Predicting random level and seasonality of hotel prices. A structural equation growth curve approach

Germà Coenders, Josep Maria Espinet and Marc Saez†

Departament d'Economia,
Universitat de Girona

Girona, March 2001

Abstract

This article examines the effect on price of different characteristics of holiday hotels in the sun-and-beach segment, under the hedonic function perspective. Monthly prices of the majority of hotels in the Spanish continental Mediterranean coast are gathered from May to October 1999 from the tour operator catalogues. Hedonic functions are specified as random-effect models and parametrized as structural equation models with two latent variables, a random peak season price and a random width of seasonal fluctuations. Characteristics of the hotel and the region where they are located are used as predictors of both latent variables. Besides hotel category, region, distance to the beach, availability of parking place and room equipment have an effect on peak price and also on seasonality. 3-star hotels have the highest seasonality and hotels located in the southern regions the lowest, which could be explained by a warmer climate in autumn.

Keywords: Hedonic functions, hotel pricing, growth curve models, random-effect models, structural equation models.

JEL classification: C33, L11, L83.

† Address: Departament d'Economia. Universitat de Girona, Campus de Montilivi, 17071 Girona, Spain. E-mail: germa.coenders@udg.es

1. Introduction

The aim of this article is to study the effect on prices of the different characteristics or attributes of a holiday hotel in the sun-and-beach segment in Spain. This country constitutes the second tourist destination worldwide (Departament d'Indústria, Comerç i Turisme, 1999) and nowadays attracts mostly sunseaking tourists, though the latest policies of both the industry and the government tend to foster other assets of the country such as culture, history, gastronomy, landscape and monuments.

The relevant attributes of holiday hotels include, among others, category, services available to guests, attributes of the region in which they are located and, given the particular segment aimed at, likely climate. From the supply's perspective the hotel's attributes on the one hand have an impact on cost and on the other hand make it possible to differentiate the offer and thus gain some bargaining power in front of tour operators. Besides, the application of cost-oriented pricing is common in the hotel sector (Chias, 1996; Witt & Moutinho, 1994). Among cost oriented models we can find the application of a given margin above the costs (known as mark-up pricing both in the economy and management literatures), or of a given percentage of investment (which is known as Hubbart's formula). Even if hotels were, as sometimes suggested, price compliant and prices were only demand driven, (Taylor, 1995), the hotel attributes would still affect the price imposed by the tour operator.

The main difficulty facing research on the value of attributes is that their price is unobserved as they are not separately traded in any market. Only the overall prices of hotel rooms including particular combinations of attributes are observed. Our analysis draws upon the hedonic-price tradition of fitting statistical models to estimate the effect of attributes on price. Early theoretical developments in hedonic prices are those of Rosen (1974), Halvorsen and Pollakowski (1981) and Cassel and Mendelsohn (1985). Empirical applications in the tourist sector are found in Sinclair et al. (1990), Clewer et al. (1992), Jaime-Pastor (1999) and Espinet (1999).

The product a given hotel is offering can be regarded as a set of attributes, which can consist of services (e.g. swimming pool, garden, television in the room) or characteristics (e.g. category, region in which it is located, number of rooms). Thus, the hedonic price function for each hotel is represented as:

$$P_i = P(q_{i1}, q_{i2}, q_{i3},...,q_{ik},,..., q_{im})$$

where i = 1,...,n represents the hotel and q_{ik} (k=1,...,m) each of its attributes.

The study of hotel room pricing is quite complex because of seasonality, different price regimes (full board, half board, bed & breakfast), and discounts and supplements on various grounds (additional bed for kids, single room, view to the sea, additional room equipment such as air conditioning, television, or mini bar). In addition, three types of prices are relevant in the Spanish tourist lodgement market:

- Prices standing on hotel guides published by official institutions such as the Spanish Tourism Office. These are official prices which are seldom paid.

- Prices paid when the room is reserved by the traveller directly.
- Prices appearing on the catalogues of the tour operators. This coincides with the amount most tourists pay as tour operators constitute the most frequent distribution channel for tourist hotels in the studied market segment (Espinet, 1999).

In Spain, tour operator prices are not systematically collected by any official tourist or statistical office. A comprehensive data base of tour operator catalogue prices in the whole area of study was gathered for the first time and is used in this article. The data base covers the 5 major tourist regions in the continental Spanish Mediterranean coast, accounting for 79% of beds offered. From north to south they are Costa Brava, Costa del Maresme, Costa Daurada, Costa Blanca and Costa del Sol. All zones represent well the sun-and-beach segment which is aimed at. Hedonic functions are estimated by means of latent trajectory models, also known as growth curve models, a particular case of random effects models, also known as mixed models, hierarchical models or multilevel models. General references for random effect models are Laird and Ware (1982) and Bock (1989). For latent trajectory models see Meredith and Tisak (1989). The model will be used with the aims of:

- Estimate price level in each zone, controlling for differences in hotel characteristics across zones.
- Estimate seasonality curves in each zone, controlling for differences in hotel characteristics across zones.
- Find the contribution of the hotel characteristics to price, within the economic hedonic prices tradition extended to collect both the effect on level and on seasonality.
- Estimate the contribution of weather in the different zones on level and on seasonality.

2. Data

Hotel prices have been obtained daily from the catalogues of 9 Spanish tour operators from May to October 1999. Foreign tour operators were disregarded as their offer of Spanish hotels tends to be much narrower than that of Spanish operators. Besides, Espinet (1999) shows price differences to be very small with respect to foreign operators. The 9 operators were selected on the basis of size and singularity, thus, the largest ones were selected on its own right, and there were representatives of operators with and without their own network of travel agencies and of operators with remarkably high and remarkably low prices. Prices were in all cases expressed in ESP per day and person in a double room with full board. Fortnight averages were computed and, in order to remove the effect of different hotels being offered by different sets of tour operators, prices were taken as the corrected hotel means in an additive analysis of variance model where hotels and operators were crossed fixed factors. The natural logs of the averages of the two prices registered every month were considered as endogenous variables. For May only the second fortnight was considered and for October only the first. This resulted in the listwise missing value rate dropping from 12% to 8%. The final listwise sample size was 471. Exploratory analyses were carried out and no outliers were detected.

The explanatory variables considered where those that were statistically significant in Espinet et al. (forthcoming):

- Size: number of rooms, log transformed.
- Category: dummy: H1: 1-star hotels; H2: 2-star hotels, H4: 4-star hotels. 3-star hotels are the reference category.
- Beach: dummy: 1 for hotels located right in front of the sea.
- Room: dummy: 1 for hotels whose rooms are equipped with at least one of the following without price surcharge: television, air conditioning or mini-bar.
- Parking: dummy: 1 for hotels with parking place.
- Sport: dummy: 1 for hotels offering at least one of the following sport facilities: tennis, squash, golf or mini-golf, with or without extra payment.
- Town: dummy. Towns with fewer than 15 hotels were grouped into an "other" category. Neighbour towns with more than 15 and fewer than 30 hotels were merged if they were not significantly different according to a multivariate analysis of variance model of all price variables as dependent and the town and category as predictors.

