

A Bivariate Integer Valued Allocation Model for Guest Nights in Hotels and Cottages*

Kurt Brännäs and Jonas Nordström

Department of Economics, Umeå University
SE-90187 Umeå, Sweden

Umeå Economic Studies 547, 2000

Abstract

The number of Norwegian guest nights in Swedish hotels and cottages is studied. Aggregation of an integer-valued AR(1) model and a two-stage demand model underlies the empirical results. The parameters in the model are check-out probability, mean check-in and the probability of selecting the hotel alternative. These parameters are specified to depend on economic variables implied by demand analysis. Via the probability of selecting a hotel, the empirical results indicate a substitution towards less expensive accommodation as the Swedish price level increases. For the check-out probability, an increase in the cottage price reduces the probability for staying another night in cottage, whereas an increase in the hotel price indicates a decrease in the check-out probability for hotel.

Keywords: Integer-valued time series, demand analysis, tourism, accommodation, hotel, cottage.

JEL Classification: C32, C35, C51, D12.

*The financial support from the Wallander Foundation is gratefully acknowledged.

1. Introduction

This paper models the number of Norwegian guest nights in Swedish hotels and cottages. We adopt a modelling strategy that builds both on how time series arise and on demand analytical results for two-stage decision making. By departing from a simple time series model for daily guest nights, i.e. from the micro level, and aggregating, temporally as well as cross-sectionally, we obtain a simple and operational time series model for aggregate guest nights. This modelling approach renders parameters that are easy to interpret. The parameters are mean check-in, the daily check-out probability, and the daily probability of selecting hotel rather than cottage. These parameters are next set to depend on economic variables implied by demand analysis. In this analysis, potential guests are in a first stage assumed to choose which countries to visit, and then conditionally on the first stage allocation problem they decide on the form of accommodation.

The approach results in a demand model that is nonlinear in economic variables and also dynamic. This blends previous economic and time series approaches in a new way. Previous studies on international tourism have mainly tried to explain the number of arriving guests to a country or the expenditures spent by the tourists (Crouch, 1994). By using accommodation statistics we can extend the analysis and also account for the length of the stay. The adopted two-stage decision process also makes it possible to incorporate data on guest night numbers and tourism expenditures within a country in the same allocation model. Although the same explanatory variables that we use have appeared in some other studies, the formal derivation of the demand functions have seldom been shown. Other studies of international tourism based on economic demand models are, e.g., Adamowicz (1994), Melenberg and van Soest (1996), and Nordström (1999). Studies employing pure time series analytical models are, e.g., Garcia-Ferrer and Queralt (1997) and Gustavsson and Nordström (1999), while studies using mixed economic – time series approaches different from ours include Turner, Kulendran and Fernando (1997), and Witt and Witt (1995).

The underlying integer-valued model for the number of guest nights in a region's

hotels or cottages is an integer-valued autoregressive (INAR) model of order one. This count data model decomposes today's number of guest nights in hotels and cottages into those that stay over from yesterday and those guests that arrive to the region in the intervening time. The INAR model was introduced by McKenzie (1985) and has later been elaborated upon by Al-Osh and Alzaid (1987), Brännäs (1995) and others. In a previous paper, Brännäs, Hellström and Nordström (1999, hereafter BHN) considered the national and monthly number of guest nights in hotels and demonstrated that the aggregation approach yields an integer-valued moving average (INMA) model of order one. BHN also considered estimation and forecasting issues and gave a substantial application to a series of Norwegian guest nights in Swedish hotels. The present paper extends on the work of BHN in the sense of studying two series jointly as well as by explicitly focusing on the allocation between the two types of accommodation.

In Section 2 the basic stochastic and micro level model is presented along with the model for a monthly time series at the national level. Also discussed is how one may include explanatory variables into the stochastic model. Section 3 discusses the demand analytic framework, and Section 4 presents the data. Estimation issues are considered in Section 5. The empirical results are given in Section 6. A concluding discussion finishes the paper.

2. Model and Aggregation

We consider a model for the number of hotel (y_H) and cottage (y_C) visitors for a region at a certain day t :

$$y_{Ht} = \alpha_H \circ y_{Ht-1} + \beta_H \circ \epsilon_t \quad (1a)$$

$$y_{Ct} = \alpha_C \circ y_{Ct-1} + \beta_C \circ \epsilon_t. \quad (1b)$$

The numbers of visitors in each category consist of those staying from the previous day ($\alpha_i \circ y_{it-1}, i = H, C$), and those having a check-in ($\beta_i \circ \epsilon_t$), in the period $(t-1, t)$. Here, the integer-valued random variable ϵ_t is the number of arriving guests to the region, e.g., from some specific country. The mean number of arriving guests at time t is $E(\epsilon_t) = \lambda'$. The guests are allocated to either of the two types of dwelling

by the thinning operation $\beta_H \circ \epsilon_t = \sum_{i=1}^{\epsilon_t} v_i$. The independent random variables v_i take the value $v_i = 1$ if guest i , of totally ϵ_t new guests, decides to stay in a hotel and $v_i = 0$ otherwise. The $v_i = 0$ event then corresponds to a check-in to a cottage. The probability for $v_i = 1$ is β_H , so that $\Pr(v_i = 0) = 1 - \beta_H = \beta_C$. In Appendix A1 we give some detailed properties of the model. Brännäs and Hellström (2000) obtain related properties of the model under weaker assumptions.

Granger and Morris (1976) and others have established that cross-sectional aggregation over N AR(1) model leads to an analogous model. BHN demonstrated a related result for the INAR(1) models in (1a)-(1b). It follows directly that the cross-sectionally aggregated day-to-day model is bivariate INAR(1). In the aggregate model $\varepsilon_t = \sum_{j=1}^N \epsilon_{jt}$, with mean $\lambda = N\lambda'$, represents the number of arriving guests at the national level. The cross-section, aggregation step could have been avoided by directly focusing at national day-to-day guest nights. Brewer (1973) and others, and BHN for the INAR model, have demonstrated that temporal aggregation of AR(1) and INAR(1) models yield ARMA(1,1) and INARMA(1,1) models, respectively. In addition, BHN gave arguments suggesting that the INARMA(1,1) practically reduces to an INMA(1) model at the new time scale if temporal aggregation is over as long a horizon as a month. Applying an INAR(1) model directly to monthly guest night numbers is likely to be of little interest as month-to-month check-out probabilities $(1 - \alpha)$ can be expected to be equal to one. In this instance the day-to-day movements are those of major interest.

