

**Panel data and tourism demand.
The case of Tenerife
by
Francisco J. Ledesma-Rodríguez*
Manuel Navarro-Ibáñez*
Jorge V. Pérez-Rodríguez**
DOCUMENTO DE TRABAJO 99-17**

October, 1999

* University of La Laguna.

** University of Las Palmas de Gran Canaria.

Acknowledgements: Preliminary versions of this paper have been presented at the XXIII Simposio de Análisis Económico (Barcelona, December 1998), at the 47th International Atlantic Economic Conference (Vienna, March 1999), and at the VI Jornadas de Economía Internacional (Valencia, June 1999). The authors would like to thank the participants for their helpful comments.

Los Documentos de trabajo se distribuyen gratuitamente a las Universidades e Instituciones de Investigación que lo solicitan. No obstante están disponibles en texto completo a través de Internet: <http://www.fedea.es/hojas/publicaciones.html#Documentos de Trabajo>

These Working Documents are distributed free of charge to University Department and other Research Centres. They are also available through Internet: <http://www.fedea.es/hojas/publicaciones.html#Documentos de Trabajo>

Resumen

En este trabajo se hace un estudio de la demanda de servicios turísticos de la isla de Tenerife. Para ello, se llevan a cabo diversas estimaciones aplicando la técnica de panel de datos tanto a modelos de carácter estático como de naturaleza dinámica. En general, los resultados reflejan una reducida sensibilidad del número de turistas alojados frente al tipo de cambio y al coste del viaje. La elasticidad demanda-renta muestra la naturaleza de bien de lujo del producto turístico. Además, los gastos de promoción y en infraestructuras aparecen como significativos, aunque su influencia es reducida. Por último, se realizan diversos ejercicios de simulación y predicción.

PALABRAS CLAVE: demanda de servicios turísticos; datos de panel.

JEL: F19, D12

1. Introduction

The development of International Trade Theory has greatly improved our understanding of the nature of commercial flows. Nowadays, we also know more clearly the causes of the international trade of goods. These causes have grown out of the introduction of imperfect competition into the analysis. Likewise, the theory has gone into more detail in order to examine the effects of international commercial policies applied to the exchange of goods under different market structures. Furthermore, international economists have begun to recognize the importance of location to set the patterns of world trade (Krugman, 1991).

The increasing interest in unknown aspects, or aspects not considered before in international trade, has not been carried into a field which has been kept in the background: international trade in services. This deficiency has made it more difficult to approach the subject from an empirical standpoint. For example, in the index of authors and subjects of the *Handbook of International Economics*, edited by Grossman and Rogoff (1995), the subject "tourism" does not appear¹. Very few international economists have centered their efforts in the analysis of tourism. A greater concentration in this field was to be hoped for, given the growing importance of the services sector in both national economies and international exchanges. Moreover, the bulk of the intellectual work has been directed to the analysis of certain services, i.e., those more closely related to traded goods (transport, insurance, legal services, etc...).

In relation to the reasons for international trade in services, as it was shown by the theoretical study by Deardoff (1985), the principle of comparative

¹ Sapir and Winter (1994) have also called attention about the few references to "services trade" in the *Handbook of International Economics* edited by Jones and Kenen in 1984.

advantage is completely applicable under conditions of perfect competition. In general, the empirical studies support this theoretical result².

Sapir and Winter (1994), following Bhagwati, proposed a four-way classification of international transactions in services, attending to the existence or not of mobility of the suppliers and buyers. Many educational and health services require, as is the case of tourism, the movement of consumers. So, four types of international transactions in services can be distinguished:

“1. Immobile users in one nation obtain services produced by immobile providers located in another nation. This occurs, for instance, in financial services and professional services, where transactions flow via telecommunication networks.

2. Mobile users from one nation travel to another nation to have services performed. This situation is most frequent in tourism, education, health care, ship repair and airport services.

3. Mobile providers from one nation travel to another nation in order to perform services. This situation occurs in certain business services, such as engineering, where frequent or close interaction is not required.

4. Providers from one nation establish a branch in another nation in order to perform services. This is the most common pattern of international service competition, involving frequent and close interaction between buyers and sellers. It is the dominant type in most services, including accounting, advertising, banking, consulting services and distribution.” (Sapir and Winter, 1994: p. 275)

Tourists at a holiday resort need to buy the basket of goods and services they would usually buy where they live; they also need goods and services

² In relation to tourism activity in Tenerife, the great influence of the natural resources endowment on market results is quite obvious.

appropriate for the type of tourism services they are consuming. Due to the wide variety of goods and services which are demanded by tourists at a resort, it is very difficult to define a “tourism” sector in the economy. On the contrary it is more adequate to refer to tourism as an activity which requires the contributions of many different industries. This generates a lot of difficulties in the measure of the real contribution of tourism to GDP. We cannot easily separate the demand of residents from the demand of non-residents in a great variety of sectors.

Furthermore, the concept of the tourist activity from the demand-side has to be completed by eliminating all other motives to travel. The decision to travel may be due to other reasons (health, education, work, etc...) which are quite distinctive from the movement founded solely on pleasure travel.

Tourism, as an activity that demands from many different industries, gives market power to those who have been able to join different goods and services in a package and thus offer inclusive tours (IT) to potential consumers. ITs are the most common product on islands. In fact, the sun and beach segment of the market on islands has always been dominated by tour operators (TO). The role played by TOs reduces the information asymmetries common to services markets. This is why the problems of moral hazard and adverse selection that consumers face have made necessary the use of reputation to convey quality. This is something which is better accomplished by selected suppliers.

The movements of tourists produce effects on the physical and cultural milieu of the receptive societies. These external effects, which can be positive and/or negative, have as a consequence many inefficiencies³.

³ This is why the correct measurement of tourism requires the complete valuation of all the costs generated by this activity.

Another characteristic of tourism, and of services in general, is the impossibility of stocking; production must be necessarily equal to the rate of sales at every moment. This characteristic is usually associated with the existence of high fixed costs for the firms which provide the tourist services. This is why it is so important to measure the degree of utilization (i.e., occupancy rate) of the productive assets.

The production of tourist services needs, as does any other good or service, the combination of several productive factors, more or less specialized. In any case, in the production function of tourism, public goods and services have a greater weight than in other non-tourism markets. Precisely, personal security as well as social and political stability of the resorts (jointly with the quality of transport, etc...) have a central role in the choice among tourist destinations.

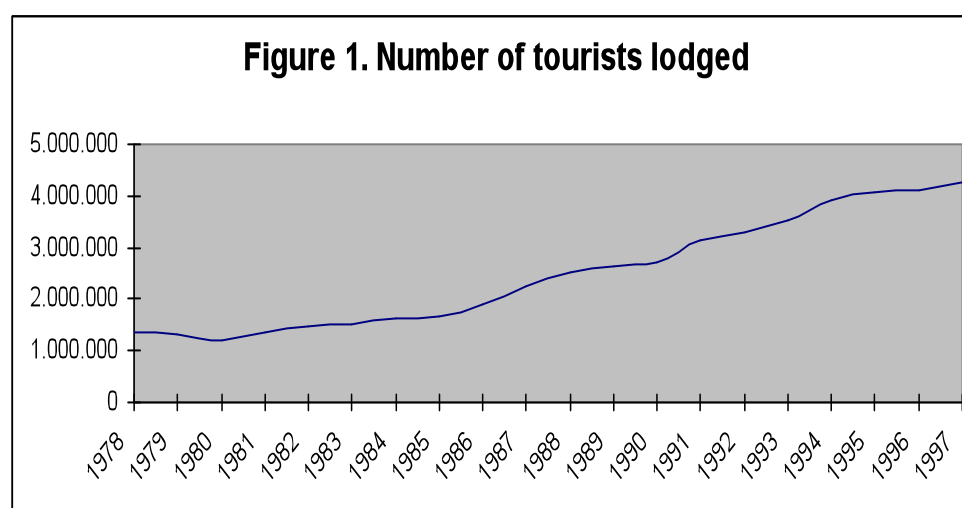
Tourism activity is, for everything mentioned so far, quite difficult to define. We must always take into account the great diversity of tourist products: the Paris product is very different from the Tenerife product.

In section 2 we describe the characteristics of the tourism in Tenerife. In section 3 we carry out some estimations for the tourism demand in Tenerife, using different panel data techniques. Moreover, we present the results of some prediction exercises and simulate some scenarios. In the last section, we indicate the main conclusions.

2. Tourism activity in Tenerife

Tenerife is one of the most important destinations for European sun and beach tourism. In fact, this activity has grown annually in the island at a rate of 10.8% during the last twenty years: Tenerife has become one of the main suppliers of this type of tourism; the island represents almost 20% of the lodging offered during winter in Europe. In the summer, Tenerife's share falls to around 7% of the market, which is still quite significant (Navarro and Becerra, 1998).

