

Sørensen, Nils Karl

Conference Paper

Modelling and seasonal forecasting of monthly hotel nights in Denmark

42nd Congress of the European Regional Science Association: "From Industry to Advanced Services - Perspectives of European Metropolitan Regions", August 27th - 31st, 2002, Dortmund, Germany

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Sørensen, Nils Karl (2002) : Modelling and seasonal forecasting of monthly hotel nights in Denmark, 42nd Congress of the European Regional Science Association: "From Industry to Advanced Services - Perspectives of European Metropolitan Regions", August 27th - 31st, 2002, Dortmund, Germany, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/115617>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Paper prepared for ERSA 2002
Dortmund, Germany, September 27-31, 2002

Modelling and Seasonal Forecasting of Monthly Hotel Nights in Denmark¹

Nils Karl Sørensen

*Southern University Denmark,
Grundtvigs allé 150, DK-6400 Sønderborg, Denmark.
Phone (+45) 65 50 12 28, e-mail: nks@sam.sdu.dk*

Abstract

The paper examines foreigners demand for hotel nights in Denmark by nationality using monthly time series covering nearly 30 years. Inbound hotel demand is assumed to depend on economic variables like the foreign and Danish CPI and the exchange rate, as well as variables measuring the climate. It is investigated whether the climate last year has a significant influence on this year demand for hotels. Further, the relative importance between economic and climatic variables are considered.

Univariate as well as multivariate ECM models are examined, and special attention is devoted to an evaluation of the forecast performance of the various models. Over time people's preferences have changed, and this implication for modelling hotel demand is taken into account.

Key words: Tourism, seasonality, forecasting, multivariate estimation, cointegration.

Theme: Tourism and regional development, econometrics.

1. Introduction

The growth of the world-wide tourism industry over the past few decades has generated great interest in tourism demand modelling and forecasting. Interesting contributions by Witt and Witt (1992,1995) and Song and Witt (2000) have shown that no single forecasting method performs consistently best across different situations.

This result is in accordance with recent econometric methodology. Here it is advocated that the »classical textbook« method of econometrics, i.e. the direct estimation of a theoretical model, and then diagnostic checking its validity by testing for autocorrelation, heteroskedasticity etc. is fundamentally wrong, and give no description of what is going on in the real world. Further, in order to make the model operational the applied statistic material often imposes severe limitations on the models ability to verify a given theory. Instead of pretending that model building takes place in this way it should be accepted that empirical models try to describe a data generating process (DGP). It is not possible ever to find the true DGP, but one can hope to obtain a better approximation than provided by a theoretical model.

The obvious presence of seasonality in tourism with peaks during the holiday season clearly makes it difficult for econometricians to estimate and evaluate models. This problem is further complicated by the fact that peoples preferences for holidays have changed. Today people are more inclined to separate their holidays into several sub-periods. Using monthly observations Sørensen (1999) using data covering the period from 1970.1 to 1996.12 found that a varying and changing seasonal component is a common phenomenon in many time series for hotel nights in Denmark.

The purpose of the present paper is to focus on four forecasting models for monthly time series. Statistics on hotel nights in Denmark are divided into 11 nationalities covering more than 90 percent of total hotel nights. Three of the models or representations attempts to model DGP by use of statistics for hotel nights only, whereas the final representation models tourism demand by use of an error correction mechanism (ECM) model.

Three of the forecasting models are univariate. The first model assumes that seasonality is governed by stochastic trends at the seasonal frequencies, and was introduced by Hylleberg, Engle, Granger and Yoo (1990) (HEGY). The second model assumes approximate deterministic seasonality, which is termed the »seasonal cycle«. This approach was introduced by Barsky and Miron (1989). The third univariate model is a traditional demand set up, where hotel nights are assumed to be a function of economic as well as climatic variables. Although obvious to examine the impact of a weather index in relation to tourism, this approach has only been considered earlier by Barry and O'Hagan (1972).

The paper is organized as follows. Section 2 gives a preliminary analyse of the evolution of hotel nights in Denmark by nationality, Section 3 deals with stochastic seasonality, i.e. the HEGY-approach, Section 4 with deterministic seasonality, and Section 5 presents the ECM demand model. In Section 6 the forecasting performance of the four different models is evaluated, and various forecasting measures are discussed. Finally, the conclusion is drawn in Section 7.

2. Preliminary Analyses of Data

A data bank has been set up covering monthly series from 1970.1 to 2000.12 by use of material from the regular publications on hotel nights supplied by Statistics Denmark. All series contain 372 observations. Data are divided into 11 different nationalities. In addition, the total number of hotel nights are considered.

Table 1. Summary statistics on hotel nights in by nationality. 1970.1 to 2000.12.