Temporal aggregation (cf. the Appendix) gives us a bivariate INMA(1) model at the national and monthly (t is now and hereafter referring to month) level for the aggregate hotel (z_H) and cottage (z_C) series:

$$z_{Ht} = \beta_H \circ \xi_t + \theta_H \beta_H \circ \xi_{t-1} \quad (2a)$$

$$z_{Ct} = \beta_C \circ \xi_t + \theta_C \beta_C \circ \xi_{t-1}. \quad (2b)$$

Here, $\theta_i \approx \alpha_i / (1 - \alpha_i)$, $i = H, C$, and $E(\xi_t) = \lambda = Nh\lambda'$ corresponds to the national and monthly mean number of arrivals to hotels and cottages (N is the number of regions, h is the number of days in the month, and λ' is the regional mean arrival). The model in (2a)-(2b) is a simplified representation that has the same

first moment properties as a true temporal aggregation model. The advantage with (2a)-(2b) is that it is readily interpretable in terms of the interesting parameters α_i (through θ_i), β_i , and λ . For estimation purposes we need the conditional expectation $E(z_{it}|\mathbf{z}_1, \dots, \mathbf{z}_{t-1}) = N\lambda'\beta_i[h(1 + \theta_i) - \theta_i] + \theta_i\beta_i\varepsilon_{t-1}, i, j = H, C$, and with $\mathbf{z}_t = (z_{Ht}, z_{Ct})$. Neither second order moments nor a specific distribution will be used for estimation.

The INMA(1) model is obtained under quite stringent assumptions, e.g., of constant parameters. While this time invariance may be a reasonable approximation within the month it most likely is too simplified over the annual horizon. Notably, there are strong seasonal patterns in guest night series, so that accounting for this feature by differencing appears a reasonable approach. However, our main interest is in how the mean check-in (λ), the check-out (α_i or θ_i) and hotel (β_H) probabilities could be related to economic variables. Studies based on demand analysis strongly suggest that the income level of the country of origin, prices and exchange rates are important determinants to the mean check-in (e.g., Nordström, 1999). Note that the parameters of the model are then time indexed in correspondence with the appropriate lag in the $\{\xi_t\}$ sequence. In addition, as parameters characterize the occurrences in time intervals ending at the indexed time, variables should be lagged to measure levels at the initial time point. This has the added advantage of avoiding potential endogeneity problems.

For sufficiently large z_{Ht} and z_{Ct} values the λ_t will be far away from zero so that empirically $\hat{\lambda}_t > 0$. When this is the case we may prefer to use a linear form $\lambda_t = \mathbf{w}_t\boldsymbol{\gamma}$ instead of an exponential one. The linear form can be expected to simplify estimation. Though \mathbf{w}_t is indexed by t it should contain lagged values on explanatory variables as λ_t (and the other parameters) characterize events in the time interval $(t - 1, t)$. Since $\alpha_{it} = \theta_{it}/(1 + \theta_{it}) \in [0, 1]$, $\theta_{it} \geq 0, i = H, C$, and an exponential representation $\theta_{it} = \exp(\mathbf{u}_t\boldsymbol{\delta}_i)$ is then appropriate. Note that this results in a logistic distribution function $\alpha_{it} = 1/[1 + \exp(-\mathbf{u}_t\boldsymbol{\delta}_i)]$ (cf. Brännäs, 1995). For β_{Ht} we may use a related logistic distribution function, i.e. $\beta_{Ht} = 1/[1 + \exp(-\mathbf{x}_t\boldsymbol{\rho})]$ for each month t . Which economic variables to include in the \mathbf{w}_t , \mathbf{u}_t and \mathbf{x}_t vectors is discussed in the next section. Each vector could also include non-economic variables,

such as monthly dummy variables. The parameter vector $\boldsymbol{\rho}$ is unknown as is $\boldsymbol{\gamma}$ and $\boldsymbol{\delta} = (\boldsymbol{\delta}'_H, \boldsymbol{\delta}'_C)'$.

3. Demand Analysis

In this section we consider demand functions for guest-nights in hotels and cottages at the national level and the preferences that may have generated them. Since we later will consider the Norwegian tourism demand in Sweden, we use a theoretical framework that is appropriate for international tourism. From market surveys we know that international tourists often make one longer trip and several shorter trips. This makes us believe that tourists have a preference structure associated with product differentiation. That is, they like to visit many different places, each possessing some unique properties that distinguishes it from other destinations, e.g., with respect to the supply of recreation facilities, the weather, or the distance the visitor has to travel to reach it.

Beginning with the work of Spence (1976) and Dixit and Stiglitz (1977) such preferences can be represented by a two-level utility function

$$U(\mathbf{q}) = F[F^1(\mathbf{q}_1), \dots, F^m(\mathbf{q}_m)]. \quad (4)$$

Here, the subutility function $F^j(\mathbf{q}_j)$ consists of goods and services (such as guest nights in hotels and cottages) consumed in the j th country (out of m countries). Although the preference structure is based on the assumption of taste for variety, this does not mean that the individual tourist necessarily would like to stay in both hotels and cottages during their stay in Sweden. All we require is that the population is composed of tourists with different preferences for staying in hotels and cottages, i.e. that there exist a taste for variety in aggregate demand.