The values of the explanatory variables were extracted from the hotel guide published by the Spanish Tourism Office and from the tour operator catalogues. In case of conflict, the information of the hotel guide prevails, except if more than two catalogues coincide in the availability of a service that is not included in the hotel guide.

Table 1 shows the distribution of the listwise sample of hotels for all zones. Table 2 shows the proportion of hotels being in front of the beach, having special room equipment, parking place, sport facilities and belonging to each of the categories, in each zone and overall. It can be seen that the mix of hotel characteristics is quite different for the different zones. In general, Costa Daurada and Costa del Sol seem to have the highest proportions of hotels of a high category and with additional services, though some services do not follow this general pattern. Table 3 shows the mean of the log numeric variables, per zones and overall. As regards prices, these means are depicted graphically in Figure 1, which shows large differences in price level across zones, that cannot be directly interpreted as the hotel characteristic mix is heterogeneous across zones. Figure 2 represents the differences with respect to August which can be interpreted as percentage price reductions with respect to high season. This Figure is useful for viewing seasonality and shows that Costa Blanca and Costa del Sol, the southernmost areas with the warmest weather, have a distinct profile, although interpretation must wait until the effect of different hotel characteristic mix is controlled for.

	count	perct.
Costa Brava	138	29.3
Costa del Maresme	60	12.7
Costa Daurada	71	15.1
Costa Blanca	121	25.7
Costa del Sol	81	17.2
Total	471	100.0

Table 1: Sample sizes in each zone

	Brava	Maresme	Daurada	Blanca	Sol	Overall
Beach	.25	.33	.23	.22	.35	.27
Room	.49	.47	.85	.78	.83	.67
Parking	.75	.58	.83	.55	.77	.69
Sport	.24	.23	.32	.22	.51	.29
H1	.11	.07	.00	.11	.00	.07
H2	.14	.28	.14	.25	.09	.18
H3	.67	.58	.72	.52	.67	.63
H4	.08	.07	.14	.12	.25	.13

Table 2: Nominal variables: proportions of hotels with given characteristics

	Brava	Maresme	Daurada	Blanca	Sol	Overall
Ln(size)	4.72	5.05	5.12	4.71	5.04	4.88
$\ln(Y_1)$ (May)	8.31	8.00	8.32	8.39	8.63	8.35
$\ln(Y_2)$ (June)	8.46	8.17	8.53	8.52	8.68	8.49
$\ln(Y_3)$ (July)	8.83	8.61	8.89	8.77	8.98	8.82
$\ln(Y_4)$ (August)	8.93	8.69	9.01	8.98	9.17	8.97
$\ln(Y_5)$ (Sept.)	8.48	8.18	8.57	8.72	8.87	8.58
$\ln(Y_6)$ (Octob.)	8.28	7.94	8.26	8.49	8.63	8.35

Table 3: Continuous variables: means

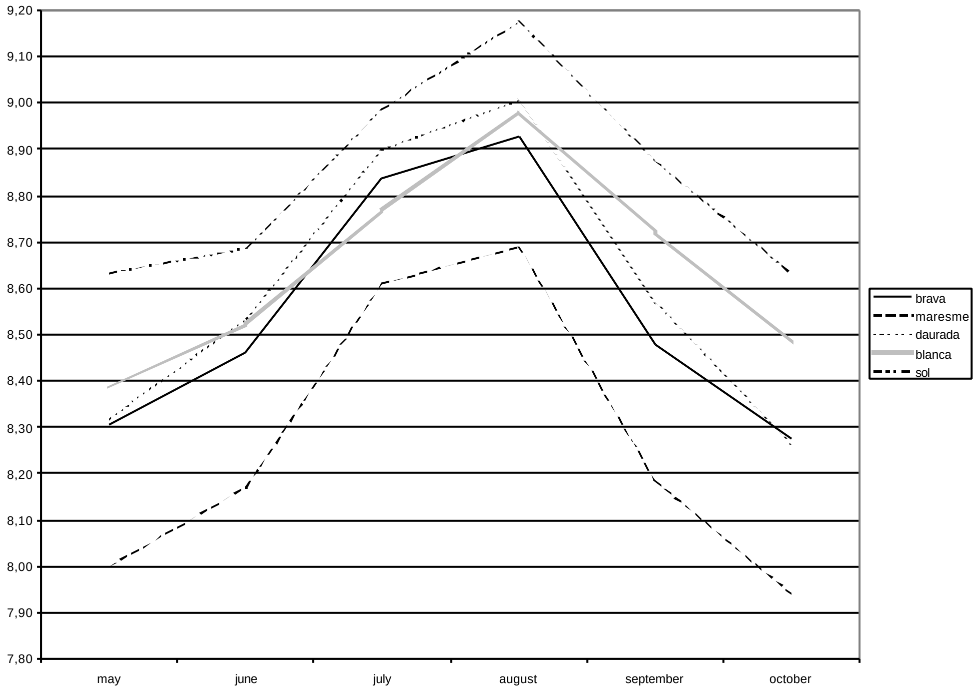

Figure 1: Average log prices in each zone and month

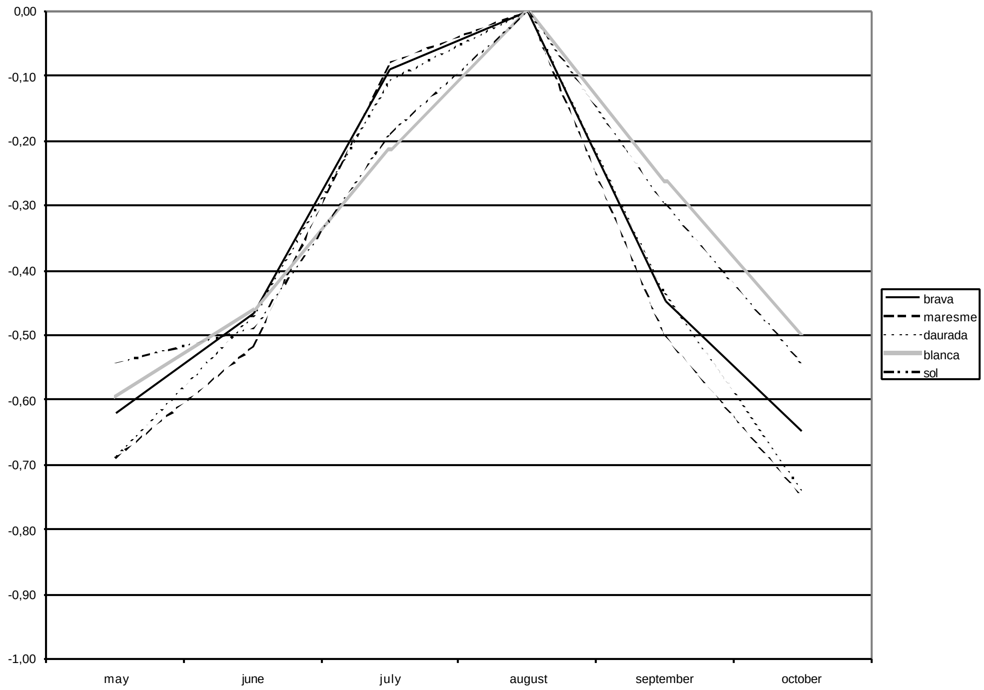

Figure 2: Average differences in log prices with respect to the peak month

3. Model

Growth curve models or latent trajectory models can be used to represent the individual evolution in time. Each individual is allowed to have its own evolution curve as a function of time, without assuming that the intercept or slope of the curve is equal for all individuals. The simplest such model would be a linear growth model:

$$Y_{it}=L_i+T_iC_t+D_{it}$$

where:

- i represents the subject and t time.
- Y_{it} is the variable whose evolution is studied, measured at time t for subject i.
- L_i is the fitted value for subject i at t=1, that is, the random intercept of the growth curve for subject i.
- T_i is the random slope of the trend growth curve for subject i.
- C_t represent time, e.g. 0,1,2,3,...,T-1 for a linear growth curve with measurements taken at regular intervals.
- D_{it}: Disturbance term for subject i at period t.

Meredith and Tisak (1989) showed that these models can be fitted as a particular case of structural equation model with latent variables with mean structures (e.g. Sörbom, 1974; Bollen, 1989; Batista-Foguet and Coenders, 2000). For this kind of models, coefficients varying randomly across subjects are specified as latent variables, also known as factors. In our case, one factor must be specified for L_i and one for T_i. The coefficients C_t are then interpreted as constrained factor loadings. The loadings of L_i are implicitly constrained to one. D_{it} is interpreted as uniqueness. The individual effects are represented by the factor scores and the average effects by the mean of the factors. Applications of structural equation models to growth curves can be found in psychology (McArdle & Epstein, 1987), education (Willett & Sayer, 1993) and health (Muthén, 2000).

In our case we suggest using this type of model for economic panel data of hotels instead of individuals and to represent seasonality instead of trend. The model is then specified as follows:

$$Y_{itk}=L_{ik}+S_{ik}C_{tk}+D_{itk}$$

where:

- i represents the hotel; t time (month); and k the zone (1:Costa Brava, 2: Costa del Maresme, 3: Costa Daurada, 4: Costa Blanca, 5: Costa del Sol). Random effects vary across hotels, while zones are treated as fixed effects by fitting the same model to multiple groups (e.g. Sörbom, 1974). All parameters are allowed to vary across zones and thus have a k subscript. As an alternative, a three level model such as described in Muthén (1997) could be considered by treating the group or zone as random. However, this was not advisable in our case because only 5 zones are available which cannot be considered to be a random sample of all possible tourist regions but virtually constitute the whole population.

- Y_{itk} : natural log of price for the ith hotel, tth month and kth zone. These prices are represented by T=6 endogenous variables, corresponding to the log price during a given month.
- S_{ik}: seasonality amplitude for hotel i in zone k. Random effect varying across hotels represented by a latent variable with loadings C_{tk} on the Y_{itk} variables.
- C_{tk}: seasonality profile for month t in zone k. C_{4k} was set to 0 in order to set August (the peak month with the most revenues and occupation, and the closest to a market equilibrium situation, as most hotels are full) as reference for the price level. The form of seasonality may be quite irregular, so that the loadings are unconstrained, unlike the case is for the linear model described above, for which the loadings would be constrained to 0,1,2,..., T-1. The profile of seasonality is constrained to be the same for all hotels in the same zone. This implies for instance that peaks are located in the same month for all hotels. Only the amplitude of the fluctuations is allowed to vary across hotels. C_{tk} is interpreted as the percentage change in price with respect to August, that is the peak month.
- L_{ik}: Price level for hotel i at t=4 (August) in zone k. Random effect varying across hotels represented by a latent variable with unit loadings on the Y_{itk} variables.
- D_{itk}: Disturbance term for hotel i during month t in zone k represented by the unique variances in the latent variable model, which can be interpreted as variation in monthly prices that is not explained. In this model, the source of unexplained variance is imposing prices to depend only on level and seasonality, seasonality being constrained to have a constant profile varying only in strength.

The model is extended to include J numeric or dummy predictor or explanatory variables (those described in the data section). If time varying, these predictors can have an effect on Y_{itk} as is common in random-effect models. More interestingly, if they are time invariant, they can be used to predict the random effects (variables) L and S:

$$S_{ik} = 1 + \beta_{s1k} X_{1ik} + ... + \beta_{sjk} X_{Jik} + U_{sik}$$
$$L_{ik} = \beta_{l0k} + \beta_{l1k} X_{1ik} + ... + \beta_{ljk} X_{Jik} + U_{lik}$$

For each zone, the model includes the parameters listed below. Some parameters can be constrained to be equal across zones but are all identified when unconstrained:

- T-1 $C_1 ... C_T$ seasonality coefficients (C_4 is set to 0), that is, percentage reductions for months outside the peak season.
- T Variances of $D_1 ... D_T$ and T-1 covariances between D_t and D_{t-1}. D is thus assumed to follow a heteroskedastic first order moving average process. Twice the standard deviation of D_t can be interpreted as the as the maximum (except for 5% extreme cases) percent variation in prices above or below what is predicted by peak level and seasonality.

- 2J β_{sjk} and β_{ljk} regression slopes.
- 2 variances and one covariance for the L and S random effects (latent variables). More precisely, these variances and covariance refer to the disturbance terms U_s and U_l . Twice the standard deviation of U_s can be interpreted as the maximum percent variation in seasonality that can occur above or below what is predicted. Twice the standard deviation of U_l can be interpreted as the maximum percent variation in the August price level that can occur above or below what is predicted.
- 2 intercept terms for the L and S random effects (latent variables). The intercept for peak level is β_{l0k} and can be interpreted as the expected log of price in August for the reference group of the dummy predictors and the value zero of the numeric predictors. The intercept for seasonality is constrained to 1. This constraint does not reduce the generality of the model but is used only for identification and clarity purposes and makes $C_1...C_T$ apply to the reference group of the dummy predictors and the value zero of the numeric predictors. In this way, for all months other than August, $E(Y_{itk})= \beta_{l0k} +C_{tk}$ for the reference group of the dummy predictors and the value zero of the numeric predictors. The expected price for any month and any values of the predictors is $E(Y_{itk})= \beta_{l0k} +\beta_{l1k}X_{1ik}+...+\beta_{ljk}X_{Jik} +(1 +\beta_{s1k}X_{1ik}+...+\beta_{sjk}X_{Jik})C_{tk}$. Note the interpretation of the β_{ljk} coefficients as percentage change in the August price level for a unit increase in the predictor or for belonging to the category indicated by the dummy variable. Note the interpretation of the β_{sjk} coefficients as percentage change in the seasonal indices for a unit increase in the predictor or for belonging to the category indicated by the dummy variable. Positive β_{sjk} coefficients indicate wider seasonal fluctuations. The ability to predict the width of seasonality is one of the key advantages of this model.