Nationality:	Share	Growth 2000	Nights 1993	Coefficient of variation, CV			
	Average Percent	Index	Relative to home	1970.1 - 1979.12	1980.1 - 1989.12	1990.1 - 2000.12	1970.1 - 2000.12
Total	100.0	186		0.49	0.46	0.44	0.51
Danes	49.6	197	1.089	0.25	0.24	0.27	0.36
Foreigners, total	50.4	175		0.72	0.71	0.67	0.73
Swedes	11.1	250	0.165	0.65	0.65	0.65	0.74
Norwegians	6.3	479	0.182	0.91	1.19	1.15	1.30
Finns	1.0	688	0.015	0.59	0.79	0.75	0.74
Germans	12.8	158	0.021	1.10	0.96	0.87	0.98
English	3.3	192	0.005	0.42	0.50	0.43	0.48
Dutch	1.5	203	0.012	0.83	0.68	0.88	0.92
French	0.9	93	0.001	0.71	0.57	0.46	0.59
Italians	1.1	201	0.002	0.62	0.94	0.98	0.99
Americans (USA)	4.7	41	0.001	0.86	0.75	0.70	0.87
Japanese	1.0	205	0.008	0.46	0.52	0.65	0.59

Note: The percentage shares are computed on the mean values for the full period. The computations in the second and third columns are on annual data additively aggregated from the monthly observations. In the second column the growth index is with 1970 set equal to 100. In the third column the figures measure the number of hotel nights relative to the total population in the home country. In the four last columns CV is the coefficient of variation computed using the non-transformed data and defined as: $CV = \text{standard deviation} / \text{mean}$.

In 2000 the total number of hotel nights in Denmark amounted to about 13.3 million,

equivalent to an increase of about 80% since 1970. With regard to nationality, half the hotel nights are attributable to Danes, and half to foreign nationals. The graph in Figure 1 displays the long-run transformation y_{it} of hotel nights by Danes (solid line) and foreigners respectively. The division of hotel nights between Danes and foreigners has remained relatively constant throughout the period. Notice also that the number of hotel nights has been affected by external events such as the slump in the mid-1970s. From the mid-1980s the pattern of hotel nights consumed by foreigners seems to have become more volatile, perhaps due to variations in taste or fluctuations in the exchange rate for some important foreign visitors. Tables 1 give some summary statistics.

The most important foreign consumers of hotel nights in Denmark are Germans, Swedes and Norwegians; in other words people from the neighbouring countries. Tourists from these three nations account for more than half the foreign hotel nights. Overall, the number of hotel nights consumed by foreigners has doubled over the period. Interesting exceptions are observed for Norwegians, who show an increase of over 300%, and the decreased significance of hotel nights consumed by Americans.

3. Stochastic Seasonality

A time series model with seasonal unit roots is an approximation that allows for changes in the seasonal pattern, i.e. integration at the seasonal frequencies. A test for seasonal unit roots in the quarterly case is developed by HEGY, and extended to the monthly case in Franses (1991).

For a given time series variable x_t a univariate model integrated at all the seasonal frequencies as well as the long-run frequency is $(1-B^{12})x_t = \epsilon_t - i.i.d(0, F^2)$, $t = 1, 2, \dots, T$, where B is the lag operator defined as $B^n x_t = x_{t-n}$. It can be proved that an AR(p) process of the form $N(B)x_t = \epsilon_t$, using a proposition given by HEGY defining the form of $N(B)$, in the monthly case can be written as

$$(1) \quad (1-B^{12})x_t = y_{8t} = B_1 y_{1t-1} + B_2 y_{2t-1} + B_3 y_{3t-1} + B_4 y_{3t-2} \\ + B_5 y_{4t-1} + B_6 y_{4t-2} + B_7 y_{5t-1} + B_8 y_{5t-2} \\ + B_9 y_{6t-1} + B_{10} y_{6t-2} + B_{11} y_{7t-1} + B_{12} y_{7t-2} + N(B)y_{8t} + \epsilon_t + \epsilon_{t-1}$$

In this equation x_t is a linear combination of the variables y_i , $i=1, \dots, 8$. These variables are all various transformations of x_t . In all there are 12 coefficients B_j , $j=1, \dots, 12$ to be estimated by applying OLS to equation (1). The coefficients of the B 's correspond to the 12 solutions of the equation $(1-B^{12})=0$ all lying on the unit cycle.

The test for monthly seasonal unit roots can then be performed as a test of whether or not the coefficients are lying on the unit cycle. There will be *no* seasonal unit root if the

coefficient to a given B is significantly different from zero. If the coefficient is insignificant, a unit root is present showing a changing seasonal component.

Consider as an example a situation in which *all* the B 's are significantly different from zero. Consequently, no seasonal unit roots are present, and the data have a deterministic or constant seasonal pattern. Only in this case is the dummy variable model appropriate; ie by applying the filter $(1-B^{12})$ the time series x_t will become stationary.

The transformations $y_i, i=1, \dots, 8$ of x_t remove the seasonal unit roots at certain frequencies while preserving them at other frequencies. For example $y_{1t} = (1+B)(1+B^2)(1+B^4+B^8)x_t = (1+B+B^2+\dots+B^{11})x_t$ is a transformation which removes the seasonal unit roots and preserves the long-run or zero-frequency unit root.

The remaining transformations are $y_{2t} = -(1-B)(1+B^2)(1+B^4+B^8)x_t$, which preserves the frequency 6/12 corresponding to a six month period; $y_{3t} = -(1-B^2)(1+B^4+B^8)x_t$, which retains the frequency 3/12 (9/12) corresponding to a four month period; $y_{4t} = -(1-B^4)(1-B+B^2)(1+B^4+B^8)x_t$, which leaves behind the frequency 5/12 (7/12); while $y_{5t} = -(1-B^4)(1+B+B^2)(1+B^4+B^8)x_t$, $y_{6t} = -(1-B^4)(1-B+B^2)(1-B^2+B^4)x_t$, $y_{7t} = -(1-B^4)(1+B+B^2)(1-B^2+B^4)x_t$ retains the frequencies 1/12 (11/12), 4/12 (8/12) and 2/12 (10/12). Finally $y_{8t} = (1-B^{12})x_t$. A survey of the data transformations and test hypothesis is given in Table 2.