To represent the preference ordering for the subutility we employ the Stone-Geary utility function

$$F^j(\mathbf{q}_j) = \prod_{i=1}^n (q_{ji} - \kappa_{ji})^{\varphi_{ji}}, \quad q_{ji} > \kappa_{ji}, 1 > \varphi_{ji} > 0, \sum_{i=1}^n \varphi_{ji} = 1,$$

where $q_{ji}, i = H, C$, represents the number (since tourism consumption spreads over a large number of goods and services n is generally larger than 2) of guest nights

in hotels and cottages in a specific country, and κ_{ji} is a parameter. The reason for choosing a quasi homothetic utility function is primarily due to the fact that most studies on international tourism demand find an income or expenditure elasticity larger than one. By applying this specification the elasticities are allowed to be larger than one, but will tend to unity as the income or expenditures increase.

The preferences for the first stage allocation problem, i.e. the decision for which country to visit can then be represented by a Cobb-Douglas aggregator utility function, such that the utility function (4) can be written as

$$U(\mathbf{q}) = \prod_{j=1}^m \left[\prod_{i=1}^n (q_{ji} - \kappa_{ji})^{\varphi_{ji}} \right]^{\phi_j}, \quad 1 > \phi_j > 0, \quad \sum_{j=1}^m \phi_j = 1.$$

The aggregator utility function translates all subutility levels, i.e. tourism in different countries, to an overall utility or welfare level. By defining appropriate quantity, q_j , and price, p_j , indices (Anderson, 1979) for the goods and services in each country we can solve the first stage allocation problem, i.e. for the demand of tourism in each country, by maximizing $F(\mathbf{q}) = \prod_{j=1}^m (q_j - \kappa_j)^{\phi_j}$, subject to $M = \sum_{j=1}^m p_j q_j$, where M is total expenditures on outbound tourism.¹ This results in the following demand equations

$$q_j = \kappa_j + \frac{\phi_j}{p_j} \sum_{k=1}^m p_k \kappa_k + \frac{\phi_j}{p_j} M, \quad j = 1, \dots, m \quad (5)$$

and corresponds to the variables in the λ function.²

Conditional on the allocation in the first stage we can solve the second stage allocation problem, i.e. the demand for each commodity within a country, by maximizing $F^j(\mathbf{q}_j) = \prod_{i=1}^n (q_{ji} - \kappa_{ji})^{\varphi_{ji}}$ subject to the expenditures that have been allocated to vacation in country j , i.e. $\phi_j M = M_j = \sum_{i=1}^n p_{ji} q_{ji}$ (Pollak, 1971). The resulting demand equation is given by

$$q_{ji} = \kappa_{ji} + \frac{\varphi_i}{p_{ji}} \sum_{l=1}^n p_{jl} \kappa_{jl} + \frac{\varphi_i}{p_{ji}} M_j, \quad i = 1, \dots, n. \quad (6)$$

¹It is by no means obvious that the demand for outbound tourism should solely be explained by the prices and expenditures on outbound tourism. Depending on the purpose of the analysis, one could also include prices and expenditures on domestic tourism or recreation in the demand for outbound trips.

²The estimated model is actually a three-stage, or higher, allocation model, where the allocation of total expenditures on all goods and services is not estimated.

This depends only on the prices of the goods and services within the country and total expenditure allocated to it, and corresponds to the explanatory variables in the β function. Hence, the consumption of guest nights in hotels and cottages in one country is influenced by the prices of tourism in the other countries only through their effects on country expenditure.

Conditional on the first-stage allocation problem, we can also derive the variables that enter the α -functions. Since this is the same maximization problem as the second stage allocation problem shown above, the same variables will enter both the β and α functions. The average duration $[1/(1 - \alpha_i)]$, $i = H, C$, of stay is thus determined by the expenditures allocated to tourism in country j , $\phi_j M$, and the prices in country j , p_{jl} , $l = 1, \dots, n$, as shown in (6). Thus the probability of staying another night in a hotel, α_H , depends not only on the hotel price, but on all prices for tourism goods within the country.

As the expected duration of stay is $1/(1 - \alpha_i)$, the demand for guest nights in hotels and cottages corresponds to $q_{ji} = \beta_i/(1 - \alpha_i)$, $i = H, C$. Marginal effects can be obtained by calculating $\partial q_{ji}/\partial x_k$ for different choices of x_k . Given the specification discussed in the previous section, $\beta_H/(1 - \alpha_i) = [1 + \exp(\mathbf{u}\boldsymbol{\delta}_i)]/[1 + \exp(-\mathbf{x}\boldsymbol{\rho})] = [1 + \theta_i]/[1 + \exp(-\mathbf{x}\boldsymbol{\rho})]$, for $i = H, C$, it appears reasonable to set $\mathbf{u} = \mathbf{x}$. The functional form suggests that it may be difficult to estimate the parameter vectors $\boldsymbol{\delta}_i$ and $\boldsymbol{\rho}$ separately.

4. Data

The two series for Norwegian guests in Swedish hotels and cottages, 1986:1–1999:9, are presented graphically in Figure 1. As is obvious there are strong seasonal patterns in both series, both have major peaks in July and their troughs are in the fall-winter period. The numbers of guests are much lower in the early 1990s than in the late 1980s and 1990s, and the direct correlation between the two series is 0.93. The number of cottage guests are much smaller throughout the period, cf. Figure 2, and there is a strong seasonal pattern also in the percent that stay in cottages. The series suggests that leisure travel is more frequent in the typical vacation weeks in

the summer and in the spring. The latter period corresponds to the skiing periods in the mountains.

Nordström (1996) finds that tourism consumption spreads over a large number of goods and services. We therefore take the consumer price indices (CPI) as proxies for tourism prices in each country. As the purchasing power of the visitors from Norway, is affected by the exchange rate, we adjust the CPI of country k with the exchange rate (ER_{Nk}) between the Norwegian krona (NOK) and the currency of country k , such that

$$p_k = \text{CPI}_k / ER_{Nk}, \quad k = 1, \dots, m. \quad (7)$$

From OECD statistics (1992) one can see that Norwegians have mainly allocated their guest nights to the other Scandinavian countries and to Germany. Therefore, we include the prices of tourism in Denmark, Finland and Germany, in addition to the Swedish tourism price in equation (5). Since data for Norwegian's expenditure on outbound travel is not available, we have used Norwegian gross domestic product (GDP) as a proxy. On a yearly basis the Norwegian GDP and gross domestic income, which can be seen as the total expenditure allocated between all goods in a first step (cf. footnote 2) have a correlation of 1.00. The GDP variable is expressed in real values and based on quarterly observations.

