Modelling Risk in the Spanish Tourism Industry

Bartolomé, A[1], M. McAller [2], V. Ramos[1], J. Rey-Maquieira [1]

[1]Department of Applied Economics, University of the Balearic Islands, Spain
[2] School of Economics and Commerce, University of Western Australia, Australia

Keywords: *Conditional volatility, tourism demand, financial econometrics, risk, daily arrivals*

EXTENDED ABSTRACT

According to United Nations World Tourism Organization (UNWTO), from 2005 to 2006, international tourist arrivals increased by 5% worldwide, notwithstanding "terrorism, natural disasters, health scares, oil prices, exchange rate fluctuations and economic and political uncertainties" (UNWTO 2006). In the case of Spain, in 2006 the growth rate was 4.7%, but this rate varied between the five main tourism regions Spain is especially known for its standardized sun and sand tourism. Despite the general belief that this type of tourism is no longer attractive to new tourists, it is one of the main pillars of the Spanish economy. In fact, the tourism industry represents approximately 9.8% of the labor force.

The main purpose of this paper is to measure and model the risk involved in the Spanish tourism industry. This paper uses monthly data in international tourist arrivals to Spain during the last 10 years, from January 1996 to April 2007. In 2006 Spain received a total of 58.5 million international tourists. This data has been classified into five sub-samples, which correspond to the five major tourism regions. Over 83% of total international tourist arrivals stay in one of these five main Spanish regions.

The tourist country of origin determines the market specialization of a destination. It is in this sense that a destination which attracts a higher number of tourist nationalities has a greater diversification of demand. In the case of the Balearic Islands, 73.4% of total tourist arrivals in 2006 originated from Germany and the UK. The Balearic Islands are the destination with the highest market specialization in Spain. In contrast, Madrid has almost 75% of its international tourist arrivals that originate from countries other than the two main markets, namely France and the UK. Consequently, Madrid is a region with a higher market diversification, or lower market specialization.

The seasonality pattern is clearly identified for every region. The Gini coefficient has been used to measure seasonality. The Balearic Islands have the highest seasonality (in the sense that the difference in monthly arrivals between the low and high seasons is the greatest) and the Canary Islands has the lowest. Valencia, Andalusia and Catalonia have a growing number of international tourist arrivals during the winter months. In the case of Catalonia, the Christmas peak is clearly identified, probably due to the popular skiing resorts. As expected, the Canary Islands follow a different seasonality pattern, and after a few years of declining number of arrivals, seem to have recovered since 2005. January is typically the month with the minimum tourism activity.

1. INTRODUCTION

In the last few decades, tourism has become one of the most prominent engines for the Spanish economy. Since the 1960's, the number of tourist international arrivals has increased considerably. In fact, in 2006 Spain was the second largest country in the world in the number of international tourist arrivals, after the leader France ahead of USA, China, Italy and UK. Furthermore, it was also the second largest country in terms of tourism receipts. Only during the last decade, has the total number of tourists who visited Spain risen from 39.5 million in 1997 to 58.5 million in 2006.

2. DATA

The data analyzed in this paper have been obtained from the "Instituto de Estudios Tursíticos" (IET, 2006a, 2006b, 2007). This national institution analyzes the tourism industry in Spain and, through its website, publishes data gathered by "Frontur". Frontur denotes the statistical studies based on the use of periodic surveys completed at Spanish frontiers by a sample of international visitors. Thus, this paper is based on total international tourists who arrive by road, rail, sea or air. All domestic tourism is excluded from the sample.

Frontur describes a visitor as any person arriving in a country other than the usual place of residence, for any reason apart following an occupation remunerated from within the country visited. An international tourist is a temporary visitor staying for at least 24 hours in the country visited, for which the purpose of the journey can be classified under the headings of either leisure or business, family, mission or meeting (IET, 2006a). The data are monthly international tourist arrivals to the five main tourist regions in Spain, which account for more than 84% of total international tourist arrivals, from January 1997 to April 2007, giving a total of 126 observations for each of the five regions, as well as the total for Spain. Due to space restriction, this paper presents the results corresponding to the three main tourist regions, that is, Catalonia, the Balearic Islands and the Canary Islands.