Table 2. Testing for seasonal unit roots in monthly data.

No.	Frequency		Transformation	H ₀ : Unit Root	H ₁ : No Unit Root
0	0	long-run	$y_{1t} = (1+B+B^2+\dots+B^{11})x_t$	$B_1 = 0$	$B_1 < 0$
1	6/12	semi-annual	$y_{2t} = -(1-B)(1+B^2)(1+B^4+B^8)x_t$	$B_2 = 0$	$B_2 < 0$
2	3/12 (9/12)	quarterly	$y_{3t} = -(1-B^2)(1+B^4+B^8)x_t$	$B_3 = 1, B_4 = 0$	$B_3 < 0, B_4 \dots 0$
3	5/12 (7/12)	monthly	$y_{4t} = -(1-B^4)(1-B+B^2)(1+B^4+B^8)x_t$	$B_5 = 1, B_6 = 0$	$B_5 < 0, B_6 \dots 0$
4	1/12 (11/12)	monthly	$y_{5t} = -(1-B^4)(1+B+B^2)(1+B^4+B^8)x_t$	$B_7 = 1, B_8 = 0$	$B_7 < 0, B_8 \dots 0$
5	4/12 (8/12)	monthly	$y_{6t} = -(1-B^4)(1-B+B^2)(1-B^2+B^4)x_t$	$B_9 = 1, B_{10} = 0$	$B_9 < 0, B_{10} \dots 0$
6	2/12 (10/12)	monthly	$y_{7t} = -(1-B^4)(1+B+B^2)(1-B^2+B^4)x_t$	$B_{11} = 1, B_{12} = 0$	$B_{11} < 0, B_{12} \dots 0$

Note: Critical values can be obtained from P.H. Franses and B. Hobijn, 'Critical values for unit root tests in seasonal time series', *Journal of Applied Statistics*, Vol 24, 1997, pp 25-47.

The test for seasonal unit roots at the relevant seasonal frequencies can then be performed either as t-tests for the estimates of B_3, \dots, B_{12} or as joint F-tests for sets of parameters, ie B_3 and B_4 etc. The t-ratios corresponding to the estimates of B_i and B_j follow the Dickey-Fuller distributions. All critical values of the test have a non-standard distribution and have to be simulated by Monte Carlo experiments. Critical values are obtained from Franses and Hobijn (1997). Equation (1) is an extension of the Dickey-Fuller auxiliary regression for a zero-frequency unit root with augmented lagged values of the left-hand

side variable.

Finally, two more points concerning equation (1) should be made. First, ϵ_t is the included deterministic component such as a constant, a trend, and seasonal dummy variables. The test can be performed without these components, but the critical values will then change. Second, equation (1) has to be augmented by lagged y_{8t} . This is done in order to make the residuals white noise, and leaves the asymptotic distribution unaffected. The power is negatively affected if too many nuisance parameters are used in the augmentation. In the present case, a strategy to determine $N(B)$ has been to start with 24 lags equalling two years, and then test down the significant augmented variables by carefully inspecting the test statistics for autocorrelation.

Results of the applying the auxiliary regression (1) is displayed in Table 3. It is evident from the material that a varying and changing seasonal pattern is a common phenomena both with regard to counties and with regard to nationality. The seasonal pattern has been quite varying in six of the counties considered and for six nationalities.

Regarding nationality the seasonal pattern has been most varying for all the Scandinavian nationalities, for Englishmen, and for Italians. The most stable pattern is observed for Germans, Americans, and Japanese using hotels in Denmark.

Finally, it should be noticed that the seasonal pattern of the aggregated series for hotel nights in Denmark seems to preserve the seasonal unit roots found at county level, whereas the picture is unclear with regard to nationality.

Summary of HEGY-tests.

Nights by county:	Unit roots at:	Nights by nationality:	Unit roots at:
Total	0, 2, 4	Germans	0, 2, 4
Danes	0, 2, 4, 5, 6	Englishmen	0, 1, 2, 4, 6
Foreigners, total	0, 2, 4	Dutchmen	0, 1, 2, (3), 5
Swedes	0, 2, 4, (6)	Frenchmen	0, 2, 4
Norwegians	0, (1), 2, 4	Italians	0, 1, 2, 4, 6
Finns	0, (1), 2, 4	Americans (USA)	0, 2, 4
		Japanese	1, 2, 4

4. Deterministic Seasonality

Barsky and Miron (1989) analyses the economics of the seasonal cycle by considering a deterministic seasonal model of the form

$$(2) \quad (1-B^{12})x_t = y_{8t} = R_1D_{1,t} + R_2D_{2,t} + \dots + R_{11}D_{11,t} + R_{12}D_{12,t} + \sum_{n=1}^m a_n y_{8t} + \epsilon_t$$

where $\epsilon_t \sim iid(0, F^2)$, $t = 1, \dots, T$, and D_1, \dots, D_{12} are 12 seasonal dummy variables. In order to render the residuals white noise augmented values of y_{8t} is included. Since the R_s parameters are assumed fixed over the sample seasonality is of a constant non-changing nature (rest to appear).

5. A Seasonal ECM-model of Tourism Demand

In this section a more traditional way of describing the DGP of hotel nights is used namely estimation of a time series demand model of the form $HN = f(TP, W, Z)$, where HN is the number of hotel nights for a given nationality, TP is tourist prices, W is a weather index, and Z is a vector capturing omitted variables like e.g. the income level which is not available at the monthly frequency².

