Modeling and Forecasting of International Tourism Demand in ASEAN Countries

Asrin Karimi, Pouya Faroughi and Khalid Abdul Rahim

Faculty of Economics and Management, University Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia
Department of Statistics, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran

Abstract: This study attempts to find the best model to forecast international tourism demand using a series of key macroeconomic variables in ASEAN countries. Generally, we find that generalized Poisson regression model is the best one for estimating long-run international tourism demand. In addition, we find that inflation and real exchange rate have negative relationship with international tourism demand. On the other hand, foreign direct investment and openness of trade have positive relationship with international tourism demand. Cointegration test result shows that there is a long-run relationship between variables.

Keywords: Tourism Demand, Economic Growth, Macroeconomic Indicators, Panel Poisson Regression, Modeling

Introduction

Today, tourism industry plays a key role in economic growth. The World Travel and Tourism Council (WTTC) estimated that tourism industry generated 10% of global GDP and 277 million jobs (1 in 11 job opportunity) in 2014.

Considerable number of literatures has been published on the relationship between tourism and economic growth. Surveys have shown that there was a positive and strong relationship between tourism and economic growth (Cortes-Jimenez and Pulina, 2010; Adnan Hye and Ali Khan, 2013; Tang and Abosedra, 2014; Pablo-Romero and Molina, 2013; Al-Mulali et al., 2014; Bouzahzah and El Menyari, 2013; Jalil et al., 2013). Sgroi et al. (2014) shown that rural communities improve economic growth. Tudisca et al. (2014) mentioned sustainable industries such as tourism industry increase economic growth.

ASEAN tourism ministers reported that this region received 99.2 million tourists in 2013. It shows an increase of 11.73% from 2012. Figure 1 shows that there is a clear increasing trend of international tourism arrival for a set of ASEAN countries during the period from 1995 to 2013 (World Bank).

Previous studies pointed out that tourism industry will be one of the three major industries which have direct effect on the world economy into the next century (Naisbitt, 1994).

In general, therefore it seems that an accurate estimation of international tourism demand has important economic consequences for the relevant industries, policy makers and governments in destination countries to implement long-term polices and plans to reduce the risk of failure or increase the possibility of achieving desired goals.

In recent years, there has been a growing interest in analyzing and forecasting tourism demand (Zhou-Grundy and Turner, 2014; Witt and Witt, 1995; Song and Li, 2008; Claveria and Torra, 2014; Peng et al., 2014). Researchers have used different quantitative methods for forecasting tourism demand. Extrapolative or time series methods are one of the common and useful methods which numerous studies have used. These methods consider historical patterns in a data series to forecast future values. They do not cover casual relationships (Frechtling, 1996). Previous studies utilized different extrapolative methods in modeling and forecasting demand in tourism.

Chan et al. (1999) considers Gulf War as an example and he found that Naïve is the best model for forecasting unstable data. Witt et al. (1994) compared annual data and seasonal data for forecasting and modeling international tourism demand. In this study, Naïve method was used. SES model (Chen et al., 2008; Witt et al., 1994), SMA model (Makridakis et al., 1998; Hu et al., 2004; Lim and McAleer, 2008), Box-Jenkins model (Makridakis and Hibon, 1979), ARIMA (Kim et al., 2011; Goh and Law, 2002; Preez and Witt, 2003), Holt's DES model (Lim and McAleer, 2001; Makridakis et al., 1998), BSM model (Greenidge, 2001; Gonzalez and Moral, 1995; Turner and Witt, 2001; Kulendran and Witt, 2003) are other time series methods used by scholars.
Casual econometric methods are other quantitative forecasting methods. These methods are based on mathematical cause and effect relationships (Frechtling, 1996). These methods illustrate how explanatory variables affect the response variable (tourism demand) over time.

