

Article

Analyzing the Influence of Transportations on Chinese Inbound Tourism: Markov Switching Penalized Regression Approaches

Woraphon Yamaka ¹, Xuefeng Zhang ² and Paravee Maneejuk ^{1,*}¹ Center of Excellence in Econometrics, Chiang Mai University, Chiang Mai 50200, Thailand; woraphon.econ@gmail.com or woraphon.yamaka@cmu.ac.th² Faculty of Economics, Chiang Mai University, Chiang Mai 50200, Thailand; xuefengzhang999@gmail.com or xuefeng_zhang@cmu.ac.th

* Correspondence: mparavee@gmail.com or paravee.m@cmu.ac.th

Abstract: This study investigates the nonlinear impact of various modes of transportation (air, road, railway, and maritime) on the number of foreign visitors to China originating from major source countries. Our nonlinear tourism demand equations are determined through the Markov-switching regression (MSR) model, thereby, capturing the possible structural changes in Chinese tourism demand. Due to many variables and the limitations from the small number of observations confronted in this empirical study, we may face multicollinearity and endogeneity bias. Therefore, we introduce the two penalized maximum likelihoods, namely Ridge and Lasso, to estimate the high dimensional parameters in the MSR model. This investigation found the structural changes in all tourist arrival series with significant coefficient shifts in transportation variables. We observe that the coefficients are relatively more significant in regime 1 (low tourist arrival regime). The coefficients in regime 1 are all positive (except railway length in operation), while the estimated coefficients in regime 2 are positive in fewer numbers and weak. This study shows that, in the process of transportation, development and changing inbound tourism demand from ten countries, some variables with the originally strong positive effect will have a weak positive effect when tourist arrivals are classified in the high tourist arrival regime.

Keywords: transportation modes; tourism demand; markov switching model; ridge and lasso estimation; China



Citation: Yamaka, W.; Zhang, X.; Maneejuk, P. Analyzing the Influence of Transportations on Chinese Inbound Tourism: Markov Switching Penalized Regression Approaches. *Mathematics* **2021**, *9*, 515. <https://doi.org/10.3390/math9050515>

Academic Editors:
José Álvarez-García, Oscar V. De la Torre-Torres and María de la Cruz del Río-Rama

Received: 8 February 2021

Accepted: 25 February 2021

Published: 2 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As travel and tourism is one of the world's largest economic sectors, it has played an essential role in contributing to job creation and economic growth and creating prosperity worldwide [1–3]. In China, tourism has become an important contributor to the domestic economy since implementing reform and opening-up policies in the early 1980s. Among global destinations, China ranks fourth in total inbound tourists (Forbes, 2019). China Tourism Academy (2020) reported that the Chinese tourism industry generates 79.87 million jobs and 10.94 trillion yuan in revenue, accounting for 10.31% and 11.05% of total employment, and China's GDP, respectively. Therefore, the Chinese government takes tourism development into account when making essential policies on economic growth.

One of the critical factors driving tourism development is transportation, of which its importance to tourism development has been widely recognized [4–7]. Since the Chinese reform and opening-up policies were first put into action, the issue of investment in transportation has received considerable attention as a prerequisite for tourism development, and transportation development has been placed at a high priority until today. There is a general agreement that tourism expands more when there are better transportation systems. Recently, China has provided advanced tourism services for international tourists inbound and developed/constructed and improved the domestic transportation infrastruc-

ture, including international airports, high-speed railways, super-highways, expressways, and tunnel-roads to link tourists with various tourist attractions.

With the importance of the travel and tourism sector, the links between transportation and tourism have attracted more attention and debate among scholars in various fields [4–8]. Nevertheless, their empirical findings seem to reveal ambiguous results. Some studies indicated that transportation has a positive impact on tourism; some argued that there is no transportation contribution to tourism development, while others mentioned that the magnitude of the transportation effects is different in different periods. Since previous research results did not reach a consistent conclusion, we doubt that these ambiguous results are due to the use of different transportation indicators or modes, such as railway transportation [9–12], road transportation [13,14], air transportation [6,7,15], and maritime transportation [16,17]. In the case of maritime transportation, given the small number of ocean liners currently traveling to and from China, this transportation may weak impact the international tourism demand. However, this transportation allows tourists to travel to several port destinations, enjoy various amenities provided onboard, and participate in leisure programs. Park, Lee, Moon and Heo [18] mentioned that Chinese maritime transportation had grown rapidly, especially in the past decade, with the number of cruise tourists rising noticeably.

Therefore, it is more suitable to investigate the impact of different transportation modes on tourism development. In addition, it is a fact that these transportation modes have different times and costs of construction, while the government budget is limited. Hence, understanding the heterogeneous impacts of transportation modes may lead to efficient budget policies. Unfortunately, very few studies are available for China, which econometrically analyze the impact of different transportation modes on tourism development.

This study aims to fill this gap in the literature by ascertaining precisely which transportation modes play an important role in the number of inbound visitors. To accomplish our goal, four different modes of transportation, railway, land, air, and maritime, along with the macroeconomic determinants as control variables, are considered in this study. We analyze these factors' influence on the number of visitors from China's ten main international tourism sources, namely South Korea, Japan, Russia, USA, Mongolia, Malaysia, Philippines, Singapore, India, and Canada. The government can more confidently plan and develop tourism-related transportations based on a deeper understanding of tourism demand. From the methodological point of view, there has been no consensus regarding whether transportation has an asymmetric effect on tourism development. We suspect that the influence of transportation on tourism may not be stable over time, and the impact is unlikely to maintain its linear state with the expansion of transportations. Although transportations have affected tourism development in China, their structures and purposes differ markedly. Therefore, we suspect there is a nonlinear relationship between the scale of transportation and tourism development. In the literature, machine learning approaches and nonlinear econometric approaches can capture the nonlinear relationship between input and output variables. Huang, Ma, and Hu [19] mentioned that machine learning approaches are a subfield of computer science that evolved from the study of pattern recognition and computational learning theory in artificial intelligence. They also suggested that these approaches could provide high forecasting performance.

Although the machine learning approaches enable us to model the nonlinear relationship between transportation modes and China's international tourism, the economic interpretation of the weight parameters in the machine learning models is limited and difficult. Hence, in this study, we instead employ the nonlinear econometric model called the Markov-switching regression (MSR) model to verify the nonlinear impact of different transportation modes on China's top ten inbound tourists. The study also takes the advantages of this model in allowing us to analyze and compare the business cycle patterns of China's main sources of international tourism. The MSR model's ability to detect the business cycle dates is described by well-known statistical institutions (namely the Economic Cycle Research Institute, ECRI, and the National Bureau of Economic Research,

NBER) [20]. Due to the different transportation impacts on tourism, the government and authorities should consider which type of transportation policy to implement, in order to manage structural changes, highlighting the importance of investigating the inbound tourism market business cycle. We also compared the business cycle patterns of China's ten main sources of international tourism. The results may reveal the reasons why significant differences arose in the structural changes in international tourism's when transportation infrastructures are developed.

The rest of the paper is organized as follows, with our literature review in Section 2. Section 3 describes the data and the methodology for empirical models. Section 4 presents the empirical results and conclusions with a discussion of major findings and policy implications are presented in Section 5.

2. Literature Review

Transportation and tourism are closely related to economic activities; the literature has increased its attention in relation to transportation's vital role in tourism development. In the last two decades, many studies have investigated the impact of various transportation modes, namely road, air, maritime, and railway, on tourism development, but controversial results exist on whether these transportation modes promote tourism. Some studies have revealed the positive influence on tourism, where some indicated the negative effect of transportation on tourism development, and others suggested that transportation has a supporting role at different magnitudes and at different times.

Many studies revealed that transportation plays a vital role in tourism development [5,21]. Khadaroo and Seetanah [6] considered air, land, and sea transports as factors affecting the tourism demand of Mauritius's island. They found that tourists from Asia, Europe, and America are susceptible to the island's transportation. Whereas, Pagliara, Mauriello, and Gomez [22]; Albalade and Fageda [9]; Yin, Pagliara, and Wilson [10]; Yin et al. [12]; and Pagliara, Mauriello, and Russo [11] have produced evidence of the significant effect of high-speed rail on tourist outcomes. Yin et al. [12] documented that high-speed rail improves the accessibility of the tourist to the destination due to the reduction of the tourist's traveling time. Button and Taylor [23] put air transportation in the tourism demand regression and conclude that it is essential in attracting international tourists. Air transport has expanded to new destinations, as well as reduced traveling time. In recent years, global tourism has developed rapidly, and the global air transport network is the main positive factor affecting inbound and outbound tourism [24]. Kanwal et al. [14] investigated the road and transportation infrastructure construction and the community support for tourism in the context of the China-Pakistan economic corridor (CPEC). They mentioned that road infrastructure and road transportation play a great role in tourism development for providing all possible destinations' accessibility. Hardy [25] also suggested that touring routes or self-drive trails have become a more popular attraction, and a better road infrastructure could provide a more satisfactory tourism experience. Although different transportation modes are considered in the tourism literature, the previous studies concluded that improved transportation usually increases a tourist destination's accessibility and increases tourist demand.

In recent years, transportation has played a significant role in explaining China's tourism development as well. For example, Chen and Haynes [26] investigated the impact of the Chinese high-speed rail on international tourism demand using a dynamic panel regression model of 21 countries. They confirmed that high-speed rail produces a weak positive effect on tourist arrivals. Li, Yang, and Cui [27] also conducted a panel regression analysis to provide a comprehensive analysis of the impact of high-speed rail on tourist arrivals in China. They found that high-speed rail has a stronger effect on international arrivals than on domestic ones.

Moreover, they found a heterogeneous impact of high-speed rail on tourist arrivals across regions. Dong et al. [28] put the road length (km) per area as the transportation indicator in the tourism demand equation, and the result shows a significant effect of road

on tourist arrivals. Jin et al. [29] studied the impact of high-speed rail on ice and snow tourism in Northeast China. They pointed out that after the implementation of high-speed rail, tourism economic ties between cities have increased. The isochron analysis shows that the central city promotes the tourism development of the surrounding cities, the number of daily and weekend trips increases significantly, and the shortest travel time between cities is significantly shortened. However, air and maritime transportations on tourism development are found to be limited for China.

