Tourism forecasting using conditional volatility models

ABSTRACT

Conditional volatility models are used in tourism demand studies to model the effects of shocks on demand volatility, which arise from changes in political, social or economic conditions. Seasonal ARIMA models have been widely employed for forecasting purpose but little attention has been given to examining the forecast accuracy of conditional volatility models. This study investigates whether the conditional volatility models can outperform seasonal ARIMA models in predicting tourist arrivals to Australia. One key result is that seasonality exists in the volatility of tourist arrivals to Australia. Hence, this paper adds a new dimension in the literature of modelling seasonality in tourism demand, by incorporating seasonal effects into conditional volatility models. Using data on tourist arrivals from USA, UK, Japan and New Zealand to Australia, this study found that conditional volatility models outperform seasonal ARIMA models in forecasting for all countries except UK.

Keywords: forecasting; tourist arrivals to Australia; seasonal ARIMA; conditional volatility models

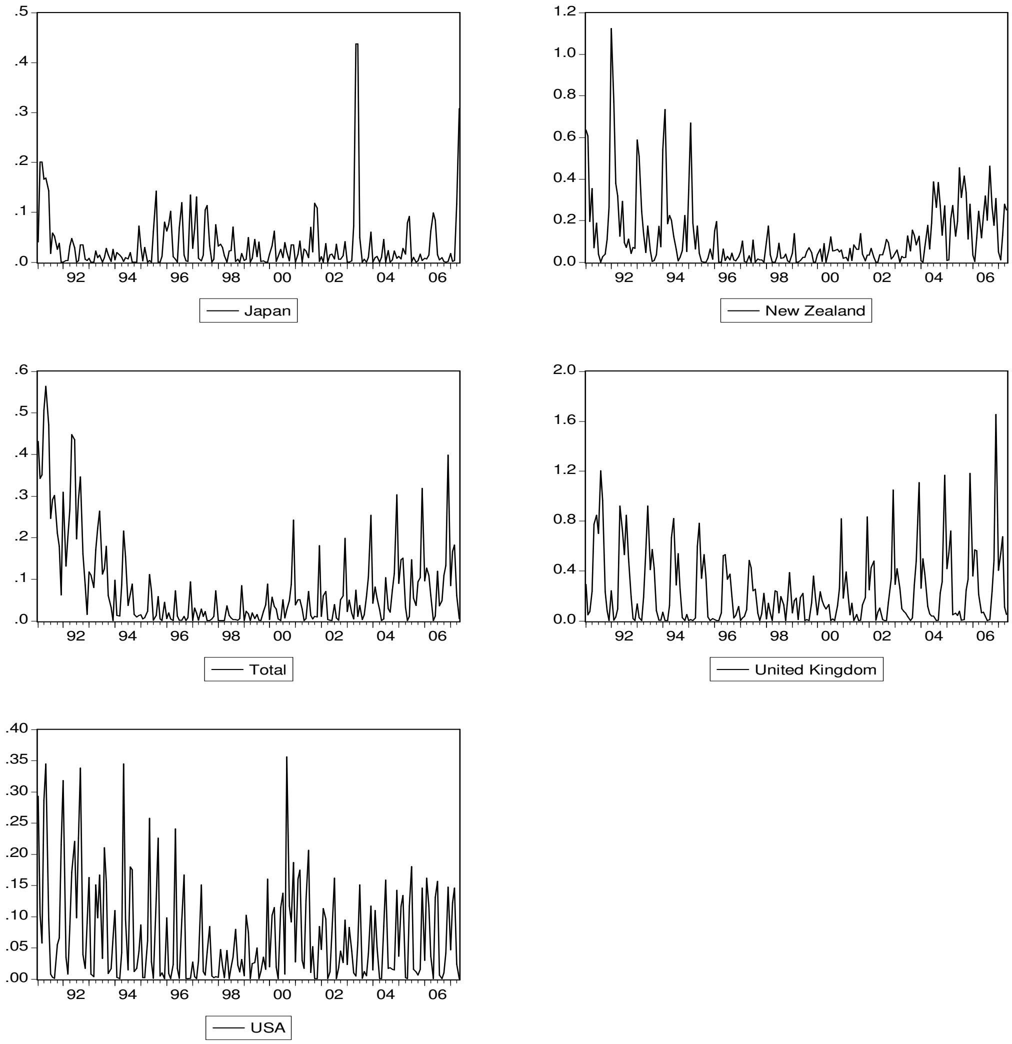
INTRODUCTION

Several empirical studies have found that tourism demand data exhibit volatility (Shareef and McAleer, 2005) and specifically, a negative shock can have more impact on the volatility in Japanese tourist arrivals to Australia than a positive shock (Chan et al., 2005). Furthermore, according to Kim and Wong (2006), the volatility in tourism demand data can be influenced by the effects of news shocks such as economic crises, outbreak of deadly diseases, natural disasters and war. Therefore, the tourism literature concludes that modelling the volatility in tourism demand is important because it can capture the occurrence of unexpected events.