Vector Autoregressive (VAR) models are preferred to the single equation. Researchers used this model for long-run and short-run forecasting (Wong et al., 2006). Hu et al. (2004) provides the VAR models analysis to forecast international tourism demand in China for the period of 1978 to 1998. Time Varying Parameter (TVP) models consider structural instability and external shocks in tourism demand analyses (Song et al., 2000). Song et al. (2008) points out TVP model have better results in short-term forecasting. The TVP model has been used in different researches, such as (Witt et al., 2003; Li et al., 2004; Song et al., 2008; Song and Wong, 2003; Song et al., 1998).

The Error Correction Models (ECM) (Kalendran and Witt, 2001; Veloce, 2004; Ouerfelli, 2008; Song et al., 2008; Choyakh, 2008; Dritsakis, 2004; Halicioglu, 2010), The Dynamic AIDS model (Durbarr and Sinclair, 2003; Li et al., 2004; De Mello and Fortuna, 2005), Gravity Model (Che, 2004; Khadaroo and Seetanah, 2008; Guo, 2007) are other different casual econometric methods which are widely used in modeling and forecasting tourism demand researches.

This paper aims to fit a suitable model to predict international tourism demand using macroeconomic determinants by focusing exclusively on ASEAN countries. Most of previous studies focused on single macroeconomic determinant, microeconomic indicators, or from single country case studies. Unlike most of these studies, we consider the fundamental macroeconomic variables such as Foreign Direct Investment (FDI), exchange rate, inflation and openness of trade. Also, we adopt macroeconomic indicators within a panel data framework.

Data

In this study we evaluate various macroeconomic variables that might have a direct effect on the international tourist demand. Variables in our regression model include: (i) Foreign Direct Investment (net inflows) as a percentage of GDP (FDI), (ii) real exchange rate (EXCHN), (iii) inflation (INF) which is measured by the annual percentage change in the cost to an average consumer of acquiring a basket of goods and services and (iv) openness of trade (OPENS) which is measured by the sum of imports and exports over GDP.

The response variable in the regression model is the number of international tourist arrivals (TOUR). We measure the rate of tourist by the number of international incoming tourists are the number of tourists who travel to other countries outside their normal habituation. ASEAN countries have been chosen for this study because of their importance on tourism destination in the world. The choice of the sample country and period depends on accessibility and availability of data on the variables. All of time series data are collected for the following countries: Indonesia, Malaysia, Philippine,
Singapore, Thailand, Vietnam. This study covers the period of 1995-2013. The data are obtained from the World Development Indicator database which is published by the World Bank.

**Empirical Methodology and Results**

The first aim of this study is to investigate on the feature of the explanatory variables (FDI, real exchange rate, inflation openness of trade) and the response variable (international tourism arrival). The next purpose is to find the best model to estimate international tourism demand.

**Panel Unit Root Test**

Before conducting the panel data regression, we conducted a panel unit root test. For this purpose we adopt three different methods, namely those of the Im, Pesaran and Shin (IPS) test (Im et al., 2003), Fisher-ADF and Fisher-PP statistics (Maddala and Wu, 1999; Choi et al., 1999). The null hypothesis of these tests is that each series in the panel is not stationary. Table 1 indicates our unit root test results. The results reveal that real exchange rate, inflation and openness of trade are not stationary at levels but they are stationary at the first difference (Table 2), therefore rejecting the null hypothesis indicate that the variables contain a panel unit root.

The results of panel unit root test indicate that time series are stationary at the first difference; thus, it seems that checking the cointegration of the series is necessary. In this study, the Pedroni panel cointegration test is employed. Pedroni (1999) was the first to develop heterogeneous panel cointegration test for series. This test allows us to accommodate individual specific fixed effects and definite trends and estimate coefficients for each series (Pedroni, 2004). In addition we use Kao residual cointegration test (Kao et al., 1999) to test the cointegration relationship in series.

Table 3 presents the panel cointegration test results. The Panel PP-statistic and Group PP-statistic strongly reject the hypothesis of no cointegration and Kao residual cointegration test also reject the hypothesis of no cointegration significantly at 5% critical value. Thus, there is a long-run relationship between the variables. This result shows that we can use the level of series for the Poisson regression model, GP and NBP regression model.