As proxies for tourist prices TP , currency exchange statistics and consumer price (CPI) statistics for the home land and the host land are used³. The price level in substitute destinations is not considered. The Danish CPI denoted PDK is taken to be a proxy for the cost of tourism. A few studies use proper tourist prices instead. The problem with using the CPI as the cost of living of tourism in the destination is that the cost of living for the local residents not always reflect the cost of living for the foreign visitors. However, Martin and Witt (1987) and Morely (1994) report no significant differences between the two types of price indices. The foreign CPI denoted PFO is taken to be a proxy for cost of living at home. The exchange rate denoted EXC is included as suggested by Sørensen (1999), who for USA and Finland observe a close similarity between the long run transformation of hotel nights and the exchange rate suggesting that these variables may be cointegrated.

A high Danish price level is expected to have a negative impact on the number of hotel nights, whereas a high foreign price level is supposed to have the opposite effect. A low exchange rate is also expected to increase the number of hotel nights. Notice, that if the purchasing power theorem is present the exchange rate is determined by the relative price ratio between the foreign and the Danish price level. This situation may lead to multicollinearity and break down of the estimation.

Table 3. HEGY-tests for monthly seasonal unit roots in hotel nights in Denmark by nationality. 1970.1 to 1997.12.

	Total	Danes	For- eigners	Swedes	Norwegi ans	Finns	Germans	English	Dutch	French	Italians	America ns	Japanese
<i>t</i> -statistic:													
$t_1: B_1$! 2.26	! 2.72	! 1.84	! 1.94	! 1.21	! 2.27	! 1.54	! 2.19	! 1.21	! 2.17	! 2.31	! 2.96	! 4.42*
$t_2: B_2$! 4.08*	! 4.42*	! 4.23*	! 5.52*	! 2.76	! 3.03*	! 3.19*	! 2.74	! 1.73	! 4.59*	! 1.74	! 4.54*	! 2.91
$t_3: B_3$	1.33	1.99	1.06	1.01	1.67	! 0.51	2.23	0.23	! 0.01	! 0.64	2.26	! 0.14	0.05
$t_4: B_4$	0.01	2.06	! 0.79	! 0.38	1.82	0.38	! 1.92	! 0.24	! 0.38	! 1.55	0.43	! 0.49	! 1.78
$t_5: B_5$! 4.01*	! 5.99*	! 4.59*	! 4.96*	! 4.14*	! 5.43*	! 5.54*	! 4.94*	! 2.45	! 4.39*	! 4.48*	! 6.77*	! 3.91*
$t_6: B_6$! 4.89*	! 5.39*	! 5.14*	! 4.49*	! 5.07*	! 5.87*	! 5.20*	! 5.31*	! 2.83	! 4.17*	! 4.78*	! 6.87*	! 3.99*
$t_7: B_7$	2.15	3.33	2.38	2.66	3.07	3.40	3.42	2.26	3.64	2.61	1.74	1.54	2.76
$t_8: B_8$! 1.61	! 2.39	! 2.26	! 2.55	! 2.35	! 2.62	! 3.30*	! 0.71	! 2.77	! 2.68	! 1.76	! 1.49	! 1.71
$t_9: B_9$! 2.17	! 2.42	! 3.24	! 2.47	! 3.78*	! 2.36	! 2.69*	! 3.30*	! 0.11	! 4.39*	! 3.39*	! 3.02*	! -5.66*
$t_{10}: B_{10}$! 3.84*	! 3.16	! 5.59*	! 3.76*	! 4.14*	! 3.59*	! 5.99*	! 5.79*	! 2.04	! 5.84*	! 4.88*	! 4.97*	! 4.44*
$t_{11}: B_{11}$! 2.08*	0.72	3.01	0.29	4.74	1.21	3.80	0.25	1.54	2.03	2.15	1.95	0.35
$t_{12}: B_{12}$! 3.27*	! 2.64	! 4.51*	! 2.41	! 4.25*	! 3.82*	! 5.47*	! 2.60	! 3.93*	! 6.47*	! 2.08	! 5.87*	! 3.66*
<i>F</i> -test:													
F: B_3 1 B_4	0.92	3.33	0.96	0.65	2.35	0.23	4.53	0.08	0.07	1.31	2.57	0.12	1.65
F: B_5 1 B_6	12.13*	18.04*	13.22*	12.43*	12.95*	17.58*	15.77*	14.35*	4.01	9.95*	11.68*	25.17*	8.39*
F: B_7 1 B_8	2.37	5.76	2.94	3.73	4.73	5.86	6.09	3.99	6.74*	3.81	1.66	1.27	4.04
F: B_9 1 B_{10}	7.44*	5.45	15.75*	7.31*	11.02*	6.71*	17.97*	16.82*	2.71	18.79*	12.42*	12.59*	17.90*
F: B_{11} 1 B_{12}	6.46*	3.73	10.54*	7.34*	14.13*	7.54*	16.07*	4.19	7.98*	22.67*	3.01	17.95*	8.39*
R ²	0.41	0.50	0.43	0.48	0.41	0.38	0.39	0.51	0.36	0.46	0.33	0.59	0.60
DW	2.02	2.00	2.03	1.96	2.12	2.06	2.05	2.00	2.03	1.97	2.00	1.93	1.99
F	0.039	0.036	0.073	0.082	0.119	0.156	0.229	0.085	0.119	0.118	0.215	0.122	0.128
Augmenta- tion	2,11, 12,14	1,2,7,10	2,12,23	2,4	2,3,12	2,5	11	1,3	1,4,12, 13,17	6	9,12,21	7	1,3,14, 24

Note: Auxiliary regression including a constant term, trend, and 11 seasonal dummies. A star indicates that the unit root hypothesis is rejected at the 5 percent level. Critical values are taken from Franses and Hobijn (1997). *F* is the standard error of estimate. *DW* is the Durbin-Watson statistic.