The shape of the zone seasonal profiles in Figure 2 suggests that southern zones (Costa Blanca and Costa del Sol), which can get good weather early in autumn, have lower price reductions in September and October. This suggests that tourists are willing to pay for good climate and this willingness to pay can also be estimated using hedonic functions in which weather variables act as explanatory. Since weather is constant or nearly constant for hotels in the same zone, the pooled data of all zones have to be analysed in order to get estimates of the effect of weather data. Here is the simplest possible pooled model, an additive model that constrains all seasonality profiles and slopes of predictors to be constant across zones, though zone dummies have an effect on price level and on the amplitude of seasonality:

$$Y_{itk}=L_{ik}+S_{ik}C_t+D_{itk}$$
$$S_{ik}= 1 +\beta_{s1}X_{1ik}+...+\beta_{sj}X_{Jik}+\delta_{s2}Z_{2k}+\delta_{s3}Z_{3k}+\delta_{s4}Z_{4k}+\delta_{s5}Z_{5k}+U_{sik}$$
$$L_{ik}=\beta_{l0} +\beta_{l1}X_{1ik}+...+\beta_{lj}X_{Jik}+\delta_{l2}Z_{2k}+\delta_{l3}Z_{3k}+\delta_{l4}Z_{4k}+\delta_{l5}Z_{5k}+U_{lik}$$

where the Z variables are zone dummies (the reference zone is Costa Brava). Note that the k subscript has been dropped from the C and β parameters. In these pooled models, the β intercepts are interpreted with respect to the reference zone and the β slopes as a sort of average effect across zones.

In order to account for different seasonality shapes across zones as suggested by Figure 2, the zone dummies are allowed to have an effect on the prices of all months except August, as the effect of this month is introduced in the L_i equation. This would be a model with interaction between zone and seasonality but with slopes of predictors still constant across zones:

$$Y_{itk}=L_{ik}+S_{ik}C_t+\gamma_{t2}Z_{2k}+\gamma_{t3}Z_{3k}+\gamma_{t4}Z_{4k}+\gamma_{t5}Z_{5k}+D_{itk}\text{ , with }\gamma_{4k}=0\text{ (august)}$$
$$S_{ik}=1+\beta_{s1}X_{1ik}+...+\beta_{sj}X_{Jik}+U_{sik}$$
$$L_{ik}=\beta_{l0}+\beta_{l1}X_{1ik}+...+\beta_{lj}X_{Jik}+\delta_{l2}Z_{2k}+\delta_{l3}Z_{3k}+\delta_{l4}Z_{4k}+\delta_{l5}Z_{5k}+U_{lik}$$

The δ_l coefficients are expected percentage price changes with respect to Costa Brava and referred to August. The γ_{t3} coefficients are differences in the seasonal profile (differences in the percentage differences with respect to August) with respect to that of the Costa Brava. Under this model, the expected log price for zone 2 is:

$$E(Y_{it2})=\beta_{l0}+\beta_{l1}X_{1i2}+...+\beta_{lj}X_{Ji2}+\delta_{l2}+C_t\,(1+\beta_{s1}X_{1i2}+...+\beta_{sj}X_{Ji2})+\gamma_{t2}$$

and for zone 1 (Costa Brava, reference):

$$E(Y_{it1})=\beta_{l0}+\beta_{l1}X_{1i1}+...+\beta_{lj}X_{Ji1}+C_t\,(1+\beta_{s1}X_{1i1}+...+\beta_{sj}X_{Ji1})$$

This model with interaction between zone and seasonality is used as framework for estimating the effect of weather. As regards the selection of an appropriate weather variable, we assumed the monthly average of maximum daily temperature to be a good indicator of attractiveness of a tourism destination in the sun-and-beach segment. These temperatures were averaged over 1961-1990 in the closest observatories to the zones (Girona, Barcelona and Malaga airports, Reus and Alicante). Of course, raising temperatures above a certain threshold may no longer be appreciated. We assumed that increases in temperature are valued when they contribute to get a temperature below 27^{o}C closer to 27^{o}C. Most people find temperatures around 25-26^{o}C very pleasant and a maximum temperature of 27^{o}C implies several hours of pleasant temperature every day. The gaps with respect to 27^{o}C were finally expressed in comparison to those of the Costa Brava, that is the reference zone in the dummy-variable model.

Average of maximum daily temperature (^oC)					
	Brava	Maresme	Daurada	Blanca	Sol
May	20.9	20.2	21.4	24.2	23.9
June	25.1	24.0	25.1	27.7	27.2
July	28.9	27.2	28.5	30.8	29.7
August	28.2	27.3	28.1	31.1	30.1
September	25.6	25.4	25.8	29.1	27.8
October	20.9	21.6	21.8	24.7	23.7
Difference with respect to 27^oC, when negative					
	Brava	Maresme	Daurada	Blanca	Sol
May	-6.1	-6.8	-5.6	-2.8	-3.1
June	-1.9	-3.0	-1.9	0	0
July	0	0	0	0	0
August	0	0	0	0	0
September	-1.4	-1.6	-1.2	0	0
October	-6.1	-5.4	-5.2	-2.3	-3.3
Difference in the difference with respect to 27^oC compared to Costa Brava					
	Brava	Maresme	Daurada	Blanca	Sol
May	0	-0.7	0.5	3.3	3.0
June	0	-1.1	0	1.9	1.9
July	0	0	0	0	0
August	0	0	0	0	0
September	0	-0.2	0.2	1.4	1.4
October	0	0.7	0.9	3.8	2.8

Table 4: Weather data per zones

The idea is to constrain the γ_{tk} effects of the zone dummies on the seasonality shape to a linear function of the difference in the gap with respect to 27^oC compared to Costa Brava. The slope of that function will be the expected increase in price for a one-degree temperature increase, when this increase contributes to bringing temperature closer to 27^oC. We do not constrain the effect on price level in August for two reasons: first, temperatures are high enough in August in all zones; second, zone dummies freely affecting peak price level can account for heterogeneity across zones that cannot be explained by hotel category and services, and that we do not want to get confounded with the effect of temperature. The first equation of the model is rewritten as:

$$Y_{itk}=L_{ik}+S_{ik}C_t+\eta_t F_{tk}+D_{itk} ,$$

with:

$$F_{1k}=-0.7Z_2+0.5Z_3+3.3Z_4+3.0Z_5$$
$$F_{2k}=-1.1Z_2+0.0Z_3+1.9Z_4+1.9Z_5$$
.....