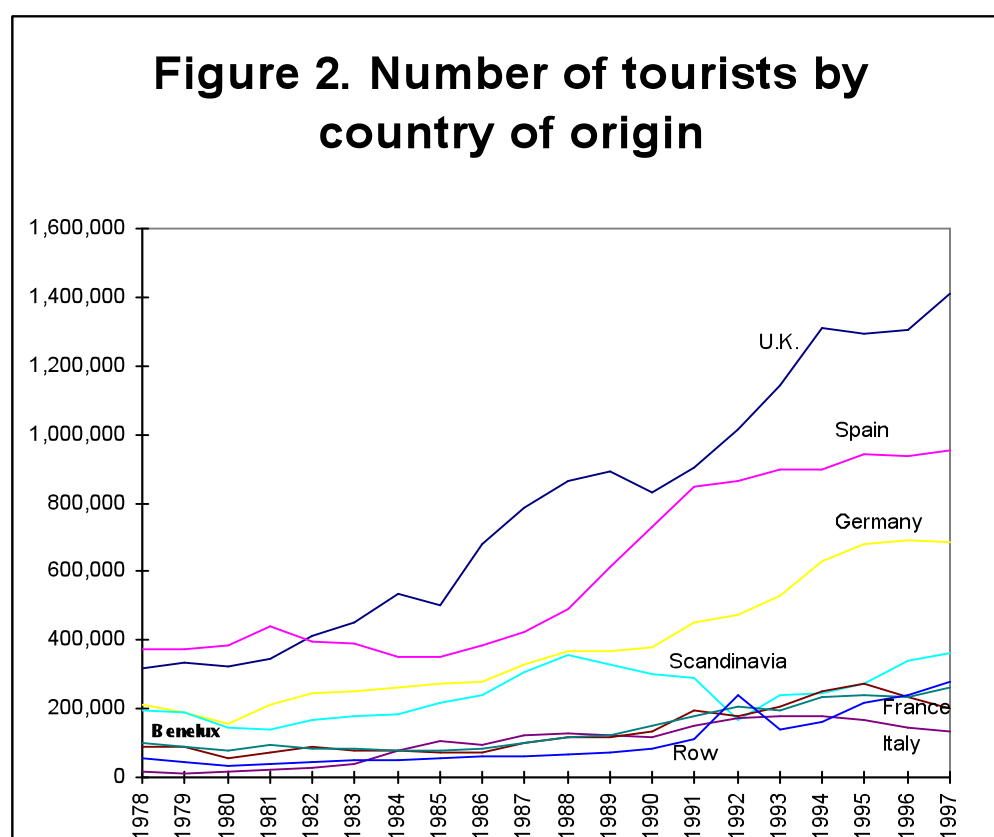
The evolution of tourism during the last two decades has specialized the island's economy even more. The share of services has grown from 61.3% of GDP in 1973 to 78.4% of GDP in 1993. Tourism has taken the leading role in the increase of the services sector's importance in the island's economy. The dynamics of Tenerife, as well as that of the rest of the Canary Islands, appear to be more dependent on the European rather than the national economy. Moreover, and probably due to its high specialization in the services sector, Tenerife shows a greater variation between the expansion and the recession periods than the regions of mainland Spain.



In relation to demand, the number of visitors lodged on the island during the last twenty years has grown from 1.35 millions in 1978 to 4.28 millions in 1997, i.e., a 10.8% average increase (Figure 1). There are three main origins of

Tenerife's tourists: Great Britain, Germany, and mainland Spain. The three represented 66.5% of the visitors in 1978 and 71.3% in 1997. The share of the Scandinavian countries (Denmark, Norway, Sweden, and Finland) has decreased from 14.3% in 1978 to 8.4% in 1997, without a reduction in the absolute number. This different evolution of the visitors makes it imperative that any study of Tenerife's tourism must take into account the diverse evolution of visitors according to their country of origin.

Tenerife presents a lesser (and even contrary) seasonality than the rest of the European resorts in the sun and beach market. In fact, its high season has been and still is, the winter. In the November-April period, Tenerife receives about 53% of its visitors. The Canary Islands as a whole do not have any significant competitor in the European market during that season.



This small seasonality of the demand has remained unchanged in the recent past. The aggregate numbers hide an important compensation among the diversity of origins, since the visitors from mainland Spain, who prefer the summer, compensate for the Germans and Scandinavians, who seem to prefer the winter. This fact justifies the study of the demand of tourism services in Tenerife adopting an annual perspective, leaving aside the seasonality problem.

The supply of lodging between 1978 and 1997 has shown a continuous growth. The number of beds has increased at an average rate of 8.2% annually. This shows the great flexibility and ability of the suppliers to follow closely the evolution of the demand for lodging. This important growth in the supply of lodging has mostly been caused by apartments, as they represent 57.2% of lodging in 1997, up from 43.5% in 1978.

Another important aspect of tourism in Tenerife is the high degree of repetition of those who go to the island. This indicates the loyalty of the consumer to the product Tenerife. Moral hazard is overcome by the reputation acquired by the island.

3. Estimation of tourism demand

In studies carried out in Spain, the output variable has been the revenue generated by tourism activity (Padilla, 1988; Buisán, 1995). The independent variables chosen have been those usually utilized in the estimation of tourist demand, i.e., a price variable and an income variable. On the other hand, the techniques of estimation used in previous studies of tourism in Spain have been quite varied. Thus, Padilla (1988) chose a transfer function, while González and Moral (1993) applied the Kalman filter.

Nevertheless, almost 70% of the studies that try to estimate tourism demand, as it is showed by Crouch and Shaw (1992), have chosen the number of visitors as the variable to explain. One of the reasons for this choice is the relative scarcity of data about the average spending of tourists. Moreover, most of the studies have used the number of people going through airports, borders, or ports, without considering the real reasons for travelling. The latter problem has been solved in this paper by only taking into account the number of visitors lodged in the tourist areas.

In the study of tourism demand we have used, as the dependent variable, the number of visitors lodged in hotels and apartments on the island. The data has

been supplied by the Cabildo of Tenerife and gives us information about visitors from thirteen countries: Germany, United Kingdom, Spain, Sweden, Norway, Finland, Netherlands, Belgium, Austria, France, Italy, Denmark and Switzerland.

The main exogenous variables are income, the price of the barrel of oil and the exchange rate. The income variable we have used is GDP per capita of each country in real terms (using 1990 constant prices). We have also introduced the price of the barrel of oil divided by the price index of each country of origin as a proxy to the cost of the trip, and the exchange rate of the peseta with respect to the currency of each country of origin. Furthermore, relative prices have been defined as the consumer price index of Tenerife divided by the index of each country of origin. The price and exchange data has been taken from IMF's International Financial Statistics and from the database Tempus of the *Instituto Nacional de Estadística* of Spain. Moreover, we have introduced the promotion expenditure of the island trying to capture the non-price competition, as well as an infrastructure variable that recognizes the relevance of public inputs in the tourists decision⁴.

The general form of the equation that gives us the number of tourists T_{it} is:

$$T_{it} = f(T_{it-1}, Y_{it}, PB_{it}, E_{it}, PE_t, INF_t)$$

where i is the country of origin and t is the year. As we can see, there are two groups of variables: those that depend both on time and the country of origin, and those that only depend on time. Y_{it} is real GDP, PB_{it} is the index price of the barrel of oil divided by the price index of each country of origin, E_{it} is the exchange rate of the peseta with respect to the currency of each country, PE_t is the Cabildo of Tenerife's expenditures for tourism promotion, and INF_t is the capital stock (BBV) in infrastructures (ports, airports, and roads). All the variables are

expressed in logarithms, which allows us to obtain the demand elasticities with respect to all the relevant variables.

In the present research, and given the richness of information provided by the different origins of the visitors to Tenerife, the econometric technique utilized has been panel data, which allows for the control of individual heterogeneity. This technique reduces the problem of colinearity, and provides more degrees of freedom, making it easier to infer the outcome when samples are small. Moreover, panel data also achieves a better representation of the adjustment dynamics, by identifying and measuring the effects which are not detected in the studies with cross-section data or with pure time series. These techniques allow for the construction and comparison of models which take into account the existence of more complex behavior (Maddala, 1993; Baltagi, 1995). There are also some limitations in the panel data technique: the design of the database, the distortions produced by the errors of measure, the selection problems, and the length of the time series.

Taking into consideration these aspects, we have built a panel with annual data starting in 1979 and going through 1997. We differentiate between two important panel data models for the empirical study of tourism demand: a static model, which considers heteroskedastic and autocorrelated disturbances, and a dynamic one, which includes a lagged endogeneous variable as a regressor.

A central objective of this paper is to ascertain the response of the decision of the tourists to several relevant variables through the correct specification of tourism demand, considering the thirteen main countries of origin of Tenerife's

⁴ In preliminary versions of this paper we have used the consumer price index CPI of Tenerife with respect to the CPI of every country of origin, as well as Tenerife's CPI with respect to alternate destinations. However, there were

tourists. We are also going to study the sensitivity of the different estimation techniques when there are specification errors, specially those related to the existence of serial correlation in the residuals, or to the presence of specific individual effects, or to both of them. This is the reason why the common element is always the one-way error component model, which only considers the specific effect for each country. We are not going to take into account the specific effect for each year due to the specification problems derived from adding 19 more time dummy variables to the model.

The rest of this section is organized as follows. First, we estimate some panel data static models. Second, we present the results of the estimation of several dynamic models. Lastly, we carry out several prediction exercises and simulate some scenarios.

3.1. Panel Data Static Model.

There are a number of ways in which to analyze the information provided by the panel data. We can estimate a fixed or a random intercept model, or a model of variables with different coefficients for each country.