There is still a controversy among the researchers despite the decisive evidence of the positive impact of transportation on tourism development. The studies of Pagliara et al. [22] and Albalade and Fageda [9] revealed a negative impact of high-speed rail on tourist arrivals. They identified several reasons why the effect of high-speed rail on tourist arrivals is not positive. First, the railway network is not appropriately designed and does not correspond to the riders' needs. Second, there is a substitution effect of air transportation. A similar finding is reached by Chen and Haynes [26]. They analyzed and found the weak impact of Chinese high-speed rail on international tourism demand. Some empirical studies have found that transportation improvement may hamper tourism development as the length of stay or traveling is shorter, resulting in reduced travel expenditure [30].

This contradiction among the researchers creates a motivation for further investigation of the transportation–tourism nexus. In this study, we attempt in this direction, and China, which is the fourth most visited country in the world, is chosen. Previous research studies have considered various empirical and methodological aspects of modeling the effect of transportation on tourism. Nevertheless, they have been unable to consider the presence of a structural shift in tourism. Specifically, the relationship between tourism and any transportation has been assumed to be linear and constant over time. Due to this aspect, we suspect that the linear regression models, which were generally used in the previous studies, may oversimplify the transportation–tourism nexus. Hence, an adequate understanding and estimations of structural shifts are essential for formulating and introducing policies on tourism development [30–32]. To deal with the possible nonlinear effect of transportation, as well as the structural change in tourism, the Markov-switching regression (MSR) model is used in this study. The nonlinear impacts of different transportation modes on the top ten inbound tourism have not been econometrically analyzed in China's case to the best of our knowledge. Specifically, this is the first attempt in determining whether a different transportation mode exerts different influences on international tourist arrivals to China.

Another contribution that can be found in this study is an econometric one. Although the MSR approach could solve the issue of structural changes, insufficient sample data (overparameterization), multicollinearity, and endogeneity problems may still occur in our model. We would like to note that, in the literature, annual transportation data and tourism data are generally collected, and there might be an interaction among the transportation modes through the “substitution effect” [9] and “complementary effect” between modes. In addition, there is a mixture of the unit of measurements for each transportation mode, which can still lead to aggregation bias even after the logged values are used in the analysis. One way to ameliorate these concerns, we contribute to the academic literature by introducing regularization methods like ridge estimation of Hoerl and Kennard [33] and lasso estimation of Tibshirani [34] to estimate all unknown parameters in the MSR model. Ridge and lasso have several advantages over the traditional unpenalized maximum likelihood as they reduce complexity and prevent over-fitting of a model. In particular, lasso estimation makes the model easier to interpret by eliminating irrelevant variables that are not associated with the response variable (variable selection). Also, it reduces the extent of the problem in enabling the maximum likelihood to work faster, making it possible to handle high-dimensional data [35]. The estimated coefficients from the MSR model based on these two methods would reveal interesting findings and destination management implications that previous studies have not explored.

3. Methodology

3.1. Markov-Switching Regression Model

There are corresponding theoretical supports to distinguish the number of international tourist arrivals into high tourist arrivals and low tourist arrivals [31,32,36]. In the literature, many studies were mostly distinguished the economic climate into two or three states. However, when there are more than two states in the Markov-switching model, the model will be too complicated to properly distinguish the characteristics of the data and make all the divided state types lack persistence and high volatility. The different states of tourism can be described by the unobserved state or regime variable, which takes the discrete values of 0 and 1 [36,37]. The model structure of the MS-regression model is defined as,

$$y_t = x_t \beta(s_t) + \sigma^2(s_t) u_t, \quad (1)$$

where y_t is the dependent variable at time t ; x_t is a $k \times 1$ vector of observed exogenous or independent variables at time t ; $\beta(s_t)$ is a $k \times 1$ vector of regime-dependent unknown coefficients. u_t is an error term at time t , which has a standard normal distribution, $u_t \sim N(0, 1)$, and $\sigma^2(s_t)$ is regime-dependent unknown variance at time t . Note that the unobserved state variable S_t is evolved according to a first-order Markov chain with the transition probabilities of,

$$Pr(s_t = 1 | s_t = 1) = p_{11} \text{ and } Pr(s_t = 2 | s_t = 2) = p_{22}, \quad (2)$$

where p_{11} is the probability of remaining in regime 1, while p_{22} is the probability of remaining in regime 2.

3.2. Estimation

In recent years, there has been significant interest in the MSR models, and it has led to the development of efficient estimation for estimating the parameters, for example, expected maximization algorithm [38], maximum likelihood estimation [39], and Bayesian estimation [40]. However, the number of parameter estimates of the MS models is quite high compared to other linear models; thus, these conventional estimations may provide excessive modeling biases and cause low predictability. In particular, when the number of estimated regime-dependent coefficients k is substantially large, the computation becomes much more challenging and difficult.

In this study, two penalty functions are proposed, one is a ridge, and the other is a lasso, for estimating all unknown parameters $\Theta(s_t) = \{\beta(s_t), \sigma(s_t), p_{11}, p_{22}\}$ in the MS regression model. It is known that y_t and x_t are directly observed but s_t is an unobserved latent state variable and is inferred by what is happening with y_t and x_t . Let $\theta = \{\beta(s_t), \sigma(s_t)\}$, this inference will take the form of two regime probabilities,

$$\pi_{1,t} = Pr(s_t = 1 | y_t, Z_{t-1}; \theta(s_t)), \quad (3)$$

$$\pi_{2,t} = 1 - \pi_{1,t} \quad (4)$$

where Z_{t-1} is the information available at time $t - 1$ as well as x_t . These two regime probabilities can be filtered by Hamilton's filter [41] as in the following steps:

- **Step 1:** We employ the following steady-state or unconditional probabilities to find the initial probability $\pi_{1,0}$ and $\pi_{2,0}$:

$$\pi_{1,0} = \frac{1 - p_{11}}{2 - p_{11} - q_{22}}, \quad (5)$$

$$\pi_{2,0} = \frac{1 - p_{22}}{2 - p_{11} - q_{22}}, \quad (6)$$

- **Step 2:** After computing the inferences for the regimes at time $t = 0$, we can update the inferences for the regimes at time $t = 1$ as,

$$\pi_{1,1} = \frac{\pi_{1,0}f(y_t|Z_{t-1};\theta(s_t = 1))}{\pi_{1,0}f(y_t|Z_{t-1};\theta(s_t = 1)) + \pi_{2,0}f(y_t|Z_{t-1};\theta(s_t =))2}, \quad (7)$$

$$\pi_{2,1} = \frac{\pi_{2,0}f(y_t|Z_{t-1};\theta(s_t =))2}{\pi_{1,0}f(y_t|Z_{t-1};\theta(s_t = 1)) + \pi_{2,0}f(y_t|Z_{t-1};\theta(s_t =))2}, \quad (8)$$

where $f(y_t|Z_{t-1};\theta(s_t = j)); j = 1, 2$ is the density function of regime j .

- **Step 3:** We then make prediction probabilities for the regimes at time $t = 2$ as:

$$\pi_{1,2} = \pi_{1,1}p_{11} + (1 - \pi_{1,1})(1 - p_{11}), \quad (9)$$

$$\pi_{2,2} = 1 - \pi_{1,2}. \quad (10)$$

- **Step 4:** We can similarly derive two regime probabilities for $t = 2, \dots, T$. The inference performed iteratively for this process is called “the filtered probabilities”. Once the filtered probabilities $\pi_{j,t}$ (for $t = 1, \dots, T; j = 1, 2$) are obtained as above, we can construct the full log-likelihood function as:

$$L(y_t|x_t;\Theta(s_t)) = \sum_{t=1}^T [\log(f(y_t|x_t;\theta(s_t = 1))\pi_{1,t}) + \log(f(y_t|x_t;\theta(s_t =))2\pi_{2,t})]. \quad (11)$$

As the normal distribution and penalty functions of lasso and ridge are considered in this study, the density function $f(y_t|x_t;\theta(s_t = j))$ can be penalized by lasso and ridge as follows:

- Ridge density

$$f(y_t|x_t;\theta(s_t)) = \left[\frac{1}{\sqrt{2\pi\sigma^2(s_t)}} \exp\left(-\frac{(y_t - x_t\beta(s_t))^2}{2\sigma^2(s_t)}\right) \right] - \lambda \sum (\beta(s_t))^2 \quad (12)$$

- Lasso density

$$f(y_t|x_t;\theta(s_t)) = \left[\frac{1}{\sqrt{2\pi\sigma^2(s_t)}} \exp\left(-\frac{(y_t - x_t\beta(s_t))^2}{2\sigma^2(s_t)}\right) \right] - \lambda \sum |\beta(s_t)| \quad (13)$$

To obtain a penalized maximum likelihood estimator $\Theta(s_t)$, we maximize the full log-likelihood function in Equation (11) with respect to $\Theta(s_t)$ for some candidates' regularization parameters λ . Hence:

$$\widehat{\Theta}(s_t)(\lambda) = \underset{\Theta(s_t)}{\operatorname{argmax}} (L(y_t|x_t;\Theta(s_t);\lambda)). \quad (14)$$

We would like to note that if λ increases, the regression coefficients shrink towards zero, while $\lambda = 0$ the traditional maximum likelihood estimation, is returned. It is worth noting that we penalize the only $\beta(s_t)$ and choose not to penalize $\sigma(s_t), p_{11}, p_{22}$. If either $\beta(s_t = 1)$ or $\beta(s_t = 2)$ is equal to zero, it corresponds to the linear relationship between x_t and y_t .

3.3. Tuning the Regularization Parameter Using Information Criteria

The selection of the regularization parameter λ is important in ensuring the consistency of the maximum penalized likelihood estimation in Equation (14). Traditionally, the Bayesian information criterion (BIC) was generally used to find the optimal $\widehat{\lambda}$. However, Chen and Chen [42] found that the BIC is too liberal for model selection when the number of the estimated parameters is large. Therefore, they introduced the extended BIC to deal

with the model with a moderate size, but a huge number of parameters. The Extended BIC can be computed as,

$$\text{EBIC}(\lambda) = -2 \left(L \left(y_t | x_t; \hat{\Theta}(s_t); \lambda \right) \right) + v(\lambda) \log(n) + 2\zeta v(\lambda) \log(P), \quad (15)$$

where $L \left(y_t | x_t; \hat{\Theta}(s_t); \lambda \right)$ is the maximum likelihood of $\Theta(s_t)$ given the candidate λ . P is the number of parameters in the model. $v(\lambda)$ is the degrees of freedom, which is measured by $T - k^*$, where k^* is the number of non-zero coefficients. ζ is the value belonging to the interval $[0, 1]$. In this study, we set $\zeta = 0.5$ [43]. To find the best λ , we consider the lowest EBIC.

