18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Modelling international tourist arrivals and volatility forTaiwan

Chang, C.-L.[1], **M. McAleer**[2] **and D. Slottje**[3]

[1]Department of Applied Economics, National Chung Hsing University, Taiwan
[2]Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, Tinbergen Institute, Rotterdam, and Department of Applied Economics, National Chung Hsing University, Taiwan
[3]FTI Consulting and Department of Economics, Southern Methodist University, USA
Email: changchialin@nchu.edu.tw

Abstract: International tourism is a major source of export receipts for many countries worldwide. Although it is not yet one of the most important industries in Taiwan (or the Republic of China), an island in East Asia off the coast of mainland China (or the People's Republic of China), the leading tourism source countries for Taiwan are Japan, followed by USA, Republic of Korea, Malaysia, Singapore, UK, Germany and Australia. These countries reflect short, medium and long haul tourist destinations. Although the People's Republic of China and Hong Kong are large sources of tourism to Taiwan, the political situation is such that tourists from these two sources to Taiwan are reported as domestic tourists. Daily data from 1 January 1990 to 30 June 2007 are obtained from the National Immigration Agency of Taiwan. The Heterogeneous Autoregressive (HAR) model is used to capture long memory properties in the data. In comparison with the HAR(1) model, the estimated asymmetry coefficients for GJR(1,1) are not statistically significant for the HAR(1,7) and HAR(1,7,28) models, so that their respective GARCH(1,1) counterparts are to be preferred. These empirical results show that the conditional volatility estimates are sensitive to the long memory nature of the conditional mean specifications. Although asymmetry is observed for the HAR(1) model, there is no evidence of leverage. The QMLE for the GARCH(1,1), GJR(1,1) and EGARCH(1,1) models for international tourist arrivals to Taiwan are statistically adequate and have sensible interpretations. However, asymmetry (though not leverage) was found only for the HAR(1) model, and not for the HAR(1,7) and HAR(1,7,28) models.

***Keywords**: Tourism demand, international tourist arrivals, heterogeneous autoregressive model (HAR), conditional volatility, asymmetry, leverage.*

1. Introduction

Taiwan (or the Republic of China) is an island in East Asia off the coast of mainland China (or the People's Republic of China). International tourism is a major source of export receipts for many countries worldwide, and Taiwan is no exception. The most well known tourist attractions in Taiwan include the National Palace Museum (Taipei), Night Markets (especially in Taipei), Taipei 101, formerly the world's tallest building, Sun Moon Lake (central highlands), and Taroko National Park (east coast).

The most important tourism source countries to Taiwan are Japan, followed by USA, Republic of Korea, Malaysia, Singapore, UK, Germany and Australia, which reflect short, medium and long haul destinations. The three most important countries during the sample period have been Japan, USA and Republic of Korea. Although the People's Republic of China and Hong Kong are large sources of tourism to Taiwan, the political situation is such that tourists from these two sources to Taiwan are reported as domestic tourists.

This paper models international tourist arrivals and volatility (or squared deviation from the mean) in international tourist arrivals to Taiwan. Daily data from 1 January 1990 to 30 June 2007 are obtained from the National Immigration Agency of Taiwan. By using daily data, we can approximate the modelling strategy and analysis to those applied to financial time series data. From a time series perspective, there are several reasons for using daily data (see, for example, McAleer (2009)). Just to mention some, daily data allow investigating whether the time series properties have changed, the time series behaviour at other frequencies can be obtained by aggregation of daily data, and the sample size is considerably increased.

2. Data

The data set comprises daily international tourist arrivals from 1 January 1990 to 30 June 2007, giving 6,390 observations, and are obtained from the National Immigration Agency of Taiwan. On an annual basis, the number of international tourist arrivals to Taiwan has shown an average growth rate of around 4% per annum from 1990 to 2007. The lowest growth rate was observed in 2003, with a decrease of 23.19% over the previous year (due to the outbreak of SARS), while the highest growth rate occurred in 2004, when there was a significant increase of 36.58% over 2003. In the sample period as a whole, there was an increase of around 75% in international tourist arrivals to Taiwan, which would seem to indicate a reasonably good performance in the tourism sector over the decade. Nevertheless, the annual average international tourist arrivals growth rate reveals that there is scope for a significant increase in international tourism to Taiwan. In order to manage tourism growth and volatility, it is necessary to model adequately international tourist arrivals and their associated volatility.

3. Unit Root Tests

Standard ADF unit root tests are obtained from the econometric software package EViews 6.0. There is no evidence of a unit root in daily international tourist arrivals to Taiwan in the model with a constant and trend as the deterministic terms, or with just a constant. These empirical results allow the use of international tourist arrivals data to Taiwan to estimate alternative univariate long memory conditional mean and conditional volatility models given in the next section.