Note that F_{1k} is the May difference in the gap with respect to 27°C for the kth zone, F_{2k} the June difference, and so on. Thus, η_t is the percentage effect on price of a one-degree increase in temperature bringing temperature closer to 27°C. It is constrained to zero in August (reference period) and in July (as temperatures are consistently over 27°C). The model is run twice, assuming η_t to be fixed for all periods and assuming it to be time varying.

The Mplus program (Muthén & Muthén, 1998) was used for estimation by full information maximum likelihood under the assumption that log prices are normally distributed conditional on predictors. However, mean and covariance scaled test statistics robust to non-normal distributions were employed (Satorra, 1992). The data were collected for virtually all hotels that appear in any tour operator catalogues and that offer full board service (70% of the offer in the studied zones). Thus, this article can be considered to be a population study, in which models and curves are used to approximate the population data rather than to estimate the parameters of an underlying population model. Traditional goodness of fit indices and tests will then be interpreted as approximation measures, and no use will be made of confidence intervals or standard errors.

4. Results of Modelling each zone separately

First, individual models were fitted for each zone without predictors or constraints across zones. These models were fitted in order to assess the feasibility of the approach and the goodness of fit of the model with two random factors for peak level and seasonality. Actually, the peak level and seasonality factors accounted for virtually all the variance in monthly prices (the minimum R^2 for the Y variables was .89, and most were above .95). However, their estimates cannot be compared, as each zone may have a particular mix of hotel characteristics, thus leading to spurious effects. The curves of fitted log prices and of estimated seasonality coefficients are identical to those that were obtained in Figures 1 and 2.

Next, individual models were fitted for each zone with all predictors and without constraints across zones. These models were then simplified:

- The log of the number of rooms was not significant in any of the zones for the L factor and was dropped. Actually, this variable does not measure any service of the hotel or any benefit for the consumer, and thus its presence in the model may be misleading.
- Only two towns were significant and all towns were dropped from all models in order to preserve the comparability of the predictor sets, as specific towns are located in only one zone. The reductions in the R^2 for the S and L factors were minor.
- The remaining variables were significant in at least one zone and were preserved in all of them in order to keep the predictor sets comparable.

The fit of the models was excellent. The minimum R^2 for the Y variables was .89, the minimum R^2 for the L factor was .52. The minimum R^2 for the S factor was rather low at .15 though we understand that seasonal profiles may be harder to predict than absolute levels of price. In addition to R^2 , structural equation models offer another type of goodness of fit measure, that compares the fit of the actual model to that of a saturated model in which all identified relationships between variables are estimated (e.g. Bollen, 1989). In our case, such a saturated model would be one in which the predictor variables would affect the 6 monthly prices directly (instead of doing so through only the two dimensions represented by the level and seasonality factors) and in which the error terms would follow a 5-lag moving average process instead of a first order moving average process. A widely used measure of discrepancy between the fit of both models is the root mean squared error of approximation (RMSEA, Steiger, 1990; Browne & Cudeck, 1993), computed from the mean and covariance scaled test statistic (Satorra, 1992). Values of RMSEA below 0.05 are considered acceptable (Browne & Cudeck, 1993). This threshold is exceeded by none of the zones. Table 5 shows the estimates and fit indices. Effects resulting in price differences above 5% are bold faced.

The first part of the Table shows seasonality coefficients for a 3-star hotel without any of the characteristics represented by the dummy variables. Percentage reductions in prices with respect to August range from 69 to 90% in May, from 52 to 62% in June, from 9 to 26% in July, from 32 to 57% in September and from 61 to 96% in October. Some of these percentages look extremely high. This is explained by the fact that the difference in natural logs is somewhere between the percentage changes computed with respect to the lower and higher prices. For instance, the predicted price in May in Costa Brava is EXP(8.728-0.694)=3,084 ESP and in August EXP(8.728)=6,173 ESP, that is about the double. If we compute the percentage change with respect to the May price we get +100% and with respect to August, -50%. The parameter estimate is -.694 which is about half way.

The seasonality coefficients are represented graphically for each zone (Figure 4), together with the estimated expected log prices (Figure 3). Both graphs refer to a 3-star hotel without any of the services represented by the dummy variables. The graph of the expected log prices is markedly different from Figure 1. The graph of seasonality coefficients has about the same shape of Figure 2 but coefficients tend to be larger in absolute value. Once more, Costa del Sol and Costa Blanca exhibit a distinctive pattern in having less marked seasonality in September and October and more marked seasonality in July.

	Brava	Maresme	Daurada	Blanca	Sol
C_{tk}					
May	**-.694**	**-.733**	**-.895**	**-.731**	**-.689**
June	**-.523**	**-.550**	**-.613**	**-.563**	**-.622**
July	**-.100**	**-.086**	**-.137**	**-.262**	**-.240**
September	**-.500**	**-.534**	**-.567**	**-.323**	**-.379**
October	**-.728**	**-.794**	**-.963**	**-.611**	**-.690**
β_{lk}					
Intercept	8.728	8.667	8.924	8.930	8.989
Beach	**.117**	-.032	**.063**	**.064**	**.061**
Room	**.087**	**.092**	.012	**.052**	**.050**
Parking	**.125**	.011	**.050**	**.089**	.046
Sport	**.101**	**.054**	**-.063**	.028	.043
H1	**-.236**	**-.232**	None in this zone	**-.384**	None in this zone
H2	**-.065**	**-.140**	**-.096**	**-.202**	**-.164**
H4	**.521**	**.393**	**.344**	**.264**	**.326**
β_{sk}					
Beach	**-.090**	-.029	.028	**-.110**	**-.128**
Room	**-.051**	-.023	**-.060**	**-.168**	**-.080**
Parking	-.021	-.029	**-.128**	**-.050**	-.039
Sport	.017	**-.058**	-.049	.042	-.036
H1	**-.138**	**-.110**	None in this zone	-.026	None in this zone
H2	**-.132**	.013	**-.279**	.041	.022
H4	**-.208**	-.046	**-.190**	**-.148**	**-.237**
covariances					
L-S	-.014	.005	-.014	-.005	-.012
D_1-D_2	.001	.000	.001	.001	.001
D_2-D_3	.001	.000	.002	.001	.001
D_3-D_4	.000	-.001	.000	.000	.000
D_4-D_5	.000	.002	-.001	.000	.000
D_5-D_6	.000	.001	-.002	.000	.000