In addition, to better understand the theoretical analysis, we summarize the important notation presented in Table 1.

Table 1. Main notation description.

Notation	Description
y_t	The dependent variable at time t
x_t	The vector of independent variables at time t
$\pi_{i,t}$	The filtered probabilities of regime i at time t
p_{ij}	The probability of switching from regime i to j
s_t	The unobserved state variable at time t
$\beta(s_t)$	The vector of regime-dependent coefficients
$\sigma(s_t)$	The vector of regime-dependent coefficients

4. Empirical Model and Data

4.1. Data Sources

In this study, the MSR model has been proposed to investigate the impacts of various transportation modes on the top ten source countries of Chinese inbound tourism. The dataset collected contains information on Chinese transportations and the number of tourist arrivals. The largest number of tourists originated from the following ten countries: South Korea, Japan, Russia, the USA, Mongolia, Malaysia, Philippines, Singapore, India, and Canada. We consider the annual time series data covering 1996–2018. The data are collected from various sources and databases (Table A1, Appendix A).

As permitted by the data availability and their tourism significance, the data required for this study is collected from many different data sources. In this study, we consider the number of tourist arrivals as the dependent variable, representing China's tourism development [4–7]. This data is collected from the CEIC database (<https://www.ceicdata.com/en> (accessed on 5 February 2021)). The core explanatory variables of interest are four transportation modes, namely road, railway, air, and maritime; (i) road transportation is proxied by the Chinese highway length [28]. This variable is used to measure the transportation and availability, and quality of internal land transportation [6,44]; (ii) railway transportation is proxied by the Chinese railway length in operation and fixed-asset investment on the railway [9,10,12]. This indicator is recognized as safety, convenience, timeliness, flexibility, and affordability of transportation for tourists [21]; (iii) maritime transportation is proxied by the Chinese coastal major port, Chinese river major port, and the river's Chinese navigable length. The first two variables are viewed as embarking and landing of passengers from ship to land. Thana [45] mentioned that port infrastructure is the initial start of the transport chain. Specifically, the travel can be started from the sea and continues with the intermodal transports (road, rail, or air). The data regarding road, rail, and air are collected from the Ministry of Transport of the People's Republic of China (http://www.epschinastats.com/db_transportation.html (accessed on 5 February 2021));

(iv) air transposition is proxied by the number of international air routes in China, the length of China's international air route, the number of Chinese airports, the number of international air passenger traffics in China, and Chinese aircraft daily utilization. Air is included in the tourism equation as it proxies greater speed, safety, service quality, and reliability [21]. As suggested by [6,15,44], air transportation plays a vital role in boosting tourism growth. For aircraft daily utilization variable, it is used to measure aircraft productivity. The air transportation data can be retrieved from the CEIC database (<https://www.ceicdata.com/en>) (accessed on 5 February 2021)).

Nevertheless, tourism has also been recognized as a volatile industry that can be affected, not only by tourism infrastructure (transportation) [5,6], but also affected by local and global economic conditions, macroeconomic policies, and security [46–48]. Although there are various amazing tourist attractions in China, Chinese inbound tourism is still highly vulnerable to these factors, and thereby, leads to challenges in establishing tourism development strategies, tourism infrastructure, as well as delivering tourism products and services at consistent standards [49]. Therefore, the economic factors are also considered as the control variables in this study. The choice of these control variables is in line with the literature. The control variables used to explain tourism arrivals, include inflation rate [50,51], as well as GDP per capita of the country of origin and China [6,52]. Although the exchange rate is another important factor affecting the tourism demand [48,53], it is essential to highlight that the unavailability of data for the period 1996–2018 represented a limitation of this study. Therefore, the ridge and lasso estimations have been employed to resolve this issue and the independent variables' multicollinearity. We note that all economic variables are collected from the World Bank database (<https://data.worldbank.org>) (accessed on 5 February 2021)). Besides, all statistical calculations and representations are done by the R program.

4.2. Empirical Model

This study intends to understand the regime-switching effect of transportation modes on international tourism development. Therefore, the two-regime MSR model is constructed, and our empirical model for each source of tourists to China can be defined,

$$\begin{aligned} \Delta \ln TOUR_{i,t} = & \Delta \ln AIR_{i,t} \beta_{i,1}(s_t) + \Delta \ln MARITIME_{i,t} \beta_{i,2}(s_t) + \Delta \ln ROAD_{i,t} \beta_{i,3}(s_t) \\ & + \Delta \ln RAILWAY_{i,t} \beta_{i,4}(s_t) + \Delta \ln CONTROL_{i,t} \beta_{i,5}(s_t) + \sigma^2(s_t) u_{i,t}, \end{aligned} \quad (16)$$

where $\ln TOUR_{i,t}$ is the log of tourist arrivals from a given individual country i to China, at time t . The vector $\ln AIR_{i,t}$ is a set of air transportation variables (the number of international air routes of China, The length of the international air route of China, the number of Chinese airports, and Chinese aircraft daily utilization), $\ln MARITIME_{i,t}$ is a vector of maritime transportation variables (Chinese coastal major port, Chinese river major port, and Chinese navigable length of the river). $\ln ROAD_{i,t}$ is a road transportation variable (Chinese highway length), $\ln RAILWAY_{i,t}$ is a set of railway transportation variables (Chinese railway length and Chinese fixed-asset investment on the railway), and $CONTROL_{i,t}$ is a set of control variables (the relative real GDP per capita between the home country and China and Chinese inflation rate). We perform natural logarithms difference transformation for all variables, except for the inflation rate (units of percentage), to reduce the possible heteroscedasticity.

The descriptive statistics of the data in this study are summarized in Table 2. Table 2 reports the summary statistics of our variables presented in Table 1. Based on the mean growth of tourist arrivals, the largest growth of tourists originated from the following ten countries: India (0.144), South Korea (0.113), Mongolia (0.094), Canada (0.093), Russia (0.092), Malaysia (0.088), the Philippines (0.081), the USA (0.080), Singapore (0.069) and Japan (0.061). According to Figure 1, we illustrate the yearly tourist arrivals in these ten countries, and it is observed that there is only Japan and Russia exhibit a weak upward trend, and the others show a persistent growth. We also observe a significant drop in tourist arrivals in 2004 and 2008, corresponding to the SARS epidemic in China, and

during 2007–2008, corresponding to the global financial crisis and during 2014–2015, which appears to be consistent with the steep visa fees, cumbersome rules for tourists, and harmful pollution. This indicates that there might exist a structural change in Chinese tourism demand.

Table 2. Descriptive statistics of variables.

Variable	Mean	Max	Min	Std	Skew	Kurt	JB–Test	ADF Test	ZA Test	Break Date
HWL	0.072	0.788	0.010	0.162	4.006	17.960	276.04	−4.767 ***	−31.241 ***	2005
RLO	0.033	0.084	0.006	0.023	1.063	2.781	4.378	−2.392	−5.718 ***	2008
RFAI	0.143	0.690	−0.294	0.238	0.882	3.286	3.063	−2.657 *	−3.857 *	2008
CPL	0.086	0.601	−0.294	0.153	1.167	7.957	28.778	−9.075 ***	−10.202 ***	2002
RPL	0.077	0.542	−0.206	0.183	0.929	3.722	3.809	−3.379 **	−4.387 **	2009
NLR	0.006	0.056	−0.009	0.013	2.447	9.599	64.708	−4.914 ***	−10.459 ***	2004
NIAR	0.113	0.466	−0.139	0.138	0.446	3.569	1.074	−5.223 ***	−6.216 ***	2011
LIAR	0.115	0.396	−0.178	0.143	−0.073	2.774	0.069	−4.79 ***	−5.642 ***	2009
APT	0.023	0.085	−0.106	0.039	−1.392	6.157	16.988	−3.769 ***	−9.366 ***	2003
ADU	0.025	0.357	−0.105	0.089	2.271	9.358	58.524	−11.839 ***	−13.776 ***	2004
GDPSK	0.052	0.289	−0.332	0.129	−1.018	4.742	0.052	−4.566 ***	−6.343 ***	2000
GDPJP	−0.001	0.129	−0.168	0.087	−0.141	1.868	−0.001	−3.853 ***	−9.806 ***	2009
GDPRU	0.091	0.379	−0.339	0.230	−0.622	2.138	0.091	−3.069 **	−4.390 **	2009
GDPUS	0.035	0.058	−0.027	0.018	−1.617	6.464	0.035	−2.673 *	−3.945 *	2009
GDPMO	0.100	0.542	−0.198	0.188	0.581	2.647	0.100	−3.225 **	−3.758 *	2002
GDPMA	0.050	0.240	−0.296	0.119	−1.134	4.424	0.050	−4.250 ***	−6.324 ***	2008
GDPPH	0.053	0.201	−0.142	0.083	−0.322	2.815	0.053	−3.591 **	−4.698 **	2007
GDPSI	0.047	0.213	−0.172	0.087	−0.344	3.420	0.047	−4.925 ***	−5.461 ***	2004
GDPIN	0.078	0.275	−0.029	0.079	0.810	3.073	0.078	−3.590 ***	−5.423 ***	2008
GDPCA	0.039	0.167	−0.144	0.082	−0.433	2.751	0.039	−3.373 ***	−7.924 ***	2007
INFR	2.202	8.313	−1.401	2.344	0.740	3.429	2.281	−1.842	−5.661 ***	2003
SK	0.113	0.567	−0.192	0.210	0.251	2.398	0.589	−4.429 ***	−6.987 ***	2008
JP	0.041	0.478	−0.229	0.150	0.793	4.443	4.407	−5.580 ***	−7.684 ***	2007
RU	0.092	0.463	−0.441	0.204	−0.613	3.438	1.626	−5.037 ***	−7.045 ***	2009
US	0.080	0.591	−0.266	0.150	1.216	7.684	26.697	−5.916 ***	−6.738 ***	2008
MO	0.094	0.512	−0.227	0.190	0.250	2.540	0.443	−5.030 ***	−6.361 ***	2004
MA	0.088	0.724	−0.273	0.192	1.278	6.515	18.105	−6.073 ***	−6.861 ***	2007
PH	0.081	0.246	−0.099	0.098	−0.193	2.095	0.927	−0.733	−10.431 ***	2008
SI	0.069	0.684	−0.239	0.164	2.063	9.895	61.879	−5.497 ***	−6.425 ***	2005
IN	0.144	0.436	−0.055	0.134	0.784	2.612	2.501	−3.226 **	−4.784 **	2006
CA	0.093	0.511	−0.209	0.139	0.705	5.226	6.657	−4.931 ***	−5.483 ***	2007

Note: JB denotes Jarque-Bera normality test, and ZA denotes Zivot and Andrews nonlinear unit root test. “*”, “**” and “***” denote rejection of the null hypothesis at the 10%, 5%, and 1% significance level, respectively.