**Estimator Models**

Panel Poisson regression, negative binomial regression and generalized Poisson techniques are used in this study. These techniques have been widely used for count data (Greene, 2008). Individual effects are extended to two full distributional assumption, random effect assumption and fixed effect assumption (Hausman et al., 1984). The individual fixed effects model considers N countries which are relative to T (time series observation). Also, it is correlated with the independent variables (Wooldridge, 2010).

In this study individual fixed effects are chosen to estimate Poisson regression model because ASEAN countries are considered.

---

**Table 1. Panel unit root test results**

<table>
<thead>
<tr>
<th></th>
<th>IPS Intercept</th>
<th>IPS Intercept + Trend</th>
<th>Fisher-ADF Intercept</th>
<th>Fisher-ADF Intercept + Trend</th>
<th>Fisher-PP Intercept</th>
<th>Fisher-PP Intercept + Trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tour</td>
<td>7.943</td>
<td>4.429</td>
<td>0.173</td>
<td>4.542</td>
<td>0.115</td>
<td>4.244</td>
</tr>
<tr>
<td>FDI</td>
<td>-4.202***</td>
<td>-4.086***</td>
<td>39.871***</td>
<td>36.769***</td>
<td>36.676***</td>
<td>44.525***</td>
</tr>
<tr>
<td>EXCHN</td>
<td>1.98</td>
<td>-0.813</td>
<td>5.111</td>
<td>17.114</td>
<td>4.314</td>
<td>9.973</td>
</tr>
<tr>
<td>INF</td>
<td>0.385</td>
<td>0.385</td>
<td>0.845</td>
<td>10.629</td>
<td>0.464</td>
<td>4.646</td>
</tr>
<tr>
<td>OPENS</td>
<td>5.779</td>
<td>0.868</td>
<td>1.936</td>
<td>11.567</td>
<td>1.888</td>
<td>13.416</td>
</tr>
</tbody>
</table>

Note: We used the Schwarz automatic selection of the lag length for the unit root test. The IPS, Fisher-ADF and Fisher-PP examine the null hypothesis of non-stationary. Probabilities for Fisher-type tests were computed by using an asymptotic $\chi^2$ distribution. All other tests assume asymptotic normality.

*Indicates that statistics are significant at the 10% level of significance.

**Table 2. First difference panel unit root test results**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tour</td>
<td>-4.355***</td>
<td>-5.331***</td>
<td>40.735***</td>
<td>49.205***</td>
<td>42.306***</td>
<td>70.343***</td>
</tr>
<tr>
<td>FDI</td>
<td>-9.010***</td>
<td>-7.084***</td>
<td>87.495***</td>
<td>60.544***</td>
<td>180.490***</td>
<td>90.235***</td>
</tr>
<tr>
<td>EXCHN</td>
<td>-4.110***</td>
<td>-1.500***</td>
<td>37.856***</td>
<td>31.393***</td>
<td>36.478***</td>
<td>54.503***</td>
</tr>
<tr>
<td>INF</td>
<td>-7.520***</td>
<td>-6.957***</td>
<td>67.875***</td>
<td>59.653***</td>
<td>101.725***</td>
<td>137.703***</td>
</tr>
<tr>
<td>OPENS</td>
<td>-6.829***</td>
<td>-6.600***</td>
<td>62.564***</td>
<td>56.548***</td>
<td>70.569***</td>
<td>101.746***</td>
</tr>
</tbody>
</table>

Note: FD denotes first difference. *** indicates statistical significance at the 1% level of significance. We used the Schwarz automatic selection of the lag length for the unit root test. The IPS, Fisher-ADF and Fisher-PP examine the null hypothesis of non-stationary. Probabilities for Fisher-type tests were computed by using an asymptotic $\chi^2$ distribution. All other tests assume asymptotic normality.
Table 3. Panel co-integration test results