4. Conditional Mean and Conditional Volatility Models

The alternative time series models to be estimated for the conditional means of the daily international tourist arrivals, as well as their conditional volatilities, are discussed below. Daily international tourist arrivals to Taiwan show periods of high volatility followed by others of relatively low volatility. One implication of this persistent volatility behaviour is that the assumption of (conditionally) homoskedastic residuals is inappropriate.

For a wide range of financial and tourism 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.

In the selected conditional volatility model, the residual series should follow a white noise process. Li et al. (2002) provide an extensive review of recent theoretical results for univariate and multivariate time series models with conditional volatility errors, and McAleer (2005) reviews a wide range of univariate and multivariate, conditional and stochastic, models of financial volatility. When international tourist arrivals data display persistence in volatility, it is natural to estimate alternative conditional volatility models.

The conditional volatility literature has been discussed extensively in recent years (see, for example, Li, Ling and McAleer (2002), McAleer (2005), and McAleer, Chan and Marinova (2007). Consider the stationary AR(1)-GARCH(1,1) model for daily international tourist arrivals to Taiwan (or their growth rates, as appropriate), y_t :

$$y_t = \phi_1 + \phi_2 y_{t-1} + \varepsilon_t , \qquad |\phi_2| < 1 \tag{1}$$

for $t = 1,...,n$, where the shocks (or movements in daily international tourist arrivals) are given by:

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

and $\omega > 0, \alpha \geq 0, \beta \geq 0$ 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 (2003). 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, $\alpha + \beta$). In equations (1) and (2), the parameters are typically estimated by the maximum likelihood method to obtain Quasi-Maximum Likelihood Estimators (QMLE) in the absence of normality of η_t , the conditional shocks (or standardized residuals).

Since the GARCH process in equation (2) is a function of the unconditional shocks, the moments of ε_t need to be investigated. Ling and McAleer (2003) showed that the QMLE for GARCH(*p*,*q*) is consistent if the second moment of ε_t 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 ε_t 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 ε_t for GARCH(1,1) is $\alpha + \beta < 1$.

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

$$h_t = \omega + (\alpha + \gamma I(\eta_{t-1}))\varepsilon_{t-1}^2 + \beta h_{t-1}, \tag{3}$$

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

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

as η_t has the same sign as ε_t. 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. For financial data, it is expected that $\gamma \geq 0$ because negative shocks increase risk by increasing the debt to equity ratio, but this interpretation need not hold for international tourism arrivals data in the absence of a direct risk interpretation. The asymmetric effect, γ, measures the contribution of shocks to both short run persistence, $\alpha + \frac{\gamma}{2}$, and to long run persistence, $\alpha + \beta + \frac{\gamma}{2}$. It is not possible for leverage to be present in the GJR model, whereby negative shocks increase volatility and positive shocks of equal magnitude decrease volatility.

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

$$\log h_t = \omega + \alpha \,|\, \eta_{t-1} \,|\, + \gamma\eta_{t-1} + \beta \log h_{t-1}, \quad |\, \beta \,| < 1 \tag{4}$$

where the parameters α, β and γ have different interpretations from those in the GARCH(1,1) and GJR(1,1) models. If γ = 0, there is no asymmetry, while γ < 0, and $\gamma < \alpha < -\gamma$ are the conditions for leverage to exist, whereby negative shocks increase volatility and positive shocks of equal magnitude decrease volatility.

5. Estimated Models and Discussion

The Heterogenous Autoregressive (HAR) model was proposed by Corsi (2004) as an alternative to model and forecast realized volatilities, and is inspired by the Heterogenous Market Hypothesis of Muller, Dacorogna, Dav, Olsen, Pictet, and Ward (1993) and the asymmetric propagation of volatility between long and short horizons. Corsi (2004) showed that the actions of different types of market participants could lead to a simple restricted linear autoregressive model with the feature of considering volatilities realized over different time horizons. The heterogeneity of the model derives from the fact that different autoregressive structures are present at each time scale (for further details, see McAleer and Medeiros (2008)). In this section the HAR model is used to model total international tourist arrivals to Taiwan, together with the three conditional volatility models discussed in the previous section.