The variables that have been included in the β and α functions are the real price of guest nights in hotels and cottages,³ cf. Figure 1, the Swedish CPI to account for all other goods and services the tourist consumes, and real Norwegian expenditures on tourism in Sweden from the balance of payments. Consequently, the $M_j = \phi_j M$ variable in (6) will be measured correctly although we have not been able to use the true variable for M in (5). All variables have been adjusted with the exchange rate (ER_{NS}) between the Swedish krona (SEK) and NOK, i.e. the variables are expressed in NOK, as in (7). As Figure 1 reveals, there is a strong seasonal pattern in both the cottage and hotel prices series, with peaks in February, March and July

³The cottage price index reported by Statistics Sweden (updated annually) have been adjusted by an additive seasonal pattern based on expenditures for nights spent in cottages. Source: the Tourist and Travel Data Base (TDB) 1990:1 - 1996:9.

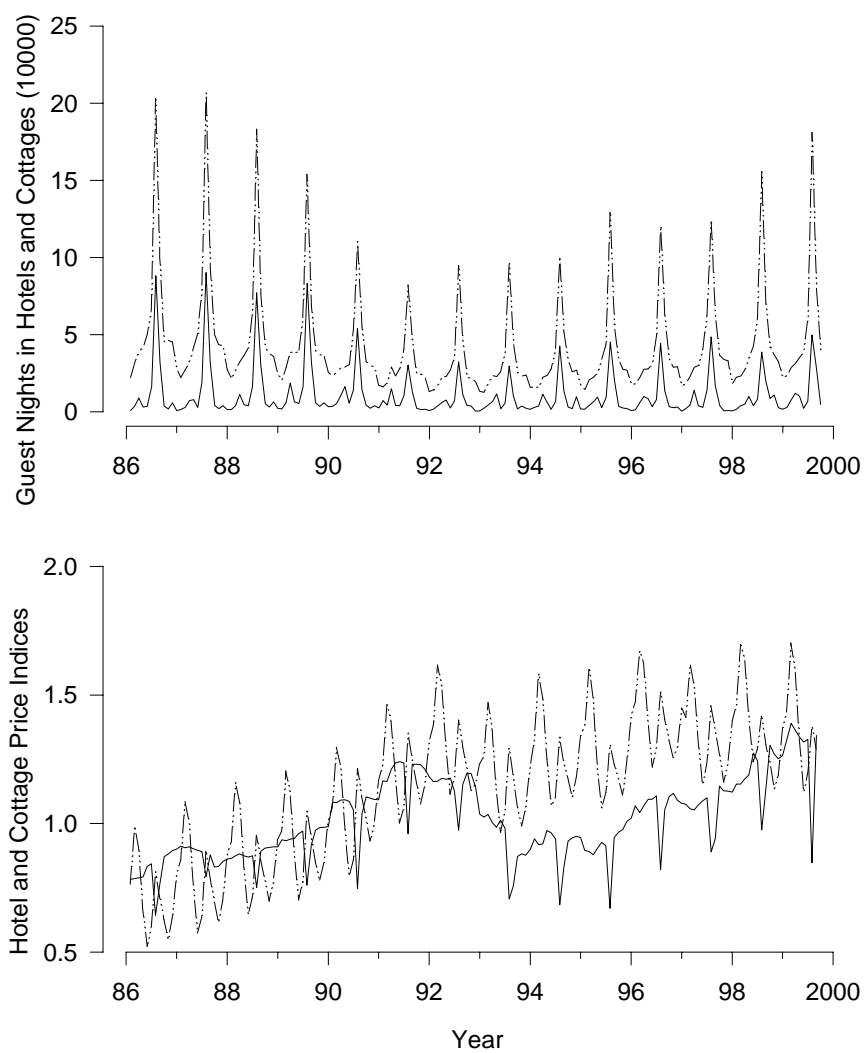


Figure 1: Norwegian numbers (in 10 000) of guests in Swedish hotels (dot-dash line) and cottages (solid line) [top], hotel price index in NOK (solid line) and cottage price index in NOK (dot-dash line) [bottom], 1986:1 - 1999:9.

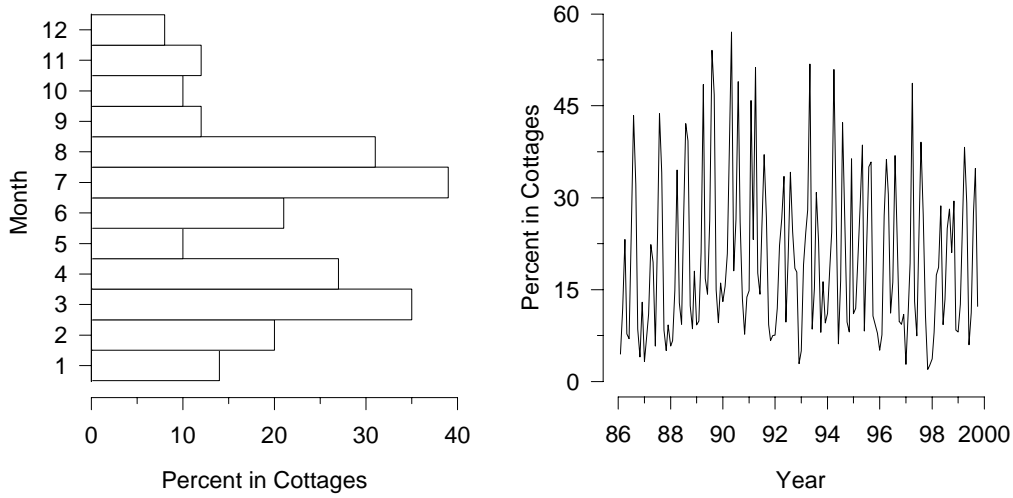


Figure 2: The percent of Norwegian guests in Swedish cottages (mean = 0.16, standard deviation = 0.08). The left panel gives averages over the entire period.

