are reported. One can conclude that *LA*, *LPR* and *LSP* are I(0), whereas *LRP* and *LER* are I(1). Again, from HEGY and ADF tests, some differences can be noticed for *LSP* in establishing the integration status.

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. 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 monthly data case, the possible cointegration between the two I(1) variables is tested. An initial unrestricted VAR(5) is first run which presented problems of non-normality and heteroscedasticity in the equation for the relative price. Two 0-1 impulse dummies are added to pick up possibly the negative effects of the first oil shock (*i.e. i1974q1* and *i1975q1*). The VAR(5) is re-estimated with these two dummies, a constant and three quarterly seasonal dummies. 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 where non-homoscedasticity appeared. Note also that, again, the coefficient of determinations for the first equation (*LRP*) is 0.99962 and, for the second equation (*LER*) equals 0.99208.

Johansen's cointegration test has given the results reported in Table 7. The test statistics suggest that the null hypothesis of the existence of one cointegrating vector cannot be rejected at the confidence level of 1%.

[11] The results for the restriction test on the coefficient is: *?^{2}*(1) = 0.64365 [0.4224]

Table 6 Augmented Dickey-Fuller Unit Root Test (1972:1 - 1994:4)					
Time Series	*ADF (1)*	*LAG*	***Time Series***	*ADF (1)*	*LAG*
			LPR(c)	- 0.64	1
LA(c)	- 3.39 **	2	DLPR(c)	- 6.37 **	0
LA(c,t)	- 4.17 **	5	LPR(c,t)	- 3.62 *	2
LA(c,s)	- 4.29 **	0	LPR(c,s)	- 0.63	1
LA(c,s,t)	- 7.51 **	0	DLPR(c,s)	- 6.25 **	0
			LPR(c,s,t)	- 3.56 *	2
LRP(c)	- 2.09	4	**LER**(c)	- 1.16	2
DLRP(c)	- 3.10 *	3	DLER(c)	- 7.74 **	1
LRP(c,t)	- 0.62	4	LER(c,t)	- 1.40	2
DLRP(c,t)	- 3.88 *	3	DLER(c,t)	- 7.77 **	1
LRP(c,s)	- 2.05	4	LER(c,s)	- 1.15	2
DLRP(c,s)	- 3.17 *	3	DLER(c,s)	- 7.33**	1
LRP(c,s,t)	- 0.44	4	LER(c,s,t)	- 1.42	2
DLRP(c,s,t)	- 4.00*	3	DLER(c,s,t)	- 7.36 **	1
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,s,t)	- 0.97	4			
DLSP(c,s,t)	- 6.17**	3			
Notes: as in Table 2. (1) ADF statistics with (c) c.v. = -2.893 at 5% and -3.503 at 1% level; with (c,t) c.v.= -3.458 at 5% and -4.059 at 1% level; with (c,s) c.v. = -2.894 at the 5% and -3.505 at 1%; with (c,t,s) c.v. = -3.46 at 5% and -4.062 at 1% level.					

Table 7 Johansen Tests for the Number of Cointegrating Vectors using Quarterly Data							
Ho	H_1	$?_{max}$	$?_{max}$(??)	C.V.(2)	$?_{trace}$	$?_{trace}$(??)	C.V.(2)
r=0	r=1	53.49**	47.61**	14.1	55.35**	49.27**	15.4
r=1	r=2	1.86	1.65	3.8	1.86	1.65	3.8
Notes: as in Table 4.							

The results of the eigenvalue and eigenvector calculations are given in Table 8.

Table 8 Eigenvalues $\hat{?}$, Eigenvectors $\hat{?}$ Weights $\hat{?}$				
Eigenvalues $\hat{?}$				
(0.4445	0.0202)			
Standardized $\hat{?}$??eigenvectors		Standardized $\hat{?}$ coefficients		
LRP	***LER***	***LRP***	-0.07	0.0006
1.00	-0.89	***LER***	-0.03	-0.0600
-0.62	1.00			

The cointegrating vector is the following:

$$CI = LRP \;-0.89\, LER$$

The coefficient for the (log) weighted exchange rate has been tested for the restriction: *? =-1*?that is accepted at the 5% level from the $?^2$ test[12]. As for the monthly case, one uses the cointegrating vector in the equation (5).

