Forecasting International Tourism Regional Expenditure

Benjamin Ognjanov
Insync Research, Melbourne, Australia

Yihong Tang
Central University of Finance and Economics, Beijing, China

Lindsay Turner
Victoria University, Melbourne, Australia

Central University of Finance and Economics, Beijing, China

The vast majority of tourism forecasting studies have centered on tourist arrivals at an aggregated level. Little research has been done of forecasting tourist expenditure at a national level let alone at a regional level. This study uses expenditure data to assess the relative economic impact of tourism into regional areas. By comparing five time-series models (the Naïve, Holt, ARMA and Basic Structural Model (BSM) with and without intervention), and three econometric models (the Vector Autoregressive (VAR) model and the Time Varying Parameter (TVP) with and without intervention), the study sought to find the most accurate model for forecasting tourism expenditure two years ahead for each of the 31 provinces of mainland China. The results show that TVP models outperform other time series and econometric models. The research also provides practical management outcomes by providing methods for forecasting tourist expenditure as an indicator of economic growth in China’s provinces. The research concludes with the findings on the most appropriate model for regional forecasting and potential new variables suitable at the regional level.

Keywords: regional forecasting, tourism expenditure, China tourism, time-series models, econometric models, model evaluation

Introduction

The academic study of demand modelling and forecasting, has resulted in widespread literature from the late 1970’s. These publications vary immensely in scope, modelling and forecasting techniques, data types, and objectives. The major literature review articles that attempt to summarize this literature include Crouch (1994), S. Witt and C. Witt (1995), Lim (1997a; 1997b; 1999), Li, Song, and Witt (2005), Song and Turner (2006), Song and Li (2008), and Goh and Law (2011).

Furthermore, the large volume of literature in the field of tourism forecasting is almost universally focused on international arrivals between nations (Song & Witt, 2005; Song & Li, 2008; Song, Li, Witt, & Fei, 2008),
and there is little research focused on regional international arrivals forecasting in underdeveloped countries (Vu & Turner, 2006; Zhou-Grundy & Turner 2014). From the major literature review articles tourism arrivals is measured as total tourist arrivals from an origin country to a destination country, and it is at times disaggregated further into holiday tourist arrivals, business tourist arrivals, tourist arrivals for visiting friends and relatives (VFR) (Turner & Witt, 2001a; 2001b; Kulendran & Wong, 2005), and in some instances tourist arrivals by air as opposed to land and sea (Coshall, 2005).

However, there is no current literature attempting to forecast the resultant tourist expenditure in China at the provincial or regional level, and very little research of this type in the past literature. Expenditure does vary as different source nationalities have different expenditure volumes and purchasing patterns. Regional areas that receive dissimilar volumes of arrivals from different international markets (a characteristic of tourism to China) will in turn receive different potential expenditure. Research on tourism expenditure has mainly focused on identifying causal variables (Wang & Davidson, 2010; Kim, Han, & Chon, 2008; Vietze, 2009) or the relationship with economic growth (Frechtling, 2006; Georgantopoulos, 2012). Literature specifically devoted to forecasting tourism expenditure is limited. These include Akal (2004), with his study on forecasting tourism revenues for Turkey and Au and Law (2000) who forecast tourist sightseeing expenditure in Hong Kong and in 2002 forecast tourist dining expenditure. Further, there has been very little empirical testing of tourism expenditure forecasting models and their accuracies since Sheldon’s (1993) study forecasting tourism expenditures and arrivals.

**Literature Review**

The literature review of research into tourism forecasting shows that researchers have used a variety of methods in forecasting tourism demand. These methods include simple time series extrapolative methods such as exponential smoothing, ARIMA models, complex econometric models, combined models, and artificial intelligence methods. Of note throughout the literature review is that researchers state that no single model outperforms others in all cases. The performance of alternative models is situation specific, and therefore it is unlikely that one model would be appropriate for all 31 provinces analyzed in this study.

The concept of regional forecasting and the forecast models that work best with regional data need to be developed and tested, and given the current knowledge that different models work differently on different time series it is necessary to do this testing in China. Also, it is significant to examine the capacity to forecast accurately in regional China as an excellent example of a developing economy, with high regional diversity and large regional scale. Although data availability will remain an important constraint, more regional data series are becoming available in both developed and underdeveloped countries to encourage further research that may be stimulated by the findings of this study.

China has significant increased industry demand for regional tourist forecasts stimulated initially from the 2008 Beijing Olympic Games, and 2010 Shanghai World Expo. The need for regional forecasting has also been accelerated by the Chinese Central Government’s initiative in developing the western and central regions, to ease social pressure and economic imbalance between the coastal developing regions and the inland and lesser developed regions. Several studies have investigated the success and impact of rural and regional tourism development in China (Gao, S. Huang, & Y. Huang, 2009; Hu, 2008; Lew & Yu, 1995; Pine, 2002; Lew, Yu, Ap, & Zhang, 2003; Li, 2008; Zhao, 2008). The research is important as the national government of China considers openness in tourism trade as one significant way that economic development can be spread.
The importance of tourism to China’s economy is significant. The direct contribution of tourism to China’s total GDP in 2015 was 2.1% while the total contribution (direct, indirect, and induced impacts) to China’s total GDP was 7.9%. Tourism is a labor-intensive sector. The total tourism contribution to China’s employment, including jobs indirectly supported by the industry, was 8.4% of total employment in 2016 (WTTC).

As China’s regional governments seek to benefit from international tourism receipts strong competition has emerged to attract foreign visitors to their regions (Vu & Turner, 2006). Regional and provincial governments have made considerable investments in updating and expanding their infrastructure to cater for international tourists. Understanding the long-term cost benefits of large capital investments in attracting international tourists, forecasting of international tourist expenditure at the regional and provincial level has become a more urgent issue. As remarked by Wang, Rompf, Severt, and Peerapatdit (2006, p. 333), tourism expenditure is “typically scrutinized by policy makers, planning officials, marketers and researchers for monitoring and assessing the impact of tourism on the local economy”.