3. CONDITIONAL MEAN AND CONDITIONAL VOLATILITY MODELS

The time series models to be estimated for the conditional means of the monthly international tourist arrivals and annual changes in tourist arrivals, as well as their conditional volatilities, are discussed below. McAleer, Chan and Marinova (2007) argue that, for a wide range of financial and other data series, time-varying conditional variances can be explained empirically through the autoregressive conditional heteroskedasticity (ARCH) model, which was proposed by Engle (1982). When the time-varying conditional variance has both autoregressive and moving average components, this leads to the generalized ARCH(*p,q*), or GARCH(*p,q*), model of Bollerslev (1986). The lag structure of the appropriate GARCH model can be chosen by information criteria, such as those of Akaike and Schwarz, although it is very common to impose the widely estimated GARCH(1,1) specification in advance.

Consider the stationary AR(1)-GARCH(1,1) model for monthly international tourist arrivals (or their monthly change, as appropriate), y_t:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t, \qquad |\phi_1| < 1 \tag{1}$$

for $t = 1,...,n$, where the shocks (or movements in monthly air passenger arrivals) are given by:

$$\varepsilon_t = \eta_t \sqrt{h_t}, \quad \eta_t \sim iid(0,1)$$
$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}, \tag{2}$$

where $\omega > 0, \alpha \ge 0$ and $\beta \ge 0$ are sufficient conditions to ensure that the conditional variance $h_t > 0$. The AR(1) model in equation (1) can easily be extended to univariate or multivariate ARMA(*p,q*) processes (for further details, see Ling and McAleer (2003a)). In equation (2), the ARCH (or α) effect indicates the short run persistence of shocks, while the GARCH (or β) effect indicates the contribution of shocks to long run persistence (namely, $\alpha + \beta$).

Ling and McAleer (2003) showed that the QMLE for GARCH(*p,q*) is consistent if the second moment of εt is finite. For GARCH(*p,q*), Ling and Li (1997) demonstrated that the local QMLE is asymptotically normal if the fourth moment of ε_t is finite, while Ling and McAleer (2003) proved that the global QMLE is asymptotically normal if the sixth moment of εt is finite. Using results from Ling and Li (1997) and Ling and McAleer (2002a; 2002b), the necessary and sufficient condition for the existence of the second moment of εt for GARCH(1,1) is $\alpha + \beta < 1$

As discussed in McAleer et al.(2007), Elie and Jeantheau (1995) and Jeantheau (1998) established that the log-moment condition was sufficient for consistency of the QMLE of an univariate GARCH(*p,q*) process. Boussama (2000) showed that the log-moment condition was sufficient for asymptotic normality. Based on these theoretical developments, a sufficient condition for the QMLE of GARCH(1,1) to be consistent and asymptotically normal is given by the log-moment condition, namely

$$E(\log(\alpha\eta_t^2 + \beta)) < 0 . \quad (3)$$

This condition involves the expectation of a function of a random variable and unknown parameters.

The effects of positive shocks (or upward movements) on the conditional variance, h_t, are assumed to be the same as the negative shocks (or downward movements) in the symmetric GARCH model. In order to accommodate asymmetric zbehaviour, Glosten, Jagannathan and Runkle (1992) proposed the GJR model, for which GJR(1,1) is defined as follows:

$$h_t = \omega + (\alpha + \gamma I(\eta_{t-1}))\varepsilon_{t-1}^2 + \beta h_{t-1}, \quad (4)$$

where $\omega > 0, \alpha \ge 0, \alpha + \gamma \ge 0$ and $\beta \ge 0$ are sufficient conditions for $h_t > 0,$ and $I(\eta_t)$ is an indicator variable defined by:

$$I(\eta_t) = \begin{cases} 1, & \varepsilon_t < 0 \\ 0, & \varepsilon_t \ge 0 \end{cases}$$

as η_t has the same sign as εt. The indicator variable differentiates between positive and negative shocks of equal magnitude, so that asymmetric effects in the data are captured by the coefficient γ, with $\gamma \ge 0$. The asymmetric effect, γ, measures the contribution of shocks to both short run persistence, $\alpha + \frac{\gamma}{2}$, and to long run persistence, $\alpha + \beta + \frac{\gamma}{2}$.

Ling and McAleer (2002b) showed that the regularity condition for the existence of the second moment for GJR(1,1) under symmetry of η_t is given by:

$$\alpha + \beta + \frac{1}{2}\gamma < 1 , \quad (5)$$

while McAleer et al. (2007) showed that the weaker log-moment condition for GJR(1,1) was given by:

$$E(\log[(\alpha + \gamma I(\eta_t))\eta_t^2 + \beta]) < 0 , \quad (6)$$

which involves the expectation of a function of a random variable and unknown parameters.