<table>
<thead>
<tr>
<th></th>
<th>Panel v</th>
<th>Panel rho</th>
<th>Panel PP</th>
<th>Panel ADF</th>
<th>Group rho</th>
<th>Group PP</th>
<th>Group ADF</th>
<th>Kao test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deterministic intercept and trend</td>
<td>2.170*</td>
<td>0.811</td>
<td>-1.326***</td>
<td>-1.258*</td>
<td>2.34</td>
<td>-2.337***</td>
<td>-0.781</td>
<td>1.735**</td>
</tr>
</tbody>
</table>

Note: Statistics are asymptotically distributed as normal. ***, ** and * rejects the null of no co-integration at the 1, 5 and 10% level, respectively.

For the formulas used in the panel co-integration test statistics, it is described in details in Pedroni (1999; 2004) and Kao et al. (1999).

* Indicates that statistics are significant at the 10% level of significance.
** Indicates that statistics are significant at the 5% level of significance.
*** Indicates that statistics are significant at the 1% level of significance.

The panel Poisson regression model with fixed effects can be used as follows:

\[ Pr(Y_i = y_i) = \frac{\exp(-\lambda_i)^{y_i}}{y_i!}, y_i = 0, 1, 2, \ldots \]

Let \( Y = (Y_1, Y_2, Y_3, \ldots, Y_n)^T \) (be the response vector where \( n \) is the sample size and \( Y_1, Y_2 \) are independent for any \( i \neq j \). If \( Y_1 \) is distributed as Poisson.

The covariates of \( \lambda_i = E(Y_i) \) for Poisson regression model can be included using log link function:

\[ \log \lambda_i = x_i^T \beta + \mu_i \]

where, \( x_i \) is the vector of covariates and \( \beta \) is the vectors of regression parameters. With mean and variance, \( E(Y) = \lambda_i \).

The Poisson regression model has been widely used for modeling count data with covariates. In this model it is assumed that conditional mean and conditional variance functions are equal. This assumption limits the applications of Poisson regression model. Guloglu and Tekin (2012), stated that unobserved heterogeneity of cross-section units is causing overdispersion. Moreover, count data is often overdispersed and Negative Binomial (NB) regression has been used for handling overdispersion whereas Generalized Poisson (GP) regression has been fitted for under- or overdispersed count data.

A two-sided Likelihood Ratio Test (LRT) are performed to test the dispersion (over- or underdispersion) in panel Poisson regression against generalized Poisson or (negative binomial) alternatives (GP or NB) (Cameron and Trivedi, 1998) where the hypothesis is:

\[ H_0: \text{Dispersion Parameter} = 0 \quad \text{Against} \quad H_1: \text{Dispersion Parameter} \neq 0 \]

Since Poisson model is nested within GP and NB models, the statistic is asymptotically distributed as a chi-square with one degree of freedom.

LRT statistic is:

\[ T = 2(\ln L_0 - \ln L_1) \]

where, \( \ln L_1 \) and \( \ln L_0 \) are the models’ log likelihood under their respective hypothesis.

The LRT for testing Poisson against GP regression models and Poisson against NB regression models are 364430.6 and 364412.6 respectively, indicating overdispersion in our data and reject null hypothesis of equality of mean and variance.

Several parameterizations were performed for the generalized Poisson and negative binomial regression models (Famoye et al., 2004; Wang and Famoye, 1997; Zamani and Ismail, 2012; Greene, 2008; Zamani et al., 2014). One of the parameterization of the GP regression model, which is used in this study, was used by Wang and Famoye (1997) for analyzing household fertility count data and by Ismail and Jemain (2007) for analyzing the Malaysian claim count data. In this study, negative binomial regression model form is used which is the most popular form. Greene (2008) used NB to panel data on health care utilization. The estimates of Poisson are suggested as initial values for fitting the GP and NB models. Table 4 shows the parameter estimates and t-ratio for the fitted models of the international tourism.