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

The final step of this analysis involves running a "pre-modelling" testing for the annual data (1972-1995). In order to obtain homogeneous results and comparisons between models, the same time series as for the previous cases will be used. In Table 9, the results from running an ADF test for each of the economic series are presented.

Table 9 Augmented Dickey-Fuller Unit Roots Test (1972-1995)					
Time Series	*ADF (1)*	*LAG*	***Time Series***	*ADF*(1)	*LAG*
LA(c) DLA(c) LA(c,t) LA(c,t)	- 1.23 - 3.35 ** - 2.24 - 3.18	0 0 1 0	**LPR**(c) DLPR(c) LPR(c,t) LPR(c,t)	- 0.16 - 6.30 ** - 3.01 - 6.30 **	0 0 0 0
LRP(c) LRP(c,t) LRP(c,t)	- 4.32 ** - 0.82 - 3.89 *	1 1 0	**LER**(c) DLER(c) LER(c,t) DLER(c,t) DDLER(c,t)	- 1.19 - 7.43 ** - 2.25 - 2.70 ** - 5.12	0 0 0 0 0
LSP(c) LSP(c,t) DLSP(c,t)	- 5.79 ** - 1.38 - 3.98 *	0 1 0			
Notes: As in Table 2. (1) ADF statistics with (*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.					

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. A comparison with Tables 2 and 6 suggests major differences in the results. These findings lead to 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 the above variables are treated as having the same integration status as suggested by the ADF test when using monthly and quarterly data. The modified series of foreign arrivals (*LA*) is treated as I(0), as well as the income proxy (*LPR*) and the substitute price (*LSP*). The relative price (*LRP*) and the exchange rate (*LER*) are treated as I(1).

Again, a Johansen cointegration analysis is undertaken for *LRP* and *LER*. An initial bivariate VAR(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. From the diagnostic tests the null hypothesis of homoscedasticity fails to be accepted at the 5% level. This finding seems to confirm those obtained for the system using data with monthly and quarterly frequency.

Table 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 10 Johansen Tests for the Number of Cointegrating Vectors using Annual Data							
Ho	H_1	?$_{max}$	?$_{max}$???)	C.V.(2)	?$_{trace}$	?$_{trace}$??)	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
Notes: as in Table 4.							

The results of the eigenvalue and eigenvector calculations are given in Table 11.

Table 11 Eigenvalues $\hat{?}$, Eigenvectors $\hat{?}$ Weights $\hat{?}$			
Eigenvalues $\hat{?}$			
(0.7867	0.0494)		
Standardized $\hat{?}$??eigenvectors		Standardized $\hat{?}$ coefficients	

LRP	***LER***	***LRP***	-0.22	0.0016
1.00	-0.88	***LER***	-0.06	-0.1485
-0.54	1.00			

Therefore, the equivalent error correction mechanism is the following:

$$CI = LRP -0.88\ LER$$

The coefficient restriction, $?$=-1, is accepted at the 5% level from the $?^{2}$ test[13]. As for the monthly and quarterly cases, the cointegrating vector is defined as in equation (5).

V. Monthly, Quarterly and Annual Modelling

In this section, the international demand for tourism in Sassari Province is estimated by employing three data frequencies. The aim is to identify similarities and differencies in terms of elasticities, and more in general, in the explicative ability of the variables included in the model. The "pre-modelling" phase allows to include in the model both short and long run information.

The monthly model for the sample period from 1972:1 to 1995:12 is defined by equation (6).

$$LA_t = a_1 + a_2\ LA_{t-1\ldots} + a_3\ LPR_{t\ldots} + a_4\ LSP_t\ ..+ a_5\ DLER_t\ ..+ a_6\ DLRP_{t..} + a_7\ CI_{\ t-1} + \\ + a_8\ Easter + a_9\ Seas + a_{10}\ Dummies + a_{11}\ Trend + e_t \qquad (6)$$

An unrestricted model 13 lags[14] is run, with '..' in the subscript indicating extra lag terms. The model includes: the dependent variable and each of the exogenous variables[15], the one lagged error correction mechanism (CI_{t-1})[16], a constant, a trend that might pick up the deviations of the (log) industrial production from the trend, 11 seasonal dummies (*Seas*), the “Easter” dummy (*Easter*) and four impulse dummies (*i1974p12, i1979p3, i1985p3* and *i1991p11*) that in the preliminary phase have managed to correct non-normality problems in the residuals.