Conventionally, volatility of tourism demand is modelled using conditional volatility models. The models that appeared in tourism literature are univariate generalised autoregressive conditional heteroscedasticity (GARCH), univariate asymmetric GARCH (or GJR), vector autoregressive moving average GARCH (VARMA-GARCH) and VARMA-asymmetry GARCH (VARMA-AGARCH) models (Chan et al., 2005; Kim and Wong, 2006; Shareef and McAleer, 2005; Shareef and McAleer, 2007).

Despite that modelling demand volatility has emerged in the literature, there are two areas which still require attention. First, seasonality exists in tourism demand volatility data. Figure 1 exhibits the volatility of tourist arrivals to Australia and find that seasonal patterns exist and vary across countries of origin. For instance, from 1991 to 2007, the highest spikes in tourist arrivals from New Zealand and United Kingdom occurred in the months of December and March, respectively. Hence, the figure suggests that incorporating seasonal dummies in the conditional volatility models are necessary to capture these effects. Second, there is a dearth of research assessing the forecast accuracy of conditional volatility models. As predicting tourism demand is very important in business planning, it is imperative to investigate whether conditional volatility models can outperform other competing models in forecasting tourism demand data.

Figure 1
Volatility of the log tourist arrivals to Australia for total and four countries (1991 to 2007)



The purposes of this paper are as follows. First, as seasonality exists in tourism demand data, this paper attempts to modify the traditional conditional volatility models by incorporating seasonal effects into the models. Thereafter, the models will be employed for the purpose of out-sample forecasting. Second, as seasonal ARIMA models have been widely employed for forecasting tourism demand to Australia (Kim, 1999; Kim and Moosa, 2001; Kulendran and Wong, 2005; and Lim and McAleer, 2002), this study investigates whether conditional volatility models can outperform seasonal ARIMA models in terms of forecasting tourist arrivals to Australia. The conditional volatility models employed in this research are univariate GARCH and GJR models.

The data employed are the logarithm of the monthly short-term tourist arrivals from Japan, New Zealand, United Kingdom, USA and all source countries to Australia from January 1991 to May 2006. The logarithm data are used because, based on Chan et al. (2005), data on logarithms of tourist arrivals to Australia are integrated to zero after taking its first difference. For forecasting purpose, the data between June 2006 and May 2007 are used.

METHODOLOGY

In a tourism demand data series, the conditional variance may not be constant. To tackle the problem of heteroscedastic conditional variances, Engle (1982) developed a volatility model which incorporated all past errors. Bollerslev (1986) further modified Engle's idea by including lagged conditional volatility into the volatility function.

Given the univariate conditional mean,

$$\begin{aligned} y_t &= E(y_t / F_{t-1}) + \varepsilon_t, \\ \varepsilon_t &= \sqrt{h_t}\eta_t, \end{aligned} \tag{1}$$

where $t = 1,\ldots,n$ and $\eta_t \sim iid(0,1)$, y_t = series of returns, F_{t-1} = past information available to time t and ε_t = error term with stochastic process, the univariate Generalised Autoregressive Conditional Heteroscedesticity (GARCH) model is given as follows:

$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} \tag{2}$$

where h_t = conditional variance, and $h_t > 0$ when $\omega > 0$, $\alpha > 0$ and $\beta > 0$. Note that α and $\alpha+\beta$ represent the short term and long term persistence of shocks to returns. ω = constant variance.