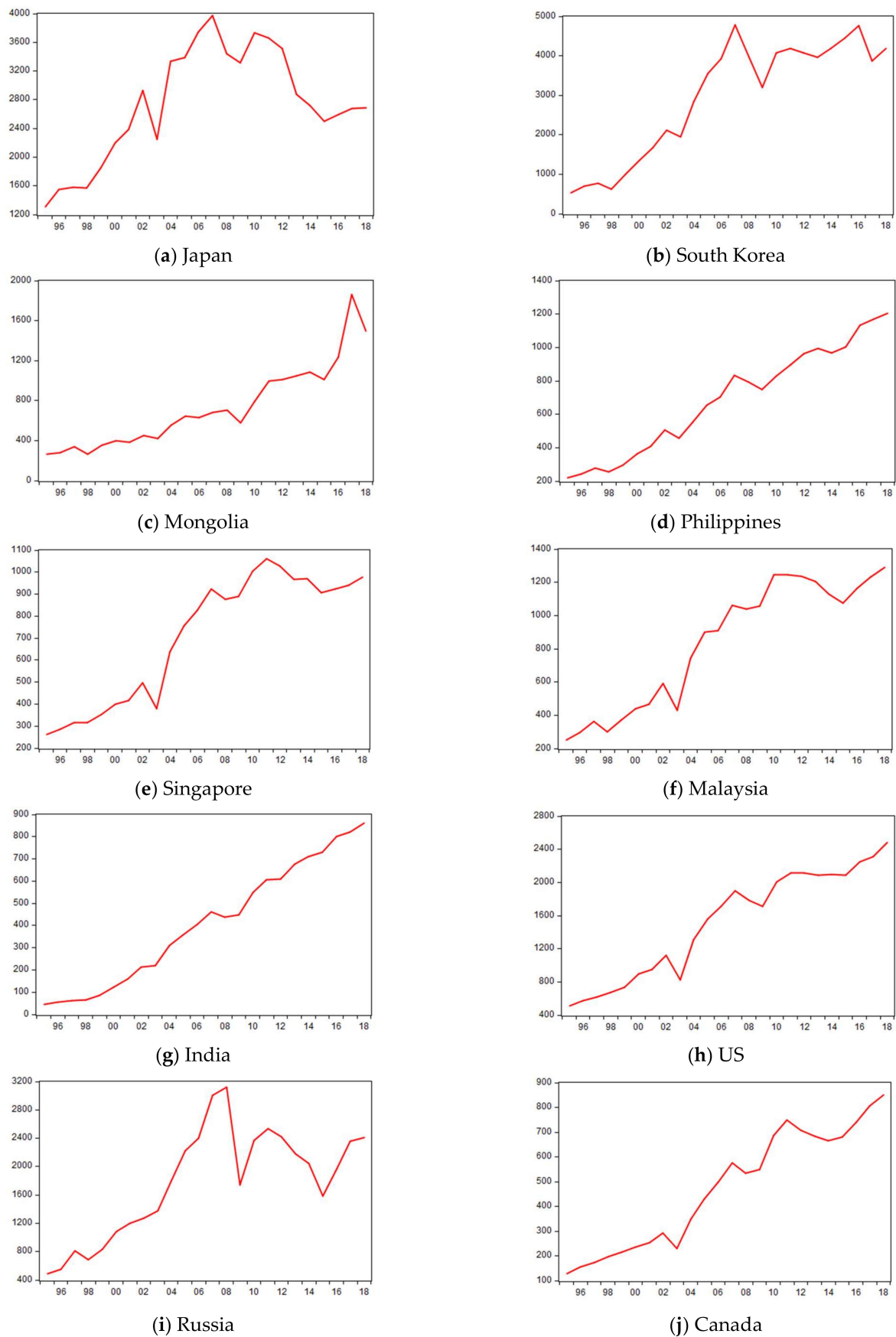


Figure 1. Tourist arrivals to China. Source: CEIC database.

4.3. Unit Root Test

Before estimating the MSR model to analyze the impact of transportation on tourism, as well as the business cycle of the Chinese tourism development, we need to check whether the data is stationary, using the unit root test as non-stationarity of the variables can lead to a spurious regression. In this study, we consider both linear and nonlinear unit root tests, namely the Augmented Dickey-Fuller (ADF) test [54] and Zivot and Andrews [55]. Note that the ADF test does not account for the potential structural breaks in the time series, and it is possible to provide inaccurate results when a structural break has occurred in the time series. Therefore, we also consider the ZA test to validate our stationary test, taking into consideration the presence of structural breaks.

In this study, we consider the ADF test that involves regressing on a lagged first difference (y_{t-1}), a linear deterministic trend, (T), and P -lagged first differences (Δy_{t-p}). It can be expressed as follows,

$$\Delta y_t = \mu_0 + \theta y_{t-1} + \pi T + \sum_{p=1}^P \mu_p \Delta y_{t-p} + \varepsilon_t \quad (17)$$

where Δ is the first difference operator. The null hypothesis of a unit root ($H_0 : \theta = 0$) is tested against the alternative that the variable is stationary ($H_1 : \theta < 0$). However, the ADF tests did not encounter possible structural breaks in estimation of unit roots. Zivot and Andrews [55] then generalized the Equation (17) as,

$$\Delta y_t = \mu_0 + \theta y_{t-1} + \pi T + \rho_1 DU_t + \rho_2 DT_t + \sum_{p=1}^P \mu_p \Delta y_{t-p} + \varepsilon_t, \quad (18)$$

where $DU_t = 1$ if $t < TB$, else $DU_t = 0$, is a dummy parameter for mean shift arising at each potential breakpoint (TB) and $DT_t = t - TB$ if $t > TB$, else $DT_t = 0$ if $t < TB$ is the trend shift variable. Among all the possible breakpoints, the optimal TB is selected when the absolute value of the t-statistic from the ADF test is minimized. The null hypothesis demonstrates that the variable encompasses a unit-root with a drift that disregards any break, whereas the alternative hypothesis illustrates that the parameter follows a stationary trend procedure with a single break that appears at an unidentified time.

The results of these two tests are reported in Table 2. The result shows that the ADF and ZA tests reject the nonstationary null hypothesis, indicating that all the time series variables are stationary. Therefore, these variables are appropriately used for studying the relationship between Chinese inbound tourism and transportation and the tourism cycle further.

5. Results and Discussion

5.1. Model Comparison

In this study, the two-regime Markov-switching lasso regression (MS-LR) and Markov-switching ridge regression (MS-RR) models are assumed to investigate the nonlinear effects of transportation modes on ten top-tourist arrivals to China. Therefore, we need to validate our MSR model's performance by comparing it with the single-regime model, say linear ridge regression (RR), and lasso regression (LR) models. In this comparison, the EBIC is used as the comparison measure, and the lowest EBIC indicates a more parsimonious model. Table 3 shows the model comparison of five different models for all ten tourism demand equations. The result confirms our hypothesis of the nonlinear effect of transportation-tourism nexus as the EBICs of the two-regime models, MS-LR and MS-RR, are lower than the single-regime models, namely linear-MLE, RR and LR. Focusing on the EBICs of our proposed models, MS-LR and MS-RR, Table 3 highlights that MS-LR is superior to MS-RR for all our arrival series, except for Russian tourists. This finding is not surprising as MS-LR is more flexible than MS-RR and can achieve robust parameter estimation and variable selection in Markov-switching regression simultaneously. However, we cannot conclude

that lasso is always the best estimation for the MSR model as a data set's characteristics can influence regularized estimator's performance. Emmert-Streib and Dehmer [56] suggested that not one method always dominates the others since they all have specific strengths and weaknesses. Therefore, the MS-RR model may be preferred in the case of Russian tourist arrivals.

Table 3. Information criterion and model comparison.

Model	Linear-MLE	RR	LR	MS-RR	MS-LR
Country	EBIC	EBIC	EBIC	EBIC	EBIC
JP	33.398	29.258	37.556	19.309	0.8153
SK	29.938	27.239	29.810	−11.785	−16.185
MO	38.650	44.060	38.092	27.520	25.662
PH	47.299	47.270	54.607	2.4229	−30.918
SI	40.415	53.763	64.560	1.4768	−12.76
MA	35.487	35.465	34.954	25.917	0.3799
IN	20.519	18.207	17.903	8.1045	−3.5109
USA	68.355	68.326	68.300	−1.5274	−18.941
RU	44.454	47.116	47.399	−2.1898	18.108
CA	24.446	26.484	25.817	8.9701	−13.276

Note: Linear-MLE is the classical linear regression estimated by Maximum likelihood estimation (MLE). The bold number indicates the lowest EBIC.

5.2. Estimation Result

We use two different econometric techniques in our estimates as the traditional MS-regression estimated by MLE can lead to a large variance when a large number of predictors and a relatively small number of observations are used. The computation is impossible if the number of parameter estimates is larger than the number of observations. Therefore, two types of penalty likelihood functions, namely ridge, and lasso are applied to our MS regression.

Indeed, it can be seen that MS-LR is superior to MS-RR for all our tourist arrival series, except for Russian tourists (see Section 4.1). Therefore, the results from the best model specification are only interpreted and reported in Table 4. We note that the ridge penalty keeps all predictors in the MS-regression model, while lasso ensures sparsity of the results by shrinking some coefficients exactly to zero. Therefore, the t-statistics is used to test the significance of each predictor in the MR-RR model.

Table 4. The results of MS regression model for 10 originating countries.

	MS-LR	MS-LR	MS-LR	MS-LR	MS-LR
Variable	JP	SK	MO	PH	SI
Intercept	2.3667	−4.7317	−0.0279	−6.9719	−2.052
HWL	0.5460	0.3353	0.4528	0.3474	0.5160
RLO	−2.5162	−0.4678	−2.1918	−1.5402	−1.678
RFAI	0	0.1417	0.4944	0.1546	0.1333
CPL	−0.1468	0.7207	0.2303	−0.1344	0
RPL	0.2966	0	0.2008	0	0.4974
NLR	0.1735	0.2552	0.3646	0.2920	0.4041
NIAR	0.3691	0	0	0.1331	0
LIAR	0	0	0	0	0

Table 4. Cont.