The objective of this study is to determine whether regional forecasting models can be used, in responding to demand from regional governments, and tourism related industry, to accurately forecast international regional tourist expenditure. In attempting to achieve this objective the study examines whether the latest forecasting techniques can accurately forecast international regional tourist expenditure, and in doing so examines possible new explanatory variables specific to regional tourism expenditure forecasting. Within this study provinces have been grouped by their traditional regional classification, which consists of six areas or regions; East Region (Shanghai, Jiangsu, Zhejiang Anhui, Fujian, Jiangxi, and Shandong), North Region (Beijing, Tianjin, Hebei, Shanxi, and Inner Mongolia), Northeast Region (Liaoning, Jilin, and Heilongjiang), Northwest Region (Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang), South Central Region (Henan, Hubei, Hunan, Guangdong, Guangxi, and Hainan), and Southwest Region (Chongqing, Sichuan, Guizhou, Yunnan, and Tibet).

The economic factors that influence tourism changes at the regional level need to be identified and examined in the development process. Although similar models and explanatory variables to those used for national forecasting may be appropriate, this assumption may not be valid.

Whilst there may be issues related to marketing, world awareness and relative regional position the bottom line for tourism is the expenditure earned by destination places. It is the economic bottom line that drives the interest in having international tourism. There are other advantages that can be classified as political, social, and humanistic related in general terms to greater understanding between cultures. However, the fundamental drive for international tourism remains economic, not social.

This research makes a significant contribution to the literature in international tourism expenditure forecasting at a regional and provincial. Furthermore, the development of causal modelling at the regional level requires the study of new representative measures that account for a changing regional structure. As such, the research looks at the question of finding new causal variables within an inevitably restrictive data environment.

Forecasts of tourist expenditure may be of greater importance than forecasts of arrivals when assessing the economic impacts of tourism. At the regional level there can be wide variations in the concentration of different source markets that can lead to varying lengths of stay, and differing levels of expenditure. It can be argued that forecasting tourism nights is superior to arrivals because it includes length of stay and implies expenditure
levels (Luzzi & Flückiger, 2003). However, in many cases data for nights are derived from hotel accommodation records that miss other forms of accommodation, and under-represent arrival numbers.

The most common methods used to collect tourist expenditure data are through bank records of foreign exchange transactions, surveys of tourists, and tourism establishments. In 1994 China ceased using foreign exchange records because it underestimated receipts, and moved to sample surveying. The data on regional tourist expenditure are collected annually and published by the National Tourism Administration’s publication “The Yearbook of China Tourism Statistics”. There is a time lag of almost 18 months in the release of the publication containing the latest provincial level data. A shortcoming of the provincial level data is that it fails to distinguish expenditure by visitor type, and expenditure does vary between different business and holiday based tourism.

In addition, although expenditure data are reported at the provincial level it does not apportion expenditure by country of origin. Both visitor type and expenditure per country of origin are only reported at an aggregated national level. To overcome this shortfall in the data, expenditures by visitors from each of the 13 source countries were calculated using the number of foreign visitors by source country for each province by the average number of days they visited the province and the average spend per day per visitor. The estimates were then compared with the published aggregated data to ensure that the results were not over or under estimated for each province.

Eight different methodologies for forecasting tourism expenditure for each of the 31 provinces and six regions in China have been tested and compared within this study. These include five time-series models; Naïve 1, Holt, ARMA, Basic Structural Model (BSM) with and without interventions, and three econometric models; the Time Varying Parameter model (TVP) with and without interventions and the Vector Autoregressive model (VAR). These models are like those used in the Zhou-Grundy and Turner (2014) study to forecast tourist arrivals, although in that case a Neural model is included, and the VAR model is not included. Naïve 1 forecasts are used in both studies and in previous literature as a benchmark measure to compare the primary forecasting models against a simple extrapolation of previous year numbers. Song and Li (2008) cited numerous studies that have utilized the naïve model as a benchmark for forecasting accuracy evaluation, whereby the accuracy of any model is expected to improve upon a naïve forecast. The interventions used in the BSM and TVP models were input as “outliers” occurring at a particular time for each province or region and measured as the value one, and zero for non-events. Thirteen source country annual data sets to each of the 31 Chinese provinces and six regions were utilized in this study, a total of 481 analyses (403 for province and 78 for regions) for each of the eight forecasting models (a total of 3,848 analyses).

Two of the main objectives of this study were to examine whether it is possible to accurately forecast tourism expenditure of international tourists to China at the regional level and whether it could be extended to include forecasting tourism expenditure at the provincial level. The forecasting horizon is short at two years ahead, and this is necessary due to the short data time frame available for the model estimation period 1994 to 2014.

**Time Series Models**

Naïve 1: The Naïve 1 method simply states that the forecast value for this period (t) is equal to the available actual value of the last period (t-1).

Holt’s Method: Holt’s method is a refinement of the popular single exponential smoothing model which considers any trend in the data. When the data exhibit either an increasing or decreasing trend over time, simple
exponential smoothing forecasts tend to lag behind observations. Double exponential smoothing is designed to address this type of data series by taking into account any trend in the data.

ARMA: An autoregressive moving average (ARMA) process consists of both autoregressive and moving average terms. If the process has terms from both an AR(p) and MA(q) process, then the process is called ARMA(p, q).

BSM: The basic structural model (BSM) is constructed by decomposing a time series into its trend, seasonal, cycle, and irregular components. The irregular component represents the transitory variations in demand which are not explained by other components.

BSM with intervention: A BSM model with intervention effects incorporated into the model framework.