for the cottage prices and troughs in July for the hotel price series. In September 1999 the price was 960 SEK per night in a double room. For cottages the average price in 1999 was 465 SEK per night.⁴

5. Estimation

To estimate the parameters of the bivariate INMA(1) model, a nonlinear, conditional least squares estimator based on prediction errors will be employed. By this there is no need to make explicit distributional assumptions. It should be remarked that national and monthly guest night numbers are usually large, so that maintaining a count data interpretation is of little practical interest for estimation. The count data setting has a key role only for providing the interpretational aspects of the

⁴Data sources: Accommodation statistics and accommodation prices for hotel and cottages (Statistics Sweden). Consumer price indices (OECD, Main Economic Indicators). Exchange rates, defined as monthly averages of the market rates, and Norwegian tourism expenditures in Sweden (the Swedish Central Bank). Norwegian GDP (International Financial Statistics, International Monetary Fund). Norwegian gross domestic income (Statistics Norway).

model parameters. These interpretations are not affected by the sizes of guest night numbers.

First, we discuss the case with explanatory variables in the levels model, and later we proceed to give some remarks on constant parameter models. The least squares estimator for the model in levels, i.e. $z_{it} = \beta_{it} \circ \xi_t + \theta_{it-1} \beta_{it-1} \circ \xi_{t-1}$, minimizes the criterion function

$$S = \sum_{t=2}^T e_{Ht}^2 + e_{Ct}^2,$$

where the prediction errors are $e_{it} = z_{it} - E(z_{it} | \mathbf{z}_1, \dots, \mathbf{z}_{t-1}) = z_{it} - \beta_{it} N h \lambda'_t - \theta_{it-1} \beta_{it-1} \times N(h-1) \lambda'_{t-1} - \theta_{it-1} \beta_{it-1} \varepsilon_{t-1}$ ($i, j = H, C$). This is a straightforward approach in all respects but for how to determine the $\{\varepsilon_t\}$ sequence. If ε_t is determined from each equation separately there would unfortunately be two different sequences. Instead, write the levels model on the alternative form $\beta_{it} \circ \xi_t = z_{it} - \theta_{it} \beta_{it-1} \circ \xi_{t-1}$, so that $\xi_t = \sum_{i=1}^2 E(\beta_{it} \circ \xi_t | \xi_t, \xi_{t-1}) = (z_{Ht} + z_{Ct}) - [\theta_{Ht} \beta_{Ht-1} + \theta_{Ct} (1 - \beta_{Ct-1})] \xi_{t-1}$, since $\beta_{Ht} + \beta_{Ct} = 1$, for all t . This expression in ξ_t can be used to recursively calculate the entire sequence. The ξ_t expression is also what one gets for a continuous variable model, for which no expectation operations are required (see also Brännäs and Hall, 1998). Using the definition of ξ_t and replacing $\varepsilon_{t-k}, k > 0$, with expected values $N \lambda'_t$, gives

$$\varepsilon_t = (z_{Ht} + z_{Ct}) - [\theta_{Ht-1} \beta_{Ht-1} + \theta_{Ct-1} (1 - \beta_{Ht-1})] (\varepsilon_{t-1} + N(h-1) \lambda'_t).$$

Note that the z_{it} are conditionally as well as unconditionally heteroskedastic. In addition, some of the specifications may be subject to remaining serial correlation. As we do not wish to unnecessarily embark on a full specification of the dependence structure (cf. Brännäs and Hellström, 2000) a weighted conditional least squares estimator is not feasible. For inference purposes we, however, recommend accounting for heteroskedasticity and serial correlation by using a robust sandwich estimator of the covariance matrix (e.g., BHN).

For a constant parameter model neither β_H nor β_C can be estimated directly. As the prediction error then contains a leading term $\lambda_i = \beta_i \lambda$ we may estimate β_i by $\hat{\lambda}_i / (\hat{\lambda}_H + \hat{\lambda}_C)$ and λ by $\hat{\lambda}_H + \hat{\lambda}_C$ once least squares estimates of λ_H and λ_C are

available. As we can not enforce the parameter restrictions across equations, the two equations can be estimated separately by standard software.

6. Results

The estimation results are presented Table 1. The models are estimated jointly by the conditional least squares estimator, with seasonal dummy variables included in the linearly specified λ function. The fully parameterized model in the first column gives α values that appear too small, in most instances equal to zero for α_C . Therefore, a less parameterized model is presented in the second column of Table 1.

Comparing the two models we see that the estimated parameters in the λ function for the fully parameterized model are larger and have smaller standard errors. Turning to the point estimates, Table 1 reveals that increases in the German (p_G), Finnish (p_F) and Swedish (p_S) price levels will reduce the number of hotel and cottage guest nights. The negative effects of the German and Finnish price levels are surprising, but can be a result of fewer transit visitors in Sweden. About one fifth of the Norwegian visitors pass through Sweden to reach their main destinations. Note, however, that the Finnish price effect is not significant. The negative effect of the Swedish price level is the expected one. As implied by the economic model the estimated coefficient ϕ_i for the income variable (M_N) is positive.

Continuing with the estimates of the β function, i.e. the probability of selecting the hotel alternative, we find that the price effects from hotels (p_H) and cottages (p_C) are positive. While the positive effect from an increase in the cottage price is expected, the positive own price effect is not. Additionally, a higher price level in Sweden reduces the probability of choosing the hotel alternative, whereas a higher expenditure level (M_{NS}), according to the fully parameterized model, will increase the probability of choosing a hotel. Compared to the price effects the expenditure effect is small.