Table 4. Estimation of seasonal monthly deterministic model of hotel nights in Denmark by nationality. 1970.1 to 1997.12.

	Total	Danes	For- eigners	Swedes	Norwegi ans	Finns	Germans	English	Dutch	French	Italians	America ns	Japanese
<i>Coefficient:</i>													
Jan. R_1	.012	.005	.019	.001	.013	! .009	.032	.010	.024	.003	.025	! .015	! .004
Feb. R_2	.019**	.011	.020	.010	.027	! .009	.029	.020	.018	.024	.042	! .013	.018
March R_3	.020**	.012	.027*	.012	.029	! .009	.027	.004	.013	.010	! .001	! .018	.003
April R_4	.014*	.009	.021	.009	.045*	.001	.036	! .001	.018	! .003	! .019	! .032	! .011
May R_5	.017*	.012	.021	.023	.072***	.002	.034	.006	.023	! .013	! .020	! .027	.014
June R_6	.012	.003	.019	.010	.072***	.002	.018	.008	! .019	! .001	! .004	! .023	.027
July R_7	.016*	.005	.027*	.016	.101***	.012	.007	.001	.006	! .020	.034	! .028	.026
Aug. R_8	.013	.004	.017	! .006	.056**	.004	.013	.012	.046*	! .033	.034	! .029	! .009
Sep. R_9	.013	.014*	.005	.004	.020	.002	.014	.008	.001	.001	.019	! .039	! .002
Oct. R_{10}	.027**	.018**	.035**	.029*	.039*	.021	.092**	.017	.036	.022	.018	! .035	! .001
Nov. R_{11}	.021**	.012	.027*	.010	.035	.005	.004	.025	.018	.022	.022	! .008	.009
Dec. R_{12}	.012	! .004	.031*	! .009	.018	.001	.075*	.018	.012	.027	.063	! .006	! .002
R ²	0.25	0.36	0.32	0.35	0.35	0.19	0.33	0.37	0.25	0.32	0.21	0.44	0.47
DW	1.97	2.08	1.86	2.13	1.75	2.04	1.92	2.09	2.03	2.02	2.00	2.12	2.06
F	0.044	0.040	0.078	0.090	0.123	0.176	0.239	0.094	0.126	0.131	0.230	0.141	0.145
Augmen- tation	1,2, 12	1,3,10, 12,13	1,2,4, 11,12, 24	1,2, 12,13	3,4,9, 11,12, 23,24	1,2,5, 12,13	3,10,11, 12,15, 23,24	1, 12,13, 24	1,4,12, 13,17, 20,24	1,3, 12,14, 24	1,4,12, 13,16,24	1, 12,13, 24	1, 12,13

Note: A *** indicates that the variable is significant at the 1 percent level, a ** at the 5 percent level, and a * at the 10 percent level. DW is the Durbin-Watson statistic. Information with regard to the augmentation of the estimates is explained in the text. T-statistics are omitted due to lack of space. F is the standard error of estimate.

As proxy for the weather index W four variables enter the demand function unweighted. $TEMP$ is the monthly mean temperature, SUN is the monthly number of bright sunshine hours, $CLOUD$ is the monthly average cloud coverage, and $RAIN$ is the monthly relative humidity. High temperature and many sunshine hours are expected to increase the number of hotel nights, whereas cloudy and rainy weather is expected to decrease the number of hotel nights. For these variables multicollinearity problems may appear between $TEMP$ and SUN , and also between $CLOUD$ and $RAIN$.

ECM-model

Over the long run, i.e. in a steady state the relation between hotel nights and for example the exchange rate is constant on any given growth path. This is the golden rule of growth known from growth theory, and should also be expected to be valid with regard to the other variables included in this model formulation. In the short run, however, fluctuations around the growth path may exist. The error correction mechanism (ECM) captures these problems. Based on the long run relation

$$(3) \quad \log HN_t = \alpha_0 + \alpha_1 \log PDK_t + \alpha_2 \log PFO_t + \alpha_3 \log EXC_t \\ + \alpha_4 \log TEMP_t + \alpha_5 \log SUN_t + \alpha_6 \log CLOUD_t + \alpha_7 \log RAIN_t$$

an ECM-model, written with a one year lag for a given foreign nationality takes the form

$$(4) \quad \Delta \log HN_t = \beta_0 + \beta_1 \Delta \log PDK_t + \beta_2 \Delta \log PFO_t + \beta_3 \Delta \log EXC_t \\ + \beta_4 \Delta \log TEMP_t + \beta_5 \Delta \log SUN_t + \beta_6 \Delta \log CLOUD_t + \beta_7 \Delta \log RAIN_t \\ + \beta_8 (\log HN_{t-12} - \beta_9 \log PDK_{t-12} - \beta_{10} \log PFO_{t-12} - \beta_{11} \log EXC_{t-12} \\ - \beta_{12} \log TEMP_{t-12} - \beta_{13} \log SUN_{t-12} - \beta_{14} \log CLOUD_{t-12} - \beta_{15} \log RAIN_{t-12}) \\ + \epsilon_t \\ + \sum_{i=1}^n a_i \log X_{t,i}$$

where $\epsilon_t \sim \text{nid}(0, F^2)$, $t = 1, \dots, T$, and Δ is the 12th difference. In order to render the residuals white noise the endogenous variable is augmented.