The quarterly model for the sample period from 1972:1 up to 1995:4 is defined by equation (7):

$$LA_t = a_1 + a_2\ LA_{t-1\ldots} + a_3\ LPR_{t\ldots} + a_4\ LSP_t\ ..+ a_5\ DLER_t\ ..+ a_6\ DLRP_{t..} + a_7\ CI_{\ t-1} + \\ + a_8\ LW + a_9\ Seas + a_{10}\ i1985p3 + a_{11}\ Trend + e_t \qquad (7)$$

The explanatory variables included in the model are: the above mentioned dependent and exogenous variables, the cointegrating vector (CI_{t-1}), the (log) weather variable (*LW*), a time trend and, finally, 3 quarterly seasonal dummies (*Seas*). An impulse dummy, *i1985q1*, is also added after inspecting the

---

[13] The results for the restriction test on the coefficient is: $?^{2}(1) = 1.5489\ [0.2133]$

[14] Note that a 13 lag model is accepted by the joint *F*-test and it is also suggested by the HQ criterion.

[15] The exogeneity condition for the economic explanatory variables is based on the following assumptions. North of Sardinian tourism is only a relatively small fraction of the origin countries' income, and it can be argued that the current value of the income variable (*LPR*) is not influenced by the current value of the endogenous variable (*LA*). Moreover, in Sassari Province registered accommodation, prices are determined on an annual basis and published at the beginning of the year. Similarly, exchange rates are determined by domestic and international economic conditions. Therefore, there are valid reasons to assume that the explanatrory variables may be regarded as predetermined in the development of an intemational tourism demand model for Sassari Province.

[16] 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*.

residuals[17]. Note that an initial 5 lag model could be reduced to a 4 lag model in accordance to the joint *F*-test and SC criterion.

The annual data used in this study that covers a period of 24 years (1972-1995) is defined by equation (8).

$$LA_t = a_1 + a_2\, LA_{t-1}\ldots + a_3\, LPR_t\ldots + a_4\, LSP_t\ ..+ a_5\, DLER_t\ ..+ a_6\, DLRP_t.. + a_7\, CI_{\ t-1} + + a_8\, LW + a_9\, Trend + e_t \quad (8)$$

The initial model, with one lag, is estimated by regressing the logarithm of the modified series of arrivals (*LA*) on the logarithm of the following variables: *LPR*, *LSP*, *DLRP*, *DLER*, the first lag of the cointegrating vector (CI_{t-1}), the weather variable (*LW*), a time trend (*TREND*). 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 and quarterly data.

After a general-to-specific simplification, as "an efficient way to find a congruent encompassing model" (Mizon 1996, p.123), one has obtained the final parsimonious models as reported in Table 12. In Table 12, short and long run dynamics and diagnostic statistics are presented.

Considering the monthly model in Table 12, 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 satisfactory form specification. However, it has to be noted that the null hypothesis of homoscedasticity for the disturbances is marginally rejected using the White test 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[18], and the final results are reported in Table 12. The lag coefficients of the foreign arrivals, as explanatory variables, present a positive sign. This indicates that foreign tourists are possibly 'psychocentric' (Sinclair and Stabler, 1997) and that the Province of Sassari is viewed as a desirable destination area. This is also consistent with the adjustment of the dependent variable to changes in the right hand side variables. Moreover, the coefficients for the long run elasticities are jointly statistically significant as inferred by the Wald test.

In the quarterly model case, the inclusion of the lagged dependent variable turns out to be statistically not significant, suggesting that the international demand is not influenced by its own history. The final model obtained can, therefore, be considered as a static model. It explains almost 99% of the variation in the number of foreign arrivals. The diagnostic statistics suggest no problems in the residuals. In addition, the same model is re-estimated using 1995(3) to 1995(4) as forecasting sample data; the Chow prediction test statistic do not reject the null hypothesis of parameter constancy.

As for the quarterly model, a final annual static model is achieved. 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 by the Chow test that denotes the coefficients are not constant over the sample period.

---

17 This 0-1 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.

18 Such a correction has been run using Microfit 4.0.

As a final note, by employing the Box and Cox (1974) test, preliminary investigation has shown that the logarithm form has to be adopted in each of the three data models (see Pulina, 2002).