		MS-LR	MS-LR	MS-LR	MS-LR	MS-LR
	Variable	JP	SK	MO	PH	SI
Regime 1 (low)	APT	0	1.3957	2.4714	1.5534	0
	ADU	0.1986	0.0794	1.7347	1.0068	0.6104
	GDPO	0.2285	0.3635	0.3877	0.4170	0.0343
	INFR	0	0	0	0	0
Regime 2 (high)	Intercept	6.2188	0.9847	0.7960	0.8477	0.1548
	HWL	0	0.0296	0.0005	0.0851	0.0003
	RLO	0	0	0	0	0
	RFAI	0	0	0	0	0
	CPL	0	0	0	0	0
	RPL	0	0	0	0	0.0003
	NLR	0	0	0	0.0009	0
	NIAR	0	0	0	0	0
	LIAR	0	0	0	0	0
	APT	0	0	0	0	0
	ADU	0.6790	0.0087	0.0029	0.0856	0.0001
	GDPO	0.0289	0.0370	0	0.0978	0.0006
	INFR	0	0	0	0	0
	EBIC	0.8153	−16.184	25.662	−30.918	−12.759
	Log-likelihood	35.592	44.092	23.168	51.459	42.379
	λ	0.4533	0.9983	5.1494	3.4413	2.1494
		MS-LR	MS-LR	MS-LR	MS-RR	MS-LR
	Variable	MA	IN	USA	RU	CA
Regime 1 (low)	Intercept	−2.6692	−0.1205	−2.0443	−0.0896 (1.0294)	−2.4930
	HWL	0.3282	0.2906	0.3817	0.8521 *** (0.1294)	0.1366
	RLO	−2.1511	−1.0364	−1.3312	−1.3656 *** (0.5122)	−2.1939
	RFAI	0.1788	0	0	0.2529 ** (0.1018)	0
	CPL	0	−0.4016	−0.2699	−1.0930 *** (0.1821)	0
	RPL	0.5330	0.0372	0.2603	0.0591 (0.1079)	0.3688
	NLR	0.1646	0.1845	0.2363	0.4235 *** (0.1024)	0.1964
	NIAR	0.5302	0.7749	0	0.2045 (0.1615)	0.5374
	LIAR	0	−0.7648	0.2125	0.4705 *** (0.1446)	0.5195
	APT	0.6818	0.6321	0	−0.7590 ** (0.3052)	1.1724
	ADU	0.4253	0.5202	0.7107	0.8758 ** (0.4131)	0.2245
	GDPO	0.2967	0.5561	0.0371	0.3007 *** (0.0655)	0.0835
	INFR	0	0	0	0.0573 *** (0.0095)	0
Regime 2 (high)	Intercept	0.1610	0.3718	0.1969	0.0312 *** (0.0006)	0.2571
	HWL	0.0001	0.0034	0.0357	0.0639 *** (0.0020)	0.0013
	RLO	0	0	0	0.0433 (0.0277)	0

Table 4. Cont.

	Variable	MS-LR	MS-LR	MS-LR	MS-RR	MS-LR
		MA	IN	USA	RU	CA
Regime 2 (high)	RFAI	0	0	0	0.0325 ** (0.0139)	0
	CPL	0	0	0	0.0489 (0.0485)	0
	RPL	0.0055	0	0	0.0434 (0.0374)	0.0051
	NLR	0	0.0006	0.0749	0.0337 *** (0.0071)	0
	NIAR	0	0.0219	0	0.0337 *** (0.0071)	0
	LIAR	0	0	0	0.0619 *** (0.0037)	0
	APT	0	0	0	0.0466 *** (0.0050)	0
	ADU	0.0006	0.0044	0.0655	0.0713 *** (0.0009)	0.0010
	GDPO	0.0016	0.0006	0	0.0290 *** (0.0016)	0
	INFR	0	0	0	0.2322 *** (0.0014)	0
EBIC		0.3799	−3.5109	−18.941	−2.1898	−13.276
Log-likelihood		35.810	37.755	45.470	37.094	42.638
λ		0.1494	0.1049	0.4413	0.4413	0.1989

Note: GDPO is the relative Gross Domestic Product of home country and China. For the ridge estimation, “***” and “**” denote rejection of the null hypothesis at the 5%, and 1% significance level, respectively. The parentheses () is the standard error. There is no standard error for the lasso estimation and statistic inference as the insignificant parameters are already eliminated by the penalty term.

From the above analysis results, transportation has a regime-switching effect on the top ten source countries of Chinese inbound tourism. It can be noted from the estimated intercept terms of regimes 1 and 2, these values are significant differences, and the value of the intercept term of regime 1 is lower than regime 2. Therefore, we can interpret regime 1 as a low tourist arrival regime, while regime 2 is a high tourist arrival regime.

For the low tourist arrival regime (regime 1), the first conclusion to be drawn is that transportation is confirmed to have played an important role together with the control variables. This is more pronounced for the case of Mongolia (MO), the Philippines (PH), Malaysia (MA), and India (IN) as ten out of eleven transportation indicators are not eliminated from the models. These results indicate that the development of all Chinese transportation infrastructures, railway, maritime, air, and roads could attract more tourists from Mongolia, the Philippines, Malaysia, and India to China. Second, the Chinese railway length in operation (RLO) shows a strongly negative impact on inbound tourism from all ten countries. Surprisingly, the net impact of Chinese railway length on tourism does not seem to be consistently positive. This result is similar to Pagliara et al. [22] and Albalade and Fageda [9], who suggested that the railway network is not appropriately designed and does not correspond to the traveler’s needs. The second is the substitution effect of air transportation. This is to say, the substitution effect of railway on aviation has decreased the number of tourist arrivals in China, which has indirectly resulted in negative effects on inbound tourism. However, if we consider the Chinese fixed-asset investment on the railway (RFAI), the effect of this variable is significant, with the expected positive sign. That is, the higher investment on the railway leads to higher Malaysian, South Korea, Russian, Singaporean, Mongolian, and Philippines tourist arrivals. Third, the maritime indicators, maritime transportation infrastructures have shown a mix of positive and negative impacts on tourism demand. It is found that the Chinese river port and the navigable length of the river contribute a positive impact on tourism demand, but it is harmful to the Chinese coastal port. It is surprising that the Chinese coastal port’s operation negatively impacts tourists from Japan, the Philippines, India, Russia, and the USA. The possible reason is that tourists and businesses employ water transportation to travel, import, and export, respectively, thus, increasing the operation of Chinese coastal ports may increase traffic intensity and extend the time cruisers remain at port destinations. Lau and Yip [57]

mentioned that the coastal ports require advanced port facilities to fulfill the increasing size of cruise ships and the increased cruiser transit. Therefore, inefficient port facilities cannot generate higher tourist arrivals but can reduce coastal tourism. In addition, maritime transport is sometimes related to the marine environment's threat as pollution threatens society's interests in various fields concerning the marine environment, ranging from human existence to recreation, including tourism [58].

Fourth, considering air transportation, airport infrastructure variables (NIAR, LIAR, APT, ADU), they have been a relatively important tourism generation element, especially true for the daily aircraft utilization (ADU). We find that daily aircraft utilization shows a positive impact on all sources of Chinese international tourism. Note that this variable can be viewed as the measurement of aircraft productivity; thus, this result provides evidence that aircraft productivity plays a vital role in boosting Chinese tourism growth. This interpretation is consistent with the results revealed by Xie and Tveterås [44], Yin et al. [12], Khadaroo and Seetanah [6], Eric, Semeyutin, and Hubbard [15]. Fifth, based on the magnitudes of transportation coefficients, we find that ADU performs the highest positive impact on Chinese tourism demand, while RLO contributes to the highest negative impact. The logarithm coefficients of aircraft daily utilization across source countries are positive and range from the lowest of 0.0794 in the South Korean tourists to the highest of 1.7347 in Mongolian tourists. For RLO, the logarithm coefficients of Chinese railway length in operation across source countries are negative and range from the lowest of -2.5162 in the Japanese tourists to the highest of -0.4678 in South Korean tourists. Overall, tourists are particularly sensitive to railway infrastructure and air productivity, as judged by their coefficients.

In relation the control variables, we find that the relative GDP of the originating country variable has the strongest impact on foreign tourists for China. The elasticities range from 0.0343 in Singaporean tourists to the highest of 0.5561 in Indian tourists, which can be interpreted as a 1% increase in the GDPO is associated with an increase of around 0.0343 to 0.5561 in foreign tourist arrivals.

In the case of the high tourist arrival regime (regime 2), a brief interpretation of the estimation results in Table 4 is given below for greater clarity. First, transportation is also confirmed to play an essential role in this regime. However, the effect is less pronounced than that in regime 1. Second, we observe that the HWL and ADU variables are key determinants attracting more tourists from all countries. Bao et al. [59] suggested that tourists spend most of their time on the roads during high travel periods, resulting in declining quality of the vacation and unpleasant travel experience. Therefore, the development of new highways could relieve traffic congestion, as well as provide an opportunity to explore new destinations during the high tourist arrival regime. For ADU, we expect that the supply of airport infrastructure is still limited and is not consistent with the demand of international tourists during the high tourism regime; hence, aircraft productivity improvement would enhance capacity utilization. Consequently, it makes economic sense to expand aircraft productivity to gain more tourist arrivals from these ten countries. Third, it is observed that maritime transportation shows a weak positive impact on seven out of ten source countries (the Philippines, Singapore, Malaysia, India, USA, Russia, and Canada) as the relatively small coefficient values of RPL and NLR are presented. Fourth, among all the control variables, it is found that the relative country's GDP has a significant impact on the number of foreign tourists from all countries, except for the USA and Canada. A possible explanation for this phenomenon is that economic expansion may generate a positive economic climate that favors international tourism activities. However, in the case of the USA and Canada, the economic growth-tourism link is quite limited. This minor role of the originating country's GDP is not surprising if we consider that the USA and Canada tourists, which are high-income tourists, are not dependent on the price levels in China [8]. Van Can [60] confirmed that the choice of high-income travelers does not depend on price.

Comparing these two regimes, we observe that the estimated coefficient is more significant in regime 1 (low tourist arrival regime). The estimated coefficients in regime 1 are all positive (except for railway length in operation). In contrast, the estimated coefficients in regime 2 are less positive and weak, which reveals that Chinese transportation has an asymmetrical impact on China's inbound tourism. This study shows that in the process of transportation development and inbound tourism demand from ten countries, some variables' original strong positive effects transform into weak positive effects. These results can be explained by the fact that the number of tourism arrivals is greater in regime 2 than in regime 1. Therefore, transportation may not be the key driving tourist arrivals in the second regime. Furthermore, there is evidence that aircraft productivity looks more promising than other variables in regime 1, while both aircraft productivity and highway road are more pronounced in regime 2. This result suggests that different development plans are required for attracting Chinese inbound tourism.