In short, the coefficients can be given the following interpretation: β_0 is the constant term. β_1 through β_{15} are the impact effects, i.e. short run elasticities of hotel nights with regard e.g. the Danish price level β_1 , etc. β_8 is the error correction term measuring the feedback effect, i.e. the speed of adjustment. Finally, β_9 through β_{15} are the long run responses, i.e. the long run elasticities. Notice that β_8 is not restricted to unity. For the long run analysis the β_9 through β_{15} coefficients should be used, whereas the remaining β coefficients should be used for the short run analysis.

Because \mathcal{S}_8 is not restricted to any specific value, a two-step procedure needs to be undertaken in order to estimate all the coefficients in (4). First the long run coefficients are estimated. Second, by use of values of the error correction form the short run coefficients are estimated, see Doornik and Hendry (1996) or Sørensen (1996) chapter 4 for a detailed discussion.

Full model estimates by OLS on the elasticities in (4) are given in Tabel 5a and 5b. To obtain these results a general to specific modelling strategy was conducted along the criteria for model evaluation suggested by Doornik and Hendry (1996). First, the long run relation (3) was estimated, and then the long run solution was found by use of the "dynamic analysis" routine in PCGIVE giving the \mathcal{S}_8 -coefficient reported in the table. This input was then used in the second step where the dynamic part of (4) was estimated. The final lag augmentation of the endogenous variable is reported in the bottom row. Besides from the coefficients some goodness-of-fit statistics on autocorrelation are reported. The Durbin Watson statistic (DW), and as an alternative an LM-test for 1st and 2nd order autocorrelation, which has an F-distribution is reported, see Doornik and Hendry (1996) for a further discussion.

6. Evaluation of Forecasting Performance

In order to evaluate the ex post forecasts over the period from 1998.1 to 2000.12, i.e. 36 months or 3 years of the three presented models consider Table xx presenting *the mean absolute percentage error* (MAPE) and *the root mean squared percentage error* RMSPE along with a number of other statistics. MAPE and RMSPE are suggested as evaluation criterions by Song and Witt (2000). MAPE is defined as

$$MAPE = \frac{\sum_{t=0}^m \frac{|e_t|}{Y_t}}{m}$$

where Y_t is the observed value, \mathcal{Y}_t is the forecasted value, $e_t = Y_t - \mathcal{Y}_t$ is the forecasted error, and m is the length of the forecasting horizon. RMSPE is defined as (rest to appear)

$$RMSPE = \sqrt{\frac{1}{m} \sum_{t=1}^m \left(\frac{e_t}{y_t} \right)^2}$$

Table 5a. Results from estimation of ECM-models of hotel nights in Denmark by nationality.

	USA		Japan		Sweden		Norway		Finland	
	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run
	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$
Constant	0.017 (0.016)		0.019 (0.016)		0.011 (0.012)		0.047*** (0.018)		0.026 (0.019)	
PDK, Danish price	! 0.258 (0.373)	0.918*** (0.455)	! 0.101 (0.329)	0.337 (2.250)	! 0.316* (0.184)	! 0.653*** (1.428)	! 0.240 (0.348)	7.927** (5.127)	! 1.581*** (0.512)	! 0.630*** (0.680)
PFO, foreign price	! 1.456*** (0.453)	! 1.814*** (0.555)	! 0.406 (0.621)	! 3.234*** (3.102)	0.417** (0.224)	0.062* (1.509)	0.687* (0.354)	! 7.541* (5.222)	! 0.084 (0.468)	0.706 (0.624)
EXC, exchange rate	0.096 (0.065)	0.914*** (0.131)	0.019 (0.072)	1.730 (1.634)	0.218*** (0.069)	1.373*** (1.014)	0.320** (0.139)	0.847 (2.081)	0.273* (0.141)	1.183*** (0.195)
TEMP, temperature	0.074* (0.042)	! 0.010*** (0.193)	0.149** (0.051)	1.789*** (1.106)	0.038 (0.030)	0.785*** (0.238)	0.071* (0.041)	0.986*** (0.616)	0.009 (0.052)	! 0.193** (0.164)
SUN, sunny hours	0.095*** (0.032)	0.651*** (0.104)	! 0.012 (0.035)	! 2.611*** (1.914)	! 0.016 (0.021)	0.295* (0.145)	! 0.011 (0.029)	0.613 (0.401)	0.008 (0.038)	0.136* (0.086)
CLOUD	0.231*** (0.064)	0.531*** (0.167)	0.096 (0.069)	1.888*** (1.661)	! 0.017 (0.045)	! 0.920*** (0.635)	0.078 (0.057)	1.492 (1.166)	0.167** (0.079)	0.611** (0.236)
RAIN	0.001 (0.011)	0.073* (0.046)	! 0.014 (0.012)	! 0.698*** (0.570)	0.009 (0.070)	0.050* (0.133)	! 0.019* (0.011)	! 0.261 (0.252)	! 0.025 (0.014)	! 0.016 (0.680)
ECM	0.027* (0.016)		0.026* (0.014)		0.008 (0.015)		0.025* (0.014)		0.036** (0.014)	
<i>Short run</i>	Diagnostics	Augmentation	Diagnostics	Augmentation	Diagnostics	Augmentation	Diagnostics	Augmentation	Diagnostics	Augmentation
R ²	0.52	H: 1,12	0.48	H: 1,7,8,12	0.43	H:1,2,11,12	0.35	H: 9,10,12	0.32	H: 2,5,12
F	0.128	PK: 12	0.143	P: 3,5	0.084	T: 12	0.122	T: 1,2,6	0.161	PK: 9
DW	2.03	E: 9	1.82	T: 3,12	2.00	S: 10	1.78	R: 1,8	1.80	E: 9
AR	0.72	R: 3	3.31***	R: 4	0.94	C: 2,10	2.40**		0.78	T: 6,10
ARCH	1.91*		0.92		1.10		3.60***		2.05**	C: 12
Normality	15.70***		1.52		0.06		8.47**		119.74***	