Focusing next on the estimation results for the explanatory variables in the θ_H and θ_C functions, we find that an increase in hotel price increases θ_H and then also

Table 1: Estimation results for Norweigan guest night series (level) (H indicates hotel and C cottage). Seasonal dummies included in the λ function are not reported.

	Est.	s.e.	Est.	s.e.
λ function				
Price index (p_G)	-0.110	0.005	-0.056	0.017
Price index (p_F)	-0.009	0.006	-0.022	0.022
Price index (p_D)	0.442	0.029	0.197	0.089
Price index (p_S)	-0.202	0.009	-0.115	0.038
Income (M_N)	0.055	0.004	0.049	0.014
β_H function				
Hotel price (p_H)	1.120	0.061	3.032	0.644
Cottage price (p_C)	0.760	0.051	1.089	1.072
Price index (p_S)	-3.736	0.149	-7.239	2.341
Expenditure (M_{NS})	0.001	0.000	-0.0004	0.001
Constant	3.088	0.092	4.811	1.204
θ_H function				
Hotel price (p_H)	-8.187	0.919	0.719	0.174
Cottage price (p_C)	-13.834	0.255	-	
Price index (p_S)	14.314	0.840	-1.806	0.569
Expenditure (M_{NS})	-0.002	0.000	-	
Constant	2.632	0.427	-0.365	0.596
θ_C function				
Hotel price (p_H)	-7.412	0.610	-	
Cottage price (p_C)	-22.800	0.486	-21.621	1.157
Price index (p_S)	26.644	0.683	17.483	2.994
Expenditure (M_{NS})	0.001	0.000	-	
Constant	-2.565	0.259	-1.088	2.040
LB(12, Hotel)	76.9		53.8	
LB(12, Cottage)	99.9		87.0	
R^2 (Hotel)	0.96		0.95	
R^2 (Cottage)	0.84		0.82	

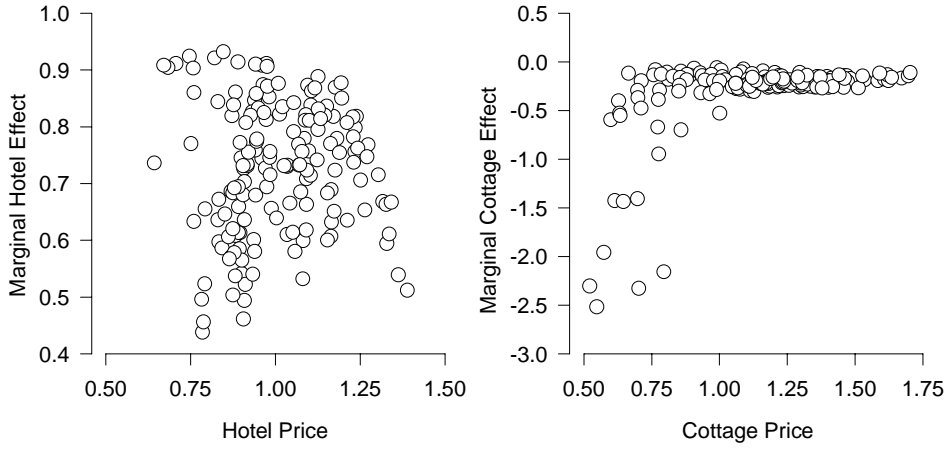


Figure 3: Marginal effects of hotel price (right) and cottage price (left) on $q_{jit} = \beta_{it}/(1 - \alpha_{it})$ for hotels and cottages. Based on all observations and the model in column two of Table 1.

the probability of remaining another night in a hotel (α_H).⁵ This unexpected effect is only present in the less parameterized model. An increase in cottage price has an expected negative effect in the θ_C functions. The Swedish price level has a negative effect on the probability of staying another night in a hotel for the less parameterized model, and a positive effect on the θ_C function. For the less parameterized model the mean values $\bar{\alpha}_C$, $\bar{\alpha}_H$ and $\bar{\beta}$ over the total sample are 0.06, 0.18 and 0.78, respectively. As Figure 2 show, the mean of β is 0.84. For both models there is substantial but rather similarly shaped remaining autocorrelation, with a large positive peak at lag 12. By applying the less parameterized specification the autocorrelation is slightly reduced, to the cost of a marginally lower explanatory power for the cottage series.

Hitherto we have focused on the direct effects within each function. In Figure 3 the marginal effects of hotel and cottage prices on $q_{ji} = \beta_i/(1 - \alpha_i)$ for hotels and cottages is shown. From the figure we see that the marginal effect varies more for cottages due to a change in the own price. However, most of the effects lie between -0.3 and -0.15 with a mean value of -0.2 as the box plot in Figure 4

⁵Note that the sign interpretations for the θ and α functions are the same.

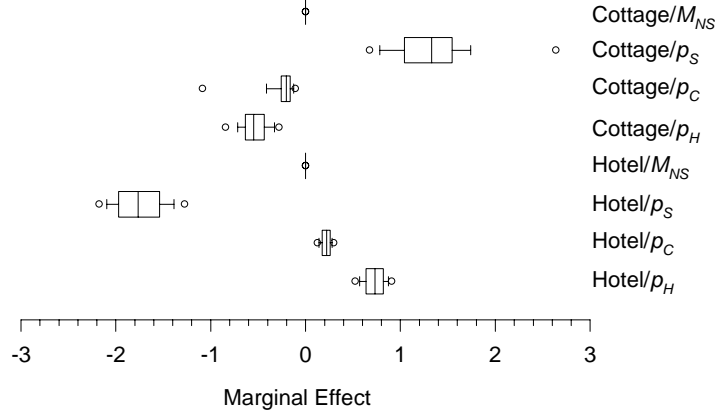


Figure 4: Box plot of marginal effects of hotel price, cottage price, Swedish price level, and expenditure on $q_{jit} = \beta_{it}/(1 - \alpha_{it})$ for hotels and cottages. Based on all observations and the model in column two of Table 1. Outlying markers correspond to the 5th and 95th percentiles.

reveals. Figure 3 also indicates that the marginal effect from an increased cottage price is larger when the cottage price is low, i.e. in off peak periods when the cottage demand is low. For hotels no such clear pattern is manifest.