Suppose that we have a panel of N countries. We observe the endogenous variable, T_{it} , and a vector of explanatory variables, x_{it} , in each time period. Let us consider the following linear equation, which is a static panel model represented by:

$$\begin{aligned} T_{it} &= \alpha + \beta_i' x_{it} + \varepsilon_{it}, & i = 1, \dots, N; t = 1, \dots, T \\ \varepsilon_{it} &= \mu_i + u_{it}, \end{aligned} \quad [1]$$

problems of significance and the results were quite implausible.

where ε_{it} is a zero mean residual and μ_i are the unobservable individual specific effects (tourist preferences,...) that are invariant over time t for the country i and are distributed as $N(0, \sigma_\mu^2)$. u_{it} denotes the remainder disturbances, i.e., a vector of possibly serially correlated disturbances that are distributed as $N(0, \sigma_u^2)$. The model can be estimated using OLS, and the resulting coefficient estimates and standard errors will be consistent if x_{it} is exogenous and ε_{it} is homoskedastic and serially uncorrelated [i.e., $E(\varepsilon_{it}/x_{it})=0$, $E(\varepsilon_{it}^2)=\sigma_\varepsilon^2$ and $E(\varepsilon_{it}\varepsilon_{js})=0$ for all $i \neq j$ or $t \neq s$]. In most empirical applications using panel data these conditions are not satisfied. In particular, if there is unobserved individual heterogeneity, then the errors are likely to be correlated throughout time for each individual, invalidating the assumption that $E(\varepsilon_{it}\varepsilon_{jt})=0$ for all $t \neq s$. In this sense, if the individual effects are random with respect to the observed explanatory variables, $E(\mu_i/x_{it})=0$, then OLS provides consistent but inefficient parameter estimates. The Generalized Least Squares (GLS) estimator provides both efficient and consistent estimates. This is called the random effects (RE) estimator, i.e., the Balestra-Nerlove estimator.

On the other hand, if $E(\mu_i/x_{it}) \neq 0$, then the individual effects are correlated with the explanatory variables, and neither the OLS or the RE estimator will be consistent. The traditional approach to overcome this problem is to eliminate the individual effects from the sample by transforming the data into deviations from the individual means. Given that the individual effects are correlated with the explanatory variables, individual constants exist, and the estimators are called fixed effects (FE) [See Appendix I for an overview of these methods]⁽⁵⁾. This is the classical perspective on panel data models.

⁵ “Unfortunately, the OLS coefficient estimates from the transformed data (FE estimator) have two important defects: (1) all time-invariant variables are eliminated by the transformation, and (2) under certain circumstances,

We will now present some other methods of estimation which, based on FE and RE, allow for heteroskedasticity and autocorrelated disturbances overcoming some of the problems due to the erroneous specification. Thus, we will be able to obtain unbiased, efficient, and consistent estimations of the parameters and their standard errors.

3.1.1. *Heteroskedasticity and Contemporary Correlation.*

The assumption of homoskedasticity and non-contemporary correlation, as in equation [1], can be too restrictive for many economic relationships. This is why we are going to estimate the parameters of a system of equations using OLS, where all observations are given equal weights⁶. For example: a) Cross-section Weighted Regression where we will estimate a feasible GLS specification assuming the presence of cross-section heteroskedasticity using estimated cross-section residual variances. The equation weights are the inverses of the estimated equation variances, and are derived from unweighted estimation of the parameters of the system⁷. And b) Seemingly Unrelated Regression (SUR), or Zellner's method, that is a feasible GLS specification correcting for both cross-section heteroskedasticity and contemporary correlation in the errors across equations with an estimated cross-section residual covariance matrix, which is based upon

the FE estimator is not fully efficient since it ignores variations between individuals in the sample”(Hausman and Taylor, 1981).

⁶ If there are no restrictions in the system, these methods are identical to estimate each equation using single-equation ordinary least squares or 2SLS. The use of system estimation techniques has a problem; that the poor estimates for the misspecification equation may “contaminate” estimates for other equations.

⁷ This method yields identical results with unweighted single-equation least squares if there are no cross-equation restrictions.

parameter estimates of the unweighted system. This specification is sometimes referred to as the Parks estimator.

We also employ some procedures that include iterations to convergence to control the feasible GLS estimation. For weighted least squares and SUR, there is an additional estimation that involves the procedure for computing the GLS weighting matrix and the coefficient vector: the method of iterating over coefficients and the one-step weighting matrices. This method carries out a first-stage estimation of the coefficients using the identity matrix. It uses the starting values obtained from OLS and iterates until the coefficients converge. If the model is linear, this procedure involves a single OLS regression. The residuals from this first-stage iteration are used to form a consistent estimate of the weighting matrix. In the second-stage of the procedure, we use the estimated weighting matrix in forming new estimates of the coefficients.

The specification issue is whether the conditional mean of the μ_i can be regarded as independent of the x_{it} 's, i.e., whether $E[\mu_i/x_{it}] = 0$. A natural test of the null hypothesis of independent μ_i 's is to consider the difference between the two estimators, $\hat{q} = \hat{\beta}_{FE} - \hat{\beta}_{GLS}$. The specification test is $m = \hat{q}' [\text{var}(\hat{q})]^{-1} \hat{q}$, where $\text{var}(\hat{q}) = \text{var}(\hat{\beta}_{FE} - \hat{\beta}_{GLS})$. If the RE specification is adequate the two estimators should be near each other, rather than differencing widely as has been reported sometimes in the literature, as a virtue of the RE specification.

Table 2 shows the results of the estimation of equation [1], considering only the variables that present a time and cross-sectional behavior. We can see the estimations of the demand elasticities using the two types of techniques: FE and RE.

In all the models, we consider identical slope parameters for all the equations, i.e., we accept the hypothesis that all the parameters are constant for every country of origin, such that:

$$\beta_{i,1} = \beta_{i,2} = \dots = \beta_{i,13} = \beta_i, \quad i = 1, 2, 3$$

Table 2 also shows the determinant value of the estimated residuals from each model $\left| \hat{\Sigma}_u \right|$, the log-likelihood, and the Wald test for the joint significance of FE, $W(\mu_1 = \dots = \mu_{N-1} = 0)$. Moreover, the table presents the Hausman test for FE (FE-OLS, FE-GLS y FE-SUR) *versus* RE.

Table 2. - Estimate of the double-log panel model with heteroskedasticity and

Contemporaneous correlation. Restricted coefficients.

Sample period 1979-1997. Equation:

$$\log T_{it} = \alpha + \beta_1 \log Y_{it} + \beta_2 \log PB_{it} + \beta_3 \log E_{it} + \varepsilon_{it}$$

$$\varepsilon_{it} = \mu_i + u_{it}$$

Parameters	FE estimator			RE estimator
	OLS	GLS ^b	SUR ^b	GLS
α				6.9005 (6.04)
β_1	2.3595 (6.97)	3.4096 (13.9)	1.6277 (96.7)	0.9029 (4.12)
β_2	-0.3815 (-5.76)	-0.0832 (-1.77)	-0.3467 (-63.9)	-0.6439 (-11.4)
β_3	0.1345 (1.26)	0.1888 (2.38)	0.2320 (20.7)	0.1906 (1.93)
$\left \hat{\sum u} \right $	1.12e-25	1.72e-25	1.38e-26	1.22e-25
Log L	237.31	186.86	245.42	
W($\mu_1=\dots=\mu_{N-1}=0$)	221.09 [0.00]	283.97 [0.00]	1385.08 [0.00]	
Hausman test	20.676 [0.00]	34.577 [0.00]	87.54 [0.00]	

Note: In parenthesis appear robust t-values and in brackets appear p-values.

Superscript “a” represents the classic method and “b” represents the same method but with iteration to convergence. We control the iterative process by specifying convergence criterion and the maximum number of iterations.

Given these results, we can say that the FE model is better than the RE, as can be derived from the Hausman tests; furthermore, the best statistical representation of the behavior of the tourists lodged in Tenerife is the estimation provided by FE-SUR. This conclusion is based on the criterium of the minimum value of the disturbances determinant and the greatest value of the log-likelihood. In this form, the estimation of the parameters of the model is carried in an efficient and robust manner, since it iterates towards convergence and there is simultaneously heteroskedasticity and contemporary correlation of the errors (Park estimator). From these estimations we can observe that the number of tourists is quite sensitive to income, indicating the nature of luxury product of tourism. Moreover, this variable shows a small elasticity with respect to the exchange rate and the cost of the trip. The latter result is similar to those obtained in the empirical literature (Crouch, 1994).

We can consider that individual heterogeneity exists. Thus, μ_i capture the effects of the non-observed or omitted variables (the expenditure for tourism promotion, the capital stock in infrastructure, etc.) which are all clearly correlated with GDP.

Although the permanent effects have been eliminated by the estimation of FE, we cannot be sure that the model is completely free of specification errors.

3.1.2. *Error dynamics with AR(1) disturbances.*

There are two ways of including dynamic elements in panel data models: by introducing autocorrelation in the errors and by adding lagged dependent variables as regressors. These two avenues are commonly known as “error dynamics” and “equation dynamics”. Following Lillard and Wallis (1978), we generalize the error component by assuming that the remainder disturbances (u_{it}) follow an stationary autorregressive process of order one AR(1), in the form:

$$u_{it} = \rho u_{it-1} + v_{it}, |\rho| < 1, v_{it} \approx N(0, \sigma_v^2)$$

In this way, a static model can exhibit dynamic errors, indicating the existence of serial correlation of the disturbances between two different time periods. Model [1] can be too restrictive for some economic relations since it shows a constant variance.