The simpler model that captures the asymmetric impacts of good and bad news is the GJR(1,1) model developed by Glosten et al. (1993) as follows:

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

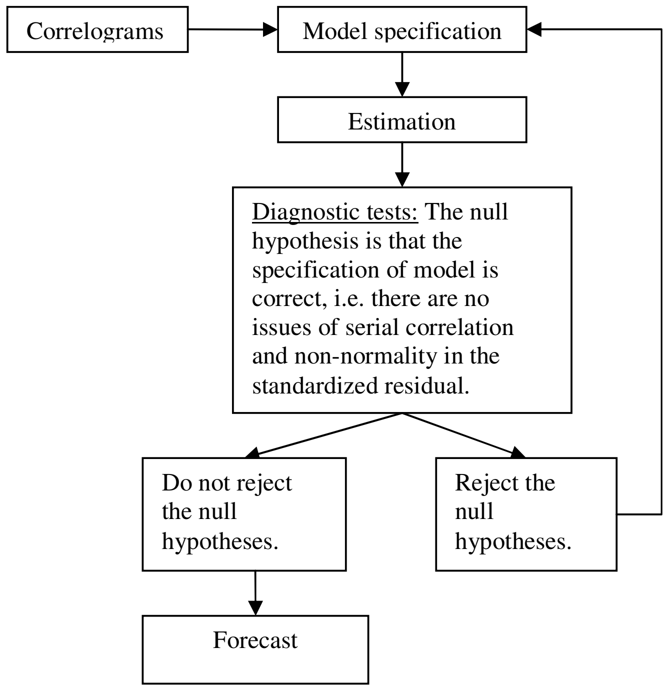
where $I(\eta_t) = \begin{cases} 1, \varepsilon < 0 \\ 0, \varepsilon \geq 0 \end{cases}$, $h_t > 0$ when $\omega > 0, \alpha \geq 0, \alpha + \gamma \geq 0$ and $\beta \geq 0$.

In terms of the persistence of shocks, the impact of short run positive and negative shocks is α and $\alpha + \gamma$ respectively. When η_t follows a symmetric distribution, the short-run persistence of shocks is $\alpha + 0.5\gamma$ and long run persistence of shocks is $\alpha + 0.5\gamma + \beta$.

To incorporate seasonal effects into the GARCH and GJR models, this paper attempts to include seasonal dummies in Equation (2) and (3).

Figure 2 illustrates the methodology of specification of conditional volatility models.

Figure 2
Flow chart of methodology



EMPIRICAL RESULTS

Estimates of GARCH and GJR models

In Tables 1 and 2, AR(1) coefficients in the conditional mean of GARCH and GJR models are statistically significant for Japan, New Zealand and Total, indicating a high persistence of tourist visitations to Australia. However, for moving average components, only the data on tourist arrival from New Zealand show highly significant. This implies that unexpected shocks in the previous period can influence the arrivals of New Zealand tourists in the current period. Furthermore, Tables 1 and 2 also reveal that tourist arrivals to Australia are highly seasonal.

In terms of the conditional volatility estimates for the GARCH model (Table 3), the ARCH effect (or the short-run shock persistence) is positive and only significant for New Zealand. This shows that a positive shock will increase the variation of tourist arrivals from New Zealand and vice-versa for negative shocks. Conversely, the β (or GARCH effects) are significant for all countries, except Japan. The signs of β are positive, as anticipated, for UK, USA and Total, but not for New Zealand. This study also discovers that seasonal effects exist in the conditional volatility of Japan, New Zealand, UK and USA (Table 3), which indicating that the variations in tourist arrivals from these source countries to Australia can be caused by seasonality.

The conditional volatility estimates for GJR model in Table 4 reveal that α (or ARCH effects) for Japan and Total are significant but the signs are not consistent with expectations. For the GARCH effects, the results for USA and Total are significant. However, the sign of β for USA is negative, which does not satisfy the condition of positivity of conditional variance. In addition, the threshold effects for New Zealand and Total are significant at 5% and 1%, respectively, and the impacts of a negative shock on the variations of both sets of data are similar. This implies that a negative shock will result in less fluctuation in Total and New Zealand tourist arrivals to Australia. This study could not generate reliable estimates for seasonal dummies in the conditional volatility of GJR model, which requires further investigation.

In this research, the estimation of GARCH and GJR models has undergone rigorous diagnostic tests to ensure there are no issues of serial correlation and non-normality in the estimated standardized residuals (Table 5). Given this fact, the conditional mean and variance estimates in Table 1 to 4 are considered to be robust.

Forecast accuracy of conditional volatility models

By comparing forecast accuracy between conditional volatility and ARIMA models, Table 6 shows that the forecast errors of GARCH and GJR models for Japan, UK and Total are lower than for the ARIMA model. This implies that conditional volatility models perform better in forecasting for these data series. For New Zealand data, ARIMA model outperforms GARCH models but under-performs GJR models. For UK data, both GARCH and GJR models generate less accurate forecasts than ARIMA models.