Note: The number in the parenthesis is the standard error. A *** indicates that the variable is significant at the 1 percent level, a ** at the 5 percent level, and a * at the 10 percent level. The significance of the short run estimates are tested by use of a t-test, whereas the significance of the long run estimates is tested by use of a F-test. F is the standard error of estimate. DW is the Durbin-Watson statistic. AR is a LM test of 1st to 7th order autocorrelation with an F-distribution. ARCH is a test for autoregressive conditional heteroscedasticity also of 7th order. Finally, a P-test for normality is presented. See also Doornik and Hendry (1996).

Table 5b. Results from estimation of ECM-models of hotel nights in Denmark by nationality.

	Germany		England		Netherlands		France		Italy	
	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run
	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$
Constant	0.021 (0.031)		0.013 (0.010)		0.047*** (0.016)		0.031** (0.013)		0.141*** (0.028)	
PDK, Danish price	! 5.282** (2.217)	6.011** (3.634)	! 0.500** (0.214)	! 3.679*** (1.333)	! 0.768** (0.353)	9.214* (8.285)	1.540*** (0.513)	1.564*** (1.525)	1.028 (0.828)	! 6.145 (5.349)
PFO, foreign price	6.364*** (1.969)	! 6.851*** (4.389)	0.534 (0.417)	3.214*** (1.079)	0.669* (0.398)	! 8.070 (9.009)	! 1.653*** (0.494)	! 1.465 (1.480)	! 1.325 (0.948)	4.316*** (3.626)
EXC, exchange rate	! 0.067 (0.494)	! 5.245*** (3.994)	! 0.106* (0.060)	0.393 (1.079)	0.048 (0.297)	! 11.030 (8.742)	0.531*** (0.203)	1.669** (0.897)	0.654*** (0.181)	1.024 (1.142)
TEMP, temperature	0.016 (0.081)	2.225*** (0.839)	0.013 (0.030)	! 0.579*** (0.339)	0.033 (0.043)	0.901 (0.954)	0.047 (0.044)	! 0.096 (0.242)	! 0.058 (0.072)	! 0.291 (0.561)
SUN, sunny hours	! 0.007 (0.007)	! 1.335*** (0.708)	0.028 (0.022)	0.201 (0.152)	0.028 (0.032)	! 0.969*** (1.608)	! 0.011 (0.029)	0.419*** (0.133)	0.005 (0.052)	0.228*** (0.555)
CLOUD	0.171 (0.133)	! 0.189 (0.881)	0.074 (0.046)	1.024*** (0.554)	0.101 (0.063)	! 3.161*** (3.863)	! 0.018 (0.059)	! 0.291*** (0.286)	0.133 (0.109)	1.334 (1.075)
RAIN	! 0.013 (0.021)	! 0.339* (0.202)	! 0.009 (0.008)	! 0.196** (0.115)	! 0.013 (0.011)	! 0.030 (0.299)	0.017 (0.011)	0.046 (0.055)	0.009 (0.018)	0.068*** (0.308)
ECM	0.007 (0.012)		0.002 (0.008)		0.005 (0.009)		0.041** (0.021)		0.049*** (0.017)	
<i>Short run</i>	Diagnostics	Augmentation	Diagnostics	Augmentation	Diagnostics	Augmentation	Diagnostics	Augmentation	Diagnostics	Augmentation
R ²	0.29	H: 10,12	0.40	H: 1,12	0.23	H: 1,4,12	0.38	H: 1,10,12	0.29	H: 11,12
F	0.241	PK: 1	0.090	P: 3,4	0.130	T: 1	0.125	E: 12	0.217	PK: 11
DW	1.93	P: 6,9	1.92	E: 7	1.93	S: 8	1.92	T: 7,11,12	1.77	P: 6,7
AR	2.47**	S: 1	1.87*	T: 6,11	1.13	C: 8	1.62	C: 8	1.90	S: 3,6
ARCH	0.17	C: 1	0.89		0.68		0.87	R: 7	2.68**	C: 6,8,12
Normality	395.01***	R: 2	10.00***		38.04***		10.31***		3.47	

Note: The number in the parenthesis is the standard error. A *** indicates that the variable is significant at the 1 percent level, a ** at the 5 percent level, and a * at the 10 percent level. The significance of the short run estimates are tested by use of a t-test, whereas the significance of the long run estimates is tested by use of a F-test. F is the standard error of estimate. DW is the Durbin-Watson statistic. AR is a LM test of 1st to 7th order autocorrelation with an F-distribution. ARCH is a test for autoregressive conditional heteroscedasticity also of 7th order. Finally, a P-test for normality is presented. See also Doornik and Hendry (1996).

Table 5c. Results from estimation of ECM-models of hotel nights in Denmark by nationality.