Table 5 shows the probabilities of being in regime 1 and regime 2 of the best-fit MS-model (presented in Table 4) for the 10 source countries. The first conclusion is that the transition probabilities for all source countries are similar and range from 0.3114 (Japan) to 0.6298 (Singapore) for p_{11} and range from 0.945 (1USA) to 0.9937 (Malaysia) for p_{22} . Specifically, when the current state of Chinese tourism demand is regime 1 at time t , the probabilities of Chinese tourism demand remaining in regime 1 at time $t + 1$ range from 0.3114 to 0.6298. Likewise, if the current state of Chinese tourism demand is regime 2 at time t , the probabilities of Chinese tourism demand remaining in regime 2 at time $t + 1$ range from 0.9451 to 0.9937. These results suggest that the chances of switching from one regime to the other are not high in the Chinese tourism market. Also, the duration for which Chinese tourism remains in regime 1 ranges from the lowest of 1.4522 months in the Japanese tourists to the highest of 2.7012 in the Singaporean tourists, and that for regime 2 ranges from the lowest of 18.2149 months in the Japanese tourists to the highest of 158.302 in the Singaporean tourist. This suggests that the duration of the high tourist arrival regime is approximately 10–100 times the length of the low tourist arrival regime. Based on these findings, we can conclude that the Chinese tourism market has a high tendency to remain in the high tourist arrival regime and a low tendency to shift from the high tourist arrival regime to the low tourist arrival regime during 1995–2018. This implies that the expansion effect on Chinese tourism produced by the transportation modes is promising as they keep the Chinese tourist arrivals to maintain high most of the time.

Table 5. The transition probabilities of the MS-model for ten countries.

Country	JP	SK	MO	PH	SI
p_{11}	0.3114 (1.4522)	0.4104 (1.6961)	0.4113 (1.6986)	0.3321 (1.4972)	0.6298 (2.7012)
p_{22}	0.9921 (126.5823)	0.9899 (99.0099)	0.9921 (126.5823)	0.9560 (27.7272)	0.9922 (128.2051)
Country	MA	IN	USA	RU	CA
p_{11}	0.6145 (2.5940)	0.6121 (2.5779)	0.6247 (2.6645)	0.4127 (1.7027)	0.4134 (1.7047)
p_{22}	0.9937 (158.7302)	0.9931 (144.9275)	0.9451 (18.2149)	0.9899 (99.0099)	0.9913 (114.9425)

Note: the parentheses () denotes Expected duration for which the Chinese tourism market remained in Regime 1 or 2.

Finally, we then investigate the tourism cycle for all ten source countries using the obtained filtered probabilities ($\pi_{1,t}$ and $\pi_{2,t}$). The plot of the filtered regime probabilities tells us when the tourist arrivals follow the same behavior, which is either the probability of tourist arrivals is high (regime 2) or low (regime 1). The results of the filtered probabilities of the high tourist arrival regime are illustrated in Figure 2. From this figure, it can be seen that the evolution of filtered probabilities shows similar time-varying paths for each of the

ten tourist arrival series, and the possibility of being in a high-tourism state is persistent and longer in regime 2. This indicates that Chinese transportations resulted in the structural change of the overall Chinese tourism demand. Moreover, we observe a sudden drop in the probabilities of staying in this regime around 2015–2016 for all tourist arrivals (except South Korea and the Philippines). This data point may denote the emergence of severe pollution and the introduction of censorship and complicated visa requirements, which made traveling to and within China riskier, and more expensive, respectively.

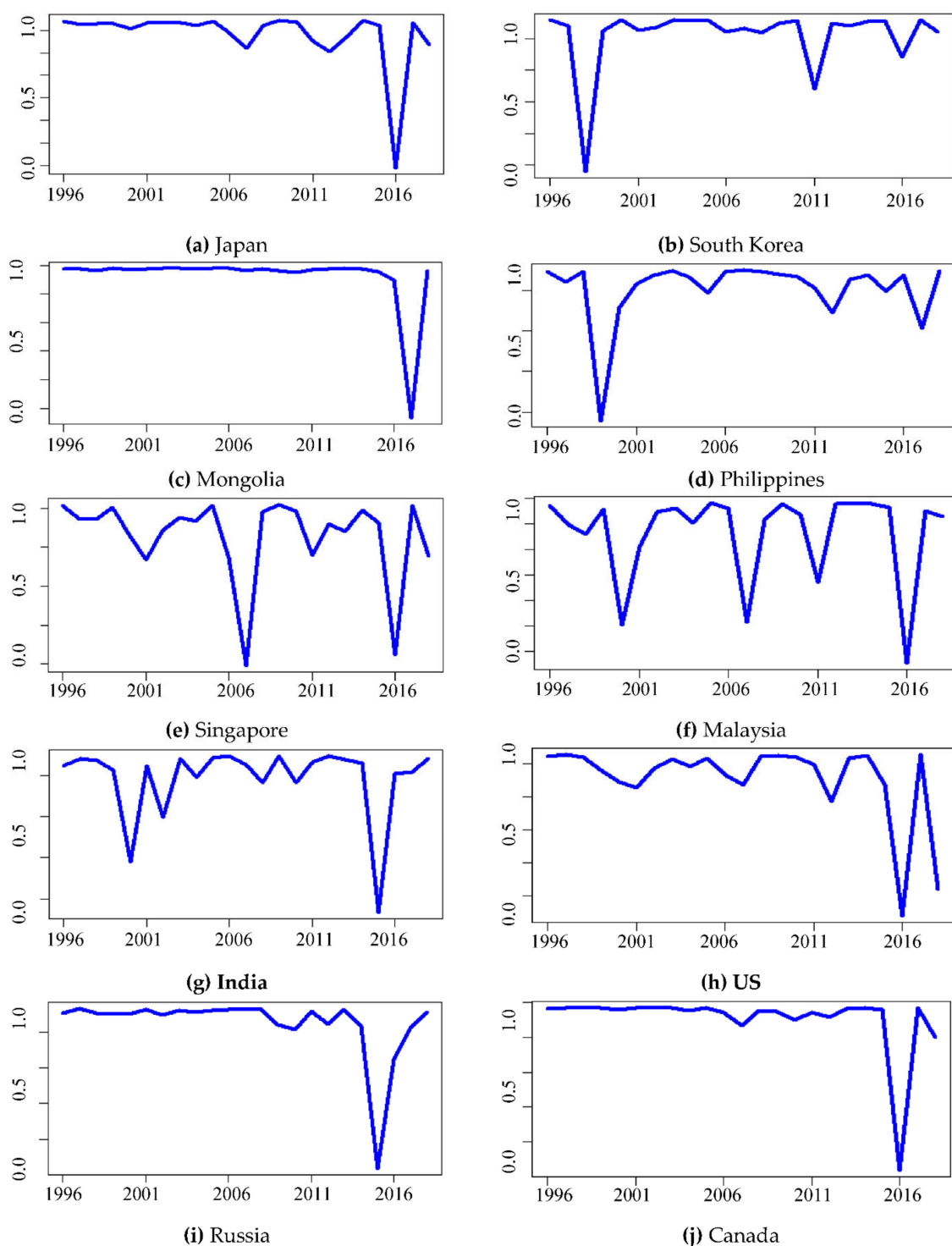


Figure 2. Filtered Probability of the high tourism arrival regime (regime 2) for ten source countries. Source: Author's calculations.

Moreover, there is a massive decline around 1997–1998 in the probabilities of South Korean, Malaysian, Indian, Philippines tourist arrivals, which appears to coincide with the Asian financial crisis period. Furthermore, when one looks at all probabilities of Malaysian and Singapore tourists, these two tourist sources' probabilities are low in 2007–2008, which appears to be the time of the global financial crisis originated from the USA. As a result of the 2007–2008 US crisis, the global economy exerted adverse spillover effects on Asian tourism [61].

Throughout these empirical and statistical facts regarding international tourist arrivals in China, our empirical results are quite consistent with the historical situation. In other words, our models are well-suited for such empirical studies. We can observe a significant structural change of South Korean, Malaysian, Indian, Philippines tourists during 1997–1998, 2007–2008 for Malaysian and Singapore tourists, and 2015–2016 for all tourist arrivals (except for South Korea and the Philippines). These data points correspond to external factors (economic crisis) and internal factors (Chinese environment and visa policy), which could steadily push China down on the list of top tourist destinations.

This study implies that there exists two regimes of international tourism arrivals to China. Unlike the results of existing literature, we show that the impacts of transportation modes are different among each other and between two regimes. Our evidence, obtained from filtered probability, reveals that unobserved events (such as economic crisis and domestic regulations) play an important role in China's international tourist arrivals. These results provide useful information for the policymaker to prepare a proper tourism policy for high tourist arrivals and low tourist arrivals.

6. Discussion, Conclusions, and Suggestions

This empirical study explores the influence of various transportation modes, namely air, road, maritime, and railway, on top-ten international Chinese inbound tourism from 1995 to 2018 using the Markov-switching regression (MSR) models. There are two distinctive features of this paper. First, this paper compares the magnitude of the impact of various transportation modes when international Chinese tourism corresponds to low tourist arrival regime and high tourist arrival regime. Specifically, the impact of various transportation modes can have a structural shift through the MSR models. Secondly, through this model, this paper runs Ridge and Lasso estimations of parameters in 10 sources tourist arrival equations after conducting the comparison of linearity against alternative switching models. We note that the Ridge and Lasso are proposed for incorporation in the MSR models to reduce complexity and prevent over-fitting of the MSR model. Also, the techniques help reduce the size of the problem to enable maximum likelihood to work faster, making it possible to handle high-dimensional data [62]. Thirdly, the cycle of 10 international Chinese inbound tourists is detected and illustrated.