Figure 4 presents the marginal effects on $q_{ji} = \beta_i/(1 - \alpha_i)$ for hotels and cottages for each of the explanatory variables in the α and β functions. The marginal effect of expenditure (Cottage/ M_{NS} and Hotel/ M_{NS}) is small. Which indicate that the Norwegians tend to increase their demand for other goods, than guest nights in hotels and cottages, as they allocate more money to trips to Sweden. The Swedish price level has a negative marginal effect on hotel (Hotel/ p_S) and a positive effect on cottage (Cottage/ p_S) guest nights. Thus, conditional on the decision to travel to Sweden, the Norwegians will increase their stay in cheaper accommodation (cottages) as the Swedish price level increases. As revealed earlier in Figure 3, the marginal own price effect on cottage (Cottage/ p_C) is negative. The marginal effect of an increase in the hotel price on cottages is also negative, since the probability of

choosing the hotel alternative increases as the hotel price increases.

For hotels, we find a positive marginal effect from both cottage and hotel prices. While the positive marginal effect from the cottage price (Hotel/p_C) is expected, the positive own price effect (Hotel/p_H) is not. This may be a result of omitted explanatory variables, since many of the hotel guests are business visitors.

7. Conclusion

The paper has adopted a novel strategy of first obtaining a dynamic and stochastic model structure by aggregation from a simple hotel or cottage specific model (a measurement model), and second used standard demand analytical results to explain some of the parameters of the stochastic model (the structural part). The obtained model and the estimation results provide a richer picture of international guest nights than previously used model approaches.

The derived parameters in the aggregated measurement model are the monthly mean check-in, the probability of selecting hotel rather than cottage and the daily check-out probability. In the analysis, the visitors were first assumed to decide upon which country to visit, and then in a second stage decide the form of accommodation. To estimate the parameters, a nonlinear, conditional least square estimator based on prediction errors was employed. The estimation results indicate that most of the explanatory variables are significant. The estimation also indicates that the results are robust, with the exception of the parameters in the hotel check-out probability.

In the decision process for which country to visit the own price has a strongly negative and significant effect. The price level in the destination country can thus be considered as one of the more important factors in the decision process for which country to visit.

Conditional on the decision to visit Sweden, the largest marginal effect on the demand for hotel and cottage guest nights comes from the overall price level within Sweden, with a positive effect on cottages and a negative effect on hotels. The small marginal effect from the expenditure variable indicates that the Norwegians tend to increase their consumption of other goods, than guest nights, as they allocate

more money for trips to Sweden. However, the expenditure variable also includes expenditures from same day visitors and border trade, which can have increased relative to expenditures from over night visitors during the sample period.

The marginal effect from an increase in the own price is negative for cottage and positive for hotel. While the negative effect on cottage is expected the positive effect on hotel is not, and may be a result of omitted explanatory variables. A large part of the hotel visitors during the fall, winter and spring are business visitors. From the time series plot of the hotel price, one can assume a high positive correlation between the price for hotel accommodation and the business trend in Sweden.

Although the adopted technique has been applied to the allocation problem for different sorts of accommodation, there are a number of other areas where this modelling approach can be of interest. For example, the procedure can be applied to a number of different stock-keeping problems, and to modelling the number of students on different courses at a university. The empirical illustration also indicates that the measurement model can easily be extended with a large number of explanatory variables.

Appendix A1: Bivariate INAR(1) Model Properties

For the INAR(1) model of Section 2, let $E(\epsilon_t) = \lambda$, $V(\epsilon_t) = \delta$ and $E(\epsilon_t \epsilon_s) = \lambda^2$, $t \neq s$. Then the following properties hold: $E(\beta \circ \epsilon_t) = \beta\lambda$, $V(\beta \circ \epsilon_t) = \beta\delta + \beta(1 - \beta)\lambda$ and $E(\beta \circ \epsilon_t)\beta \circ \epsilon_s) = \beta^2\lambda^2$, $t \neq s$. The thinning operations involving α_i , $i = 1, 2$, are analogously defined. For the model at the daily and regional level we can demonstrate the following properties:

$$\begin{aligned} E(y_{it}) &= \beta_i \lambda / (1 - \alpha_i) \\ V(y_{it}) &= (\alpha_i \beta_i \lambda + \delta) / (1 - \alpha_i^2) \\ Cov(y_{1t}, y_{2t}) &= \beta_1 \beta_2 \delta / (1 - \alpha_1 \alpha_2) \geq 0 \\ Corr(y_{it}, y_{it-k}) &= \alpha_i^k \geq 0, \quad i = 1, 2, \quad k \geq 0 \\ Cov(y_{1t}, y_{2t-k}) &= \alpha_1^k Cov(y_{1t}, y_{2t}) \geq 0, \quad k \geq 0 \\ Cov(y_{1t-k}, y_{2t}) &= \alpha_2^k Cov(y_{1t}, y_{2t}) \geq 0, \quad k \geq 0 \end{aligned}$$

Hence, there is both serial as well as cross correlation in the day-to-day series.