Table 12 Estimated Models Using Monthly, Quarterly and Annual Data

Frequency	Estimated Model
Monthly (short-run elasticity)	LA_t = -2.754 + 0.125 LA_{t-1}+ 0.103 LA_{t-2}+ 0.134 LA_{t-3}+ 0.108 LA_{t-11} (-1.40) (2.40) (2.86) (3.55) (2.20) + 2.569 LPR_{t-3} - 2.007 LPR_{t-7} - 4.746 $RLSP_t$ + 5.214 LSP_{t-11}+ (3.92) (-3.24) (-2.67) (2.36) - 4.533 LSP_{t-12} - 0.338 CI_{t-1} + 0.426 Easter + 1.556 i1974p12+ (-2.07) (-2.07) (6.07) (25.75) - 0.575 i1979p3 + 0.675 i1985p3 - 0.595 i1991p11 + 0.257 Jan + (-8.02) (9.61) (-12.97) (3.24) + 0.631Feb + 1.057 Mar + 1.728 Apr + 2.764 MJJ + 2.507Aug + (5.39) (6.35) (7.94) (7.94) (11.39) + 2.354 Sep + 1.190 Oct + 0.017 Nov (13.71) (8.29) (0.21)
R??=0.98004 F(24, 249) = 559.53 DW = 1.91 AR _F(7,242) = 1.046 ARCH_ F(7, 235) = 0.421 NORM_Chi^2(2) = 3.805 HETER_F(34,214) = 1.646 * RESET_F(1,248) = 1.533 CHOW_F(8,241) = 1.165	
Monthly (long-run elasticity)	LA = -5.204 -0.6388 CI +0.8054 Easter +1.062 LPR -8.967 RLSP + (-8.15) (-10.85) (4.95) (2.62) (-2.47) +1.286 LSP +5.223 MJJ +2.94 i1974p12 - 1.087 i1979p3 + (4.27) (12.02) (6.01) (2.74) +1.276 i1985p3 -1.125 i1991p11 +0.4865 Jan +1.193 Feb (3.15) (-2.82) (3.07) (4.86) +1.997 Mar +3.265 Apr +4.737 Ago +4.449 Sep +2.249 Oct (6.68) (8.48) (14.46) (12.72) (8.54) +0.03249 Nov (0.20)
WALD test Chi^2(18) = 635.14 [0.0000] **	
Quarterly	LA_t = -0.470 + 0.788 LPR_t + 1.424 LSP_t - 4.224 $DLER_{t-4}$ - 0.429 CI_{t-1} + (-0.19) (2.47) (6.01) (-3.79) (-2.02) + 0.551 LW_t + 0.853 i1985q1 - 0.648 GFM + 1.494 AMG + (2.09) (2.09) (-6.69) (20.10) + 2.354 LAS (12.70)
R??=0.9858 F(9, 81) = 624.87 DW = 1.74 AR _F(7,74) = 0.880 ARCH_ F(7, 67) = 0.820 NORM_Chi^2(2) = 0.486 HETER_F(14,66) =0.707 RESET_F(1,80) = 0.233 CHOW_F(2,79) = 1.247	
Annual	LA_t = -5.09 + 2.342 LPR_t + 1.547 LSP_{t-1} - 0.729 CI_{t-1} - 0.021 Trend (-1.23) (3.68) (4.44) (-2.70) (-1.79)
R??=0.9079 F(4, 17) = 41.985 DW = 1.42 AR _F(1,16) = 1.05 ARCH_ F(1,15) = 0.265 NORM_Chi^2(2) = 0.195 HETER_F(8,8) =0.257 RESET_F(1,16) = 5.296 * CHOW_F(1,16) = 5.383 *	

Notes: 1) *t-value* in parenthesis. 2) *RLSP* (monthly model), coefficient restriction between the first and second lag of the substitute price, *LSP*. Such a coefficient restriction has been accepted by the joint *F*-test and suggested by the SC criterion. 3) *MJJ* (monthly model), seasonal dummy created giving the value of 1 to May, June and July and the value zero to the other months. Such a coefficient restriction has been accepted by the joint *F*-test and suggested by the SC criterion.

