CIRJE-F-651

Modelling and Forecasting Daily International Mass Tourism to Peru

Jose Angelo Divino
Catholic University of Brasilia

Michael McAleer
Erasmus University Rotterdam
and Tinbergen Institute
and CIRJE, Faculty of Economics, University of Tokyo

August 2009

Modelling and Forecasting Daily International Mass Tourism to Peru

Jose Angelo Divino

Department of Economics
Catholic University of Brasilia

Michael McAleer

Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
and
Tinbergen Institute
The Netherlands
and
Center for International Research on the Japanese Economy (CIRJE)
Faculty of Economics
University of Tokyo

August 2009

The authors are most grateful to the Editor and three referees for helpful comments and suggestions. The second author wishes to thank the Australian Research Council and National Science Council, Taiwan, for financial support.

Abstract

Peru is a South American country that is divided into two parts by the Andes Mountains. The rich historical, cultural and geographic diversity has led to the inclusion of ten Peruvian sites on UNESCO's World Heritage List. For the potentially negative impacts of mass tourism on the environment, and hence on future international tourism demand, to be managed appropriately require modelling growth rates and volatility adequately. The paper models the growth rate and volatility (or the variability in the growth rate) in daily international tourist arrivals to Peru from 1997 to 2007. The empirical results show that international tourist arrivals and their growth rates are stationary, and that the estimated symmetric and asymmetric conditional volatility models all fit the data extremely well. Moreover, the estimates resemble those arising from financial time series data, with both short and long run persistence of shocks to the growth rate in international tourist arrivals.

Keywords: Daily International Tourim; Conditional Mean Models; Conditional Volatility Models.

JEL codes: C51; C53.

1. Introduction

Peru is a South American country, bordering Ecuador and Colombia to the north, Brazil to the east, Bolivia to the southeast, Chile to the south, and the Pacific Ocean to the west (see Figure 1). It is the 20th largest country in the world, with a territory of 1,285,220 km², and has a population of over 28 million (July 2007 estimate). The country is divided into two parts by the Andes Mountains, which cross the territory parallel to the Pacific Ocean. The east of the Andes up to the border with Brazil is covered by the Amazon rainforest, corresponding to about 60% of the country's territory. It is in the Peruvian territory, precisely at the mountain peak Nevado Mismi located in the Andes, that the Amazon River has its glacial source. The largest rivers in the country are integral parts of the upper Amazon Basin.

The country's climate is influenced by the proximity to the Equator, the presence of the Andes, and the cold waters from some Pacific currents. As a result of this combination, there is wide diversity in the climate, ranging from the dryness of the coast, to the extreme cold of the mountain peaks, and to heavy rainfall in the Amazon region.

Peru is one of the few areas in the world where there has been indigenous development of civilization, as the home of the Inca Empire, which emerged in the 15th Century as a powerful state and the largest empire in pre-Columbian America. The Inca Emperor was defeated in 1532 by the Spanish, who imposed colonial domination of the country. During this domination, silver mining with Indian forced labour became the basic economic activity, rendering considerable revenues for the Spanish Crown. However, the Royal income was reduced considerably over the years due to widespread smuggling and tax evasion. The Spanish Crown tried to recover control over its colonies by a series of tax reforms, which yielded numerous revolts across the continent. Finally, after successful military campaigns, Peru proclaimed independence from Spain in 1821.

Cultural diversity is one of the major attractions of Peru, and arises from a combination of different traditions over several centuries. Important contributions to its cultural diversity include native Indians, Spanish colonizers, and ethnic groups from Africa, Asia and Europe. There are also Pre-Inca and Inca cultures, with impressive achievements in architecture, such as the world famous holy city of Machu Picchu.

Peruvian cuisine is also linked to the diversity of the country, and uses many different ingredients that are combined through distinctive techniques. Climatic differences contribute to the success of the Peruvian cuisine by allowing the production and integration of a wide variety of flora and fauna in the country.

The music in Peru follows similar diversification, combining Andean, Spanish and African rhythms, instruments and expressions. In recent decades, a new ingredient given by the urbanization has influenced traditional Andean expressions and increased the musical variety.

As a result of such rich historical, cultural and geographic environments, Peru attracts short, medium and long haul tourists from all over the world to visit its territory. The major destination of the country is the region of Cuzco, accounting for about 27% of international tourist arrivals. In this region are located the city of Cuzco, which was the capital of the Inca Empire, the spiritual city of Machu Picchu, and the Sacred Valley of the Incas. Another important destination, which is visited by around 13% of international tourists, is the region of Lima, the capital of Peru, where the major attraction is the historical side of Lima. In third place, which receives around 11% of total international tourists, is the city of Arequipa, located in the valley of the volcanoes. Taking as a whole, these three regions receive around 51% of the international tourists to Peru.

UNESCO's World Heritage List includes properties that form part of the world's cultural and natural heritage with outstanding universal value. The City of Cuzco and the Historic Sanctuary of Machu Picchu were inscribed as World Heritage Sites in 1983, while the historical centre of the City of Arequipa was inscribed in 2000. There are presently seven other Peruvian sites in the World Heritage List. Owing to the destructive effects of unbridled mass tourism, the retention of Machu Picchu on the World Heritage List is a matter of great importance, even though Machu Picchu, was recently voted one of the New Seven Wonders of the World.

In 2006, Peru received about 908,000 tourists from around the world. The majority of international tourists to Peru come from North America, with around 37% of the total, and from Europe, with around 30%. In North America, the USA is the major source of tourists to Peru, accounting for about 31% of the total. In Europe, the major sources of international tourists are Spain, United Kingdom, France, Germany and Italy, with average proportions that range from 6% to 4% of the total of Europeans who visit Peru annually. South America accounts for around 22% of international tourist arrivals to Peru, with Argentina, Colombia, Chile and Brazil being the major sources, with each having shares of around 4% of the total.

International tourism has not yet achieved the status of an important economical activity in Peru. According to the Ministry of External Commerce and Tourism of Peru, the consumption of international tourism as a proportion of GDP increased from 0.5% in 1992 to 1.8% in 2005. However, after a significant increase in the late 1990's, the participation of international tourism in GDP decreased to 1.4% in 2004, and has been hovering at around 1.8% since 2003. This represents international tourism revenues of only $1.4 billion to the country on an annual basis. Consequently, there is clearly significant room for improvement in international tourism receipts. However, the potentially negative impacts of mass tourism on the environment, and hence on future international tourism demand, must be managed appropriately. In order to manage tourism growth and volatility, it is necessary to model the growth and volatility in international tourist arrivals adequately.