Forecast errors of GARCH models for UK and Total are lower than GJR models (Table 6). This outcome implies that GARCH models can predict better than GJR models for these three data series. Conversely, for Japan, New Zealand and USA data, GJR models provide more accurate forecast than GARCH models.

CONCLUSION

For the first time, this study demonstrates that seasonality exists in the volatility of tourist arrivals to Australia. To capture the seasonal effects, seasonal dummy variables were included in the conditional volatility models. Furthermore, given that seasonal ARIMA models have been widely employed in forecasting tourist arrivals to Australia, this paper intends to investigate whether conditional volatility models provide better forecasts than seasonal ARIMA models.

The results showed that seasonal dummy variables in the conditional mean of GARCH and GJR models are statistically significant for all source countries. Furthermore, the empirical results of GARCH models revealed that seasonal dummy variables were significant in the conditional variances for GARCH model for Japan, New Zealand, United Kingdom and USA data, leading to the conclusion that seasonality exists in the volatility of tourist arrivals to Australia.

In terms of comparing forecast performance of conditional volatility and ARIMA models, this study found that the former model forecast better for all source countries except UK. Furthermore, by evaluating forecast accuracy between GARCH and GJR models, the empirical results showed that GJR can generate relatively more accurate forecasts for Japan, New Zealand and USA data.

Overall, this paper concludes that conditional volatility models can outperform seasonal ARIMA models in predicting tourist arrivals to Australia. Nevertheless, future research should compare the forecast accuracy of conditional volatility models with econometric models.

Table 1
Estimation of conditional mean for GARCH model

Data	Constant	AR(1)	MA(1)	SMA(12)	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
Japan	**10.9681**	**0.7128**					**0.1089**				**0.2053**	**0.2797**				**0.102**
	0.0258	0.0382					0.0167				0.002	0.0209				0.0209
NZ	**11.1021**	**0.7747**	**-0.0419**	**0.8297**					**-0.1139**	**-0.1092**			**0.0274**		**-0.1374**	
	0.0651	0.043	0.081	0.0326					0.0319	0.0432			0.0364		0.0294	
UK	**0.6627**							**-0.7592**	**-1.4638**	**-1.7104**	**-1.0253**	**-1.3508**	**-1.1854**	**-0.5088**		**0.4481**
	0.0265							0.0352	0.0342	0.0469	0.0403	0.0367	0.0405	0.0484		0.0453
USA	**10.3865**				**-0.1689**				**-0.1479**	**0.0801**	**0.1129**		**-0.1044**			
	0.0557				0.0162				0.0313	0.0403	0.0398		0.0369			
Total	**12.9948**	**0.9812**	**-0.5983**			**0.1028**	**0.0747**	**-0.0442**	**-0.2038**	**-0.1368**	**0.0758**		**-0.0783**		**0.0781**	**0.2770**
	0.1167	0.0062	0.0761			0.0141	0.0123	0.0139	0.0135	0.0118	0.0121		0.0138		0.0117	0.0105

Note: The two entries corresponding to each variable are their estimates (in bold) and their Bollerslev and Wooldridge (1992) robust standard errors, respectively.
AR(p) denotes autoregressive at lag p.
MA(q) denotes moving average at lag q.
SMA(Q) denotes seasonal moving average at lag Q.
S1 to S12 signify seasonal dummies from January to December.
The AR, MA, SMA and seasonal dummies terms above are statistically significant at 5%.
In the interests of presentation, those insignificant regressors are not reported here.