Comparison in terms of significance of estimates of covariates between Poisson, GP rand NB regression models shows that all models provide the same significant estimates at 0.10 level. All of the covariates except inflation under negative binomial regression model are significant at level 0.05. The absolute values of t-ratio for dispersion parameter under GP and NB are 14.22 and 7.99 respectively, indicating that the dispersion parameter for both models is significant.

For a comparison of non-nested models, information criteria such as Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) can be used. AIC and BIC are given by:

- AIC = 2dim(\( \theta \)) – 2 ln (L)
- BIC = ln (n) – dim(\( \theta \)) – 2 ln (L)

where, L is the maximum likelihood function for the estimated model.

The model with the smallest AIC or BIC is the best model. AIC includes a penalty by adding the number of parameters in the criteria, while BIC includes a penalty by adding the number of parameters and the sample size in the criteria.
Based on the AIC and BIC in the Table 4 the GP regression model is the best model, followed by NB and Poisson regression models.

### Conclusion

Table 4 illustrates the Poisson, GP and NB Poisson regression results. According to AIC and BIC results, GP regression model is the best model for forecasting international tourism demand. Based on our results inflation and real exchange rate has negative relationship with international tourism demand. Panel cointegration test result shows that there is a long-run relationship between variables.

In general, our findings fall in line with previous studies. The result shows that there is a positive significant relationship between FDI, openness of trade and international tourism demand. Numerous literatures have provided empirical evidence in support of these results. A number of studies have found that governments try to attract FDI for more international tourism arrival in developing countries (Andergassen and Candela, 2009). Siddique et al. (2012) illustrated that there is a causal interaction between FDI and tourism arrival. Also, an empirical study in China showed that there is a causal relationship between tourism arrivals, FDI and economic growth from 1978 to 2005 (Tang et al., 2007).

Numerous studies have attempted to explain the relationship between tourism and international trade. Previous studies explained that tourism industry met imports’ needs and it enhance exports (Massidda and Mattana, 2012; Santana-Gallego et al., 2010). International tourism can make a lot of business opportunities by export sales and import purchases (Khan et al., 2005). Katircioglu (2009) found that there is a long-run relationship between international tourism arrival, economic growth, exports and imports in Cyprus. The author used ARDL-ECM model. Sarmidi and Saleh (2010) cross-country analysis (2009) showed a casual links between tourism and trade (export and import). Akinboade and Braimoh (2010) illustrated a causality links between real export and international tourism in long term.

Furthermore, the results indicate that there is a negative relationship between inflation, real exchange rate and international tourism demand. Chatziantoniou et al. (2013) pointed out there is reverse causality and negative effect between tourism industry and inflation.

Based on previous studies, exchange rate fluctuation plays a key role on tourism industry (Blake et al., 2008; Becken et al., 2008). Previous researches have reported that exchange rate has an adverse interaction with tourism arrival (Hanafiah and Harun, 2010). Other studies illustrated strong domestic currency has negative correlation on international tourism (Chang and McAleer, 2012). This study indicates that FDI, real exchange rate, inflation and openness of trade could be an effective tool, which can predict international tourism demand.

### Acknowledgement

The authors would like to thank the reviewer for his effort and valuable comments.

### Author’s Contributions

**Asrin Karimi:** Was in charge of: (a) implementing the unite root test and the development co-integration test, (b) developed literature review.

**Pouya Faroughi:** Developed R program for fitting GP, NB regression and contributed to the writing of the manuscript.

**Khalid Abdul Rahim:** Contributed to the writing of the manuscript, designed the research plan and carried out the overall editing of manuscript.

### Ethics

This article is original and contains unpublished materials. The corresponding author confirms that all
of the other authors have read and approved the manuscript and no ethical issues involved.

References


Guo, W., 2007. Inbound tourism, an empirical research based on gravity model of international trade. Tourism Tribune, 22: 30-34.