Econometric Models

TVP: The Time Varying Parameter (TVP) models are often used when there are structural changes in data or when constant coefficients are too restrictive. When the coefficients are estimated with the TVP, the more recent data have a stronger influence than previous data. The TVP models are normally specified in a state space form (SSF) and estimated by the Kalman filter algorithm. The state space presentation allows unobserved variables to be included into, and estimated along with, the observable variables.

TVP with intervention: A TVP model with intervention effects incorporated into the model framework.

VAR: The Vector Autoregressive (VAR) model, all variables are treated as endogenous (dependent) and each variable is specified as a linear relationship of the others, unlike single equation models where the variables are exogenous. The current values of the variables are regressed against lagged values of all the variables in the system.

Causal Variable Selection

A review of econometric studies of tourism demand by Song and Li (2008) and tourist expenditure by Wang and Davidson (2010) have shown that demand for tourism is influenced by economic, social, and psychological factors as well as trip-related and destination related variables, as important determinants of tourism demand. However, there is no current research that has identified any causal variables in empirical modelling of tourist expenditure at the regional level.

After careful examination of causal variables utilized in contemporary research and the availability of accurate and complete time-series data for the required period across all 31 provinces and six regions, this study included the following causal variables for regional and provincial forecasting; total number of foreign tourist arrivals from origin country (ARR), average length of stay in a province or region (AVE_STAY), relative price (RPRICE), foreign direct investment (FDI), gross domestic product per capita of origin country (GDP), and transport density (T_DENS). These variables are distinct from the Zhou-Grundy and Turner (2014) variables used for regional arrivals forecasting in China: Per Capita income of the source country, Gross regional product (similar to GDP), Regional Foreign Direct Investment (same variable as FDI), Gross Capital Formation, Urban Rural Ratio, Average monthly sunshine, Own price (same as RPRICE), Regional road network length (similar to T_DENS). The Urban/Rural ratio was found to be important in Zhou-Grundy and Turner’s study and this relates to the growth of cities which in turn increases arrivals. However, for expenditure this should not be a causal issue when arrivals are an explanatory variable. Sunshine, gross capital formation, and own price were the least effective causal measures in Zhou-Grundy and Turner’s study. There is little logical reason to relate expenditure
to sunshine. Foreign direct investment is intended to capture business travel, and does vary in level provincially in China, with some regions particularly designed to attract foreign investment by the central government.

The reader may query the inclusion of transport density, a supply-side variable, in a demand forecasting model. This approach is generally not supported by neoclassical economic theory. However, there have been a number of studies which have included supply-side variables in demand forecasting models. Examples include Vu’s (2010) thesis in forecasting international regional arrivals in Canada, Yang’s (2009) thesis on modelling the growth of tourism flows to China, Louw’s (2011) thesis on forecasting tourism demand in South Africa, and Zhou-Grundy and Turner (2014). Yang (2009) in his thesis on the spatial econometric approach to modelling growth of tourism flows to Chinese cities highlights a number of causal variables for regional areas including infrastructure. As such this study incorporates transport density as possibly an important causal variable for analysis. The assumption is that the more infrastructures (roads, rail, etc.) within a given province the more likely tourists are able, and want to travel within the province. In this study transport density is calculated by adding the total length of road and the total length of rail network in kilometers within a given province and dividing by the area (total sq. km.) of the province.

Lim (2006) in his study utilized relative price as a measure of own price and found that relative price is negatively related to tourism expenditure. As such a rise in relative prices means that costs in destination \(i\) are relatively more expensive for tourists from origin \(j\). Using the following formulae relative price is calculated:

\[
RP_{it} = \left( \frac{CPI_{it}}{CPI_{it}} \right) \times ER_{it} 
\]

where:

\(RP_{it}\) = relative price variable in destination \(i\) in period \(t\);
\(CPI_{it}\) = consumer price index in destination \(i\) in period \(t\);
\(CPI_{it}\) = consumer price index in origin \(i\) in period \(t\);
\(ER_{it}\) = an index of the price of origin currency in terms of destination \(i\) currency in period \(t\) (Lim, 2006).

Exchange rates have been utilized in numerous studies as a causal variable in tourism demand. However, the formulae expressed above in calculating relative price takes into consideration exchange rates.

A visitor’s length of stay at a destination country will impact their total expenditure. The longer the stay the more they will spend. As such it would be logical to assume that this variable should be incorporated in the analysis. However, there is little research that shows length of stay as a causal variable. Song et al. (2008) in their study on demand forecasting identified length of stay as a tourism demand measure utilized in only three previous studies by Crouch.

Dummy variables have been included in a number of studies in tourism demand models to capture the impacts of “one-off” events. Considering the potential effects of manmade crises, natural disasters and other one-off events which have impacted on international tourism demand it is necessary to consider the use of dummy variables for such events. These events can impact the whole country such as the Severe Acute Respiratory Syndrome (SARS) in 2003, or impact a specific region such as the Sichuan earthquake in 2008.

**Evaluation of Forecasting Accuracy**

Forecast performance has customarily been evaluated by statistical measures such as the mean absolute percentage error (MAPE) and root mean square error (RMSE). Li et al. (2005) and Song, Witt, Wong, and Wu
(2009) note that the principal measure for forecasting error magnitudes used in the overwhelming majority of tourism demand forecasting is MAPE followed by RMSE. MAPE provides a percentage measure of how large the forecast error is in comparison to the actual series value, and is unit free. As expenditure values differ considerably across provinces, MAPE has an advantage in enhancing the capacity to compare forecast accuracy of the varying expenditure values across the provinces. The MAPE is calculated as:

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|$$

where: \(Y_t\) is the actual time period \(t\); \(\hat{Y}_t\) is the forecast value for the time period \(t\).