Appendix A2: Temporal Aggregation

Consider aggregation over three time points and focus only on the MA-part. After this induction could be used, but we abstain from giving a full account. The result is based on an underlying Poisson i.i.d. assumption for $\{\epsilon_t\}$. We may write

$$\begin{aligned} \nu_t &= \epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \alpha \circ (\epsilon_{t-1} + \epsilon_{t-2} + \epsilon_{t-3}) + \alpha^2 \circ (\epsilon_{t-2} + \epsilon_{t-3} + \epsilon_{t-4}) \\ &= \xi_t^{t-2} + \alpha \circ \xi_{t-1}^{t-3} + \alpha^2 \circ \xi_{t-2}^{t-4} \end{aligned}$$

Here, $E(\xi_t^{t-s}) = V(\xi_t^{t-s}) = 3\lambda$ and hence $E(\nu_t) = 3\lambda(1 + 3\alpha + 3\alpha^2)$. The conditional expectation is $E(\nu_t | z_1, \dots, z_{t-3}) = 3\lambda + 2\alpha\lambda(1 + \alpha) + \alpha(1 + \alpha)\epsilon_{t-3}$. In this we have set $E(\epsilon_{t-4} | z_1, \dots, z_{t-3}) = \lambda$ as ϵ_{t-4} is between time points at the new time scale. Strictly, it is, however, in the information set. For the variance we need

$$\begin{aligned} Cov(\xi_t^{t-2} \xi_{t-1}^{t-3}) &= 2\lambda, \quad Cov(\xi_t^{t-2} \xi_{t-2}^{t-4}) = \lambda, \\ Cov(\xi_{t-1}^{t-3} \xi_{t-2}^{t-4}) &= 2\lambda, \quad V(\alpha^k \circ \xi) = \alpha^{2k} V(\xi) + \alpha^k (1 - \alpha^k) E(\xi). \end{aligned}$$

Then $V(\nu_t) = \lambda[3 + 7\alpha + 5\alpha^2 + 4\alpha^3]$. Writing instead on the new time scale $\omega_t = \xi_t + \alpha \circ \xi_{t-1} + \alpha^2 \circ \xi_{t-1}$ gives the same first conditional and unconditional moments,

but not the same second order moments. The difference is due to the variable dependence between the ξ components in the original model now being forced to be constant. This variation in dependence will remain for all aggregation over more than two time periods.

References

- Adamowicz, W.L. (1994). Habit Formation and Variety Seeking in a Discrete Choice Model of Recreation Demand. *Journal of Agricultural and Resource Economics* **19**, 19-31.
- Al-Osh, M.A. and Alzaid, A.A. (1987). First-Order Integer Valued Autoregressive (INAR (1)) Process. *Journal of Time Series Analysis* **8**, 261-275.
- Alzaid, A.A. and Al-Osh, M.A. (1988). First Order Integer Valued Autoregressive (INAR(1)) Process: Distributional and Regression Properties. *Statistica Neerlandica* **42**, 53-61.
- Anderson, R.A. (1979). Perfect Price Aggregation and Empirical Demand Analysis, *Econometrica* **47**, 1209-1230.
- Blundell, R., Griffith, R. and Windmeijer, F. (1999). Individual Effects and Dynamics in Count Data Models. Working Paper W99/3, Institute of Fiscal Studies, London.
- Brewer, K.R.W. (1973). Some Consequences of Temporal Aggregation and Systematic Sampling for ARMA and ARMAX Models. *Journal of Econometrics* **1**, 133-154.
- Brännäs, K. (1995). Explanatory Variables in the AR(1) Count Data Model. Umeå Economic Studies 381.
- Brännäs, K. and Hall, A. (1998). Estimation in Integer-Valued Moving Average Models. Umeå Economic Studies 477.
- Brännäs, K. and Hellström, J. (2000). Generalized Integer-Valued Autoregression. To appear in *Econometric Reviews*.
- Brännäs, K., Hellström, J. and Nordström, J. (1999). A New Approach to Modelling and Forecasting Monthly Guest Nights in Hotels. Umeå Economic Studies 503 (revised).
- Crouch, G.I. (1994). The Study of International Tourism Demand: A Survey of Practice. *Journal of Travel Research*, Spring, 41-55.
- Dixit, A. and Stiglitz, J.E. (1977). Monopolistic Competition and Optimum Product Diversity. *American Economic Review* **67**, 297-308.

- Garcia-Ferrer, A. and Queralto, R.A. (1997). A Note on Forecasting International Tourism Demand in Spain. *International Journal of Forecasting* **13**, 539-549.
- Granger, C.W.J. and Morris, M.J. (1976). Time Series Modelling and Interpretation. *Journal of the Royal Statistical Society* **A139**, 246-257.
- Gustavsson, P. and Nordström, J. (1999). The Impact of Seasonal Unit Roots and Vector ARMA Modeling on Forecasting Monthly Tourism Flows. Umeå Economic Studies 496.
- Harvey, A.C. (1981). *Time Series Analysis*. Philip Allan, Oxford.
- McKenzie, E. (1985). Some Simple Models for Discrete Variate Time Series. *Water Resources Bulletin* **21**, 645-650.
- McKenzie, E. (1986). Autoregressive Moving-Average Processes with Negative-Binomial and Geometric Marginal Distributions. *Advances in Applied Probability* **18**, 679-705.
- Melenberg, B. and van Soest, A. (1996). Parametric and Semi-Parametric Modelling of Vacation Expenditures. *Journal of Applied Econometrics* **11**, 59-76.
- Nordström, J. (1999). Unobserved Components in International Tourism Demand. Umeå Economic Studies 502.
- Nordström, J. (1996). Tourism Satellite Accounts for Sweden 1992-93. *Tourism Economics* **2**, 13-42.
- OECD. (1992). *Tourism Policy and International Tourism in OECD Member Countries*. Paris.
- Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1987). *Numerical Recipes*. Cambridge University Press, Cambridge.
- Pollak, R.A. (1971). Conditional Demand Functions and the Implications of Separable Utility. *Southern Economic Journal* **37**, 423-433.
- Pollak, R.A. and Wales, T.J. (1992). *Demand System Specification and Estimation*. Oxford University Press, Oxford.
- Spence, M.E. (1976). Product Selection, Fixed Costs, and Monopolistic Competition, *Review of Economic Studies* **43**, 217-236.
- Sveriges Turistråd. (1992). Den utländska turismen i Sverige. Rapport 1992:3 (ISSN 1101-0754), Stockholm.

- Turner, L.W., Kulendran, N. and Fernando, H. (1997). The Use of Composite National Indicators for Tourism Forecasting, *Tourism Economics* 3, 309-317.
- Witt, S.F. and Witt, C.A. (1995). Forecasting Tourism Demand: A Review of Empirical Research. *International Journal of Forecasting* 11, 447-475.