	Brava		Maresme		Daurada		Blanca		Sol	
Disturbances										
D_t	R^2	stdev	R^2	stdev	R^2	stdev	R^2	stdev	R^2	stdev
May	..98	.05	.97	.04	.97	.06	.98	.04	.98	.04
June	.97	.05	.98	.03	.97	.05	.99	.04	.98	.05
July	.95	.03	.99	.00	.91	.06	.96	.06	.96	.06
August	.98	.07	.92	.06	.95	.04	1.00	.00	1.00	.00
Septem.	.97	.05	.89	.07	.99	.03	.97	.05	.96	.06
October	.98	.05	.93	.05	.95	.08	.97	.06	1.00	.00
U										
L	.73	.15	.66	.12	.52	.13	.78	.12	.55	.16
S	.17	.19	.15	.10	.22	.22	.30	.19	.33	.22
Fit measure										
RMSEA		.000		.049		0.000		0.000		.040

Table 5: Estimates of individual models per zones. Bold faced if larger than .05

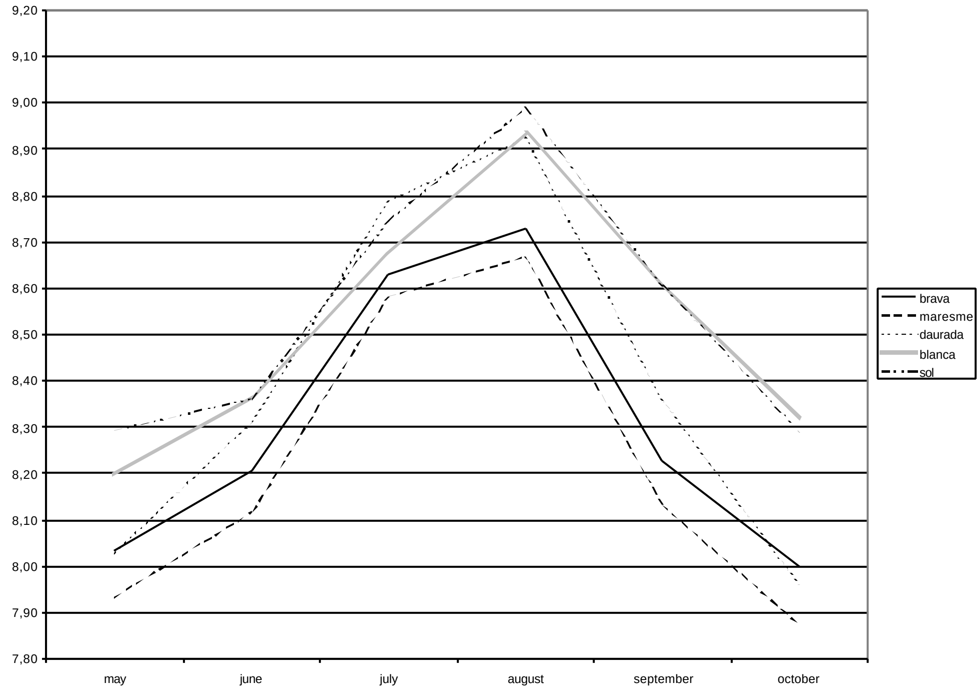

Figure 3: Fitted average log prices for a 3-star hotel without additional services

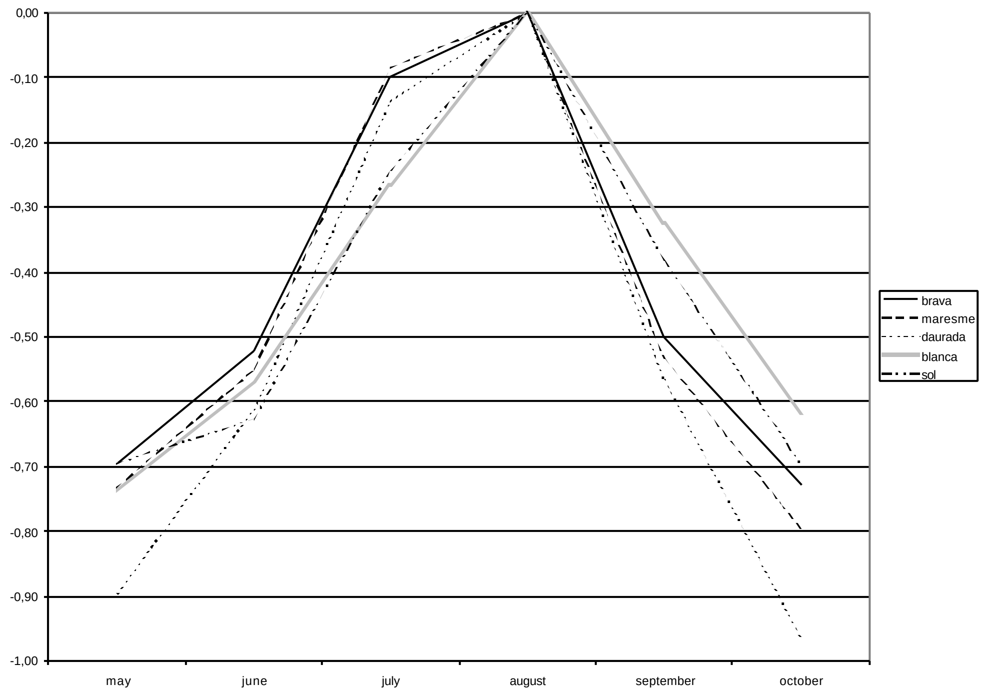

Figure 4: Seasonal indices for a 3-star hotel without additional services

The second part of the table shows the equations predicting the August price level. The intercepts refer to a 3-star hotel without any of the services represented by the dummy variables and show Costa Daurada, Costa Blanca and Costa del Sol to be more expensive than Costa Brava and Costa del Maresme. As regards slopes, 1-star hotels are on average between 23 and 38% cheaper than 3-star hotels, 2-star hotels are on average between 7 and 20% cheaper than 3-star hotels, and 4-star hotels are on average between 26 and 52% more expensive than 3-star hotels. Differences can be observed across zones and price differences between the lower categories and the reference are highest in Costa Blanca while price differences between the higher category and the reference are highest in Costa Brava. Price is not only determined by category, but other attributes also play a role. All attributes have a positive contribution on price for all zones except for a couple of anomalous results which, fortunately, are among the values in the low range. The contributions tend to be around or above 5% in the majority of zones for all attributes except sport facilities. Differences can be observed across zones and, overall, in Costa Brava all attributes have the highest effect on price.

The next part of the table shows the equations predicting the amplitude of seasonal fluctuations. Low season price reductions are smaller for 1, 2 and 4 star hotels than for 3 star hotels, which suggests that 4 star hotels are always expensive and 1 and 2 star hotels always cheap. The presence of services also tends to decrease the amplitude of seasonality. The reference hotels (3 star without special attributes) seem to be the most seasonal ones.

The standard deviations of the disturbances for the L and S factors suggest that some amount of hotel variation remains in both August price level (standard deviations around .15, that is, variations around +/-30% are possible between hotels with the same attributes and in the same zone) and amplitude of seasonality (standard deviations around .2, that is, variations around +/-40% are possible).