**Analysis**

The data series employed in the forecast analysis consists of annual observations covering 21 periods from 1994 to 2014. To evaluate the models for predictive validity, the analysis employs an ex post forecast where the data are divided into two segments (Yaffee, 2010). The first segment, or estimation segment, consists of data from 1994 to 2012. The second segment, or validation segment, includes data for 2013 and 2014. The models are then estimated on the estimation segment and the forecast is over the validation segment.

The Naïve 1 model has been chosen as a benchmark. The naïve method assumes the data will remain unchanged from the previous period to the next. As the naïve methodology is simplistic in nature it would be expected that results from more sophisticated models should provide better forecasting accuracy. As such this study requires forecasts produced by the naïve method to be the minimum benchmark forecast accuracy. Therefore, it is essential that the mean absolute percentage error (MAPE) values of all forecast models improve upon the MAPE values of the naïve model.

Lewis (1982) suggests the interpretation of MAPE as follows: less than 10% is highly accurate forecasting, 10% to < 20% is good forecasting, 20% to < 50% is reasonable forecasting, and 50% or more is inaccurate forecasting. The error ranges suggested by Lewis are considered by many researchers to be subjective, and there has been no recognized specific range for error assessment in the literature reviewed so far. However, Zhou-Grundy and Turner (2014) suggested a more succinct percentage range for evaluating the accuracy of forecasts which will be employed in this study (refer to Table 1). These ranges have been derived from previous extensive forecasting in industry completed for the Pacific Asia Travel Association from 2000 to 2010, including regional forecasting in Canada, China, Japan, Korea, Japan, Malaysia, Mexico, and the USA (Turner & Witt, 2010b). As such they are a subjective industry based scale.

| MAPE value | Level of accuracy |
|------------|------------------|
| 0% - <10%  | Very accurate    |
| 10% - <20% | Accurate         |
| 20% - <30% | Good             |
| 30% +      | Inaccurate       |

Table 2 summarizes the one step ahead and two steps ahead MAPE values for all models for tourist expenditure from the 13 source countries to the 31 Chinese provinces. There are a total of \(31 \times 13 = 403\)
forecast series for each model. For example, the TVP forecasts, two steps ahead, has 306 (75.9%) MAPE counts below 10% compared with the VAR method which only has 30 (7.4%) overall MAPE counts in this band. The results in Table 4 show that both the TVP and the TVP with interventions models provide very accurate forecasting of tourism expenditure from all 13 source countries to each of the 31 Chinese provinces in the short term (two years) for both one step ahead and two steps ahead. With both TVP methodologies having over 90% of tourism expenditure forecasts below 20% MAPE error, the TVP method forecasting accuracy for tourism expenditure at the provincial level outperformed all other evaluated models.

Table 2
One and Two Steps Ahead MAPE Comparison of All Models for Tourist Expenditure From 13 Source Countries to the 31 Chinese Provinces

| Forecast Models       | MAPE Values |          |          |          |          |          |          |          |
|-----------------------|-------------|----------|----------|----------|----------|----------|----------|----------|
|                       |             | 0<10% n | 0<20% n | 20<30% n | 30%+ n   | Total n  |          |          |
| TVP                   | One step ahead | 317     | 78.7     | 57       | 14.1     | 24       | 6.0      | 5        | 1.2      | 403      | 100.0    |
|                       | Two steps ahead | 306   | 75.9     | 70       | 17.4     | 20       | 5.0      | 7        | 1.7      | 403      | 100.0    |
| TVP + interventions   | One step ahead | 319   | 79.2     | 58       | 14.4     | 19       | 4.7      | 7        | 1.7      | 403      | 100.0    |
|                       | Two steps ahead | 295   | 73.2     | 76       | 18.9     | 26       | 6.5      | 6        | 1.5      | 403      | 100.0    |
| Naïve                 | One step ahead | 126   | 31.3     | 64       | 15.9     | 56       | 13.9     | 157      | 39.0     | 403      | 100.0    |
|                       | Two steps ahead | 67    | 16.6     | 67       | 16.6     | 63       | 15.6     | 206      | 51.1     | 403      | 100.0    |
| ARMA                  | One step ahead | 100   | 24.8     | 66       | 16.4     | 46       | 11.4     | 191      | 47.4     | 403      | 100.0    |
|                       | Two steps ahead | 44    | 10.9     | 61       | 15.1     | 64       | 15.9     | 234      | 58.1     | 403      | 100.0    |
| Holt                  | One step ahead | 76    | 18.9     | 76       | 18.9     | 50       | 12.4     | 201      | 49.9     | 403      | 100.0    |
|                       | Two steps ahead | 38    | 9.4      | 54       | 13.4     | 51       | 12.7     | 260      | 64.5     | 403      | 100.0    |
| BSM                   | One step ahead | 64    | 15.9     | 68       | 16.9     | 52       | 12.9     | 219      | 54.3     | 403      | 100.0    |
|                       | Two steps ahead | 35    | 8.7      | 43       | 10.7     | 43       | 10.7     | 282      | 70.0     | 403      | 100.0    |
| BSM + interventions   | One step ahead | 61    | 15.1     | 56       | 13.9     | 57       | 14.1     | 229      | 56.8     | 403      | 100.0    |
|                       | Two steps ahead | 32    | 7.9      | 44       | 10.9     | 39       | 9.7      | 288      | 71.5     | 403      | 100.0    |
| VAR                   | One step ahead | 51    | 12.7     | 61       | 15.1     | 46       | 11.4     | 245      | 60.8     | 403      | 100.0    |
|                       | Two steps ahead | 30    | 7.4      | 31       | 7.7      | 43       | 10.7     | 299      | 74.2     | 403      | 100.0    |

This finding is quite different from the Zhou-Grundy and Turner’s (2014) study on forecasting arrivals, even though the variables and models used are not greatly dissimilar. In Zhou-Grundy and Turner’s study the BSM and BSM with interventions were the most accurate models, although this was followed by the TVP models. In addition, the Naïve 1 method performed far better in forecasting tourism expenditure at the provincial level than it did forecasting tourist arrivals.