An alternative model to capture asymmetric behaviour in the conditional variance is the Exponential GARCH (EGARCH(1,1)) model of Nelson (1991), namely:

$$\log h_t = \omega + \alpha|\eta_{t-1}| + \gamma\eta_{t-1} + \beta \log h_{t-1}, \ |\beta| < 1 \quad (7)$$

where the parameters have a distinctly different interpretation from those in the GARCH(1,1) and GJR(1,1) models.

4. ESTIMATED MODELS

We have applied the modified unit root tests, denoted as $MADF^{GLS}$ and MPP^{GLS}, to the time series of monthly international tourist arrivals in Spain, and to the five main tourist regions. In essence, these tests use GLS de-trended data and the modified Akaike information criterion (MAIC) to select the optimal truncation lag. The asymptotic critical values for both tests are given in Ng and Perron (2001).

The results of the unit root tests are obtained from the econometric software package EViews 5.0, and are reported in Tables 1-4. The existence of a zero frequency unit root is tested for the monthly international tourist arrivals, first difference in arrivals, annual difference in arrivals, logarithm of arrivals, first difference in the log of arrivals, and the annual differences in the log of arrivals (that is, the annual growth rate) for the five regions and for Spain.

The results of the unit root test for Spain and the five major tourist regions are remarkably similar. Apart from a few exceptions, the results of both tests are strikingly similar for the six variables defined above for Spain and the five major tourist regions. For a variety of lag lengths, monthly international tourist arrivals, the difference in monthly international tourist arrivals, the logarithm of monthly international tourist arrivals and the growth rate in monthly international tourist arrivals are all found to be non-stationary, that is, integrated of order one. The yearly difference in monthly international tourist arrivals and the yearly growth rate in monthly international tourist arrivals are both found to be stationary, that is integrated of order zero. For these reasons, models of both monthly international tourist arrivals and the yearly growth rate in monthly international tourist arrivals, as well as their associated volatilities, will be estimated by maximum likelihood methods.

Models 1 and 2 below are used to estimate monthly international tourist arrivals and the yearly difference in monthly tourist arrivals, as well as their respective volatilities using the GARCH(1,1), GJR(1,1) and EGARCH(1,1) specifications:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-12} + \varepsilon_t \qquad (8)$$

The conditional mean estimates for Equation 8 for Spain in Table 5 suggest that international tourist arrivals lagged one month have a quantitatively small, though significant, effect on current monthly arrivals, while international tourist arrivals lagged twelve months (that is, the yearly lagged effect) is highly significant and essentially a unit root. The asymmetric effect, γ, in both the GJR and EGARCH models is found to be zero so that the effects of positive and negative shocks of equal magnitude on volatility are equivalent. The short run persistence of shocks in the GARCH model is significant at 0.005, while the long run persistence shocks is 0.812. The second moment condition for both GARCH and GJR is satisfied, so that the log-moment condition is also satisfied. Therefore, the QMLE are consistent and asymptotically normal, so that inferences are valid.

The estimates of Equation 8 for Andalusia delivers a similar comment to that for Spain holds regarding the impact of international tourist arrivals lagged one month and one year. The asymmetry effect for both GJR and EGARCH are found to be significant, with the estimates for EGARCH suggesting Type 2 Asymmetry associated with overbooking pressure on carrying capacity. Unlike the estimates for Spain, the short run persistence of shocks for Andalusia is very high and not significantly different from unity for both GARCH and GJR. As the log-moment condition is satisfied in both cases, the QMLE are consistent and asymptotically normal, and inferences are valid.

The estimates for the Balearic Islands in Table 5 are reasonable similar to those for Spain in that the asymmetric effects are not significant for GJR or EGARCH. However the short run persistence of shocks for the GARCH model is significant at 0.534 which is far higher than for Spain and much lower than for Andalusia. However, like the two previous sets of results, the log-moment condition is satisfied for both GARCH and GJR, so that the QMLE are consistent and asymptotically normal, and inferences are valid.

The conditional mean estimates for the Canary Islands are very similar to the previous cases, except that the effect of the yearly lag is much lower at between 0.805 and 0.883. As in the case of Spain and the Balearic Islands, the asymmetric effects are not significant for GJR or EGARCH. The short run persistence of shocks for GARCH is positive but not significant. Overall, the QMLE for the Canaries do not seem to be particularly intuitive.

Both the conditional mean and conditional volatility estimates for Catalonia are presented in Table 5. The impact of international tourist arrivals lagged one month and one year are very similar to the previous three sets of results, and as in the case of all previous results, the asymmetric effects of positive and negative shocks are insignificant. The short run persistence of shocks for GARCH is positive and similar to that of the Balearic Islands at 0.487. As both the second moment and log moment conditions are satisfied, the QMLE are consistent and asymptotically normal, and inferences are valid.