5. Results of modelling the pooled data of all zones

First, the additive model with constant seasonality profile across zones was fitted (first column of Table 6). This model gives a poor fit to the data, with a high RMSEA and D_{itk} standard deviations that are substantially higher than when modelling zones separately.

On the contrary, as expected from the pattern in Figure 4, the model with interaction between zone and seasonality (second column of Table 6) has a much better fit (RMSEA=0.007, D_{itk} standard deviations about as high as when modelling zones separately). This model is much more parsimonious than those in the previous section as it includes only one effect of each category dummy, which can be interpreted as overall or average effect of the given characteristic. Seasonality coefficients refer to Costa Brava, and have to be added to the γ coefficients to get meaningful seasonality coefficients for other zones. According to this model, situation in front of the beach leads to average peak increases of 6.9% and reduces the amplitude of seasonal fluctuations by 6.5%. Room services increase peak price by 9.8% and reduce the amplitude of seasonality by 6%. Parking place increases peak price by 8.2% and reduces the amplitude of seasonality by 3.8%. Sport facilities increase peak price by 4.7% and reduce the amplitude of seasonality by 1.2%. 1-star hotels have peak prices 29.6% lower than 3-star hotels and 8.5% narrower seasonal fluctuations. 2-star hotels have peak prices 12.4% lower than 3-star hotels and 6% narrower seasonal fluctuations. 4-star hotels have peak prices 35.9% higher than 3-star hotels and 18.7% narrower seasonal fluctuations. In August, hotels in Maresme are 21.7% cheaper than in Costa Brava, in Costa Daurada 1.9% cheaper, in Costa Blanca 4.3% more expensive and in Costa del Sol 9.5% more expensive.

	Additive	Interaction seasonality-zone	constant weather effect	time variant weather effect
C_t				
May	**-.742**	**-.701**	**-.755**	**-.734**
June	**-.566**	**-.527**	**-.571**	**-.553**
July	**-.158**	**-.106**	**-.161**	**-.158**
September	**-.461**	**-.491**	**-.458**	**-.510**
October	**-.745**	**-.735**	**-.759**	**-.804**
γ_{tk}				
May-Maresme		**-.066**		
May-Daurada		**-.080**		
May-Blanca		.012		
May-Sol		.047		
June-Maresme		-.048		
June-Daurada		-.017		
June-Blanca		-.002		
June-Sol		**-.051**		
July-Maresme		.014		
July-Daurada		-.019		
July-Blanca		**-.123**		
July-Sol		**-.101**		
Sept.-Maresme		**-.067**		
Sept.-Daurada		-.012		
Sept.-Blanca		**.161**		
Sept.-Sol		**.114**		
Oct.-Maresme		**-.093**		
Oct.-Daurada		**-.106**		
Oct.-Blanca		**.142**		
Oct.-Sol		**.078**		
η_t				
May			.017	.020
June			.017	.018
September			.017	.119
October			.017	.054

Table 6: estimates of models with pooled data. Bold faced if larger than .05

	Additive		Interaction seasonality-zone		constant weather effect		time variant weather effect	
β_l and δ_l								
Intercept	8.830		8.810		8.821		8.811	
Beach	**.068**		**.069**		**.068**		**.067**	
Room	**.096**		**.098**		**.095**		**.094**	
Parking	**.082**		**.082**		**.081**		**.082**	
Sport	.047		.047		.046		.044	
H1	**-.297**		**-.296**		**-.295**		**-.294**	
H2	**-.120**		**-.124**		**-.119**		**-.122**	
H4	**.357**		**.359**		**.358**		**.357**	
Maresme	**-.208**		**-.217**		**-.196**		**-.186**	
Daurada	-.025		-.019		-.002		.007	
Blanca	-.013		.043		.008		.021	
Sol	**.045**		**.095**		**.057**		**.092**	
β_s and δ_s								
Beach	**-.074**		**-.065**		**-.077**		**-.061**	
Room	**-.065**		**-.060**		**-.068**		**-.052**	
Parking	**-.051**		-.038		**-.052**		-.037	
Sport	-.013		-.012		-.017		-.007	
H1	**-.104**		**-.085**		**-.125**		**-.105**	
H2	**-.065**		**-.060**		**-.062**		**-.047**	
H4	**-.186**		**-.187**		**-.187**		**-.176**	
Maresme	**.115**							
Daurada	**.086**							
Blanca	**-.133**							
Sol	**-.107**							
covariances								
L-S	-.016		-.010		-.015		-.010	
D_1-D_2	.000		.001		-.001		.001	
D_2-D_3	.003		.001		.003		.002	
D_3-D_4	.002		-.001		.002		.000	
D_4-D_5	.004		.001		.003		.001	
D_5-D_6	.004		-.001		.005		-.001	
Disturbances								
D_t	R^2	stdev	R^2	stdev	R^2	stdev	R^2	stdev
May	.99	.04	.98	.06	1.00	.00	.97	.06
June	.97	.05	.97	.05	.98	.04	.97	.06
July	.91	.09	.96	.05	.91	.09	.93	.09
August	.91	.08	.98	.04	.93	.07	.97	.04
September	.92	.10	.97	.06	.91	.10	.97	.06
October	.95	.08	.98	.04	.94	.10	.98	.05
U								
L	.76	.13	.73	.15	.75	.13	.73	.14
S	.35	.18	.15	.20	.17	.20	.14	.19
Fit measure								
RMSEA		0.116		.007		0.116		.094

Table 6 continued: estimates of models with pooled data . Bold faced if larger than .05

The fit of the model including weather and with a constant effect of temperature on price (third column of Table 6) is not good, as it leads to an unacceptable RMSEA and to a substantial increase in D_{itk} standard deviations with respect to the previous model. The model with time varying temperature prices (last column in Table 6) gets somewhat more support from the data. RMSEA is still too high but D_{itk} standard deviations are nearly back to the level of the good fitting interaction model. The estimated increases in price for each additional degree of temperature are 11.9% for September and 5.4% for October. The effects in May and June are much lower, at 2% and 1.8% respectively.

6. Discussion

In this article random-effect models were fitted to study the peak level and seasonality of hotel prices and their predictors. Structural equation models were used with this purpose. The major advantage of using structural equation models is the fact that treating random effects as latent variables allows researchers to relate these effects to a set of predictor variables.

The variables which showed most important to explain level and seasonality were zone, category, closeness to the beach, room equipment and availability of parking place. Hotel category and attributes affect level in the expected way. Seasonality is higher for 3-star hotels without any of the additional attributes considered. These effects may differ from zone to zone, though the overall patterns are roughly similar.