The primary purpose of this paper is to model the growth and volatility (that is, the variability in the growth rate) in international tourist arrivals to Peru. Information from 1997 to 2007 is used on daily international arrivals at the Jorge Chavez International Airport in Lima, which is the only international airport in Peru. By using daily data, we can approximate the modelling and management strategy and risk analysis to those applied to financial time series data. Although the volatility in international tourist arrivals has been analyzed at the monthly time series frequency in Chan, Lim and McAleer (2005), Divino and McAleer (2008), Hoti, McAleer and Shareef (2005, 2007), and Shareef and McAleer (2005, 2007, 2008), to the best of our knowledge there is no other work that models daily international tourist arrivals. The paper also contributes with the recent literature applying econometric techniques on forecasting tourism demand, where important references are Athanasopoulos et al. (2009), Bonham et al. (2009), and Gil-Alana et al. (2008).

From a time series perspective, there are several reasons for using daily data as compared with lower frequency data at the monthly or quarterly levels. Among other reasons, McAleer (2008) discusses how daily data can lead to a considerably higher sample size, provide useful information on risk in finance, lead to the determination of optimal environmental and tourism taxes, enable aggregation of high frequency data to yield aggregated data with volatility, analyze time series behavior at different frequencies through aggregated data, investigate whether time series properties have changed over time, capture day-of-the-week effects through differential pricing strategies in the tourism industry, including airlines, tourist attractions and the accommodation sector, and determine optimal tourism marketing policies through exploiting day-of-the-week effects to enable tourism operators to formulate pricing strategies and tourism packages to increase tourist arrivals in periods of low demand.

The empirical results show that the time series of international tourist arrivals and their growth rates are stationary. In addition, the estimated symmetric and asymmetric conditional volatility models, specifically the widely used GARCH, GJR and EGARCH models, all fit the data extremely well. In particular, the estimated models are able to account for the higher volatility persistence that is observed at the beginning and end of the sample period. The empirical second moment condition also supports the statistical adequacy of the models, so that statistical inference is valid. Moreover, the estimates resemble those arising from financial time series data, with both short and long run persistence of shocks in the growth rate of international tourist arrivals. Therefore, volatility can be interpreted as risk associated with the growth rate in international tourist arrivals[1].

The remainder of the paper is organized as follows. Section 2 presents the daily international tourist arrivals time series data set and discusses the time-varying volatility. Section 3 performs unit root tests on both the levels and logarithmic differences (or growth rates) of daily international tourist arrivals for Peru. Section 4 discusses alternative conditional mean and conditional volatility models for the daily international tourist arrivals series. The estimated models and empirical results are discussed in Section 5. Finally, some concluding remarks are given in Section 6.

2. Data

The data set comprises daily international tourist arrivals at the Jorge Chavez International Airport, the only international airport in Peru, which is located in the city of Lima, the capital of Peru. The data are daily, with seven days each week, for the period 1 January 1997 to 28 February 2007, giving a total of 3,711 observations. The source of the data was the Peruvian Ministry of International Trade and Tourism.

Figure 2 plots the daily international tourist arrivals, the logarithm of daily international tourist arrivals, and the first difference (that is, the log-difference or growth rates) of daily international tourist arrivals, as well as the volatility of the three variables, where volatility is defined as the squared deviation from the sample mean. There is higher volatility persistence at the beginning and at the end of the period for the series in levels and logarithms, but there is a single clear dominant observation in the series in around 2000. This extreme observation is 31 December 1999, which is higher than the typical decrease in international tourist arrivals in December each year. However, this observation is not sufficiently influential to affect the empirical results as there is no

[1] See McAleer and da Veiga (2008a, 2008b) for some applications of risk modelling and management to forecast value-at-risk (VaR) thresholds and daily capital changes.

significant change when this observation is deleted from estimation and testing. An increasing deterministic trend is present for the whole period in both series.

The series in log differences is clearly trend stationary and does not show higher volatility at the beginning or end of the sample, but there is clear volatility persistence. It is interesting that the single clear dominant observation in the logarithmic series is mirrored in the log difference series.

On an annual basis, the number of international tourist arrivals to Peru has shown an average growth rate of 8.8%, as illustrated in Figure 3. The lowest growth rate was observed in 2000, with an increase of just 0.8% over the previous year, while the highest growth rate occurred in 2005, when there was a significant increase of 27.1% over 2004. In the sample period as a whole, there was an increase of around 110% in international tourist arrivals to Peru, which would seem to indicate a reasonably good performance in the tourism sector over the decade. Nevertheless, annual average international tourist arrivals of 620,000 reveal that there is scope for a significant increase in international tourism to Peru. However, the potentially negative impacts of mass tourism on the environment, and hence on future international tourism demand, must be managed appropriately. In order to manage tourism growth and volatility, it is first necessary to model growth and volatility adequately.

In the next section we analyze the presence of a stochastic trend by applying unit root tests before modeling the time-varying volatility that is present in the logarithmic and log-difference (or growth rate) series.

3. Unit Root Tests

It is well known that traditional unit root tests, primarily those based on the classic methods of Dickey and Fuller (1979, 1981) and Phillips and Perron (1988), suffer from low power and size distortions. However, these shortcomings have been overcome by modifications to the testing procedures, such as the methods proposed by Perron and Ng

(1996), Elliott, Rothenberg and Stock (1996), and Ng and Perron (2001). It is worth mentioning, however, that the modified tests are also subject to low power and size distortions under short run persistence implied by GARCH components, as shown in Cook (2006). Nevertheless, size distortions might be even greater for the traditional Dickey-Fuller test, despite the sensitivity of the modified tests to the degree of volatility in the GARCH process.

We applied the modified unit root tests, given by $MADF^{GLS}$ and MPP^{GLS}, to the time series of daily international tourist arrivals in Peru. 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 Table 1. There is no evidence of a unit root in the logarithm of daily international tourist arrivals to Peru (LY) in the model with a constant and trend as the deterministic terms, so that LY is trend stationary. For the model with just a constant, however, the null hypothesis of a unit root is not rejected at the 5% significance level. For the series in log differences (or growth rates), the null hypothesis of a unit root is rejected for both specifications under the $MADF^{GLS}$ test.