VI. 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 13 summarises the main results.

Table 13 Short and Long Run Elasticities for the International Demand of Tourism

Elasticities	Monthly Model (288 obs.)	Quarterly Model (96 obs.)	Annual Model (24 obs.)
INCOME (long run)	1.06 (2.62)	0.79 (2.47)	2.34 (3.68)
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)	- 0.43 (-2.02)	- 0.73 (-2.70)
CI (short run)	- 0.34 (-2.07)	-	-
SUB.PRICE(long run)	1.29 (4.27)	1.42 (6.01)	1.55 (4.44)
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 (Pindyck and Rubinfeld, 1991).

The long run income elasticity shows different values with respect to the data frequency that 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 numbers 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 survey by Blackwood & Partners (1994), where the tourists were asked 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.

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 of 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 has 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 to 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 negatively 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.

In general, the short run price changes do not 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 with the expected 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.

Contrasting results appear for the nominal substitute price elasticities, which present an unexpected positive sign. However, as shown in Pulina (2002), the inclusion of a real substitute price does not improve the results, given the persistence of a positive elasticity. On further investigation the inclusion of a disaggregated real substitute price, for each of the competitors, shows a better specification for France and Portugal. Further work needs to be attempted in this direction.

For the extra-economic variables some remarks are due. From Table 12, the international demand of tourism in Sassari Province is highly influenced by seasonality and periodic events. In this respect, it 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. Turning to the quarterly model, the climate variable (*LW*), turns out statistically significant and in the long run this variable has a positive impact on the international demand for tourism. This information can be used to adopt marketing policies of de-seasonalisation in Sassari Province. 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.

VII. Conclusions

In this paper, an empirical investigation of the international tourism demand to the north of Sardinia for the period between 1972 and 1995 has been presented. Several concepts have been used such as: trading-day factors, non-stationarity, cointegration and short and long run dynamics.

The main proposition under investigation is to consider whether similarities exist amongst monthly, quarterly and annual time series models. The relatively large number of observations available when using monthly data has allowed one to test the possible existence 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. 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 (T=24 in this case). Nevertheless, testing for cointegration has revealed similar findings using any of the three frequencies. Notably, by using the Johansen analysis, a cointegrating relationship has been suggested for the relative price and the weighted exchange rate. The results obtained justify, statistically, the separation of prices and exchange rate in the short run. Whereas, a real exchange rate should be used in the long run in accordance to the PPP theory.

The preliminary time series investigation and tests enabled the inclusion of both the short and long run information in the models. The "LSE general-to-specific" methodology has been followed for the model estimations. Each of the final restricted models has been evaluated in the light of the most adopted test statistics as well as in terms of diagnostic tests still much negleted in tourism literature. Notably, from a

preliminary investigation, the logarithm form has been validated by the Box-Cox test. The final restricted models have been interpreted in the light of economic theory.
By using monthly data, the seasonal pattern and other extra-economic components could be evaluated empirically. For example, the importance of the Easter break for foreigners has given new information for the private and public sector that can adopt price discrimination for tourist consumers, together with higher standard of quality of the goods and services supplied during "second holiday" periods.
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. Differences in terms of short and long run income and price elasticities as well as in terms of magnitude of the coefficients are of particular importance in tourism. From this empirical analysis, it emerges that monthly data models are more likely to reflect consumers' decisions to be taken several months in advance or sometimes at the last minute in response to "special offers".

References

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 Vaccaro, G. (1992), "*Elementi di Statistica per il Turismo*", FrancoAngeli, Milano.

Baron, R. (1989), *Travel and Tourism Data*, Euromonitor Publication Limited, Wiltshire.

Blackwood & Partners (1994), *Interviste Turisti Stranieri*, ESIT, Cagliari.

Box, G.E.P. and Cox, D.R. (1964), "An Analysis of Transformations", *Journal of the Royal Statistical Society, Series B*, 26(2), 211-43.

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.

Crouch, G.I. (1994a), "The Study of International Tourism Demand: A Survey of Practice", *Journal of Travel Research*, 32(4), 41-55.

Crouch, G.I. (1994b), "The Study of International Tourism Demand: A Review of Findings", *Journal of Travel Research*, 33(1), 12-23.

D.E.I.S. (1995), *Il Settore Turistico in Corsica e Sardegna, Un'analisi comparata*. Università degli Studi di Sassari.