In this subsection, we estimate equation [1] with the variables that show a time and cross-sectional behavior, imposing an AR(1) process for the errors. The results of the estimation of FE by non-linear least squares NLS and by non-linear

two stages least squares N2SLS (Fair, 1984, pp. 210-214) appear in table 3. In the latter case, the instruments vector is $\log Y_{it-1}, \log Y_{it-2}, \log PB_{it-1}, \log PB_{it-2}$, which contains income and the cost of the trip lagged up to two periods due to the inclusion of AR(1) in equation [1]. Table 3 also shows the value of the residual's determinant estimated by each model $\left| \hat{\Sigma}_u \right|$ and the log-likelihood. Moreover, the table presents the ρ coefficient and the t-ratio for the null hypothesis $\rho=1$ ⁸. This allow us to see the level of significance of the serial correlation by way of a stationary stochastic process AR(1) and the proximity of a unit root in said process. The estimated parameters are quite similar to those of table 2 (except for the exchange rate coefficient) indicating a smaller elasticity. The serial autocorrelation of the errors is not close to one, as the t-Student test suggests. Since the AR(1) effects are statistically significant, relevant variables are being omitted. Thus, the static model has been misspecified and we must consider a dynamic version of the model.

⁸ The statistic ρ should be verified with the unit roots tests for panel data. In the next subsection, we present the results of an asymptotic test of unit roots.

Table 3. - Estimate of the double-log panel model with AR(1) disturbances.

Restricted coefficients. Sample period 1979-1997. Equation

$$\log T_{it} = \alpha + \beta_1 \log Y_{it} + \beta_2 \log PB_{it} + \beta_3 \log E_{it} + \varepsilon_{it}$$

$$\varepsilon_{it} = \mu_i + u_{it}, u_{it} = \rho u_{it-1} + v_{it}, |\rho| < 1$$

Parameters	FE estimator with AR(1) disturbances	
	NLS	N2SLS
β_1	1.7467 (4.89)	1.4930 (3.96)
β_2	-0.3436 (-6.21)	-0.2684 (-2.44)
β_3	0.0200 (1.21)	0.0587 (0.69)
$\left \hat{\sum u} \right $	4.84e-28	1.02e-28
Log L	248.01	
$W(\mu_1 = \dots = \mu_{N-1} = 0)$	3899.14 [0.00]	3933.42 [0.00]
ρ	0.8040 (22.03)	0.6491 (11.15)
$t(\rho=1)$	-5.3705	-6.0276

Note: In parenthesis appear robust t-values and p-values in brackets.

We control the iterative process by specifying convergence criterion and the maximum number of iterations.

As can be observed, even with the introduction of autorregresive errors, the results show again a high elasticity of the number of tourists lodged with respect to income. The elasticity with respect to the cost of the trip is quite similar to the one obtained from the estimation of the model that did not take into account serially correlated errors. The number of tourists seems less sensitivity to the exchange rate than in the model presented in the previous subsection; in any case this parameter shows problems of significance.

3.2. *Dynamic model for panel data ('Equation Dynamic')*

The presence of an endogenous lagged variable in any model, as well as the existence of autocorrelated errors⁹, would imply that many economic relations among variables in a static model would become dynamic. In this way, the panel data model can also facilitates the understanding of the adjustment dynamics [Balestra y Nerlove (1966); Arellano y Bond (1991)].

If we include a new variable that recognizes the influence of the number of tourists in the past upon the evolution of the current number of tourists, the panel model becomes dynamic. Moreover, the one-period lagged number of tourist variable also exhibits the influence of past decisions on current decisions of the tourists. As we mentioned before, the high degree of repetition is a mechanism that permits suppliers to acquire a reputation that overcomes the problems derived from information asymmetries. The significance of the one-lagged dependent variable with one lag could reflect the importance of this mechanism.

In the rest of this section we carry out an analysis of the existence of unit roots; we also propose a dynamic model for the lodged tourists, which is utilized to make some exercises of prediction and simulation.

⁹ A static model, such as [1], with errors that follow an AR(1) process, can become dynamic by multiplying all the equation variables by $1-\rho L$, where L is the lag operator. For instance, the number of lodged tourists, T_{it} , becomes $(1-\rho L)T_{it}=T_{it}-\rho T_{i,t-1}$, while the explicative variables would be equal to $(1-\rho L)x_{it}=x_{it}-\rho x_{i,t-1}$.

3.2.1. *Unit roots test.*

Harris and Tzavalis (1999) introduce some asymptotic unit roots tests for panel model where residuals follow an AR(1) and the time dimension is fixed, which allows FE and deterministic individual trends. These tests employ a normalized OLS estimator of the autorregressive coefficient, which corrects itself for inconsistencies. The latter grows as the result of the inclusion of FE and individual trends in the regression model, and is only influenced by the sample size. The tests have the normal as the limiting distribution.

In general, we consider two types of data generation process (DGP): $\log y_{it} = \omega_i + \rho \log y_{i,t-1} + u_{it}$, and $\log y_{it} = \omega_i + \phi_i t + \rho \log y_{i,t-1} + u_{it}$, where y_{it} is some relevant variable; ω, ϕ y ρ are parameters; t is a trend and $u_{it} \approx Niid\left(0, \sigma_u^2\right)$. The null hypothesis is the existence of a unit root in the DGP, i.e., $\rho=1$, while the alternative hypothesis is $|\rho|<1$, i.e., the process is stationary. In the first model, the hypothesis is a non-stationary process with heterogeneous constants and the alternative hypothesis is a stationary process with heterogeneous intercepts. The second model, which includes heterogeneous fixed effects and individual trends provides a test with greater ability to distinguish between the null hypothesis that each series follows a random walk with drift and the alternative hypothesis that each series is stationary around a deterministic trend.

The results of the estimation appear in table 4 where we see the two models, one with constants for all countries and without trend, and another with a constant

and trends. The empirical quantiles are based on the limiting distribution to which the statistic built for the normalized autorregressive coefficient converges¹⁰.

Table 4. Test of unit roots in panel data with a time dimension fixed (Harris and Tzavalis, 1999). Fixed effects and individual deterministic trends. Normalized values.

Variables in logs	Heterogeneous Constants		Heterogeneous constants and trends	
Number of tourists	1.8131		0.5703	
GDP per capita	3.2474		1.3794	
Cost of the trip	1.5550		-0.7018	
Exchange rate	-2.0153		-2.2726	
Critical values	T=10 N=10	T=25 N=10	T=25 N=10	T=25 N=10
1%	-2.97	-3.14	-2.48	-2.80
5%	-2.02	-2.14	-1.77	-1.95
10%	-1.57	-1.66	-1.41	-1.52

Note: Different values of T and N are used due to the critical values of the empirical samples (T=18, N=13) are not tabulated. The empirical quantiles appear in tables 1b and 1c of Harris y Tzavalis (1999).

The critical values chosen are T=N=10, and T=25 and N=10. These values were selected to test the sensitivity of the critical values to the sample sizes nearer to T=19 and N=13, since our sample sizes have not been tabulated. In this sense, the hypothesis that the series are integrated of order 1, considering the existence of constants and individual trends when utilizing the FE estimator, is not rejected at a

¹⁰Levin and Lin (1993a), Quah (1994), and Breitung and Meyer (1994) have gone much deeper into the study of unit roots, assuming that both T and N tend to infinity. Contrary to these studies, Harris and Tzavalis (1999, pp. 206-207) assume that the panel time dimension is fixed. The normalized distribution of the statistic, when we use

the model $\log y_{it} = \omega_i + \rho \log y_{i,t-1} + u_{it}$ is $\sqrt{N}(\hat{\rho} - 1 - B_2) \xrightarrow{L} N(0, C_2)$, where

$B_2 = -3(T+1)^{-1}$, $C_2 = 3 \left(7T^2 - 20T + 17 \right) \left(5(T-1)(T+1)^3 \right)^{-1}$. On the other hand, when we use the model

$\log y_{it} = \omega_i + \phi_i t + \rho \log y_{i,t-1} + u_{it}$, the normalized distribution of the statistic is $\sqrt{N}(\hat{\rho} - 1 - B_3) \xrightarrow{L} N(0, C_3)$,

where $B_3 = -15(2(T+2))^{-1}$, $C_3 = 15 \left(93T^2 - 728T + 1147 \right) \left(112(T-2)(T+2)^3 \right)^{-1}$.

5% significance level. Therefore, the necessity to differentiate the involved series in equation [1] could be justified, in order to use stationary series.

3.2.2. *A dynamic model and time-variant regressors for lodged tourists.*

The most simple dynamic panel model with exogenous variables has the following form:

$$\begin{aligned} T_{it} &= \gamma T_{it-1} + \alpha + \beta_i' x_{it} + \varepsilon_{it}, i = 1, \dots, N; t = 1, \dots, T \\ \varepsilon_{it} &= \mu_i + u_{it} \end{aligned} \quad [2]$$

where T_{it} is a function of μ_i and T_{it-1} . T_{it-1} is correlated with the error term, and so the OLS estimator is biased and inconsistent even if u_{it} is not serially correlated. The FE estimator is biased, and its consistency will depend upon the time period. The same problem occurs with the RE estimator. If we take into account the endogeneity of the lagged dependent variable, the valid estimation method is referred to as the instrumental variables technique. This technique leads to consistent but not necessarily efficient estimates of the parameters in the model because it does not make use of all the available moment conditions, and it does not take into account the differenced structure of the residual disturbances. Some methods within the instrumental variables class are the TSLS and the 3SLS. TSLS and 3SLS are applied to the FE estimations, and can be calculated consistently using valid instruments, denoted by Z_{it} , that must satisfy the strict exogeneity condition, $E(u_{it}/Z_{it}) = 0$ for all t . If the strict exogeneity condition does not hold, then $E(u_{it}/Z_{it}) \neq 0$ and the parameters cannot be estimated consistently.