These empirical results allow the use of both levels and log differences in international tourist arrivals to Peru to estimate the alternative univariate 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. As Figure 1 illustrates, daily international tourist arrivals, logarithm of daily international tourist arrivals, and the first difference (that is, the log difference or growth rate) of daily international tourist arrivals, to Peru show periods of high volatility followed by others of relatively low volatility. An obvious implication of this persistent volatility behaviour is that the assumption of (conditionally) homoskedastic residuals is not appropriate empirically.

It is well known that, for a wide range of financial 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 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 (logarithmic) international tourist arrivals data, as well as their growth rates, display persistence in volatility, as shown in Figure 1, it is natural to estimate alternative conditional volatility models. As mentioned previously, the GARCH(1,1) and GJR(1,1) conditional volatility models have been estimated using monthly international tourism arrivals data in Chan, Lim and McAleer (2005), Hoti, McAleer and Shareef (2005, 2007), Shareef and McAleer (2005, 2007, 2008), and Divino and McAleer (2008).

Consider the stationary AR(1)-GARCH(1,1) model for daily international tourist arrivals to Peru (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 $t \in Z$, as in Mikosch and Starica (2000), where the shocks (or movements in daily international tourist arrivals) are given by:

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

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 (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, $\alpha + \beta$). The stationary AR(1)-GARCH(1,1) model can be modified to incorporate a non-stationary ARMA(*p,q*) conditional mean and a stationary GARCH(*r,s*) conditional variance, as in Ling and McAleer (2003b).

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). The conditional log-likelihood function is given as follows:

$$\sum_{t=1}^{n} l_t = -\frac{1}{2} \sum_{t=1}^{n} \left(\log h_t + \frac{\varepsilon_t^2}{h_t} \right).$$

The QMLE is efficient only if η_t is normal, in which case it is the MLE[2]. When η_t is not normal, adaptive estimation can be used to obtain efficient estimators, although this can be computationally intensive. Ling and McAleer (2003b) investigated the properties of adaptive estimators for univariate non-stationary ARMA models with GARCH(*r,s*) errors. The extension to multivariate processes is complicated.

[2] See, for example, McAleer and da Veiga (2008a, b) for the use of alternative univariate and multivariate distributions for financial data.

As 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 (2003a) 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 (2003a) 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 $\alpha + \beta < 1$ and, under normality, the necessary and sufficient condition for the existence of the fourth moment is $(\alpha + \beta)^2 + 2\alpha^2 < 1$.

A sufficient condition for the QMLE of GARCH(1,1) to be consistent and asymptotically normal is given by the log-moment condition, namely

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

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 a univariate GARCH(p,q) process (see Lee and Hansen (1994) for the proof in the case of GARCH(1,1)), while Boussama (2000) showed that the log-moment condition was sufficient for asymptotic normality.

However, this condition is not easy to check in practice, even for the GARCH(1,1) model, as it involves the expectation of a function of a random variable and unknown parameters. Although the sufficient moment conditions for consistency and asymptotic normality of the QMLE for the univariate GARCH(1,1) model are stronger than their log-moment counterparts, the second moment condition is far more straightforward to check. In practice, the log-moment condition in equation (3) would be estimated by the sample mean, with the parameters α and β, and the standardized residual, η_t, being replaced by their QMLE counterparts.

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:

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

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

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

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, $\alpha + \beta + \frac{\gamma}{2}$.

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

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

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

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

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:

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

where the parameters α, β and γ have different interpretations from those in the GARCH(1,1) and GJR(1,1) models. Leverage, which is a special case of asymmetry, is defined as $\gamma < 0$ and $\alpha < |\, \gamma \,|$.

As noted in McAleer et al. (2007), there are some important differences between EGARCH and the previous two models, as follows: (i) EGARCH is a model of the logarithm of the conditional variance, which implies that no restrictions on the parameters are required to ensure $h_t > 0$; (ii) moment conditions are required for the GARCH and GJR models as they are dependent on lagged unconditional shocks, whereas EGARCH does not require moment conditions to be established as it depends on lagged conditional shocks (or standardized residuals); (iii) Shephard (1996) observed that $|\, \beta \,| < 1$ is likely to be a sufficient condition for consistency of QMLE for EGARCH(1,1); (iv) as the standardized residuals appear in equation (7), $|\, \beta \,| < 1$ would seem to be a sufficient condition for the existence of moments; and (v) in addition to being a sufficient condition for consistency, $|\beta|<1$ is also likely to be sufficient for asymptotic normality of the QMLE of EGARCH(1,1).

Furthermore, EGARCH captures asymmetries differently from GJR. The parameters α and γ in EGARCH(1,1) represent the magnitude (or size) and sign effects of the standardized residuals, respectively, on the conditional variance, whereas α and $\alpha+\gamma$ represent the effects of positive and negative shocks, respectively, on the conditional variance in GJR(1,1).

5. Estimated Models

The conditional mean model was estimated as AR(1), ARMA(1,1), ARMA(1,2), ARMA(2,1) and ARMA(2,2) processes, with AR(1) or ARMA(1,1) generally being empirically preferred on the basis of AIC and BIC (see Table 2).

The estimated conditional mean and conditional volatility models for the logarithm of tourist arrivals and the log-difference (or growth rate) of tourist arrivals are given in Table 3. The method used in estimation was the Marquardt algorithm. As shown in the unit root tests, the logarithmic and log difference (or growth rate) series are stationary. These empirical results are supported by the estimates of the lagged dependent variables in the estimates of equation (1), with the coefficients of the lagged dependent variable being significantly less than one in each of the estimated six models. Significant ARCH effects are detected by the LM test for ARCH(1) for LY, though not for DLY. The Jarque-Bera LM test of normality rejects the null hypothesis in all six cases.

As the second moment condition holds in each case, and hence the weaker log-moment condition (which is not reported) is necessarily less than zero (see Table 2), 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 at al. (2007)).