However, although the TVP and TVP with intervention methods performed well overall, the results vary across the provinces. Table 3 shows the most accurate forecast model based on the lowest MAPE value for each province and source country for two steps ahead. For instance, the lowest MAPE error values of forecasts for Beijing province in relation to expenditure by English and German tourists were the naïve method while the
BSM forecast of expenditure by American and German tourists in Tianjin had the lowest MAPE error. In the majority of instances, it can be clearly seen that both TVP methodologies had the lowest MAPE error values.

Table 3

**Optimal Model for Tourist Expenditure From 13 Source Countries to the 31 Chinese Provinces Based on Lowest MAPE Value (Two Steps Ahead)**

| Source Country | Japan | Korea | Malaysia | Philippines | Singapore | Thailand | USA | Canada | UK | France | Germany | Russia | Australia |
|----------------|-------|-------|----------|-------------|-----------|----------|-----|--------|----|--------|---------|--------|----------|
| Beijing        | TVP   | Ti    | Ti       | TVP         | TVP       | TVP      | Ti  | Naïve  | TVP| Naïve  | Ti      | Ti     | TVP      |
| Tianjin        | TVP   | Ti    | ARMA     | Naïve       | TVP       | TVP      | Ti  | Ti     | Ti | Ti     | Ti      | Ti     | TVP      |
| Hebei          | TVP   | VAR   | TVP      | Ti          | Ti        | Ti       | Ti  | Ti     | Ti | TVP    | Ti      | Ti     | TVP      |
| Shanxi         | TVP   | TVP   | Ti       | TVP         | TVP       | TVP      | TVP | Ti     | TVP| TVP    | Ti      | Ti     | TVP      |
| Inner Mongolia | TVP   | Holt  | BSM      | T/Ti        | T/Ti      | T/Ti     | Ti  | Ti     | Ti | Ti     | TVP     | TVP    | TVP      |
| Liaoning       | T/Ti  | TVP   | TVP      | TVP         | TVP       | TVP      | TVP | T/Ti   | TVP| TVP    | TVP     | TVP    | TVP      |
| Jilin          | TVP   | BSM   | Ti       | TVP         | Ti        | T/Ti     | Ti  | Ti     | Ti | TVP    | TVP     | TVP    | TVP      |
| Heilongjiang   | T/Ti  | Ti    | T/Ti     | T/Ti        | Naïve     | ARMA     | TVP | VAR    | Ti | Bi     | TVP     | Ti     | Naïve    |
| Shanghai       | TVP   | TVP   | ARMA     | Ti          | TVP       | T/Ti     | TVP | Ti     | Naïve| TVP    | TVP     | TVP    | TVP      |
| Jiangsu        | T/Ti  | Ti    | T/Ti     | Ti          | Ti        | T/Ti     | Ti  | Ti     | Ti | Ti     | Ti      | Ti     | TVP      |
| Zhejiang       | Ti    | TVP   | Ti       | Ti          | Ti        | TVP      | TVP | Ti     | TVP| Ti     | Ti      | Ti     | TVP      |
| Anhui          | T/Ti  | TVP   | T/Ti     | T/Ti        | TVP       | TVP      | Ti  | T/Ti   | Ti | Ti     | Ti      | Ti     | TVP      |
| Fujian         | Ti    | Ti    | Ti       | TVP         | Ti        | Ti       | Ti  | TVP    | TVP| TVP    | TVP     | TVP    | TVP      |
| Jiangxi        | TVP   | TVP   | TVP      | TVP         | TVP       | TVP      | Ti  | TVP    | Naïve| TVP    | TVP     | TVP    | TVP      |
| Shandong       | Ti    | Ti    | Naïve    | Naïve       | Ti        | Ti       | Ti  | TVP    | TVP| TVP    | TVP     | TVP    | TVP      |
| Henan          | Ti    | TVP   | Ti       | Ti          | T/Ti      | Ti       | Ti  | TVP    | Ti | Ti     | Ti      | Ti     | TVP      |
| Hubei          | Holt  | ARMA  | TVP      | TVP         | TVP       | TVP      | Ti  | Ti     | Ti | TVP    | TVP     | TVP    | TVP      |
| Hunan          | Ti    | TVP   | TVP      | TVP         | TVP       | TVP      | Ti  | Ti     | T/Ti| TVP    | TVP     | Ti     | Naïve    |
| Guangdong      | VAR   | Ti    | Ti       | Ti          | Ti        | VAR     | Ti  | TVP    | TVP| Ti     | Ti      | Ti     | TVP      |
| Guangxi        | Ti    | Ti    | Ti       | TVP         | Naïve     | TVP      | TVP | TVP    | Ti | Ti     | Ti      | Ti     | TVP      |
| Hainan         | T/Ti  | TVP   | TVP      | TVP         | T/Ti      | Ti       | Ti  | ARMA   | VAR| Holt   | TVP     | T/Ti   | Holt     |
| Chongqing      | TVP   | ARMA  | T/Ti     | Holt        | T/Ti      | ARMA     | T/Ti| T/Ti   | T/Ti| T/Ti   | TVP     | T/Ti   | TVP      |
| Sichuan        | TVP   | T/Ti  | VAR      | BSM         | VAR       | Bi       | Ti  | TVP    | Bi | Ti     | TVP     | TVP    | BSM      |
| Guizhou        | Naïve | ARMA  | Naïve    | Naïve       | Naïve     | Naïve    | Holt| Naïve  | ARMA| Bi     | Ti      | Naïve  | TVP      |
| Yunnan         | TVP   | Ti    | T/Ti     | T/Ti        | Ti        | T/Ti     | Ti  | T/Ti   | Ti | Ti     | TVP     | TVP    | TVP      |
| Tibet          | Ti    | Ti    | Ti       | TVP         | TVP       | ARMA     | Ti  | TVP    | BSM| VAR    | TVP     | Naïve  | BSM      |
| Shanxi         | TVP   | Ti    | Ti       | Ti          | Ti        | TVP      | Ti  | Ti     | Ti | TVP    | Ti      | TVP    | TVP      |
| Gansu          | Ti    | T/Ti  | T/Ti     | T/Ti        | TVP       | TVP      | Ti  | Ti     | Ti | TVP    | T/Ti    | Ti     | T/Ti     |
| Qinghai        | Ti    | Ti    | TVP      | Ti          | Ti        | TVP      | Ti  | Ti     | Ti | TVP    | Ti      | TVP    | Ti       |
| Ningxia        | Naïve | Naïve | TVP      | T/Ti        | Ti        | Bi      | Ti  | TVP    | Ti | TVP    | Ti      | Ti     | TVP       |
| Xinjiang       | TVP   | TVP   | TVP      | TVP         | T/Ti      | T/Ti     | T/Ti| TVP    | Ti | Ti     | T/Ti    | T/Ti   | T/Ti     |