The Weighted–Two Least Squares (W2SLS) is an appropriate technique when some of the right-hand side variables are correlated with the error terms, and there is heteroskedasticity but no contemporary correlation in the residuals. The

Three-Stage Least Squares (3SLS) is the two-stage least-squares version of the SUR method. It is an appropriate technique when right-hand side variables are correlated with the error terms and there is both heteroskedasticity and contemporary correlation in the residuals¹¹.

Table 5. - Estimate of the double-log panel model with dynamic.

Restricted coefficients. Sample period 1979-1997. Equation
 $\log T_{it} = \gamma_1 \log T_{i,t-1} + \beta_1 \log Y_{it} + \beta_2 \log PB_{it} + \beta_3 \log E_{it} + \varepsilon_{it}$
 $\varepsilon_{it} = \mu_i + u_{it}$

Parameters	FE-2SLS	FE-W2SLS ^b	FE-3SLS ^a
γ_1	0.7876 (20.06)	0.7530 (22.40)	0.7796 (35.34)
β_1	0.5757 (2.57)	0.6506 (3.57)	0.5460 (9.00)
β_2	-0.0315 (-1.66)	-0.0803 (-2.28)	-0.0373 (-2.21)
β_3	0.1169 (1.15)	0.1765 (2.90)	0.1258 (3.47)
$\left \sum u \right $	1.80e-29	2.63e-29	1.59e-29
$W(\mu_1=\dots=\mu_{N-1}=0)$	41.228 [0.00]	739.87 [0.00]	458.06 [0.00]
$W(\gamma_1=\beta_1=\beta_2=\beta_3=0)$	1805.32 [0.00]	3217.18 [0.00]	7741.9 [0.00]
$t(\gamma_1=1)$	-5.8558 [0.00]	-7.3477 [0.00]	-9.9909 [0.00]

Note: In parenthesis appear robust t-values and p-values in brackets.

Superscript “a” represent the classic method and “b” represent the same method but with iteration to convergence. We control the iterative process by specifying convergence criterion and the maximum number of iterations.

The estimations of equation [2] for the FE model appear in table 5. We also consider the restricted parameters in all the equations of each country¹². Thus,

$$\begin{aligned} \gamma_{1,1} &= \gamma_{1,2} = \dots = \gamma_{1,13} = \gamma_1 \\ \beta_{i,1} &= \beta_{i,2} = \dots = \beta_{i,13} = \beta_i, \quad i = 1, 2, 3 \end{aligned}$$

¹¹ We apply 2SLS to the unweighted system, enforcing any cross-equation parameter restrictions. These estimates are used to form an estimate of the full cross-equation covariance matrix which, in turn, is used to transform the equations to eliminate the cross-equation correlation. 2SLS is applied to the transformed model. In the case of estimating our model using 2SLS or 3SLS, we must specify the instrumental variables to be used in estimation.

¹² Furthermore, we have to be cautious with the result obtained for γ_1 as well as taking into account the unit roots test for panel data presented in subsection 3.2.1.

Given the endogeneity of the one-period lagged lodged tourists variable, we chose an instruments vector with several lags, such as: $\log T_{it-2}, \log T_{it-3}, \log Y_{it-1}, \log Y_{it-2}, \log PB_{it-1}, \log PB_{it-2}$ (¹³). The results vary slightly with respect to the earlier estimates that appear in tables 3 and 4. Thus, income elasticity exhibits a reduction that introduces some doubts about the luxurious nature of tourism. Moreover, the elasticity with respect to the cost of the trip decreases slightly in relation to those obtained in previous subsections. The parameters of equation [2] are jointly significant in all cases, as it is shown by the test of Wald for said hypothesis [$W(\gamma_1=\beta_1=\beta_2=\beta_3=0)$]. A t-Student test rejects the hypothesis that $\gamma_1=1$, in each case (according to the p-values). The FE-W2SLS estimation obtains better results in terms of the determinant of the residuals matrix and the test of Wald for the hypothesis of the joint significance of the parameters and the fixed effects.

Table 6 shows the long-run dynamic multipliers, i.e., the long-run elasticities calculated from the short-run estimates derived from the dynamic model. The quotients are calculated from the estimated values corresponding to the short-run and the quantity $(1-0.7417)$ in the first estimate, and $(1-0.7695)$ in the second estimate. The results obviously show that the long-run elasticities are greater than the short-run; in the case of income elasticity, it is even greater than 2.

¹³ We have tried other instruments vectors. This vector however has been the most adequate in all the estimations according to the test of Sargan. Moreover, we have not increased the number of variables so as not to lose degrees of freedom.

Table 6. Long-run multipliers or elasticities. Restricted coefficients.

Use results of equation [2] (table 5).

	FE-2SLS	FE-W2SLS	FE-3SLS
GDP	2.6359	2.7754	2.9457
Petroleum	-0.1442	-0.2702	-0.1366
Exchange Rate	0.5352	0.7402	0.8816

So far, the dynamic model has used the FE estimator. Nevertheless, there is another estimation method for dynamic panel models which is the first-differences (FD) estimator and which eliminates the permanent effects. The main difference with the FE estimator is that while the latter eliminates the individual effects subtracting the time mean for each observation, the first one eliminates said effects taking first differences.

Anderson and Hsiao (1981) suggested using a FD model. FD uses predetermined variables as valid instruments and permits consistent estimations. This is justified because not all the right hand side variables are exogenous, forcing the estimation of the parameters of the new equation using instrumental variables methods. The endogeneity problem produces correlations between the lagged endogenous variable and the non-zero residuals, even though these residuals are not serially correlated.

The results obtained from applying FD with the methods of instrumental variables will be biased but consistent, although not necessarily efficient. The efficiency will depend on the use of the complete information in the system.

Table 7. - Estimate of the double-log panel model with dynamic.

Restricted coefficients. Sample period 1979-1997. Equation

$$\Delta \log T_{it} = \gamma_1 \Delta \log T_{i,t-1} + \beta_1 \Delta \log Y_{it} + \beta_2 \Delta \log PB_{it} + \beta_3 \Delta \log E_{it} + \varepsilon_{it}$$

$$\varepsilon_{it} = \mu_i + u_{it}$$

Parameters	FD-2SLS	FD-W2SLS ^b	FD-3SLS ^a
γ_1	0.3320 (3.43)	0.3983 (4.57)	0.3439 (9.62)
β_1	1.3783 (3.03)	1.2275 (3.27)	1.2376 (12.71)
β_2	-0.1991 (-2.00)	-0.1325 (-1.59)	-0.1618 (-4.01)
β_3	0.2177 (1.37)	0.2902 (1.98)	0.2925 (5.28)
$ \hat{\Sigma}_u $	4.23e-28	6.79e-28	5.51e-28
$W(\gamma_1=\beta_1=\beta_2=\beta_3=0)$	74.96 [0.00]	98.86 [0.00]	932.8 [0.00]
Hausman-Taylor test (H_A)	7.8450 [0.097]	7.8450 [0.097]	1.1442 [0.89]
Hausman-Taylor test (H_B)	3.0353 [0.55]	3.0353 [0.55]	2.1571 [0.71]
$t(\gamma_1=1)$	-6.895 [0.00]	-6.909 [0.00]	-30.54 [0.00]

Note: In parenthesis appear robust t-values and in bracket appear p-values.

Superscript “a” represents the classic method and “b” represents the same method but with iteration to convergence. We control the iterative process by specifying the convergence criterion and the maximum number of iterations.

The results of the three methods of the estimation of the instrumental variables (FD-2SLS, FD-W2SLS y FD-3SLS) appear in table 7 (¹⁴). The elasticities are quite similar to those obtained from the static models. The estimation by FD-W2SLS of model [3] has the lowest value of the residuals determinant during the period studied. The estimations of the instrumental variables use Z as an instruments vector, which has been chosen among different alternatives. The chosen one has been

$$\Delta \log T_{it-2}, \Delta \log T_{it-3}, \Delta \log Y_{it-1}, \Delta \log Y_{it-2}, \Delta \log PB_{it-1}, \Delta \log PB_{it-2}.$$

¹⁴ Some papers, such as Arellano and Bond (1991), Keane and Runkle (1992), and Ahn and Schmidt (1995) have all advocated the use of the GMM methodology for the estimation of dynamic panel models, or panel models with predetermined rather than exogenous right hand side variables. This method is not applied, however, in this paper because a near singular matrix exists. Both GMM and FIML (full information maximum likelihood) cannot be applied because $N < T$.

With these estimation procedures two important hypotheses that can be tested using a Hausman-Taylor-type test. This is applied to test the validity of the instruments when we include lagged dependent variables¹⁵. The first hypothesis, which we denote by H_A , makes the group of instruments strictly exogenous. On the contrary, if we reject the strong exogeneity of the group of instruments, it will imply the within estimator is inconsistent and the standard FE and GLS estimators are inappropriate. Keane y Runkle (1992) intend to analyze the differences between FE-2SLS and FD-2SLS. If H_A is true, then (FE-2SLS is consistent), but if H_A is rejected, then FD-2SLS is consistent. Contrary to Keane and Runkle (1992) we compare the different methods, starting with the most inefficient one, in order to comply with the Hausman test requirements. Thus, we compare FD-2SLS, FD-W2SLS and FD-3SLS with the FE-2SLS, FE-W2SLS and FE-3SLS. If H_A is not rejected in this comparison, we do not reject that the set of instruments are strictly exogenous, and cannot argue against the application of the within estimator using a version of the Hausman-Taylor test.