The GARCH(1,1) estimates for the logarithm of international tourist arrivals to Peru suggest that the short run persistence of shocks is 0.118 while the long run persistence is 0.921. 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 Peru are treated asymmetrically, this can be evaluated in the GJR(1,1) model. The asymmetry coefficient is found to be positive, namely 0.309, 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 tourist arrivals and the growth rate of tourist arrivals resemble financial shocks. They can be interpreted as risk associated to tourist arrivals. Moreover, the long run persistence of shocks is estimated to be 0.857. 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.

The interpretation of the EGARCH model is in terms of the logarithm of volatility. For the logarithm of international tourist arrivals, each of the EGARCH(1,1) estimates is statistically significant, with the size effect, α, being positive and the sign effect, γ, being negative. The conditions for leverage are satisfied for LY, but not for DLY. The coefficient of the lagged dependent variable, β, is estimated to be 0.763, which suggests that the statistical properties of the QMLE for EGARCH(1,1) will be consistent and asymptotically normal.

The GARCH(1,1) estimates for the log difference (or growth rate) of international tourist arrivals to Peru suggest that the short run persistence of shocks is 0.139 while the long run persistence is 0.891, which is very close to the corresponding estimates for the logarithm of international tourist arrivals. As the second moment condition is satisfied, the log-moment condition is necessarily satisfied, so that the QMLE are consistent and asymptotically normal, and hence the symmetric GARCH(1,1) estimates are statistically significant.

The GJR(1,1) estimates for the log difference (or growth rate) of international tourist arrivals to Peru suggest that the asymmetry coefficient is positive at 0.187, which indicates that decreases in the growth rate in international tourist arrivals increase volatility. The short run persistence of positive shocks is 0.025, the short run persistence of negative shocks is 0.212 (= 0.025 + 0.187), and the long run persistence of shocks is 0.898. As the second moment condition is satisfied, the log-moment condition is necessarily satisfied, so that the QMLE are consistent and asymptotically normal. Therefore, as in the case of asymmetry in financial markets, statistical inference using the asymptotic normal distribution is valid, and the asymmetric GJR(1,1) estimates are statistically significant.

For the log difference (or growth rate) of international tourist arrivals, each of the EGARCH(1,1) estimates is statistically significant, 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 0.913, which suggests that the statistical properties of the QMLE for EGARCH(1,1) will be consistent and asymptotically normal.

Overall, the QMLE for the GARCH(1,1), GJR(1,1) and EGARCH(1,1) models for both the logarithm and log difference of international tourist arrivals, are statistically adequate and have sensible interpretations.

The estimated conditional mean and conditional volatility models for the logarithm of annualized tourist arrivals and the log-difference (or growth rate) of annualized tourist arrivals are given in Table 4. The annualized series would appear to have a unit root, whereas the growth rate does not. Significant ARCH effects are detected for LYMA, though not for DLYMA. The Jarque-Bera LM test of normality rejects the null hypothesis in only two of six cases. The GARCH(1,1) model has short run persistence of shocks of 0.15 and long run persistence of shocks of 0.75. The GJR(1,1) model does not have significant asymmetry, so that GARCH(1,1) is preferred. The second moment condition is satisfied, so the QMLE are consistent and asymptotically normal, and the log-moment condition is necessarily satisfied. The EGARCH(1,1) estimates are significant, including the asymmetry coefficient, albeit marginally. However, the conditions for leverage are not satisfied for LYMA or DLYMA. Again, the QMLE are statistically adequate, so that inferences are sensible and statistically valid.

The correlation matrix of the forecasts in logarithmic levels and logarithmic first differences (or growth rates) are given in Tables 4 and 5. The forecasts in Table 4 can be very high at 0.999, but they can also be much lower between the annualized and original data series, as depicted in Figures 4 and 6, respectively. However, all of the correlations for the forecasts in log-differences are very high in Table 5, which is captured in the annualized international tourist forecasts in Figures 5 and 7. These results suggest that annualized figures are much easier to forecast and manage than are their daily counterparts.

The forecasts presented in Figures 4 to 7 are out-of-sample dynamic forecasts derived from each estimated model reported in Tables 3 and 4. The models were estimated using daily international tourist arrivals data to Peru from 1/1/1997 to 2/28/2007. Then out-of-sample daily forecasts are calculated for the period from 1/3/2007 until 2/28/2008. Thus, Figures 4 to 7 plot the actual series and the daily forecasts one year ahead.

It is worth noting that the high volatility of the daily series makes it somewhat difficult to predict the log-level and log-difference of international tourist arrivals to Peru. In both cases, as presented in Figures 4 and 6, respectively, the forecasts are roughly able to identify a trend in the data. On the other hand, for the annualized daily series plotted in

Figures 5 and 7, respectively, the models succeed in predicting the one-year ahead annualized series. Comparing the relative performance of the alternative models, there is no significant differences in the forecasts arising from the GARCH, GJR, and EGARCH models.

6. Concluding Remarks

The rich historical, cultural and geographic diversity that arises from a combination of different traditions over several centuries has led to the inclusion of ten Peruvian sites on UNESCO's World Heritage List of properties that form part of the world's cultural and natural heritage with outstanding universal value. These sites, particularly the City of Cuzco, the Historic Sanctuary of Machu Picchu, which was recently voted one of the Seven New Wonders of the World, and the historical centre of the City of Arequipa, are the major attractions for short, medium and long haul international tourists.

As international tourism has not yet achieved the status of an important economic activity for Peru's finances, there is significant room for improvement in international tourism receipts. However, the potentially negative impacts of mass tourism on the environment, and hence on future international tourism demand, must be managed appropriately. In order to manage tourism growth and volatility, it is necessary to model growth and volatility adequately.

The paper modelled the growth and volatility (or variability in the growth rate) in daily international tourist arrivals to Peru from 1997 to 2007. There are several benefits arising from using daily data as compared with lower frequency data. Among other reasons, daily data capture day-of-the-week effects as arrival patterns on the week-end, allowing for differential pricing strategies in the tourism industry, as well as the determination of optimal tourism marketing policies to increase tourist arrivals during periods of low demand. In addition, the growth rate in daily tourist arrivals expenditure, which is of primary interest in the tourism industry, is virtually identical to the growth rate in daily tourist arrivals because the growth rate in daily spending per tourist arrivals changes very little over time.