	Danes		Foreigners		Total	
	Short run	Long run	Short run	Long run	Short run	Long run
	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$	$\$_0$ to $\$_8$	$\$_9$ to $\$_{15}$
Constant	0.029*** (0.006)		0.034*** (0.010)		0.026*** (0.007)	
PDK, Danish price	! 0.255*** (0.069)	2.208*** (0.558)	! 0.115 (0.115)	0.108*** (0.168)	! 0.205*** (0.068)	0.647*** (0.133)
PFO, foreign price						
EXC, exchange rate						
TEMP, temperature	! 0.009 (0.013)	! 0.684 (0.981)	! 0.010 (0.024)	0.785*** (0.449)	! 0.007 (0.013)	0.286** (0.274)
SUN, sunny hours	0.013 (0.010)	3.899*** (3.738)	0.005 (0.018)	0.522** (0.212)	0.001 (0.009)	0.058 (0.189)
CLOUD	! 0.019 (0.020)	! 7.438*** (7.137)	0.043 (0.037)	! 0.820*** (0.899)	! 0.038* (0.021)	! 1.598*** (0.814)
RAIN	! 0.001 (0.003)	! 0.542* (0.557)	! 0.004 (0.006)	! 0.095** (0.178)	! 0.006* (0.003)	0.082** (0.133)
ECM	! 0.005*** (0.001)		0.029*** (0.013)		! 0.038*** (0.016)	
<i>Short run</i>	Diagnostics	Augmentation	Diagnostics	Augmentation	Diagnostics	Augmentation
R ²	0.34	H: 2,3,10,12	0.38	H: 1,2,11,12	0.35	H: 1,2,11,12
F	0.040	S: 10	0.074	T: 2,4,5	0.041	T: 2
DW	1.43	C: 6,10	1.89	R: 1	1.97	
AR	5.90***		1.81*		2.11**	
ARCH	0.86		1.46		0.52	
Normality	3.74		5.34*		5.84*	

Note: The number in the parenthesis is the standard error. A *** indicates that the variable is significant at the 1 percent level, a ** at the 5 percent level, and a * at the 10 percent level. The significance of the short run estimates are tested by use of a t-test, whereas the significance of the long run estimates is tested by use of a F-test. F is the standard error of estimate. DW is the Durbin-Watson statistic. AR is a LM test of 1st to 7th order autocorrelation with an F-distribution. ARCH is a test for autoregressive conditional heteroscedasticity. Finally, a P-test for normality is presented. See also Doornik and Hendry (1996).

7. Conclusion

(To appear)

Notes

1. This research is a part of the project *Seasonal Forecasting of the Demand for Hotel and Camping Nights*, financially supported by the Danish Social Science Research Council, grant 9901857. A background material in form of the data set used and detailed plots of all series is available from the author on request. For the statistical analyses programming and estimation were undertaken in PcGive/PcFiml 9 and WinRATS 4.3.

2. Jensen (1998) presents estimates of income and price elasticities for the six most important nationalities visiting Denmark as tourists using annual observations ranging from 1960 to 1995. Using annual observations it is possible to include variables like income and transportation costs. However, using the petrol price as proxy for transportation costs does not improve the results. The influence of the climate is also discussed, but no variable is included a tested.
3. Data on exchange rates and price indices have kindly been supplied by Dan Knudsen from the Danish National Bank.

References

- Barry, K., and J. O'Hagan, 1972, An Econometric Study of British Tourist Expenditure in Ireland. *Economic and Social Review*, 3, p 143-161.
- Barsky, R.B., and J.A. Miron, 1989, The Seasonal Cycle and the Business Cycle. *Journal of Political Economy*, 97, p 503-534.
- Clements, M.P., and D.F. Hendry, 1999, *Forecasting Non-stationary Economics Time Series*. MIT Press.
- Doornik, J.A., and D.F., Hendry, 1996, *PcGive 9.0. PcFiml*. International Thompson Publishing.
- Franses, P.H., 1991, Seasonality, Non-Stationarity and the forecasting of Monthly Time Series. *International Journal of Forecasting*, 11, p 447-475.
- Franses, P.H., and B. Hobijn, 1997, Critical Values for Unit Root Tests in Seasonal Time Series. *Journal of Applied Statistics*, 18, p 25-47.
- Hylleberg, S., R.F. Engle, C.W.J. Granger, and B.S. Yoo, 1990, Seasonal Integration and Cointegration. *Journal of Econometrics*, 44, p 215-238.
- Jensen, T., 1998, Income and Price Elasticities by Nationality for Tourists in Denmark. *Tourism Economics*, 4, p 101-130.
- Martin, C.A. and S.F. Witt, 1989, Accuracy of Econometric Forecasts of Tourism. *Annals of Tourism Research*, 16, p 407-428.
- Song, H. And S.F. Witt, 2000, *Tourism Demand Modelling and Forecasting*. Pergamon Press.
- Sørensen, N.K., 1999, Modelling the Seasonality of Hotel Nights in Denmark, *Tourism Economics*, 5, p 9-25.
- Sørensen, N.K., 1996, *Empirical Studies of Trade and Seasonal Macroeconomic Analyses*. Phd-thesis 1996-1, Department of Economics University of Aarhus.
- Witt, S.F. and C.A. Witt, 1992, *Modelling and Forecasting Demand in Tourism*. Academic Press.
- Witt, S.F. and C.A. Witt, 1995, Forecasting Tourism Demand: A Review of Empirical research. *International Journal of Forecasting*, 11, p 447-475.