If a predetermined set of instruments Z exists, such that $E[u_{it}/Z_{it}] = 0$, where Z contains the lagged values of T_{it} , we try to evaluate if the individual effects are correlated with the instruments. If H_A is not correlated, we could see if the individual effects are correlated with the instruments. We call this hypothesis H_B . In this case, the traditional test is inappropriate when predetermined variables exist, so Keane and Runkle propose to analyze the differences between FD-2SLS and 2SLS. If H_B is true, both estimators are consistent; if not, FD-2SLS is the preferred estimator. We also compare the FD estimators: FD-2SLS, FD-W2SLS and FD-3SLS with 2SLS, W2SLS y 3SLS. The results do not reject the null hypothesis, so we can consider both estimators as consistent.

¹⁵ Keane and Runkle (1992) use two hypothesis tests.

We can argue that the FE procedure seems to be the most adequate due to the Hausman test and following Keane and Runkle (1992).

3.2.3. Other dynamic specifications for the number of tourists lodged: cross-sectional variant and invariant regressors.

An alternative specification of the dynamic model presented before occurs when we include cross-sectional-invariant variables. In particular, the promotion expenditure and a variable referred to the infrastructure of the island are included. The former is a variable without available data differentiated by countries of origin. The latter is a variable related to supply, which impedes the differentiation by origin of the tourists. Both variables are nearly correlated with the GDP. For this reason, these variables are used in the estimations as substitutes of the GDP, i.e., the GDP is eliminated when one of these variables is introduced in the model.

Having demonstrated that the FE method yields the most statistically adequate results, models that incorporate promotion expenditure (PE_t) and the infrastructure (INF_t) will also be estimated by FE. Table 8 shows panel estimations when we substitute both PE_t and INF_t for the income variable. PE_t and INF_t are only time-variant, not varying among countries.

The sample period has been reduced since it goes from 1984 to 1994. Given that the estimation procedures are the instrumental variables, the instruments vector chosen has been the same as the one utilized in the previous section. The results show similarities among the estimated elasticities for each one of the variables for both FE and FD.

Moreover, it can be observed that the two new variables are significant and the promotion expenditure seems to have only a small influence in the number of tourists. However, the infrastructure variable has a greater influence as the results of the FD-W2SLS estimation display.

Table 8. – Estimates of the double-log dynamic panel data model.
Restricted coefficients. Sample period 1982-1994.
Consider PE_t and INF_t as substitutes of GDP.
Estimates of FE-W2SLS y FD-W2SLS.

Parameters	FE-W2SLS ^b		FD-W2SLS ^b	
	Model with PE_t	Model with INF_t	Model with PE_t	Model with INF_t
γ_1	0.5954 (10.99)	0.8177 (16.9)	0.4393 (5.94)	0.4937 (5.63)
β_1	0.0400 (3.17)	0.0681 (1.45)	0.0653 (2.99)	0.4700 (1.59)
β_2	-0.1124 (-1.93)	-0.0868 (-2.64)	-0.2482 (-4.75)	-0.1488 (-2.07)
β_3	0.0697 (1.33)	0.2633 (3.06)	0.0487 (1.13)	0.2132 (1.14)
$\left \hat{\sum u} \right $	0	0	0	0
$W(\mu_1=\dots=\mu_{N-1}=0)$	82.534 [0.00]	28.74 [0.01]		
$W(\gamma_1=\beta_1=\beta_2=\beta_3=0)$	516.01 [0.00]	162.08 [0.00]	74.97 [0.00]	88.65 [0.00]
$t(\gamma_1=1)$	-7.4686 [0.00]	-3.0463 [0.00]	-7.3015 [0.00]	-5.4248 [0.00]

Note: In parenthesis appear robust t-values and in bracket appear p-values.

Superscript “a” represent the classic method and “b” represent the same method but with iteration to convergence. We control the iterative process by specifying convergence criterion and the maximum number of iterations.

3.2.4. *Prediction and simulation.*

In order to predict and simulate, we build a system of equations in which the values of the parameters are known, having been previously estimated by FE-W2SLS or FD-W2SLS. Thanks to the coefficients, we will find the unknown values of the endogeneous variable, using a dynamic method (multi-step forecasts), i.e., we use the historical variables of the dependent variables lagged to the first period of simulation. The values predicted by the model are then used. The solution method used for these linear models is the Gauss-Seidel iterative method. This method evaluates each equation in the order in which it appears in the model, and utilizes the variable's new value for the left hand side in another equation, using it as the variable's value when it appears later on. The algorithm depends on the order of the model's equations; each equation then has a different endogeneous variable.

Prediction is one of the steps of the applied econometric analysis. In this case, we predict the variable number of tourists lodged. We build a system of equations for tourism demand formed by the specifications corresponding to each country. FE and FD are analyzed, and then estimated by W2SLS and FD-3SLS, respectively. With this we would like to analyze the sensitivity of the predictions one-step ahead to the utilization of each method, even though FE-W2SLS become the preferred method of estimation.

Table 11 shows the results of two statistics, the mean absolute error (MAE) which is expressed as the number of people, and the mean absolute percentage error (MAPE). The prediction period chosen goes from 1996 through 1997.

Table 11. Statistics forecast in the panel. Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). Period 1996-1997

	Fixed Effects				First Differences			
	W2SLS ^b		3SLS ^a		W2SLS ^b		3SLS ^a	
	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE
Germany	36459	5.30	34275	4.90	20184	2.90	27393	3.90
United Kingdom	192051	14.1	110376	8.20	13248	10.1	67585	5.10
Spain	121701	12.8	60637	6.40	25191	2.60	17781	1.80
Sweden	4229	3.40	7722	6.10	9285	7.40	9836	7.90
Norway	10639	15.8	7276	10.9	11112	16.7	10522	15.8
Denmark	10848	16.8	4785	7.40	8805	14.3	8032	13.2
Finland	8791	9.50	7511	8.20	10675	11.6	10864	11.9
Netherlands	11782	11.5	12905	12.2	14980	13.4	13307	11.8
Belgium	11107	7.90	3671	2.60	9611	3.40	5027	3.70
Austria	6739	16.4	5315	13.0	3517	8.50	2865	7.10
France	53771	25.2	50221	24.4	36201	16.6	35414	16.5
Italy	42389	30.5	37405	26.7	11236	7.90	12132	8.60
Switzerland	9471	22.5	6057	14.4	2394	5.60	2591	6.10

Note: Superscript “a” represent the classic method and “b” represent the same method but with iteration to convergence. We control the iterative process by specifying convergence criterion and the maximum number of iterations.

The results show that the 3SLS method, obtained smaller MAPE numbers for both prediction periods using both FE and FD. The FD-3SLS method obtains better overall results for the predictions.

We will simulate the behaviour of the flow of the number of tourists lodged in Tenerife in alternate scenarios of the international economy. The goal is to find the most probable evolution of the future (up to the year 2005) number of tourist.

The search for unknown values of the endogeneous variables is done by the multi-step procedure of forecast, i.e., we use the past values of the endogeneous variables lagged up to the first period of simulation. Afterwards, we use the model’s predicted values. The solution method is the iterative Gauss-Seidel, which evaluate each equation in the same order in which it appears in the model.

The new value of the left-hand side variable is used in the other equation as the new value on the variable. The algorithm depends on the order of the equations in the model; in this way, each equation has a different endogeneous variable.

The 3SLS estimations show a more adequate prediction since they minimize the MAPE numbers.

The a priori scenarios are two. In the first place, a real GDP per capita growth of 4.68% (calculated from a GDP increase of 3% for every country with a population growth of 0.3%) and an increase of 5% in the quotient between the cost of the trip and the CPI of each country. The second scenario presupposes a 1.196% rate of growth in the real GDP per capita (obtained from a GDP increase of 1.5% and a population growth of 0.3%) and a 15% increase in the relative cost of the trip.

Table 12. Rate of growth simulated: Period 1998-2005.

Countries	GER	U.K.	SP	SWE	NOR	DAN	FIN	NET	BEL	AUS	FRA	ITA	SWI
Scenario 1	10.6	12.8	10.6	10.6	13.1	11.6	11.1	11.2	10.7	10.7	10.3	10.1	7.3
Scenario 2	2.8	5.0	2.8	2.7	5.3	3.7	3.3	3.4	2.9	2.9	2.4	2.3	-0.5

Note: GER: Germany, UK: United Kingdom, SP: Spain, SWE: Sweden, NOR: Norway, DAN: Danmark, FIN: Finland, NET: Netherlands, BEL: Belgium, AUS: Austria, FRA: France, SWI: Switzerland.

As can be observed, in scenario 1, in which there exists an economic expansion with a moderate increase of the cost of the trip, we obtain an annual growth rate of the number of visitors between 10 and 13% in the simulated period. On the contrary, in scenario 2, where the European economy grows slowly and the cost of the trip has a big hike, the number of tourists only rises at a rate between 2 and 5%. Thus, the number of tourists lodged in Tenerife does not seem to be very sensitive to such different scenarios.

The figures in appendix II show the evolution of the number of tourists attending to their country of origin. In scenario 1, the number of visitors increases at a growing rate. This optimistic scenario may generate some difficulties in relation to a possible lack of capacity of the supply, i.e., to problems related to satisfy the size of demand. On the other hand, the more pessimistic scenario would not avoid the rise of the tourism demand, in spite of the slight decrease in the case

of Switzerland. However, the number of tourists would increase at a decreasing rate.

4. Conclusions

In this paper, we have studied the demand elasticities for the tourists lodged in Tenerife. To do this we have used different estimation methods that belong to the econometrics of panel data.