*Legend: T = TVP, Ti = TVP with intervention, B = BSM, Bi = BSM with intervention*
Results for Regions

Table 4 summarizes MAPE values for all models for tourist expenditure from the 13 source countries to the six Chinese regions for both one step ahead and two steps ahead. There are a total of $6 \times 13 = 78$ forecast series for each model. For example, the TVP forecast has 46 (59.0%) of the best MAPE counts below 10% compared with the BSM with intervention method which only has 4 (5.1%) overall MAPE counts in this band for two steps ahead. The results are again repeated for the 13 source countries to each of the six Chinese regions. The TVP and TVP with intervention methods of forecasting tourism expenditure to the regions had over 85% of forecasts below 20% mean MAPE error.

Table 4
One and Two Steps Ahead MAPE Comparison of All Models For Tourist Expenditure From 13 Source Countries to the Six Chinese Regions

| Forecast Models | MAPE values | Total |
|-----------------|-------------|-------|
|                 | 0% - < 10% | 10% - < 20% | 20% - < 30% | 30% + | n | % |
| **One step ahead** | | | | | | |
| TVP             | 55          | 70.5 | 19 | 24.4 | 2 | 2.6 | 2 | 2.6 | 78 | 100.0 |
| Two steps ahead | 46          | 59.0 | 25 | 32.1 | 5 | 6.4 | 2 | 2.6 | 78 | 100.0 |
| TVP + interventions | 55          | 70.5 | 16 | 20.5 | 4 | 5.1 | 3 | 3.8 | 78 | 100.0 |
| Two steps ahead | 46          | 59.0 | 21 | 26.9 | 8 | 10.3 | 3 | 3.8 | 78 | 100.0 |
| Naïve           | 26          | 33.3 | 16 | 20.5 | 6 | 7.7 | 30 | 38.5 | 78 | 100.0 |
| Two steps ahead | 21          | 26.9 | 15 | 19.2 | 11 | 14.1 | 31 | 39.7 | 78 | 100.0 |
| ARMA            | 20          | 25.6 | 18 | 23.1 | 7 | 9.0 | 33 | 42.3 | 78 | 100.0 |
| Two steps ahead | 17          | 21.8 | 12 | 15.4 | 12 | 15.4 | 37 | 47.4 | 78 | 100.0 |
| Holt            | 14          | 17.9 | 18 | 23.1 | 6 | 7.7 | 40 | 51.3 | 78 | 100.0 |
| Two steps ahead | 9           | 11.5 | 14 | 17.9 | 10 | 12.8 | 45 | 57.7 | 78 | 100.0 |
| BSM             | 12          | 15.4 | 14 | 17.9 | 8 | 10.3 | 44 | 56.4 | 78 | 100.0 |
| Two steps ahead | 5           | 6.4 | 11 | 14.1 | 10 | 12.8 | 52 | 66.7 | 78 | 100.0 |
| VAR             | 13          | 16.7 | 12 | 15.4 | 6 | 7.7 | 47 | 60.3 | 78 | 100.0 |
| Two steps ahead | 5           | 6.4 | 12 | 15.4 | 13 | 16.7 | 48 | 61.5 | 78 | 100.0 |
| BSM + interventions | 7          | 9.0 | 13 | 16.7 | 14 | 17.9 | 44 | 56.4 | 78 | 100.0 |
| Two steps ahead | 4           | 5.1 | 3 | 3.8 | 17 | 21.8 | 54 | 69.2 | 78 | 100.0 |

As with the provinces, the most accurate forecasting model based on the lowest MAPE error value for each region and source country varies for two steps ahead. Table 5 shows the best performing model based on the lowest MAPE error for each region and source country. The ARMA model performed well in the North Region in forecasting tourist expenditure by Korean, American, and Russian tourists. Again, as with the provincial results, both TVP methodologies provided the most accurate forecast results.

Although Tables 3 and 5 show that the models performed differently in each of the provinces and regions by source country it should be emphasized that they were based on the lowest MAPE value and in many cases the TVP with and without intervention models returned MAPE errors approaching those that performed slightly better. Taken into consideration if one was to select a single model to forecast all provinces or all regions then the TVP or the TVP with intervention would be the most appropriate selection based on the forecast results in total (refer to Table 2 for provinces and Table 4 for regions), for forecasting expenditure whilst Zhou-Grundy and Turner found that the BSM and BSM with interventions are best for forecasting arrivals.
It should be noted that 2013 saw a significant decline in the number of tourist arrivals from the major source countries. For example, there was an 18.2% decrease in the number of Japanese tourists visiting China in 2013, a major source of tourist arrivals and revenue. In addition, many provinces saw declines in foreign tourist arrivals. Heilongjiang saw a 26.4% decline in tourist arrivals with a 27.7% decline in revenue. Similarly, Liaoning and Beijing also saw declines in tourist arrivals of 14.6% and 10.1% respectively. This downturn has had a significant impact on the forecasts from the various models evaluated in this research as the test period included 2013. Further forecasting beyond this point in the future would normally consider this downturn and treat it with an intervention for the TVP and BSM with intervention models tested here. As such the forecasting was made more difficult but the results show that the TVP methodology was able to more accurately forecast the results, even with a downturn in tourist arrivals and expenditure.