Table 2
Estimation of conditional mean for GJR model

Data	Constant	AR(1)	AR(2)	MA(1)	SMA(12)	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
Japan	**10.9206**	**0.7452**						**0.1298**				**0.2017**	**0.2460**				**0.1149**
	0.0279	0.0365						0.0144				0.02	0.0203				0.0087
NZ	**11.2708**	**0.9806**		**-0.5892**	**0.6422**	**-0.3861**	**-0.4488**	**-0.1534**		**-0.1499**	**-0.0763**			**0.0608**		**-0.1157**	
	0.2383	0.0134		0.0615	0.0515	0.0297	0.0281	0.0373		0.032	0.034			0.0255		0.0321	
UK	**0.6634**								**-0.7522**	**-1.4594**	**-1.7065**	**-1.0066**	**-1.3557**	**-1.2038**	**-0.5364**		**0.4348**
	0.0251								0.0359	0.0342	0.0481	0.0368	0.0355	0.0405	0.0454		0.0453
USA	**10.3884**	**-0.1249**	**0.8082**	**0.9957**		**-0.1873**				**-0.1776**		**0.1705**		**-0.1761**			
	0.0548	0.0445	0.0328	0.0237		0.0221				0.0134		0.0218		0.0218			
Total	**12.8855**	**0.9803**		**-0.5114**			**0.0967**	**0.0744**	**-0.0448**	**-0.2090**	**-0.1346**	**0.0772**		**-0.0715**		**0.0820**	**0.2803**
	0.0935	0.0032		0.0639			0.0124	0.0106	0.0131	0.0125	0.0111	0.0114		0.0121		0.0093	0.0088

Note: The two entries corresponding to each variable are their estimates (in bold) and their Bollerslev and Wooldridge (1992) robust standard errors, respectively.
AR(p) denotes autoregressive at lag p.
MA(q) denotes moving average at lag q.
SMA(Q) denotes seasonal moving average at lag Q.
S1 to S12 signify seasonal dummies from January to December.
The AR, MA, SMA and seasonal dummies terms above are statistically significant at 5%.
In the interests of presentation, those insignificant regressors are not reported here.

Table 3
Estimation of conditional volatility for GARCH model

Data	ω	α	β	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
Japan	**0.0065**	**0.1405**	**0.2332**	**-0.0247***	**-0.0120***	**-0.0228***	**-0.0055***		**-0.0342***	**-0.0220***	**-0.0208***	**-0.0106***	**-0.0144***	**-0.0167***	**-0.0119***
	0.0042	0.1852	0.4589	0.0159	0.0158	0.0157	0.0175		0.0276	0.0155	0.0152	0.0151	0.0157	0.0159	0.0163
NZ	**0.0084***	**0.2763***	**-0.3518**$^{+}$		**-0.0612***	**-0.0460***	**-0.0563***	**-0.0678***	**-0.0484***	**-0.0626***	**-0.0664***	**-0.0613***	**-0.0628***	**-0.0666***	**-0.0602***
	0.0017	0.0767	0.146		0.0132	0.0134	0.0127	0.0119	0.0137	0.0128	0.0120	0.0124	0.0123	0.0120	0.0125
UK	**0.0034***	**0.0337**	**0.9232***				**-0.029***								
	0.0011	0.0188	0.05				0.0047								
USA	**0.0287***	**0.0366**	**0.3319**$^{+}$	**-0.0390***		**-0.0355***									
	0.0057	0.0514	0.1564	0.0049		0.004									
Total	**0.0006**	**0.0832**	**0.6913**$^{+}$												
	0.0007	0.0824	0.3119												

Note: ω = constant variance, α = ARCH effect, β = GARCH effect. The two entries corresponding to each variable are their estimates (in bold) and their Bollerslev and Wooldridge (1992) robust standard errors, respectively. S1, S3 and S4 are seasonal dummies for January, March and April, respectively.

*denotes significant at 1%.

\+ denotes significant at 5%.

In the interests of presentation, some of the insignificant seasonal dummy variables are not reported here.

Table 4
Estimation of conditional volatility for GJR model

Data	ω	α	γ	β
Japan	**0.0058**	**-0.1089***	**0.2852**	**0.3593**
	0.0031	0.0189	0.2233	0.3658
NZ	**0.0074***	**-0.0104**	**0.4018+**	**-0.1403**
	0.0014	0.0443	0.1669	0.0973
UK	**0.0087**	**0.2976**	**-0.2840**	**0.4876**
	0.0062	0.1545	0.1469	0.3276
USA	**0.0355***	**0.0409**	**-0.0631**	**-0.9958***
	0.0037	0.0442	0.0538	0.0288
Total	**0.0002+**	**-0.1075***	**0.2573***	**0.9108***
	0.0001	0.0294	0.0816	0.0562

Note: ω = constant variance, α = ARCH effect, γ = threshold effect, β = GARCH effect. The two entries corresponding to each variable are their estimates (in bold) and their Bollerslev and Wooldridge (1992) robust standard errors, respectively.
* denotes significant at 1%.
\+ denotes significant at 5%.