We estimate a static panel data model with and without dynamics errors. From these estimations we consider the need to analyze a dynamic model by introducing the endogenous lagged variable. Moreover, we estimate a dynamic model both with and without the application of first differences to the variables of the model.

The results point out that the number of visitors lodged in Tenerife exhibits a high elasticity with respect to the real income per capita, showing the luxurious nature of tourism. For this reason, tourism policies should take into account the high sensitivity of this demand to the economic cycle. Furthermore, the exchange rate and the cost of the trip have a significant influence in the number of visitors but this variable is inelastic with respect to both price variables. The introduction of the endogenous lagged variable as an explanatory variable and its significance could indicate the importance of the reputation captured by the high degree of repetition of the tourists lodged in Tenerife.

For its part, the level of the infrastructures and the promotion expenditure show the importance of the public inputs and the non-price competition in the tourism activity.

Finally, we carry out prediction exercises and simulate some scenarios. A very optimistic scenario could lead to difficulties in relation to a possible lack of the capacity to lodge the growing number of visitors.

References

- Ahn, S. and Schmidt, P. (1995): “Efficient Estimation of Models for Dynamic Panel Data”. *Journal of Econometrics*, 68, 5-27.
- Anderson, T.W. and Hsiao, C. (1981): “Estimations of Dynamics Models with error Components”. *Journal of the American Statistical Association*, 76, 598-606.
- Arellano, M. and Bond, S. (1991): “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations”. *Review of Economic Studies*, 58, 277-298.
- Balestra, P. and Nerlove, M. (1966): “Pooling Cross-section and Time Series Data in the Estimation of a Dynamic Model: the Demand for Natural Gas”. *Econometrica*, 34, 585-612.
- Baltagi, B.H. (1995). *Econometric Analysis of Panel Data*. John Wiley and Sons.
- Breitung, J. And Meyer, W. (1994): “Testing for Unit Roots in Panel Data: Are Wages on Different Bargaining Levels Cointegrated?”. *Applied Economics*, 26, 353-361.
- Buisán, A. (1995). “Principales determinantes de los ingresos por turismo”. Documento de Trabajo nº9502, Servicios de Estudios del Banco de España.
- Crouch, G.I. (1994). “The Study of International Tourism Demand: A Review of Findings”. *Journal of Travel Research*, pp. 12-23.
- Crouch, G.I. y R.N. Shaw (1992). “International Tourism Demand: A Meta-analytical Integration of Research Findings”, en P. Johnson y B. Thomas (eds.), *Choice and Demand in Tourism*, pp. 175-207, Mansell, London.
- Deardoff, A. (1985). *Comparative Advantage and International Trade and Investment in Services*, en R.M. Stern (ed.), *Trade and Investment in Services: Canada/USA Perspectives*, Ontario Economic Council, pp. 39-71.
- Dickey, D.A. y Fuller, W. (1981): “Likelihood Ratio Statistics for Autorregressive Time Series with a Unit Root”. *Econometrica*, 49, 1057-1072.
- Fair, R. (1984): *Specification, Estimation, and Analysis of Macroeconometric Models*. Harvard University Press.
- González, P. y P. Moral (1993). “An Analysis of the International Tourism Demand in Spain”. Documento de trabajo nº9310, Facultad de Ciencias Económicas, Universidad del País Vasco.
- Harris, R. and Tzavalis, E. (1999): “Inference for Unit Roots in Dynamic Panels where the Time Dimension is Fixed”. *Journal of Econometrics*, 91, 201-226.
- Hausman, J.A. and Taylor, W.E. (1981): “Panel Data and Unobservable Individual Effects”. *Econometrica*, 49, 1377-1398.
- Jafari, J. (1977). “Editor’s Page”, *Annals of Tourism Research*, Vol. 5, pp. 6-11.

- Keane, M.P. and Runkle, D.E. (1992): “On the Estimation of Panel Data Models with Serial Correlation when Instruments are not Strictly Exogenous”. *Journal of Business and Economics Statistics*, 10, 1-9.
- Krugman, P.R. (1991). “Increasing returns and economic geography”, *Journal of Political Economy*, 99:183-199.
- Levin, A. And Lin, C-F. (1993a) “Unit Root Test in Panel Data: Asymptotic and Finite Sample Properties”. Discussion Paper 92-93, UCSD.
- Lillard, L. And Willis, R. (1978): “Dynamic Aspects of Earning Mobility”. *Econometrica*, 46, 985-1012.
- Maddala, G.S. (ed.) (1993). *Econometrics of Panel Data*, Vol. I y II. Elgard Publishing Limited.
- Navarro, M. y M. Becerra (1998).”Algunas reflexiones sobre el turismo en Tenerife”, ponencia presentada en el Seminario Red ISA-UNESCO. Ponta Delgada, Azores.
- Padilla, R. (1988). “La demanda de servicios turísticos en España”, *Investigaciones Económicas*, Vol. XII, nº1, pp. 133-157.
- Quah, D. (1994): “Exploiting Cross Section Variation for Unit Root Inference in Dynamic Data”. *Economic Letters*, 44, 9-19.
- Sapir, A. y C. Winter (1994). “Services Trade”, en D. Greenaway y L.A. Winters (eds.), *Surveys in International Trade*, Blackwell, Oxford, pp. 271-302.
- Wallace, T.D. and Hussain, A. (1969): “The Use of Error Components Models in Combining Cross-Section with Time Series Data”. *Econometrica*, 37, 57-72.

Appendix I

Panel Data Estimation Methods

1. Fixed Effects

The class of models that can be estimated using panel data can be written as $y_{it} = \alpha_{it} + x'_{it}\beta_i + u_{it}$, where y_{it} is the dependent variable, x_{it} and α_{it} and β_i are k -vectors of non-constant regressors and parameters for $i = 1, 2, \dots, N$ cross-sectional units or individuals, observed for dated periods $t = 1, 2, \dots, T$. We can view these data as a set of cross-section specific regressions so that we have N cross-sectional equations: $y_i = \alpha_i + x'_i\beta_i + u_i$, with T observations. For simplicity, we represent the stacked model

$$y = X\beta + \varepsilon.$$

The Fixed Effects are computed by subtracting the “within” mean from each variable, and we can write:

$\tilde{y} = \tilde{X}\beta + \tilde{\varepsilon}$, where $\tilde{y} = Qy$, $\tilde{X} = QX$, $\tilde{\varepsilon} = Q\varepsilon$ and $Q = I_{NT} - P$ with $P = I_N \otimes \bar{J}_T$ and $\bar{J}_T = J_T / T$ in the Baltagi's notation. Q is a matrix which obtains the deviations from individual means. P and Q are symmetric idempotent matrices (i.e., $P' = P$), orthogonal (i.e., $PQ = 0$) and they sum to the identity matrix ($P + Q = I_{NT}$). The i -th typical elements in the transformed data are give by following equation:

$$y_i - \bar{y}_i = \beta_i' \left(x_i - \bar{x}_i \right) + \left(\varepsilon_i - \bar{\varepsilon}_i \right)$$

where $\bar{y}_i, \bar{x}_i$ are sample means of i -th individual. The Fixed Effects estimator allows differ between cross-section countries by estimating different constants for each cross-section. So, the classical estimator applied to transformed data is

$$\hat{\beta}_{OLS} = (\tilde{X}'\tilde{X})^{-1} \tilde{X}'\tilde{y} \text{ and the variance estimator is given by } \text{var}(\hat{\beta}_{OLS}) = \hat{\sigma}_u^2 (\tilde{X}'\tilde{X})^{-1}$$

where $\hat{\sigma}_u^2$ is the residual variance in the stacked system $\hat{\sigma}_u^2 = \frac{e'e}{NT - N - K}$, where $e'e$ is the SSR from the fixed effects model.

In general, OLS is appropriate when the residuals are contemporaneously uncorrelated, and time-period and cross-section homoskedastic, where $\Omega = \sigma^2 I_N \otimes I_T = \text{diag}(\sigma^2, \dots, \sigma^2) \otimes I_T$.

The Fixed Effects are estimated in a second step following to

$$\hat{\alpha}_i = \frac{\sum_{i=1}^N \left(\bar{y}_i - \bar{x}_i' \hat{\beta} \right)}{N}.$$

We employ others estimation methods for fixed effects:

1) Under the assumption that there is heteroskedasticity, but no serial correlation or contemporaneous correlation in the residuals, the weighted least squares estimator is efficient, and the variance estimator consistent. The weighted least squares estimator is given by $\hat{\beta}_{GLS} = (\tilde{X}' \hat{\Omega}^{-1} \tilde{X})^{-1} \tilde{X}' \hat{\Omega}^{-1} \tilde{y}$ where $\hat{\Omega} = \text{diag}(\sigma_1^2, \dots, \sigma_N^2) \otimes I_T$. A consistent estimator of $\hat{\Omega}^{-1}$ and σ_i^2 are $\hat{\Omega}^{-1} = \text{diag}(1/\sigma_1^2, \dots, 1/\sigma_N^2) \otimes I_T$, $\sigma_i^2 = (y_i - x_i' \hat{\beta}_{GLS})(y_i - x_i' \hat{\beta}_{GLS})/T$, respectively for all i . The estimator for the coefficient variance matrix is: $\text{var}(\hat{\beta}_{GLS}) = (\tilde{X}' \hat{\Omega}^{-1} \tilde{X})^{-1}$

Weighted LS estimates will be identical to equation by equation weighted OLS if there are no cross-equation restrictions.