Overall, the impacts of transportation modes on top-ten international Chinese tourist arrivals are mainly nonlinear, and the MSR models capture two tourist arrivals regimes. In relation to the magnitude of the effects of transportation modes on the Chinese tourism market, we observe that the aircraft productivity shows the highest positive contribution to Chinese tourism demand, while the railway length in operation contributes the highest negative impact. Surprisingly, while railway transport infrastructure is commonly believed to impact the tourism demand positively, our study obtains inconsistent evidence in favor of a positive relationship between railway transport infrastructure and tourist outcomes. As suggested by Albalade and Fageda [9], railway transportation may be more competitive in terms of travel time, frequencies, and comfort, but not necessarily in terms of price. Hence, visitors are more sensitive to price than time; the overall competitiveness of railway transportation concerning aviation may not positively affect tourist outcomes. In addition, Pagliara et al. [22] and Albalade and Fageda [9] also suggested that the railway network is not appropriately designed and does not correspond to the visitor's needs. In the view of Chinese railway transportation, the first suggestion may not be the key source of the negative outcome of tourism demand in China as the Chinese railway transportation price

is relatively low compared to other countries. We expect that an inappropriate design of the railway network in China is the main problem. At the end of 2018 there was 131,000 km of a railroad across China, and 29,000 km made up a high-speed railroad, which is the largest high-speed rail system in the world [63]. However, the railway network, especially the traditional railway, may fail to fully transport all passengers to their destinations as the railway network in most places is not in a good condition, and thus, access to destinations is problematic. Moreover, the railway network expansion may lead to long-time travel and an unpleasant train journey. For aircraft productivity, the result shows a strong positive impact for all tourist arrival series, signifying that aircraft productivity is one of the main factors affecting the Chinese tourism demand.

By comparing high and low tourist arrival regimes, we find a stronger impact of transportation modes in low tourist arrival regime than high tourist arrival regime. A strong positive effect of some variables will transform into a weak positive effect. This indicates that transportation is a key driver boosting the international tourist arrivals in the low tourist arrival regime. We also find a significant structural change of South Korean, Malaysian, Indian, Philippines tourists during 1997–1998 corresponding to the Asian financial crisis, during 2007–2008 corresponding to the global financial crisis for Malaysian and Singapore tourists, and during 2015–2016 corresponding to Chinese visa restriction and pollution for all tourist arrivals (except for South Korea and the Philippines).

The impacts of transportations, especially highway road and aircraft productivity on Chinese tourism, become more salient when tourist arrival is classified in the low tourist arrival regime. Thus, the expansion of highway road construction and aircraft productivity improvement could generate sizable fiscal revenues from the international tourist arrival. In addition, by continuously improving aircraft productivity in the high tourist arrival regime, China could attract more international tourist arrivals. This suggests that China's government should continue to pursue innovations and technologies in its aircraft productivity, as well as simplify procedures and improve efficiency. Specifically, if tourism demand is growing rapidly, but capacities are limited, it makes economic sense to expand stretched capacities or increase air transportation productivity in China.

Furthermore, the structural change of the Chinese tourism demand is another critical issue observed in this study. If it is ignored, this may result in inappropriate transportation development, causing the government's policy to be misdirected. The Chinese government should be aware of the factors affecting structural change and incorporate them into the policy-planning.

Some limitations suggest directions for future work. Firstly, we use annual data, while monthly or quarterly data may be better for understanding the structural change in tourism demand. Future research may consider collecting monthly or quarterly data to capture better the structural changes in the tourism cycle. Secondly, the variables analyzed in this paper are transportation and some macroeconomic factors (control variables). Future research can introduce other variables when investigating this topic, such as the number of five-star hotels, exchange rate, and population. In addition, this paper adopts the number of arrivals of international tourists from the source country. Future research may consider international tourist receipts. In addition, as suggested by Thai, Wu, and Xiong [64] and Stai et al. [65], social characteristics and personalization are the important key factors determining human preference. Therefore, further research may consider these variables and investigate their effects on tourists' preferences for transportation modes.

Author Contributions: Conceptualization, W.Y., and P.M.; data curation, X.Z.; formal analysis, P.M.; methodology, W.Y.; visualization, X.Z.; writing—original draft, W.Y., X.Z. and P.M.; writing—review and editing, W.Y., and P.M. All authors have read and agreed to the published version of the manuscript.

Funding: Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Thailand.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in the empirical analyses are available online at the CEIC database, World Bank database, and http://www.epschinastats.com/db_transportation.html (accessed on 5 February 2021). The data, however, are available upon request.

Acknowledgments: The authors would like to express their gratitude to Laxmi Worachai for her help and constant support. The authors are also grateful to the editors and three anonymous referees for their comments and suggestions. This work was supported by the Center of Excellence in Econometrics, Faculty of Economics, Chiang Mai University, Thailand.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Variable description and sources.

Variable		Description	Unit	Source
Road Transportation	HWL	Chinese highway length	Kilometer	Ministry of Transport of the People's Republic of China
Railway transportation	RLO	Chinese railway length in operation (including both full operation and temporary operation).	Kilometer	Ministry of Transport of the People's Republic of China
	RFAI	Chinese fixed-asset investment on the railway	RMB (million)	Ministry of Transport of the People's Republic of China
Maritime Transportation	CPL	Chinese coastal major port: length of a quay in operation	Meter	Ministry of Transport of the People's Republic of China
	RPL	Chinese river major port: length of a quay in operation	Meter	Ministry of Transport of the People's Republic of China
	NLR	Chinese navigable length of the river	Kilometer	Ministry of Transport of the People's Republic of China
Air Transportation	NIAR	The number of international air routes in China	Route	CEIC database
	LIAR	The length of the international air route of China	Kilometer	CEIC database
	APT	The number of Chinese airports	Airport	CEIC database
	ADU	Chinese aircraft daily utilization. Aircraft daily utilization is presented in block hours per day. This indicator is calculated by dividing aircraft block hours by the number of aircraft days assigned to service on air carrier routes.	hour per day	CEIC database
Control Variable	GDPSK	Relative real GDP per capita of South Korea	millions of US dollars	World Bank
	GDPJP	Relative real GDP per capita of Japan	millions of US dollars	World Bank
	GDPRU	Relative real GDP per capita Russia	millions of US dollars	World Bank

Table A1. Cont.

Variable		Description	Unit	Source
Control Variable	GDPUS	Relative real GDP per capita the US	millions of US dollars	World Bank
	GDPMO	Relative real GDP per capita Mongolia	millions of US dollars	World Bank
	GDPMA	Relative real GDP per capita Malaysia	millions of US dollars	World Bank
	GDPPH	Relative real GDP per capita Philippines	millions of US dollars	World Bank
	GDPSI	Relative real GDP per capita Singapor	millions of US dollars	World Bank
	GDPIN	Relative real GDP per capita India	millions of US dollars	World Bank
	GDPCA	Relative real GDP per capita Canada	millions of US dollars	World Bank
	INFR	Chinese inflation rate	Percentage	World Bank
Tourist Arrival	SK	South Korean tourist arrival	person (thousand)	CEIC database
	JP	Japanese tourist arrival	person (thousand)	CEIC database
	RU	Russian tourist arrival	person (thousand)	CEIC database
	US	US tourist arrival	person (thousand)	CEIC database
	MO	Mongolian tourist arrival	person (thousand)	CEIC database
	MA	Malaysian tourist arrival	person (thousand)	CEIC database
	PH	Philippines tourist arrival	person (thousand)	CEIC database
	SI	Singaporean tourist arrival	person (thousand)	CEIC database
	IN	Indian tourist arrival	person (thousand)	CEIC database
	CA	Canadian tourist arrival	person (thousand)	CEIC database

References

1. Dritsakis, N. Tourism development and economic growth in seven Mediterranean countries: A panel data approach. *Tour. Econ.* **2012**, *18*, 801–816. [\[CrossRef\]](#)
2. Chiang, G.N.; Sung, W.Y.; Lei, W.G. Regime-switching effect of tourism specialization on economic growth in Asia Pacific countries. *Economies* **2017**, *5*, 23. [\[CrossRef\]](#)
3. Valadkhani, A.; O'Mahony, B. Identifying structural changes and regime switching in growing and declining inbound tourism markets in Australia. *Curr. Issues Tour.* **2018**, *21*, 277–300. [\[CrossRef\]](#)
4. Kitamura, R. The effects of added transportation capacity on travel. A review of theoretical and empirical results. *Transportation* **2009**, *36*, 745–762. [\[CrossRef\]](#)
5. Van Truong, N.; Shimizu, T. The effect of transportation on tourism promotion: Literature review on application of the Computable General Equilibrium (CGE). *Model. Transport. Res. Proc.* **2017**, *25*, 3096–3115. [\[CrossRef\]](#)
6. Khadaroo, J.; Seetanah, B. Transport infrastructure and tourism development. *Ann. Tour. Res.* **2007**, *34*, 1021–1032. [\[CrossRef\]](#)
7. Tang, C.; Weaver, D.; Lawton, L. Can stopovers be induced to revisit transit hubs as stayovers? A new perspective on the relationship between air transportation and tourism. *J. Air Transp. Manag.* **2017**, *62*, 54–64. [\[CrossRef\]](#)
8. Guirao, B.; Campa, J.L. The effects of tourism on HSR: Spanish empirical evidence derived from a multi-criteria corridor selection methodology. *J. Transp. Geogr.* **2015**, *47*, 37–46. [\[CrossRef\]](#)
9. Albalade, D.; Fageda, X. High speed rail and tourism: Empirical evidence from Spain. *Transp. Res. Part A Policy Pract.* **2016**, *85*, 174–185. [\[CrossRef\]](#)
10. Yin, P.; Pagliara, F.; Wilson, A. How does high-speed rail affect tourism? A case study of the capital region of China. *Sustainability* **2019**, *11*, 472. [\[CrossRef\]](#)
11. Pagliara, F.; Mauriello, F.; Russo, L. A Regression tree approach for investigating the impact of high speed rail on tourists' choices. *Sustainability* **2020**, *12*, 910. [\[CrossRef\]](#)