Estimates of the conditional mean and conditional volatility for Valencia are not presented for space restrictions. A similar comment to the previous five sets of results applies to the impact of international tourist arrivals lagged one month and one year. Although the asymmetric effect in the GJR model is not significant the asymmetric effect for EGARCH displays Type 3 Asymmetry, namely tourism saturation in the high season. As the short and long run persistence for GARCH are not intuitive, even though the QMLE are consistent and asymptotically normal, the EGARCH model is preferred.

5. REFERENCES

Bollerslev, T. (1986), Generalised autoregressive conditional heteroscedasticity, *Journal of Econometrics*, 31, 307-327.

Boussama, F. (2000), Asymptotic normality for the quasi-maximum likelihood estimator of a GARCH model, *Comptes Rendus de l'Academie des Sciences,* Serie I(131), 81-84(in French).

Elie, L. and T. Jeantheau (1995), Consistency in heteroskedastic models, *Comptes Rendus de l'Académie des Sciences,* Série I, 320, Série I, 320, 1255-1258 (in French).

Engle, R.F. (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica,* 50, 987-1007.

Glosten, L., R. Jagannathan and D. Runkle (1992), On the relation between the expected value and volatility of nominal excess return on stocks, *Journal of Finance,* 46, 1779-1801.

IET (2006a), *Encuesta de Gasto Turístico (EGATUR)*. Madrid: Instituto de Estudios Turísticos.

IET (2006b), *Moviminetos Turísticos en Fronteras (FRONTUR)*. Madrid: Instituto de Estudios Turisticos.

IET (2007), *Balance del Turismo en España en 2006*. Madrid: Instituto de Estudios Turísticos.

Jeantheau, T. (1998), Strong consistency of estimators for multivariate ARCH models, *Econometric Theory,* 14, 70-86.

Lee, S.W. and B.E. Hansen (1994), Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator, *Econometric Theory,* 10, 29-52.

Ling, S. and W.K. Li (1997), On fractionally integrated autoregressive moving-average models with conditional heteroskedasticity, *Journal of the American Statistical Association,* 92, 1184-1194. .

Ling, S. and M. McAleer (2002a), Necessary and sufficient moment conditions for the GARCH(r,s) and asymmetric power GARCH(r,s) models, *Econometric Theory,* 18, 722-729.

Ling, S. and M. McAleer (2002b), Stationarity and the existence of moments of a family of GARCH processes, *Journal of Econometrics,* 106(1), 109-117.

Ling, S. and M. McAleer (2003), Asymptotic theory for a vector ARMA-GARCH model, *Econometric Theory,* 19, 278-308.

McAleer, M., F. Chan and D. Marinova (2007), An econometric analysis of asymmetric volatility: theory and application to patents, *Journal of Econometrics,* 139, 259-284.

Ng, S. and P. Perron (2001), Lag length selection and the construction of unit root tests with good size and power, *Econometrica,* 69, 1519-1554.

Table 1. Unit Root Tests for Spain

Variables	$MADF^{GLS}$		MPP^{GLS}		Lags
	Z=(1, t)	Z=(1)	Z=(1, t)	Z=(1)	
y_t	-3.498**	6.117	0.046	0.526	11
Δy_t	-1.039	-0.461	0.138	0.102	12
$\Delta_{12}y_t$	-4.041***	-3.936***	28.443***	-24.955***	4
$Log(y_t)$	-2.622	6.730	0.045	0.668	12
$\Delta log(y_t)$	-1.856	-0.158	0.164	0.021	12
$\Delta_{12}log(y_t)$	-4.051***	-3.528***	-24.805***	-15.265***	4

Notes: y_t denotes monthly international tourist arrivals to Spain.
(1,t) and (1) denote the presence of an intercept and trend, and intercept, respectively.
(***), (**) and (*) denote the null hypothesis of a unit root is rejected at the 1%, 5% and 10% levels, respectively.