The empirical results showed that the time series of international tourist arrivals and their growth rates are stationary. In addition, the estimated symmetric and asymmetric conditional volatility models, namely the widely used GARCH, GJR and EGARCH models, all fit the data extremely well. In particular, the estimated models were able to account for the higher volatility persistence observed at the beginning and end of the sample period for both the logarithm and log difference (or growth rate) of international tourist arrivals. The empirical second moment condition also supported the statistical adequacy of the models, so that statistical inferences were valid. Moreover, the estimates resembled those arising from financial time series data, with both short and long run persistence of shocks to the growth rates of international tourist arrivals. Therefore, volatility can be interpreted as risk associated with the growth rate in international tourist arrivals.

The forecasts of daily international tourist arrivals to Peru suggested that the tourism influx to the region is likely to be very small in the years ahead. This finding points to the need for a much wider development strategy of the sustainable tourist industry to the region. Given the historical, natural, and cultural importance of Peru, appropriate marketing strategies should be directed toward attracting a greater number of international tourists to the country. The development of a sustainable tourism industry is essential to income generation, job creation, and economic growth of Peru. Rational exploration of tourism activity would help to bring economic progress to Peru without negatively affecting the natural environment and the lives of local communities.

Extensions of the models and data used in the paper to the multivariate level using modern systems methods is a topic of current research. For a theoretical comparison of alternative dynamic models of conditional correlations and conditional covariances, see McAleer et al (2008). The alternative conditional volatility models can also be used to forecast value-at-risk thresholds. A panel data analysis of temporal and spatial aggregation of alternative tourist destinations, incorporating conditional volatility models, could also be a useful direction of research.

References

Athanasopoulos, G., Ahmed, R. A., & Hyndman, R. J. (2009). Hierarchical forecasts for Australian domestic tourism. International Journal of Forecasting, 25, 146–166.

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

Bonham, C., Gangnesa, B., & Zhoub, T. (2009). Modeling tourism: A fully identified VECM approach. International Journal of Forecasting, 25, 531-549.

Boussama, F. (2000). Asymptotic normality for the quasi-maximum likelihood estimator of a GARCH model. Comptes Rendus de l'Academie des Sciences, Serie I, 331, 81-84 (in French).

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

Cook, S. (2006). The robustness of modified unit root tests in the presence of GARCH. Quantitative Finance, 6, 359-363

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74, 427-431.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica, 49, 1057-1072.

Divino, J. A., Farias, A., Takasago, M., & Teles, V. K. (2007). Tourism and economic development in Brazil. Unpublished paper, Centro de Excelencia em Turismo (CET-UnB). University of Brasilia, Brazil.

Divino, J. A., & McAleer, M. (2008). Modelling and forecasting sustainable international tourism demand for the Brazilian Amazon. Environmental Modelling & Software. (Forthcoming).

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

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient tests for an autoregressive unit root. Econometrica, 64, 813-836.

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

Gil-Alana, L., A., Cunado, J., & Gracia, F. P. (2008). Tourism in the Canary Islands: forecasting using several seasonal time series models. Journal of Forecasting, 27, 621 – 636.

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

Hoti, S., McAleer, M. & Shareef, R. (2005). Modelling country risk and uncertainty in small island tourism economies. Tourism Economics, 11, 159-183.

Hoti, S., McAleer, M., & Shareef, R. (2007). Modelling international tourism and country risk spillovers for Cyprus and Malta. Tourism Management, 28, 1472-84.

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

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

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

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

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

Ling, S., & McAleer, M. (2003b). On adaptive estimation in nonstationary ARMA models with GARCH errors. Annals of Statistics, 31, 642-674.

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

McAleer, M. (2008). The Ten Commandments for optimizing value-at-risk and daily capital charges. Journal of Economic Surveys. (Forthcoming)

McAleer, M., Chan, F., Hoti, S., & Lieberman, O. (2008). Generalized autoregressive conditional correlation. Econometric Theory, 24, 1554-1583.

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

McAleer, M., & da Veiga, B. (2008a). Forecasting value-at-risk with a parsimonious portfolio spillover GARCH (PS-GARCH) model. Journal of Forecasting, 27, 1-19.

McAleer, M., & da Veiga, B. (2008b). Single-index and portfolio models for forecasting value-at-risk thresholds. Journal of Forecasting, 27, 217-235.

Mikosch, T. & Starica, C. (2000). Limit theory for the sample autocorrelations and extremes of a GARCH (1; 1) process. Annals Statistics, 28, 1427-1451.

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

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

Perron, P., & Ng, S. (1996). Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. Review of Economic Studies, 63, 435-463.

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. Biometrika, 75, 335-346.

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

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

Shareef, R., & McAleer, M. (2008). Modelling international tourism demand and uncertainty in Maldives and Seychelles: a portfolio approach. Mathematics and Computers in Simulation, 78, 459-68.

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In O.E. Barndorff-Nielsen, D.R. Cox and D.V. Hinkley (eds.), Statistical Models in Econometrics, Finance and Other Fields, Chapman & Hall, London, pp. 1-67.