Table 5
Diagnostic tests of GARCH and GJR models

		Diagnostic tests on standardised residuals	
Data	**Models**	**Ho: No serial correlation[a]**	**Ho: Normality[b]**
Japan	GARCH	1.5844 [0.208]	1.1684 [0.5576]
	GJR	2.3054 [0.129]	1.0414 [0.5941]
NZ	GARCH	5.5763 [0.018]	2.6336 [0.268]
	GJR	4.7289 [0.03]	0.1829 [0.9126]
UK	GARCH	1.2448 [0.265]	3.7821 [0.1509]
	GJR	0.3926 [0.531]	2.6552 [0.2651]
USA	GARCH	1.8159 [0.178]	1.1445 [0.5642]
	GJR	2.1282 [0.145]	7.6344 [0.022]
Total	GARCH	3.0007 [0.083]	0.2166 [0.8974]
	GJR	2.1198 [0.145]	1.0073 [0.6043]

Note: (a) The test statistics are obtained from Q-statistics of collegram of standardized residuals.
(b) The test statistics are based on Jarque-Bera of normality tests.
Figures in brackets are p-value.

Table 6

Summary of forecast errors for ARIMA and conditional volatility models

		Models		
Data	**Forecast error**	**ARIMA**	**GARCH**	**GJR**
Japan	RMSE	1.4716	0.1724	0.1603
	MAE	1.2446	0.1298	0.1228
	MAPE	11.3687	1.1947	1.1263
	Theil coefficient	0.0634	0.0078	0.0073
NZ	RMSE	0.32	0.3474	0.1727
	MAE	0.2821	0.2725	0.1379
	MAPE	2.5636	2.5284	1.268
	Theil coefficient	0.0148	0.0158	0.0079
UK	RMSE	0.5494	1.1127	1.1486
	MAE	0.4507	0.7565	0.775
	MAPE	4.1948	7.1065	7.2719
	Theil coefficient	0.0266	0.0528	0.0545
USA	RMSE	3.0687	0.244	0.2388
	MAE	2.6824	0.2012	0.1976
	MAPE	25.7913	1.9662	1.9279
	Theil coefficient	0.131	0.0118	0.0116
Total	RMSE	0.1729	0.0695	0.0846
	MAE	0.1466	0.0549	0.0682
	MAPE	1.14	0.4326	0.5328
	Theil coefficient	0.0068	0.0027	0.0033

ACKNOWLEDGEMENTS

The author is grateful to two anonymous reviewers for their comments and STCRC for their financial support to attend CAUTHE 2008 conference in Gold Coast, Australia.

REFERENCES

Bollerslev, T., (1986) 'Generalized Autoregressive Conditional Heteroskedasticity', *Journal of Econometrics,* 31 (3), 307–327.

Bollerslev, T., & Wooldridge, J. M., (1992) 'Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariance' *Econometric Review*, 11 (2), 143-173.

Chan, F., Lim, C., & McAleer, M., (2005) 'Modelling multivariate international tourism demand and volatility' *Tourism Management*, 26 (3), 459-471.

Glosten, L.R., Jagannathan, R., & Runkle, D., (1993) 'On the Relation between the Expected Value and the Volatility of the Normal Excess Return on Stocks' *Journal of Finance*, 48 (5), 1779–1801.

Kim, J. H., & Moosa, I., (2001) 'Seasonal behaviour of monthly international tourist flows: specification and implications for forecasting models' *Tourism Economics*, 7 (4), 381-396.

Kim, S. S., & Wong, K. K. F., (2006) 'Effects of news shocks on inbound tourist demand volatility in Korea' *Journal of Travel Research*, 44 (4), 457-466.

Kulendran, N., & King, M. L., (1997) 'Forecasting international quarterly tourist flows using error-correction and time-series models' *International Journal of Forecasting*, 13 (3), 319-327.

Kulendran, N., & Wong, K. K. F., (2005) 'Modelling seasonality in tourism forecasting' *Journal of Travel Research*, 44 (2), 163-170.

Lim, C., & McAleer, M., (2002) 'Time series forecasts of international travel demand for Australia' *Tourism Management*, 23 (4), 389-396.

Engle, R. F., (1982) 'Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation' *Econometrica*, 50 (4), 987-1008.

Shareef, R., & McAleer, M., (2005) 'Modelling international tourism demand and volatility in small island tourism economics' *International Journal of Tourism Research*, 7 (6), 313-333.

Shareef, R., & McAleer, M., (2007) 'Modelling the uncertainty in monthly international tourist arrivals to the Maldives' *Tourism Management*, 28 (1), 23-45.