2) If all the right-hand side regressors are assumed to be exogenous, and the error variance matrix is given by $\Omega = \Sigma \otimes I_T$, where Σ is a symmetric matrix of contemporaneous correlations, then the SUR method can be appropriated. SUR weighted least squares (sometimes referred to as the Parks estimator) is the feasible GLS estimator when the residuals are both cross-section heteroskedastic and contemporaneously correlated. The Zellner's SUR estimator is given by $\hat{\beta}_{OLS} = (\tilde{X}' (\hat{\Sigma}^{-1} \otimes I_T) \tilde{X})^{-1} \tilde{X}' (\hat{\Sigma}^{-1} \otimes I_T) \tilde{y}$ where $\hat{\Sigma}^{-1}$ is a consistent estimate of Σ and its typical elements are $\sigma_{ij} = (y_i - x_i' \hat{\beta}_{OLS})(y_j - x_j' \hat{\beta}_{OLS})/T$.

3) If some of the variables in X are endogenous then the 2SLS is an estimation method that is appropriate. Write $y = z\delta + u$, where $z = [Y \ X_1]$ is a matrix of predetermined variables: endogenous variables, Y , and exogenous variables, X_1 ; and $\delta' = [\gamma' \ \beta_1']$ is a vector of parameters of these variables.

We can transform the model to get $\tilde{y} = Z\delta + \tilde{u}$, premultiplying by Q matrix all variables and where $Z = \tilde{z}$. We can derive the estimator by employing a set of instruments or in an equivalent way to the single equations 2SLS method by regressing in two steps. In this way, in the first stage we would regress the right-hand side endogenous variables Y on all exogenous variables X and get the fitted

values $\hat{Y} = \tilde{X}_1 (\tilde{X}_1' \tilde{X}_1)^{-1} \tilde{X}_1' \tilde{Y}$. In the second stage, we regress $\tilde{y}$ on $\hat{Y}$ and $\tilde{X}$ to get $\hat{\delta}_{TSLs} = (\hat{Z}' \hat{Z})^{-1} \hat{Z}' \tilde{y}$, where $\hat{Z} = [\hat{Y} \ \tilde{X}]$.

If we want to assume that there is heteroskedasticity, we can employ the Weighted 2SLS. This method applies the weights in the second stage so that $\hat{\delta}_{WTSLs} = (\hat{Z}' \hat{\Omega}^{-1} \hat{Z})^{-1} \hat{Z}' \hat{\Omega}^{-1} \tilde{y}$, where the elements of diagonal matrix $\hat{\Omega}^{-1} = \text{diag}(1/\hat{\sigma}_1^2, \dots, 1/\hat{\sigma}_N^2) \otimes I_T$ are estimated in the usual fashion using the residuals from 2SLS.

If we use to iterate the weights, Σ is estimated at each step using the coefficients and residuals.

4) Finally, 2SLS is not fully efficient because this estimator does not take account of the covariances between residuals. In this sense, 3SLS is a method that estimates all of the coefficients of the model, then forms weights and reestimates the model using the estimated weighting matrix. The first two stages of 3SLS are the same as in 2SLS. In the third stage, we apply feasible generalized least squares (FGLS) to the equations in the system in a manner analogous to the SUR estimator. SUR uses the OLS residuals to obtain a consistent estimate of the cross-equation covariance matrix Σ . This covariance estimator is not, however, consistent if any of the right-hand side variables are endogenous.. The estimator is $\hat{\delta}_{3SLS} = (Z' (\hat{\Sigma}^{-1} \otimes P_{\tilde{X}}) Z)^{-1} Z' (\hat{\Sigma}^{-1} \otimes P_{\tilde{X}}) \tilde{y}$, where $P_{\tilde{X}} = X(X'X)^{-1}X'$ and $\tilde{X}$ is a set of instruments.

3SLS uses the 2SLS residuals to obtain a consistent estimate of Σ , where $\hat{\Sigma}$ has typical element $\hat{\sigma}_{ij} = (\tilde{y}_i - Z_i' \hat{\delta}_{TSLs})(\tilde{y}_j - Z_j' \hat{\delta}_{TSLs})/T$.

2. Random Effects

The classic random effects model or variance components model assumes that the term α_{it} is the sum of a common constant α and a time-invariant cross-section specific random variable u_{it} that is uncorrelated with the residual ε_{it} . Instead of treating μ_i as a fixed constant, this specification assumes that μ_i is drawn from an i.i.d. distribution, $\mu_i \sim N(0, \sigma_\mu^2)$, and is uncorrelated both with the ε_{it} and x_{it} .

The specification then becomes: $y_{it} = x_{it}' \beta_i + u_{it}$, $u_{it} = \mu_i + \varepsilon_{it}$, so that $E[u] = 0$ and the covariance matrix is block diagonal $\Omega = \Sigma \otimes I_T$. Here, the appropriate estimator

is generalized least squares $\hat{\beta}_{GLS} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y$ which can be expressed in least squares form by transforming the variables by $y_i - \lambda \bar{y}_i = \beta_i' \left(x_i - \lambda \bar{x}_i \right) + \left(u_i - \lambda \bar{u}_i \right)$ and the running ordinary least squares where $\lambda = 1 - \frac{\sigma_u}{\sqrt{\sigma_u^2 + T\sigma_\mu^2}}$. Usually the variances σ_μ^2 (the between groups variance) and σ_u^2 are not known, so consistent estimates are derived from initial least squares estimates to form $\hat{\lambda}$ (See Wallace and Hussain (1969)). This estimator is asymptotically efficient and if iterated to convergence, it yields the maximum likelihood estimates.

Appendix II.

Figure A.II.1. Evolution of the number of tourists lodged by countries. Scenario 1.

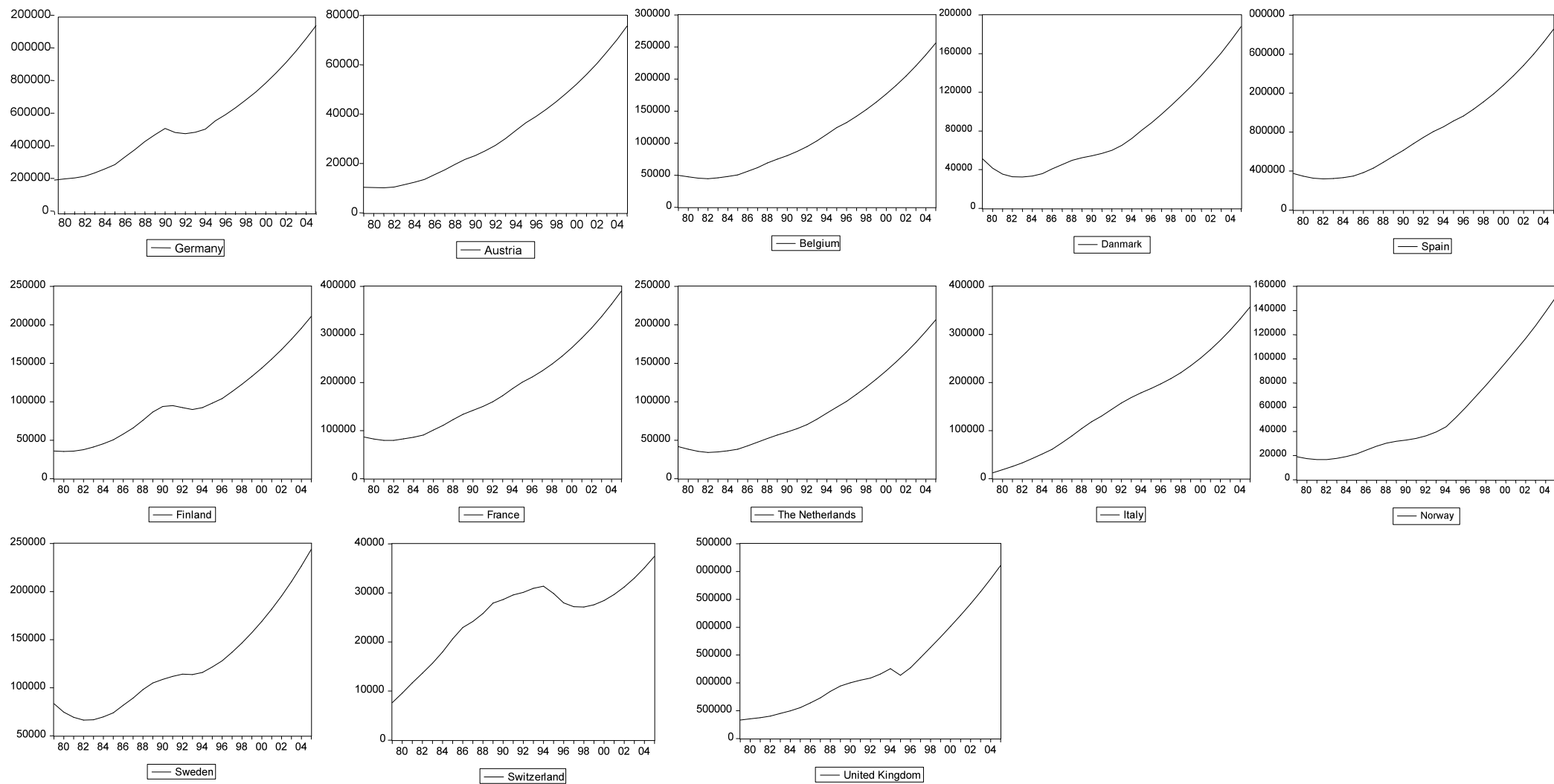


Figure A.II.2. Evolution of the number of tourists lodged by countries. Scenario 2.