The alternative HAR(*h*) models to be estimated to capture long memory are based on the following:

$$y_{t,h} = \frac{y_t + y_{t-1} + y_{t-2} + ... + y_{t-h+1}}{h} \tag{5}$$

where typical values of *h* are one (daily data), seven (weekly data), and 28 (monthly data). In the empirical application, the three models to be estimated for international tourist arrivals to Taiwan are as follows:

$$y_t = \phi_1 + \phi_2 y_{t-1} + \varepsilon_t \tag{6}$$

$$y_t = \phi_1 + \phi_2 y_{t-1} + \phi_3 y_{t-1,7} + \varepsilon_t \tag{7}$$

$$y_t = \phi_1 + \phi_2 y_{t-1} + \phi_3 y_{t-1,7} + \phi_4 y_{t-1,28} + \varepsilon_t \,. \tag{8}$$

which will be referred to as the HAR(1), HAR(1,7) and HAR(1,7,28) models, respectively.

The conditional mean estimates in Tables 1-3 show that the HAR(1), HAR(1,7) and HAR(1,7,28) estimates are all statistically significant, such that the long memory properties of the data are captured adequately. As the second moment condition is less than unity in each case, and hence the weaker log-moment condition (which is not reported) is necessarily less than zero (see Tables 1-3), the regularity conditions are satisfied, and hence the QMLE are consistent and asymptotically normal, and inferences are valid. The EGARCH(1,1) model is based on the standardized residuals, so the regularity condition is satisfied if $|\, \beta \,| < 1$, and hence the QMLE are consistent and asymptotically normal (see, for example, McAleer et al. (2007)).

The GARCH(1,1) estimates in Tables 1-3 for the HAR(1), HAR(1,7) and HAR(1,7,28) models of international tourist arrivals to Taiwan suggest that the short run persistence of shocks lies between 0.254 and 0.285, while the long run persistence lies between 0.236 and 0.432. As the second moment condition, $\alpha + \beta < 1$, is satisfied, the log-moment condition is necessarily satisfied, so that the QMLE are consistent and asymptotically normal. Therefore, statistical inference using the asymptotic normal distribution is valid, and the symmetric GARCH(1,1) estimates are statistically significant.

If positive and negative shocks of a similar magnitude to international tourist arrivals to Taiwan are treated asymmetrically, this can be evaluated in the GJR(1,1) model. The asymmetry coefficient is found to be positive and significant for HAR(1), namely 0.317, which indicates that decreases in international tourist arrivals increase volatility. This is a similar empirical outcome as is found in virtually all cases in finance, where negative shocks (that is, financial losses) increase risk (or volatility). Thus, shocks to international tourist arrivals resemble financial shocks, and can be interpreted as risk associated with international tourist arrivals. Although asymmetry is observed for the HAR(1) model, there is no evidence of leverage. As the second moment condition, $\alpha+\beta+\frac{1}{2}\gamma<1$, is satisfied, the log-moment condition is necessarily satisfied, so that the QMLE are consistent and asymptotically normal. Therefore, statistical inference using the asymptotic normal distribution is valid, and the asymmetric GJR(1,1) estimates are statistically significant. However, in comparison with the HAR(1) model, the estimated asymmetry coefficients for GJR(1,1) are not statistically significant for the HAR(1,7) and HAR(1,7,28) models, so that their respective GARCH(1,1) counterparts are to be preferred. These empirical results show that the conditional volatility estimates are sensitive to the long memory nature of the conditional mean specifications.

The interpretation of the EGARCH model is in terms of the logarithm of volatility. For international tourist arrivals, each of the EGARCH(1,1) estimates is statistically significant for the HAR(1) model, with the size effect, α, being positive and the sign effect, γ, being negative. The coefficient of the lagged dependent variable, β, is estimated to be 0.122, which suggests that the statistical properties of the QMLE for EGARCH(1,1) will be consistent and asymptotically normal. As in the case of the GJR(1,1) model, the estimated asymmetry coefficients for EGARCH(1,1) are not statistically significant for the HAR(1,7) and HAR(1,7,28) models. These empirical results show that the volatility in the shocks to international tourist arrivals to Taiwan are sensitive to the long memory nature of the conditional mean specifications.

Acknowledgements

For financial support, the first author wishes to thank the National Science Council, Taiwan, and the second author wishes to thank the Australian Research Council.

References

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

Corsi, F. (2004), A simple long memory model of realized volatility," Unpublished paper, University of Southern Switzerland.

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.

Li, W.K., S. Ling and M. McAleer (2002), Recent theoretical results for time series models with GARCH errors, *Journal of Economic Surveys*, 16, 245-269. Reprinted in M. McAleer and L. Oxley (eds.), *Contributions to Financial Econometrics: Theoretical and Practical Issues*, Blackwell, Oxford, 2002, pp. 9-33.

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), Stationarity and the existence of moments of a family of GARCH processes, *Journal of Econometrics*, 106, 109-117.