12. Yin, P.; Lin, Z.; Prideaux, B. The impact of high-speed railway on tourism spatial structures between two adjoining metropolitan cities in China: Beijing and Tianjin. *J. Transp. Geogr.* **2019**, *80*, 102495. [\[CrossRef\]](#)
13. Teye, V.B. Land transportation and tourism in Bermuda. *Tour. Manag.* **1992**, *13*, 395–405. [\[CrossRef\]](#)
14. Kanwal, S.; Rasheed, M.I.; Pitafi, A.H.; Pitafi, A.; Ren, M. Road and transport infrastructure development and community support for tourism: The role of perceived benefits, and community satisfaction. *Tour. Manag.* **2020**, *77*, 104014. [\[CrossRef\]](#)
15. Eric, T.N.; Semeyutin, A.; Hubbard, N. Effects of enhanced air connectivity on the Kenyan tourism industry and their likely welfare implications. *Tour. Manag.* **2020**, *78*, 104033. [\[CrossRef\]](#)
16. Ishiguro, K.; Inamura, H. Identification and elimination of barriers in the operations and management of maritime transportation. *Res. Transp. Econ.* **2005**, *13*, 337–368. [\[CrossRef\]](#)
17. Cottam, H.; Roe, M.; Challacombe, J. The impact of transitional changes of maritime passenger transport on Croatian tourism development. *Mar. Policy Manag.* **2007**, *34*, 185–202. [\[CrossRef\]](#)
18. Park, S.; Lee, W.S.; Moon, J.; Heo, J. Examination of Chinese cruise tourists' attributes using a Choice Experiment. *Sustainability* **2019**, *11*, 3621. [\[CrossRef\]](#)
19. Huang, X.L.; Ma, X.; Hu, F. Machine learning and intelligent communications. *Mobile Netw. Appl.* **2018**, *23*, 68–70. [\[CrossRef\]](#)
20. Cologni, A.; Manera, M. The asymmetric effects of oil shocks on output growth: A Markov–Switching analysis for the G-7 countries. *Econ. Modell.* **2009**, *26*, 1–29. [\[CrossRef\]](#)
21. Yang, Y.; Li, D.; Li, X. Public transport connectivity and intercity tourist flows. *J. Travel Res.* **2019**, *58*, 25–41. [\[CrossRef\]](#)
22. Pagliara, F.; La Pietra, A.; Gomez, J.; Vassallo, J.M. High Speed Rail and the tourism market: Evidence from the Madrid case study. *Transp. Policy* **2015**, *37*, 187–194. [\[CrossRef\]](#)
23. Button, K.; Taylor, S. International air transportation and economic development. *J. Air Transp. Manag.* **2000**, *6*, 209–222. [\[CrossRef\]](#)
24. Khan, S.A.R.; Qianli, D.; SongBo, W.; Zaman, K.; Zhang, Y. Travel and tourism competitiveness index: The impact of air transportation, railways transportation, travel and transport services on international inbound and outbound tourism. *J. Air Transp. Manag.* **2017**, *58*, 125–134. [\[CrossRef\]](#)
25. Hardy, A. An investigation into the key factors necessary for the development of iconic touring routes. *J. Vacat. Mark.* **2003**, *9*, 314–330. [\[CrossRef\]](#)
26. Chen, Z.; Haynes, K.E. Impact of high-speed rail on international tourism demand in China. *Appl. Econ. Lett.* **2015**, *22*, 57–60. [\[CrossRef\]](#)
27. Li, L.S.; Yang, F.X.; Cui, C. High-speed rail and tourism in China: An urban agglomeration perspective. *Int. J. Tour. Res.* **2019**, *21*, 45–60. [\[CrossRef\]](#)
28. Dong, D.; Xu, X.; Yu, H.; Zhao, Y. The impact of air pollution on domestic tourism in China: A spatial econometric analysis. *Sustainability* **2019**, *11*, 4148. [\[CrossRef\]](#)
29. Bazin, S.; Beckerich, C.; Delaplace, M. High speed railway, service innovations and urban and business tourism's development. In *Economics and Management of Tourism: Trends and Recent Developments*; Collecção, M., Sarmento, M., Matias, A., Eds.; Universidade Lusitana Editora: Lisboa, Portugal, 2011.
30. Yamaka, W.; Pastpipatkul, P.; Sriboonchitta, S. Business cycle of international tourism demand in Thailand: A Markov-switching Bayesian vector error correction model. In *International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making*; Springer: Cham, Switzerland, 2015; pp. 415–427.
31. Hsu, P.P. Examination of Taiwan's travel and tourism market cycle through a two-period Markov regime-switching model. *Tour. Manag.* **2017**, *63*, 201–208. [\[CrossRef\]](#)
32. Jin, S.; Yang, J.; Wang, E.; Liu, J. The influence of high-speed rail on ice–snow tourism in northeastern China. *Tour. Manag.* **2020**, *78*, 104070. [\[CrossRef\]](#)
33. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **1970**, *12*, 55–67. [\[CrossRef\]](#)
34. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Methodol.* **1996**, *58*, 267–288. [\[CrossRef\]](#)
35. Soh, A.N.; Puah, C.H.; Arip, M.A.; Kuek, T.H. Oil price and Fijian tourism cycle: A Markov regime-switching model. *Int. J. Energy Econ. Policy* **2019**, *9*, 188–192. [\[CrossRef\]](#)
36. Pastpipatkul, P.; Yamaka, W.; Sriboonchitta, S. Effect of quantitative easing on ASEAN-5 financial markets. In *Causal Inference in Econometrics*; Springer: Cham, Switzerland, 2016; pp. 525–543.
37. Akintug, B.; Rasmussen, P.F. A Markov switching model for annual hydrologic time series. *Water Resour. Res.* **2005**, *41*, W09424. [\[CrossRef\]](#)
38. Krolzig, H.M. *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 454.
39. Kim, C.J.; Piger, J.; Startz, R. Estimation of Markov regime-switching regression models with endogenous switching. *J. Econom.* **2008**, *143*, 263–273. [\[CrossRef\]](#)
40. Kim, C.J.; Nelson, C.R. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*; MIT Press Books: Cambridge, MA, USA, 1999; Volume 1.
41. Hamilton, J.D. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econom. J. Econom. Soc.* **1989**, 357–384. [\[CrossRef\]](#)

42. Chen, J.; Chen, Z. Extended Bayesian information criteria for model selection with large model spaces. *Biometrika* **2008**, *95*, 759–771. [\[CrossRef\]](#)
43. Foygel, R.; Drton, M. Extended Bayesian information criteria for Gaussian graphical models. In *Advances in Neural Information Processing Systems*; MIT Press Books: Cambridge, MA, USA, 2010; pp. 604–612.
44. Xie, J.; Tveterås, S. Economic drivers for the Chinese tourists. *Scand. J. Hosp. Tour.* **2020**, *20*, 110–125. [\[CrossRef\]](#)
45. Thana, E. Maritime transport and tourism sustainable development. *Eur. Sci. J.* **2013**, *9*, 1473–1483.
46. Assaf, A.G.; Josiassen, A. Identifying and ranking the determinants of tourism performance: A global investigation. *J. Travel Res.* **2012**, *51*, 388–399. [\[CrossRef\]](#)
47. Zidana, R. Macroeconomic Determinants of Tourism Sector Performance in Malawi. *Int. J. Soc. Sci. Hum. Res.* **2015**, *3*, 481–491.
48. Suh, H.; Kim, S.B. The Macroeconomic Determinants of International Casino Travel: Evidence from South Korea's Top Four Inbound Markets. *Sustainability* **2018**, *10*, 554. [\[CrossRef\]](#)
49. Song, H.; Li, G.; Witt, S.F.; Athanasopoulos, G. Forecasting tourist arrivals using time-varying parameter structural time series models. *Int. J. Forecast.* **2011**, *27*, 855–869. [\[CrossRef\]](#)
50. Tang, C.F. An exploration of dynamic relationship between tourist arrivals, inflation, unemployment and crime rates in Malaysia. *Int. J. Soc. Econ.* **2011**, *38*, 50–69. [\[CrossRef\]](#)
51. Athari, S.A.; Alola, U.V.; Ghasemi, M.; Alola, A.A. The (Un) sticky role of exchange and inflation rate in tourism development: Insight from the low and high political risk destinations. *Curr. Issues Tour.* **2020**, *1*, 1–16. [\[CrossRef\]](#)
52. Massidda, C.; Mattana, P. A SVECM analysis of the relationship between international tourism arrivals, GDP and trade in Italy. *J. Travel Res.* **2013**, *52*, 93–105. [\[CrossRef\]](#)
53. Tang, J.; Sriboonchitta, S.; Ramos, V.; Wong, W.K. Modelling dependence between tourism demand and exchange rate using the copula-based GARCH model. *Curr. Issues Tour.* **2016**, *19*, 876–894. [\[CrossRef\]](#)
54. Dickey, D.A.; Fuller, W.A. Likelihood ratio statistics for autoregressive time series with a unit root. *Econom. J. Econom. Soc.* **1981**, 1057–1072. [\[CrossRef\]](#)
55. Zivot, E.; Andrews, D.W.K. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *J. Bus. Econ. Stat.* **2002**, *20*, 25–44. [\[CrossRef\]](#)
56. Emmert-Streib, F.; Dehmer, M. High-dimensional LASSO-based computational regression models: Regularization, shrinkage, and selection. *Mach. Learn. Knowl. Extract.* **2019**, *1*, 21. [\[CrossRef\]](#)
57. Lau, Y.Y.; Yip, T.L. The Asia cruise tourism industry: Current trend and future outlook. *Asian J. Shipp. Logist.* **2020**, *36*, 202–213. [\[CrossRef\]](#)
58. Churchill, R.R.; Lowe, A.V. *The Law of the Sea*; Manchester University Press: Manchester, UK, 1999.
59. Bao, Y.; Xiao, F.; Gao, Z.; Gao, Z. Investigation of the traffic congestion during public holiday and the impact of the toll-exemption policy. *Transp. Res. Part B Methodol.* **2017**, *104*, 58–81. [\[CrossRef\]](#)
60. Van Can, V. Estimation of travel mode choice for domestic tourists to Nha Trang using the multinomial probit model. *Transp. Res. Part A Policy Pract.* **2013**, *49*, 149–159. [\[CrossRef\]](#)
61. Sio-Chong, U.; So, Y.C. The impacts of financial and non-financial crises on tourism: Evidence from Macao and Hong Kong. *Tour. Manag. Perspect.* **2020**, *33*, 100628.
62. Jin, F.; Lee, L.F. Lasso maximum likelihood estimation of parametric models with singular information matrices. *Econometrics* **2018**, *6*, 8. [\[CrossRef\]](#)
63. Wei, S.; Teng, S.N.; Li, H.J.; Xu, J.; Ma, H.; Luan, X.L.; Xu, C. Hierarchical structure in the world's largest high-speed rail network. *PLoS ONE* **2019**, *14*, e0211052. [\[CrossRef\]](#)
64. Thai, M.T.; Wu, W.; Xiong, H. (Eds.) *Big Data in Complex and Social Networks*; CRC Press: Boca Raton, FL, USA, 2016.
65. Stai, E.; Kafetzoglou, S.; Tsiropoulou, E.E.; Papavassiliou, S. A holistic approach for personalization, relevance feedback & recommendation in enriched multimedia content. *Multimed. Tools Appl.* **2018**, *77*, 283–326.