Dickey, D.A. and Fuller, W.A.(1981), "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, 49(4), 1061-1072.

Doornik, J.A. and Hendry, D.F. (1996), *PC FIML 9.0 - Interactive Econometric Modelling of Dynamic Systems*, Institute of Economics and Statistics, University of Oxford.

Doornik, J.A. and Hendry, D.F. (1996), *PC GIVE 9.0 - Empirical Econometric Modelling Using*, International Thomson Business Press, London.

Engle, R.E. and Granger, C.W.J. (1987), "Co-Integration and Error Correction: Representation, Estimation, and Testing", *Econometrica*, 55(2), 251-276.

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.

García-Ferrer, A. and Queralt, R.A. (1997), "A Note on Forecasting International Tourism Demand in Spain", *Departamento de Anàlisis Econòmico, Madrid*.

Gonzàles, P. and Moral, P. (1995), "An Analysis of the International Tourism Demand in Spain", *International Journal of Forecasting*, 11, 233-251.

Gonzàles, P. and Moral, P. (1996), "Analysis of Tourism Trends in Spain", Annals of Tourism Research, 23(4), 739-754.

Gray, H.P. (1966), "The Demand for International Travel by the United States and Canada", *International Economic Review*, 7(1), 83-92.

Hargreaves, C.P. (1994), *Nonstationary Time Series Analysis and Cointegration*, New York, Oxford University Press.

Hylleberg, S., Engle, R.F., Granger, C.W.J., and Yoo, B.S. (1990), "Seasonal Integration and Cointegration", *Journal of Econometrics,* 44, 215-238.
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 Juselius, K. (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. and Witt, S.F. (2001), "Cointegration Versus Least Squares Regression", *Annals of Tourism Research*, 28(2), 291-311.
Lanza, A. and Urga, G.(1995), "Tourism Specialisation and Economic Growth: A Cross Country Study Using Multivariate Cointegration", *Discussion Paper*, LBS, 10, 1-17.
Lim, C. (1997), "Review of International Tourism Demand Models", *Annals of Tourism Research*, 24 (4), 835-849.
Malacarni, L. (1991), "Analisi Econometrica della Domanda Turistica in Italia", *Politica del Turismo*, 2, 19-29.
Martin, C.A. and Witt, S.F. (1987), "Tourism Demand Forecasting Models: Choice of Appropriate Variable to Represent Tourists Cost of Living", *Tourism Management*, 8, 223-245.
McIntosh, R.W., Goeldner, C.R., and Ritchie, J.R.B. (1995), *Tourism. Principles, Practices, Philosophies,* New York, Wiley.
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.
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.
Pesaran, M.H. and Pesaran, B. (1997), *Working with Microfit 4.0*, O.U.P.
Pulina, M. (2002), Modelling Demand for Tourism in Italy, P.h.D. thesis: Southampton University.
Pindyck, R.S. and Rubinfeld, D.L.(1991), *Econometric Models and Economic Forecasts*, Singapore, International Editions.
Sinclair, M.T. (1998), "Tourism and Economic Development. A Survey", *Journal of Development Studies*, 34(5), 1-51.
Sinclair, M.T. and Stabler, M. (1997), *The Economics of Tourism*, Routledge: London and New York.
Solinas, C. (1992), *Il Turismo nel Nord Sardegna. Una Valutazione dell'Impatto Economico nella Spesa Turistica*, no. 99, CIRIEC.
Song, H., Romilly, P., and Liu, X. (2000), "An Empirical Study of Outbound Tourism Demand in the UK", *Applied Economics*, 32, 611-624.
Syriopoulos, T.C. (1995), "A Dynamic Model of Demand for Mediterranean Tourism", *International Review of Applied Economics*, 9(3), 318-336.
Uysal, M. and El Roubi, M.S. (1999), "Artificial Neural Networks Versus Multiple Regression in Tourism Demand Analysis", *Journal of Travel Research*, 38, 111-118.
Vogt, G.M. and Wittayakorn, C. (1998), "Determinants of the Demand for Thailand's Exports of Tourism", *Applied Economics*, 30, 711-715.
Witt, S. and Witt, C.A. (1992), *Modelling and Forecasting Demand in Tourism*. London, Academic Press.