**Causal Variables**

The research aimed to answer the question, what variables can be currently used to achieve an accurate econometric forecast at a regional level? In addition, could those variables also be apt for econometric forecasting at a provincial level? Table 6 shows the percentage of frequency for each of the independent variables as found from the TVP model for each province and region from the 13 source countries.

Table 6 shows that all the causal variables examined in this study have contributed to forecast accuracy particularly with the TVP and TVP with intervention models. The selection of the causal variables in this study may suggest why the TVP model performed significantly better than the other time series models under investigation. Current and past research studies have generally concluded that the BSM and BSM with causal variables, commonly referred to as Structural Time Series Models (STSM) or Causal Structural Model (CSM), in the short-run are far superior to econometric models (Kulendran & Witt, 2003; Blake, Durbary, Eugenio-Martin, Gooroochurn, Hay, Lennon, & Yeoman, 2006; Shen, Li, & Song, 2009; Zhou-Grundy & Turner, 2014). However, the findings from this research reflect a different picture, that the econometric TVP model is the more accurate model in forecasting tourism expenditure for Chinese provinces and regions in the short term. An explanation for the poorer performance of econometric models in previous research may possibly be due to less significant causal variables, given the analytic advantages of the varying parameters within the TVP model are similar in all cases.
Conclusion and Limitations

Regardless of the data being limited to annual data and narrow in time series length, the results show that very accurate forecasts (defined as below 10% error) are obtainable for tourism expenditure at both China’s regional and provincial level in the short term (two years ahead). Although the time series models performed poorly against the naïve method, most notably the BSM with intervention method and VAR model, the empirical results demonstrate that for the two-year ahead forecasting horizon, the TVP with and without
intervention generated the most accurate forecasts of tourism expenditure at both the provincial and regional level.

The most accurate time series models under comparison were the ARMA and Holt models, based on MAPE error values for two steps ahead. However, there is a limitation in using annual data and an ARIMA and BSM model using seasonal data may make for a stronger comparison. The VAR model performed poorly when compared with other methods. This may be due partly to the short forecast period and also to the small data set. The data set employed in this research limited the number of lags possible for each variable and the number of variables that can be employed in each model. The naïve method proved more accurate than the time series models under investigation in forecasting expenditure at both the provincial and regional levels (refer to Table 4 and Table 6).

The analysis of the causal variables shows that they have a varying level of importance in modelling tourism expenditure for each of the provinces and regions (refer to Table 6). The importance of the variables seems to be specific to each province and region. This is reflected in Table 6 (Causal variables as percentage of frequency by province and region for 13 source countries). The table is an important finding for forecasters and policy makers as well as provincial and regional governments wishing to gain a greater share of tourism receipts in China. It provides a selection of key variables suitable for forecasting tourism expenditure within each province and region. They also provide guidance in variable selection to future research in forecasting regional tourism expenditure in other countries.

This study aids future researchers in providing the results from comparative model selection, for provincial and regional forecasting should they wish to compare other forecasting methods.

The limitation to annual data is a very important data constraint because seasonality remains unmeasured. Additionally, as in other countries the expenditure data in China are dependent on sample surveys. As with most sample surveys of this nature sample size is often relatively small especially when viewed at the provincial level. The expenditure by visitor type, tourist, visiting family and friends, businessperson, and so forth, and source country was not available at the regional or provincial level (only available at an aggregated sum for each province and region), consequently expenditure estimates used in this study were based on visitors generally rather than visitor type. The significance of this is that different visitor types have different spending patterns and can have an impact on the results.

As data become more available further research could be undertaken to examine the medium- and long-term forecasting capabilities of the models presented. Most notably, a comparison of the TVP with interventions forecasting performance for the short term could be made against its forecasting capabilities for periods greater than two years. In addition, should expenditure data be provided by visitor type and by province become available, further research could be made to estimate whether the assumptions of this research, that is treating all visitors alike in their expenditure patterns, is a valid assumption.

References

Akal, M. (2004). Forecasting Turkey’s tourism revenues by ARMAX model. Tourism Management, 25(5), 565-580.
Au, N., & Law, R. (2000). The application of rough sets to sightseeing expenditures. Journal of Travel Research, 39, 70-77.
Blake, A., Durbary, R., Eugenio-Martin, J., Gooroochurn, N., Hay, B., Lennon, J., Sinclaira, M. T., Sugiyartoe, G., & Yeoman, I. (2006). Integrating forecasting and CGE models: The case of tourism in Scotland. Tourism Management, 27(2), 292-305.
Coshall, J. T. (2005). A selection strategy for modelling UK tourism flows by air to European destination. Tourism Economics, 11, 141-158.
Crouch, G. I. (1994). The study of international tourism demand: A review of practice. *Journal of Travel Research*, 33, 41-54.

Frechtling, D. C. (2006). An assessment of visitor expenditure methods and modes. *Journal of Travel Research*, 20, 1-10.

Gao, S., Huang, S., & Huang, Y. (2009). Rural tourism development in China. *International Journal of Tourism Research*, 11, 439-450.

Georgantopoulos, A. G. (2012). Forecasting tourism expenditure and growth: A VAR/VECM analysis for Greece at both aggregated and disaggregated levels. *International Research Journal of Finance and Economics*, 96, 155-167.