Figure 1 – Map of Peru

Source: Wikipedia

Figure 2 – International Tourist Arrivals and Volatility

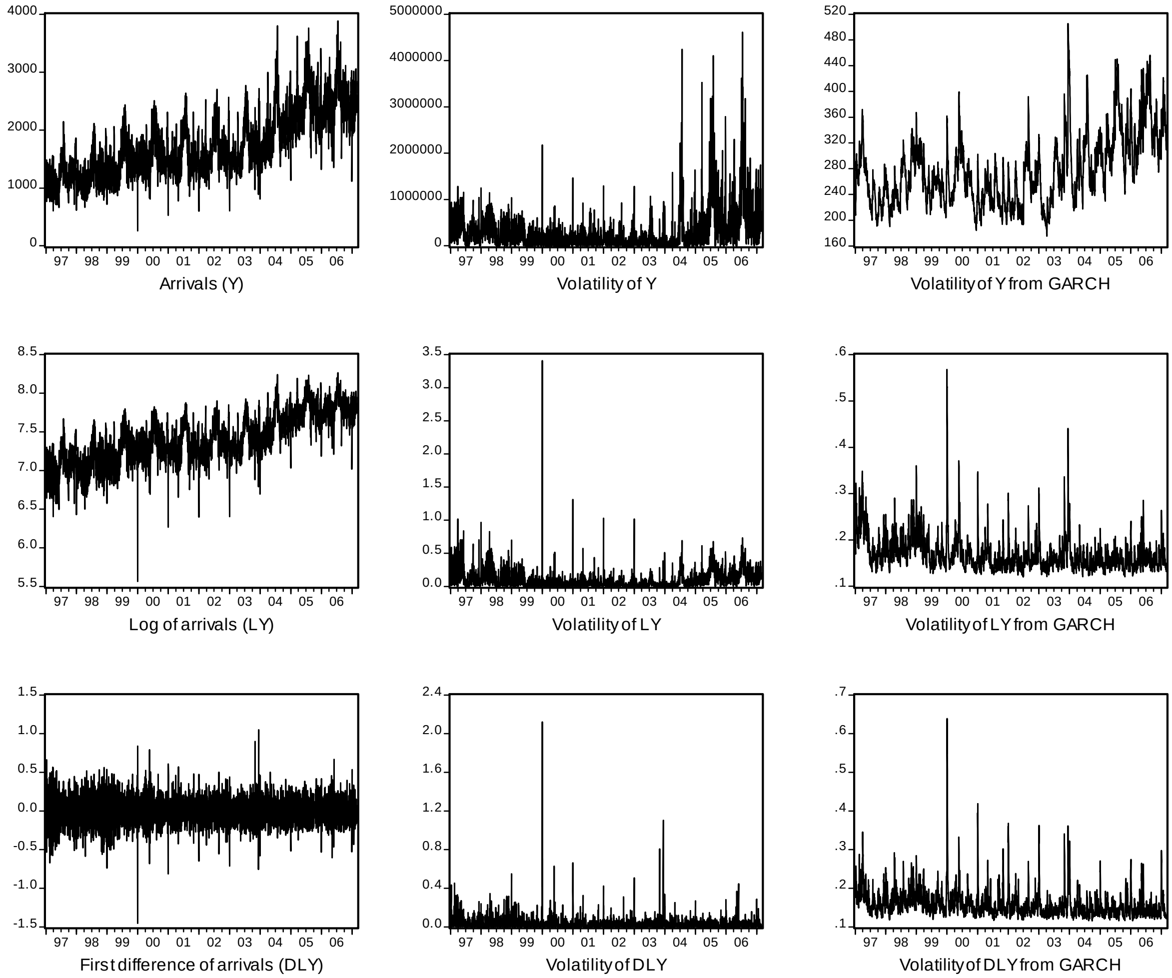

Figure 3 - International Tourist Arrivals to Peru

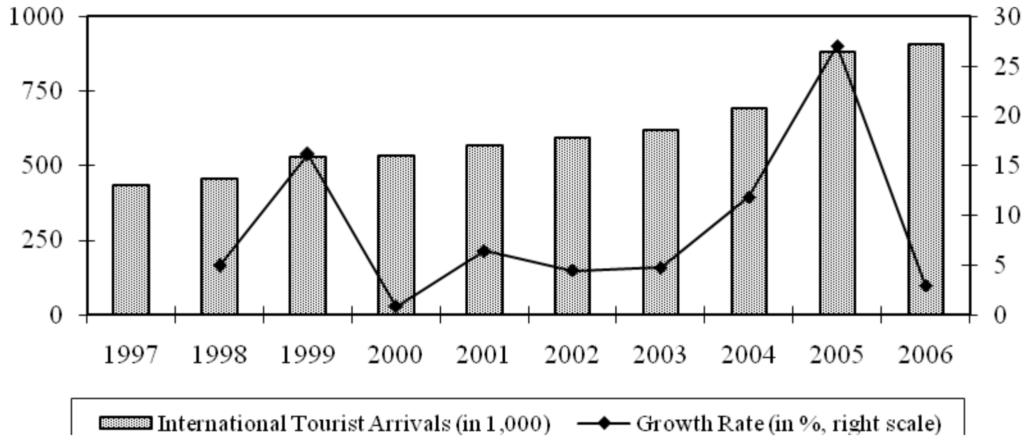

Figure 4
Forecasts of International Tourist Arrivals to Peru in Log-Levels

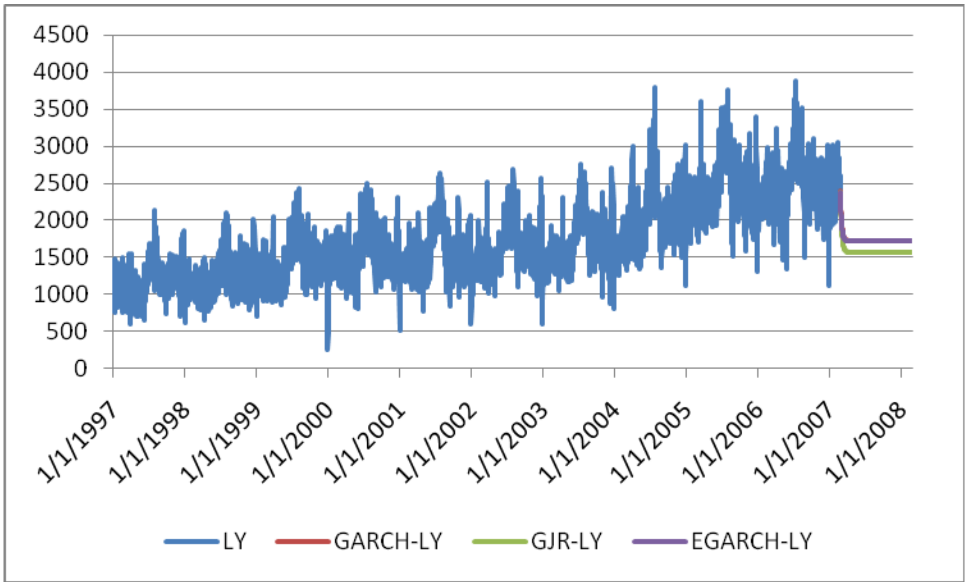

Figure 5
Forecasts of Annualized International Tourist Arrivals to Peru in Log-Levels