Table 2. Unit Root Tests for Balearic Islands

y_t	-2.969*	1.112	0.144	0.062	12
Δy_t	-1.389	-0.288	0.162	0.093	12
$\Delta_{12}y_t$	-3.831***	-3.851***	-16.934***	-16.721***	2
$Log(y_t)$	-2.577	0.999	0.023	0.334	12
$\Delta log(y_t)$	-3.878***	0.041	0.343	-0.001	12
$\Delta_{12}log(y_t)$	-1.098	-0.665	-2.653	-0.346	12

Table 3. Unit Root Tests for Canary Islands

y_t	-1.704	-1.564	-0.282	-0.240	12
Δy_t	0.251	-1.103	1.432	1.937	12
$\Delta_{12}y_t$	-3.524**	-2.261**	-26.252***	-7.514*	4
$Log(y_t)$	-1.880	-1.673*	-0.365	-0.291	12
$\Delta log(y_t)$	0.137	-1.129	1.226	1.640	12
$\Delta_{12}log(y_t)$	-3.525**	-2.361**	-26.165***	-8.812**	4

Table 4. Unit Root Tests for Catalonia

y_t	-3.168**	5.442	-0.025	0.819	11
Δy_t	-0.770	-0.562	0.200	0.179	12
$\Delta_{12}y_t$	-4.202***	-3.866***	-36.083***	-27.825***	3
$Log(y_t)$	-2.912*	6.938	0.032	0.857	11
$\Delta log(y_t)$	-2.694	-0.018	0.231	-0.014	12
$\Delta_{12}log(y_t)$	-3.787***	-3.226***	-19.733***	-10.722**	2

Table 5. Conditional Mean and Conditional Volatility Models for Equation (8)

Parameters	Spain			Balearics		
	GARCH	GJR	EGARCH	GARCH	GJR	EGARCH
ϕ_0	142915 (86429)	144333 (67867)	105737* (65168)	3712* (10745)	3740* (11170)	3690* (7237)
ϕ_1	0.021* (0.031)	0.018* (0.027)	0.021* (0.020)	0.013* (0.019)	0.013* (0.019)	0.009* (0.014)
ϕ_2	0.986 (0.034)	0.990 (0.029)	0.996 (0.026)	0.993 (0.027)	0.993 (0.027)	0.996 (0.013)
ω	7.3E+9* (1.7E+10)	9.1E+8* (5.30E+9)	31.501 (0.273)	2.08E+9 (4.76E+8)	2.1E+9 (5.1E+8)	31.379 (3.359)
GARCH/GJR α	0.005* (0.034)	0.046 (0.010)	--	0.534 (0.237)	0.538* (0.290)	--
GJR γ	--	-0.079* (0.045)	--	--	-0.009* (0.463)	--
GARCH/GJR β	0.886 (0.270)	0.981 (0.094)	--	0.123* (0.104)	0.122* (0.126)	--
EGARCH α	--	--	0.401 (0.180)	--	--	0.839 (0.154)
EGARCH γ	--	--	-0.028* (0.164)	--	--	0.065* (0.071)
EGARCH β	--	--	-0.278 (0.000)	--	--	-0.450 (0.148)
Second moment	0.892	0.988	--	0.657	0.656	--
Log-moment	-0.114	-0.013	--	-1.027	-1.029	--
Parameters	**Canary Islands**			**Catalonia**		
ϕ_0	97438 (39779)	91976 (30705)	76892 (30837)	52727 (26634)	56117 (25909)	73132 (16065)
ϕ_1	0.047* (0.041)	0.085 (0.035)	0.024* (0.039)	0.053* (0.030)	0.052* (0.033)	0.016* (0.017)
ϕ_2	0.836 (0.048)	0.805 (0.038)	0.883 (0.038)	0.955 (0.022)	0.954 (0.024)	0.975 (0.017)
ω	2.70E+9 (6.24E+8)	2.79E+9 (4.92E+9)	33.391 (3.953)	7.91E+9 (2.26E+9)	7.5E+9 (2.2E+9)	35.699 (1.853)
GARCH/GJR α	0.170* (0.095)	0.263 (0.106)	--	0.487 (0.147)	0.745 (0.365)	--
GJR γ	--	-0.149* (0.130)	--	--	-0.396* (0.391)	--
GARCH/GJR β	-0.786 (0.202)	-0.852 (0.117)	--	0.015* (0.157)	0.011* (0.135)	--
EGARCH α	--	--	0.656 (0.223)	--	--	0.982 (0.162)
EGARCH γ	--	--	0.136* (0.130)	--	--	-0.030* (0.103)
EGARCH β	--	--	-0.606 (0.179)	--	--	-0.576 (0.077)
Second moment	-0.616	-0.663	--	0.502	0.559	
Log-moment	NA	NA	--	-1.957	-1.967	

Notes: Y_t is the number of passenger arrivals. Numbers in parentheses are standard errors.
(*) indicates the coefficient is not significant at 5%; otherwise, all estimates are significant at the 5% level.