Goh, C., & Law, R. (2011). The methodological progress of tourism demand forecasting: A review of related literature. *Journal of Travel and Tourism Marketing*, 28(3), 296-317.

Hu, Y. (2008). Ecotourism and sustainable development of rural tourism. *Gansu Agriculture*, 8, 22-23.

Kim, S. S., Han, H., & Chon, K. (2008). Estimation of the determinants of expenditures by festival visitors. *Tourism Analysis*, 13(5), 387-400.

Kulendran, N., & Witt, S. (2003). Leading indicator tourism forecasts. *Tourism Management*, 24(5), 503-510.

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

Lew, A., & Yu, L. (Eds.). (1995). *Tourism in China: Geographic, political and economic perspectives*. Oxford: Westview Press.

Lew, A., Yu, L., Ap, J., & Zhang, G. (Eds.). (2003). *Tourism in China*. New York: Haworth.

Lewis, C. D. (1982). *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*. London: Butterworth Scientific.

Li, G., Song, H., & Witt, S. (2005). Recent development in econometric modelling and forecasting. *Journal of Travel Research*, 44, 82-99.

Li, P. (2008). Study and analysis on alleviating poverty by tourism in poor regions. *Social Sciences Review*, 23(5), 53-54.

Lim, C. (1997a). An econometric classification and review of international tourism demand models. *Tourism Economics*, 3, 69-81.

Lim, C. (1997b). Review of international tourism demand models. *Annals of Tourism Research*, 24(4), 835-849.

Lim, C. (1999). A meta-analytic review of international tourism demand. *Journal of Travel Research*, 37, 203-220.

Lim, C. (2006). A survey of tourism demand modelling practice: Issues and implications. In L. Dwyer and P. Forsyth (Eds.), *International handbook on the economics of tourism*. Cheltenham: Edward Elgar.

Louw, R. (2011). Forecasting tourism demand for South Africa (Doctoral dissertation), North-West University.

Luzzi, G. F., & Flückiger, Y. (2003). An econometric estimation of the demand for tourism: The case of Switzerland. *Pacific Economic Review*, 8(3), 289-303.

Pine, R. (2002). China’s hotel industry: Serving a massive market. *Cornell Hotel and Restaurant Administration Quarterly*, 43(3), 61-70.

Sheldon, P. J. (1993). Forecasting tourism: Expenditures versus arrivals. *Journal of Travel Research*, 32(1), 13-20.

Shen, S., Li, G., & Song, H. (2009). Effect of seasonality treatment on the forecasting performance of tourism demand models. *Tourism Economics*, 15(4), 693-708.

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. *Tourism Management*, 29(2), 203-220.

Song, H., Li, G., Witt, S. F., & Fei, B. (2008). Tourism demand modelling and forecasting: How should demand be measured? *Tourism Economics*, 16(1), 63-81.

Song, H., & Turner, L. (2006). Tourism demand forecasting. In L. Dwyer and P. Forsyth (Eds.), *International handbook on the economics of tourism*. Cheltenham: Edward Elgar.

Song, H., & Witt, S. F. (2005). Forecasting international tourism flows to Macau. *Tourism Management*, 27, 214-224.

Song, H., Witt, S. F., Wong, K. F., & Wu, D. C. (2009). An empirical study of forecast combination in tourism. *Journal of Hospitality and Tourism Research*, 33(1), 3-29.

Turner, L., & Witt, S. (2001a). Factors influencing demand for international tourism: Tourism demand analysis using structural equation modelling. *Tourism Economics*, 7, 21-38.

Turner, L., & Witt, S. (2001b). Forecasting tourism using univariate and multivariate structural time series models. *Tourism Economics*, 7, 135-147.

Turner, L., & Witt, S. (2010a). *Asia Pacific tourism forecasts 2011-2013*. Pacific Asia Travel Association, Bangkok, Soft Cover, p. 321 (Also Chinese Edition, Beijing).

Turner, L., & Witt, S. (2010b). *Regional tourism forecasts 2011-2013*. Pacific Asia Travel Association, Bangkok, Soft Cover, p. 137 (Also Chinese Edition, Beijing).
Vietze, C. (2009). What’s pushing international tourism expenditures? *Jena Economic Research Papers, 2009-014.*

Vu, C., & Turner, L. (2006). Regional data forecasting accuracy: The case of Thailand. *Journal of Travel Research, 45*(2), 186-193.

Vu, T. T. M. (2010). Forecasting international regional arrivals in Canada (Master’s thesis). Victoria University.

Wang, Y., & Davidson, M. (2010). A review of micro analyses of tourist expenditure. *Current Issues in Tourism, 13*(6), 507-524.

Wang, Y., Rompf, P., Severt, D., & Peerapatdit, N. (2006). Examining and identifying the determinants of travel expenditure patterns. *International Journal of Tourism Research, 8*(5), 333-346.

Witt, S., & Witt, C. (1995). Forecasting tourism demand: A review of empirical research. *International Journal of Forecasting, 11*(3), 447-475.

WTTC. (2012). *Travel and tourism economic impact 2016—China.* World Travel and Tourism Council. Retrieved from http://www.wttc.org/-/media/files/reports/economic-impact-research/countries-2016/china2016.pdf, on December 9, 2016

Yaffee, R. A. (2010). Forecast evaluation with Stata. *United Kingdom Stata Users Group Conference, September 9, 2010.*

Yang, Y. (2009). A spatial econometric approach to model the growth of tourism flows to China cities (Doctoral dissertation). Hong Kong Polytechnic University.

Zhao, C. (2008). Problems and solutions of China’s sustainable development in rural tourism. *Agricultural Economy, 4*, 18-19.

Zhou-Grundy, Y., & Turner, L. W. (2014). The challenge of regional tourism demand forecasting: The case of China. *Journal of Travel Research, 53*(6), 747-759