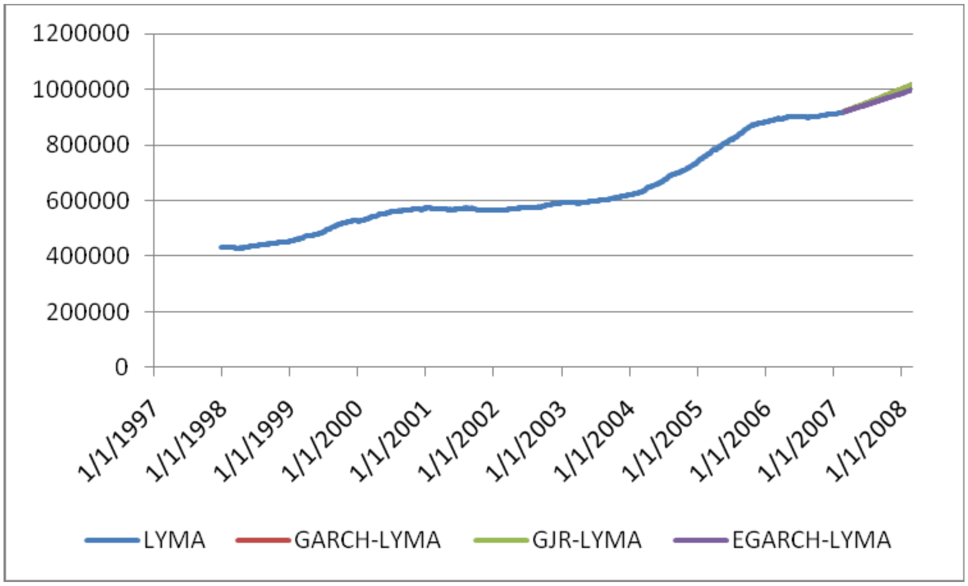

Figure 6
Forecasts of International Tourist Arrivals to Peru in First Differences

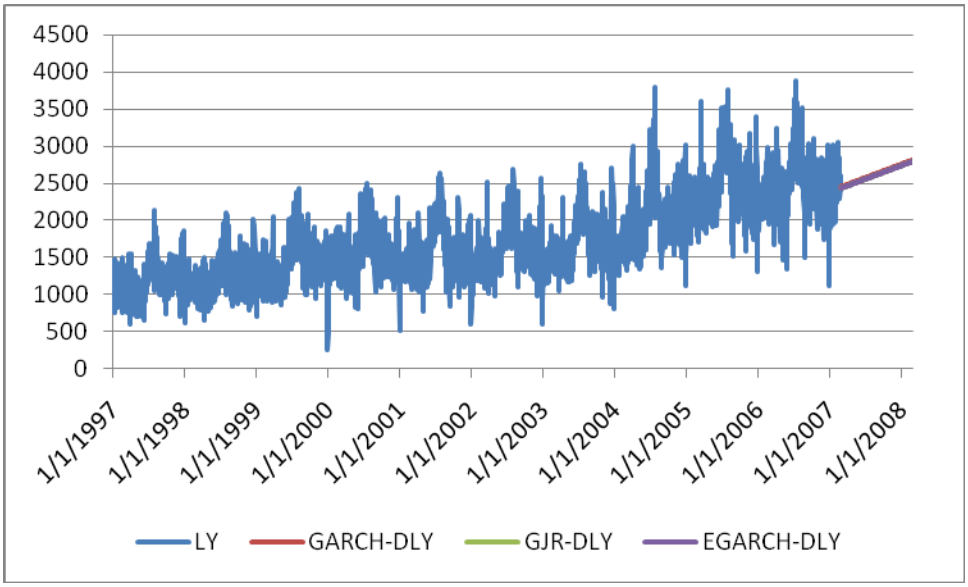

Figure 7
Forecasts of Annualized International Tourist Arrivals to Peru in First Differences

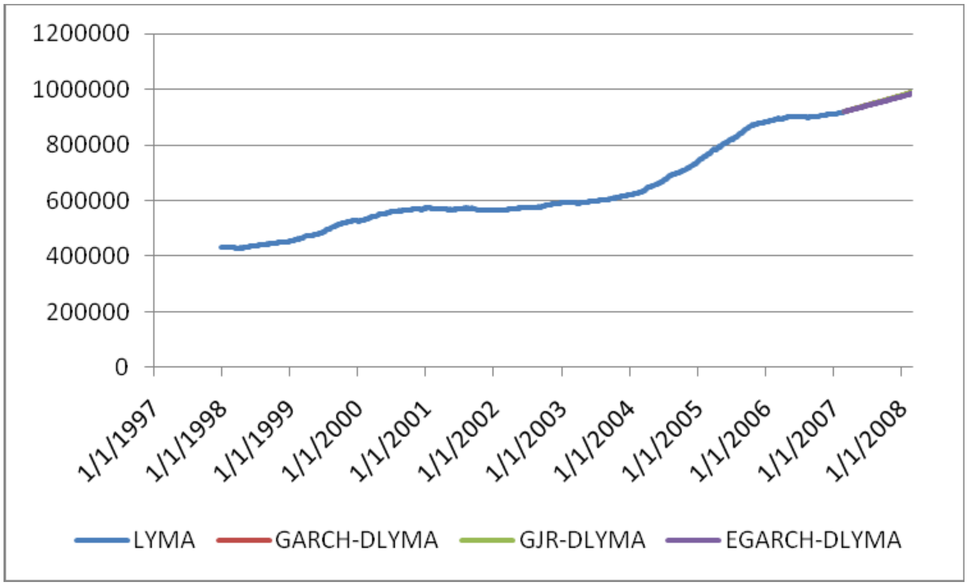

Table 1 - Unit Root Tests

Variables	$MADF^{GLS}$	MPP^{GLS}	Lags	Z
LY	-3.84**	-19.98**	27	{1, t}
LY	-0.41	-0.73	28	{1}
ΔLY	-98.41**	-2890.48**	0	{1, t}
ΔLY	-5.27**	-0.46	30	{1}

Notes:
LY is the logarithm of international tourist arrivals to Peru.
The critical values for $MADF^{GLS}$ and MPP^{GLS} at the 5% significance level are –2.93 and –17.3, respectively, when Z={1,t}, and –1.94 and –8.1, respectively, when Z={1}.
** denotes the null hypothesis of a unit root is rejected at the 5% significance level.