The effect of zone on level shows that hotel attributes are not the sole determinants of price, and that tourists do not only pay for a hotel room but also for its environment. This fact must inspire the local government policies. In Spain, local policy also plays a role in another respect, namely by limiting permits to build new hotels. Further research is needed on the effect on prices of these aspects of supply. The fact that certain tour operators specialise in certain zones might also contribute to differential zone prices, but this effect has been partialled out of the prices before fitting the model.

If the zone effect on seasonality is attributed to weather and the zone effect on level is attributed to other characteristics of the zone, then estimates of the effect of weather on price can be obtained. Of course this is done under the assumptions that weather does not affect peak level, that other characteristics of the zone do not affect seasonality, and that the effect of temperature is flat above 27^{o}C. If these assumptions do not hold, the results are questionable. Many zone-specific variables, such as environmental quality, landscape or urban services, are indeed non-seasonal. In any case, it must be admitted that the estimates of the effect of weather may be contaminated by differential low season strategies to face the diminished low-season demand. Actually, all zones have virtually no vacancy during the peak months, but different zones have different levels of vacancy in the low season. The percentage of vacancy in October 1999 was about 10% higher in Costa Brava than in Costa del Sol or Costa Blanca. Besides, 13% of Costa Brava hotels had already closed by early October, whereas only 1% of hotels in Costa Blanca and Costa del Sol had done so. If Costa del Sol and Costa Blanca manage to keep September and October prices higher, in spite of a higher supply, and yet attract more customers, then it is suggested that the model may even have subestimated the economic value of weather.

To a large extent, prices negotiated between hotels and tour operators depend on past demand (Espinet, 1999). Sun-bathing is reported as the main activity for 75.2% of non-business visitors (Departament d'Indústria, Comerç i Turisme, 1999), which provides additional support to the importance of weather variables on demand. However, availability of holidays is of course a prerequisite for there being demand. In Spain schools do not open until mid or late September, which allows families with children to enjoy some weeks of holidays in September, thus increasing demand in zones with a mild weather, and also price. This could explain the high differences in weather effect between late spring and early autumn.

The authors are currently working to enlarge the data base with prices during year 2000, which will eventually make it possible to include a trend latent variable in the model.

References

Batista-Foguet, J. M. and Coenders, G. (2000), *Modelos de Ecuaciones Estructurales [Structural Equation Models]*. La Muralla, Madrid, Spain.

Bock, R. D. (1989), *Multilevel Analysis of Educational Data*, Academic Press, San Diego.

Bollen, K. A. (1989), *Structural Equations with Latent Variables*, John Wiley & Sons, New York.

Browne, M.W. and Cudeck, R. (1993), Alternative Ways of Assessing Model Fit, in Bollen, K. A. and Long, J. S., (Eds.) *Testing Structural Equation Models*, Sage, Thousand Oaks, Ca, (pp. 136-162).

Cassel, E. and Mendelsohn, R. (1985), The Choice of Functional Forms for Hedonic Price Equations: Comment, *Journal of Urban Economics*, 18, 135-142.

Chias, J. (1996), ¿Existe o no una Estrategia de Precios en el Turismo Español? [Is there a Pricing Strategy in Spanish Tourism?] *EDITUR*, 1894, 11-15.

Clewer, A., Pack, A. and Sinclair, T. (1992), Price Competitiveness and Inclusive Tourism Holidays in European Cities, in Johnson, P. and Thomas, B. (Eds.), *Choice and Demand in Tourism*, Mansell, London, (pp 123-144).

Departament d'Indústria, Comerç i Turisme (1999), *Resum de la Temporada Turística 1999 [summary of the 1999 turist season]*. Generalitat de Catalunya, Barcelona, Spain.

Espinet, J.M. (1999), *Anàlisi dels Preus al Sector Hoteler de la Costa Brava Sud [Analysis of Hotel Prices in Southern Costa Brava]*, Unpublished doctoral dissertation, University of Girona, Girona, Spain.

Espinet, J. M., Saez, M., Coenders, G. and Fluvià, M. (forthcoming). The Effect on Prices of the Attributes of Holiday Hotels: a Hedonic Prices Approach, *Tourism Economics*

Halvorsen, R. and Pollakowski, H.O. (1981), Choice of Functional Form for Hedonic Price Equations, *Journal of Urban Economics*, 10, 37-49.

Jaime-Pastor, V. (1999), Un Análisis de los Precios Hoteleros Empleando Funciones Hedónicas [Analysis of Hotel Prices using Hedonic Functions], *Estudios Turísticos*, 139, 65-87.

Laird, N. M. and Ware, J. H. (1982), Random-Effects Model for Longitudinal Data, *Biometrics,* 38,963-974.

McArdle, J. J. and Epstein, D. (1987), Latent Growth Curves with Developmental Structural Equation Models, *Child Development*, 58, 110-133.

Meredith, W. and Tisak, J. (1989), Latent Variable Modeling in Heterogeneous Populations, *Psychometrika,* 54, 557-585.

Muthén, B. (1997), Latent Variable Modeling of Longitudinal and Multilevel data, in Raftery A., (Ed.) *Sociological Methodology 1997,* Basil Blackwell, Boston, (pp. 453-480).

Muthén, B. (2000), Integrating Person-Centered and Variable-Centered Analyses, Growth Mixture Modeling with Latent Trajectory Classes, *Alcoholism, Clinical and Experimental R esearch,* 24, 1-10.

Muthén, L.K. and Muthén, B. (1998), *Mplus User's Guide,* Muthén & Muthén, Los Angeles.

Rosen, S (1974), Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, *Journal of Political Economy*, 82, 34-55.

Satorra, A. (1992), Asymptotic Robust Inferences in the Analysis of Mean and Covariance Structures, in Marsden, P. V., (Ed.) *Sociological Methodology 1992*, Basil Blackwell, Oxford, (pp. 249-278).

Sinclair, M.T., Clewer, A. and Pack, A. (1990), Hedonic Prices and the Marketing of Package Holidays: the Case of Tourism Resorts in Malaga, in Ashworth, G. and Goodall, B., (Eds) *Marketing of Tourism Places*, Routledge, London, (pp 85-103).

Sörbom, D. (1974), A General Method for Studying Differences in Factor Means and Factor Structures Between Groups, *British Journal of Mathematical and Statistical Psychology*, 27, 229-239.

Steiger, J. H. (1990), Structural Model Evaluation and Modification. An Interval Estimation Approach, *Multivariate Behavioural Research,* 25, 173-180.

Taylor, P. (1995), Measuring Changes in the Relative Competitiveness of Package Tour Destinations, *Tourism Economics*, 1, 169-182.

Willett, J. B. and Sayer, A. G. (1993), Using Covariance Structure Analysis to Detect Correlates and Predictors of Individual Change over Time, *Psychological Bulletin*, 116, 368-381.

Witt, S.F. and Moutinho, L. (1994), *Tourism Marketing and Management Handbook*. Prentice Hall, Upper Saddle River, NJ.