Ling, S. and M. McAleer (2002b), 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, (2003), Asymptotic theory for a vector ARMA-GARCH model, *Econometric Theory*, 19, 278-308.

McAleer, M. (2005), Automated inference and learning in modeling financial volatility, *Econometric Theory*, 21, 232-261.

McAleer, M. (2009), The Ten Commandments for optimizing value-at-risk and daily capital charges, to appear in *Journal of Economic Surveys*.

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.

McAleer, M. and M. Medeiros (2008), A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries, Journal of Econometrics, 147(1), 2008, 104-119.

Muller, U., M. Dacorogna, R. Dav, R. Olsen, O. Pictet and J. ward (1993), Fractals and intrinsic time - a challenge to econometricians," in Proceedings of the XXXIXth International AEA Conference on Real Time Econometrics.

Nelson, D.B. (1991), Conditional heteroscedasticity in asset returns: a new approach, *Econometrica*, 59, 347-370.

Table 1: Estimated Conditional Mean (HAR(1)) and Conditional Volatility Models

Parameters	GARCH	GJR	EGARCH
ϕ_1	1115** (48.85)	1020** (47.22)	1004** (46.97)
ϕ_2	0.806** (0.007)	0.816** (0.007)	0.817** (0.007)
ω	868407** (24864)	807223** (25610)	11.81** (0.524)
GARCH/GJR α	0.254** (0.015)	0.155** (0.010)	--
GARCH/GJR β	-0.018 (0.015)	0.011 (0.018)	--
GJR γ	--	0.317** (0.043)	--
EGARCH α	--	--	0.483** (0.021)
EGARCH γ	--	--	-0.128** (0.016)
EGARCH β	--	--	0.122** (0.037)
Second moment	0.236	0.324	-
AIC	16.716	16.709	16.706
BIC	16.722	16.715	16.713
Jarque-Bera	690.73	814.82	782.94
[p-value]	[0.000]	[0.000]	[0.000]

Notes:
The dependent variable, TA, is international tourist arrivals to Taiwan.
Numbers in parentheses are standard errors.
The log-moment condition is necessarily satisfied as the second moment condition is satisfied.
AIC and BIC denote the Akaike Information Criterion and Schwarz Bayesian Information Criterion, respectively.
** denotes the estimated coefficient is statistically significant at 1%.

Table 2: Conditional Mean (HAR(1,7)) and Conditional Volatility Models

Parameters	GARCH	GJR	EGARCH
ϕ_1	311.34** (51.00)	311.52** (51.19)	294.32** (49.58)
ϕ_2	0.299** (0.014)	0.299** (0.014)	0.320** (0.013)
ϕ_3	0.642** (0.015)	0.642** (0.015)	0.625** (0.015)
ω	526553** (20618)	526310** (21106)	9.563** (0.430)
GARCH/GJR α	0.285** (0.015)	0.285** (0.017)	--
GARCH/GJR β	0.147** (0.022)	0.147** (0.022)	--
GJR γ	--	-0.001 (0.031)	--
EGARCH α	--	--	0.501** (0.022)
EGARCH γ	--	--	-00007 (0.015)
EGARCH β	--	--	0.271** (0.031)
Second moment	0.432	0.432	-
AIC	16.491	16.491	16.493
BIC	16.497	16.499	16.500
Jarque-Bera	914.70	913.55	889.92
[p-value]	[0.000]	[0.000]	[0.000]

Table 3: Conditional Mean (HAR(1,7,28)) and Conditional Volatility Models

Parameters	GARCH	GJR	EGARCH
ϕ_1	167.28** (54.26)	166.58** (54.59)	144.40** (52.93)
ϕ_2	0.298** (0.014)	0.299** (0.014)	0.317** (0.013)
ϕ_3	0.460** (0.021)	0.459** (0.021)	0.445** (0.020)
ϕ_4	0.208** (0.019)	0.208** (0.019)	0.208** (0.018)
ω	532729** (19854)	533665** (20228)	10.032** (0.439)
GARCH/GJR α	0.285** (0.015)	0.283** (0.017)	--
GARCH/GJR β	0.131** (0.021)	0.130** (0.021)	--
GJR γ	--	0.006 (0.031)	--
EGARCH α	--	--	0.501** (0.021)
EGARCH γ	--	--	-0.010 (0.015)
EGARCH β	--	--	0.236** (0.031)
Second moment	0.416	0.416	-
AIC	16.478	16.478	16.480
BIC	16.485	16.487	16.488
Jarque-Bera	1020.8	1026.4	1036.8
[p-value]	[0.000]	[0.000]	[0.000]