Table 2 - Information Criteria for Alternative ARMA Models

Variable	IC	AR\MA	0	1	2
LY	AIC	1	-0,6553	-0,9224	-0,9301
	BIC		-0,6519	-0,9174	-0,9234
	AIC	2	-0,8290	-0,9334	-0,9376
	BIC		-0,8239	-0,9267	-0,9293
LYMA	AIC	1	-12,268	-12,278	-12,339
	BIC		-12,264	-12,272	-12,331
	AIC	2	-12,283	-12,398	-12,401
	BIC		-12,278	-12,391	-12,392
DLY	AIC	1	-0,7938	-0,9304	-0,9358
	BIC		-0,7905	-0,9254	-0,9291
	AIC	2	-0,8236	-0,9413	-0,9425
	BIC		-0,8185	-0,9346	-0,9341
DLYMA	AIC	1	-12,282	-12,4	-12,402
	BIC		-12,278	-12,394	-12,394
	AIC	2	-12,354	-12,403	-12,407
	BIC		-12,349	-12,395	-12,398

Notes: IC denotes information criteria, AIC is the Akaike information criterion, and BIC is the Schwarz information criterion.

Table 3 – Estimated Conditional Mean and Conditional Volatility Models

Parameters	Dependent variable: LY			Dependent variable: DLY		
	GARCH	GJR	EGARCH	GARCH	GJR	EGARCH
ϕ_1	0.954* (0.07)	0.892* (0.07)	0.854* (0.06)	−0.001* (0.002)	−0.008* (0.002)	−0.01* (0.002)
ϕ_2	0.872* (0.01)	0.879* (0.009)	0.883* (0.01)	−0.464* (0.016)	−0.452* (0.02)	−0.452* (0.02)
ω	0.002* (0.000)	0.005* (0.000)	−0.976* (0.09)	0.003* (0.000)	0.003* (0.000)	−0.473* (0.06)
GARCH/GJR α	0.118* (0.01)	−0.027* (0.01)	--	0.139* (0.01)	0.025** (0.01)	--
GARCH/GJR β	0.803* (0.02)	0.688* (0.025)	--	0.752* (0.02)	0.780* (0.02)	--
GJR γ	--	0.316* (0.03)	--	--	0.187* (0.02)	--
EGARCH α	--	--	0.176* (0.02)	--	--	0.196* (0.02)
EGARCH γ	--	--	−0.245* (0.02)	--	--	−0.141* (0.01)
EGARCH β	--	--	0.763* (0.02)	--	--	0.913* (0.01)
Diagnostic						
Second moment	0.921	0.857		0.891	0.898	
ARCH(1) LM test	20.17	7.08	10.80	0.011	0.088	0.033
[p-value]	[0.000]	[0.008]	[0.001]	[0.916]	[0.766]	[0.855]
Jarque-Bera	128.48	126.49	162.25	452.21	330.35	349.97
[p-value]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]

Notes:
LY is the logarithm of international tourist arrivals to Peru, and DLY is the log difference (or growth rate).
Numbers in parentheses are standard errors.
* The estimated coefficient is statistically significant at the 1% significance level.
** The estimated coefficient is statistically significant at the 5% significance level.
The log-moment condition is necessarily satisfied as the second moment condition is satisfied.

Table 4 – Estimated Conditional Mean and Conditional Volatility Models

Parameters	Dependent variable: LYMA			Dependent variable: DLYMA		
	GARCH	GJR	EGARCH	GARCH	GJR	EGARCH
ϕ_1	-0.001* (0.000)	-0.001* (0.000)	-0.001* (0.000)	0.000* (0.000)	0.000* (0.000)	0.000* (0.000)
ϕ_2	1.000* (0.000)	1.000* (0.000)	1.000* (0.000)	0.107* (0.019)	0.106* (0.019)	0.107* (0.018)
ω	0.000 (0.000)	0.000 (0.000)	-0.618* (0.127)	0.000* (0.000)	0.000* (0.000)	-0.488* (0.111)
GARCH/GJR α	0.150* (0.058)	0.150** (0.066)	--	0.061* (0.008)	0.066* (0.009)	--
GARCH/GJR β	0.600* (0.16)	0.600** (0.15)	--	0.919* (0.011)	0.917* (0.011)	--
GJR γ	--	0.050 (0.09)	--	--	-0.012 (0.008)	--
EGARCH α	--	--	0.153* (0.017)	--	--	0.129* (0.016)
EGARCH γ	--	--	0.014** (0.007)	--	--	0.010** (0.005)
EGARCH β	--	--	0.967 (0.008)	--	--	0.975* (0.007)
Diagnostic						
Second moment	0.750	0.775		0.980	0.977	
ARCH(1) LM test	19.277	27.486	4.748	1.469	1.287	2.864
[p-value]	[0.000]	[0.000]	[0.029]	[0.226]	[0.257]	[0.091]
Jarque-Bera	12.97	20.47	0.91	2.74	1.63	0.92
[p-value]	[0.001]	[0.000]	[0.633]	[0.254]	[0.441]	[0.630]

Notes:
LYMA is the logarithm of annualized international tourist arrivals to Peru, and DLYMA is the log difference (or growth rate).
Numbers in parentheses are standard errors.
* The estimated coefficient is statistically significant at the 1% significance level.
** The estimated coefficient is statistically significant at the 5% significance level.
The log-moment condition is necessarily satisfied as the second moment condition is satisfied.

Table 5 - Correlation Matrix: Forecasts of the Series in Log-Levels

Model	GARCH-LY	GARCH-LYMA	GJR-LY	GJR-LYMA	EGARCH-LY	EGARCH-LYMA
GARCH-LY	1.000					
GARCH-LYMA	-0.323	1.000				
GJR-LY	0.999	-0.332	1.000			
GJR-LYMA	-0.323	1.000	-0.332	1.000		
EGARCH-LY	0.999	-0.310	0.998	-0.310	1.000	
EGARCH-LYMA	-0.324	0.999	-0.334	0.999	-0.312	1.000

Table 6 - Correlation Matrix: Forecasts of the Series in Log-Differences

Model	GARCH-DLY	GARCH-DLYMA	GJR-DLY	GJR-DLYMA	EGARCH-DLY	EGARCH-DLYMA
GARCH-DLY	1.000					
GARCH-DLYMA	0.999	1.000				
GJR-DLY	1.000	0.999	1.000			
GJR-DLYMA	0.999	1.000	0.999	1.000		
EGARCH-DLY	1.000	0.999	1.000	0.999	1.000	
EGARCH-DLYMA	0.999	1.000	0.999	1.000	0.999	1.000