Forecasting inbound tourists in Cambodia

著者	Tanaka Kiyoyasu
権利	Copyrights 日本貿易振興機構（ジェトロ）アジア経済研究所 / Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO) http://www.ide.go.jp
journal or publication title	IDE Discussion Paper
volume	601
year	2016-06-01
URL	http://doi.org/10.20561/00037580

INSTITUTE OF DEVELOPING ECONOMIES

IDE Discussion Papers are preliminary materials circulated
to stimulate discussions and critical comments

IDE DISCUSSION PAPER No. 601

Forecasting Inbound Tourists in Cambodia

Kiyoyasu TANAKA*

June 2016

Abstract

Forecasting tourism demand is crucial for management decisions in the tourism sector. Estimating a vector autoregressive (VAR) model for monthly visitor arrivals disaggregated by three entry points in Cambodia for the years 2006–2015, I forecast the number of arrivals for years 2016 and 2017. The results show that the VAR model fits well with the data on visitor arrivals for each entry point. *Ex post* forecasting shows that the forecasts closely match the observed data for visitor arrivals, thereby supporting the forecasting accuracy of the VAR model. Visitor arrivals to Siem Reap and Phnom Penh airports are forecast to increase steadily in future periods, with varying fluctuations across months and origin countries of foreign tourists.

Keywords: Tourism demand, visitor arrivals, forecasting, VAR, Cambodia.
JEL classification: C53, L83, Z32

---

* Research Fellow, Institute of Developing Economies; Visiting Professor, Royal University of Phnom Penh (Kiyoyasu_Tanaka@ide.go.jp)

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.

The views expressed in this publication are those of the author(s). Publication does not imply endorsement by the Institute of Developing Economies of any of the views expressed within.

INSTITUTE OF DEVELOPING ECONOMIES (IDE), JETRO
3-2-2, WAKABA, MIHAMA-KU, CHIBA-SHI
CHIBA 261-8545, JAPAN

©2016 by Institute of Developing Economies, JETRO
No part of this publication may be reproduced without the prior permission of the IDE-JETRO.

Forecasting Inbound Tourists in Cambodia

Kiyoyasu Tanaka[§]

June 2016

Abstract

Forecasting tourism demand is crucial for management decisions in the tourism sector. Estimating a vector autoregressive (VAR) model for monthly visitor arrivals disaggregated by three entry points in Cambodia for the years 2006–2015, I forecast the number of arrivals for years 2016 and 2017. The results show that the VAR model fits well with the data on visitor arrivals for each entry point. *Ex post* forecasting shows that the forecasts closely match the observed data for visitor arrivals, thereby supporting the forecasting accuracy of the VAR model. Visitor arrivals to Siem Reap and Phnom Penh airports are forecast to increase steadily in future periods, with varying fluctuations across months and origin countries of foreign tourists.

Keywords**:** Tourism demand, visitor arrivals, forecasting, VAR, Cambodia

JEL Classification**:** C53, L83, Z32

[§] Institute of Developing Economies, JETRO; address: 3-2-2 Wakaba, Mihama-ku, Chiba-shi, Chiba 261-8545, Japan; e-mail: kiyoyasu_tanaka@ide.go.jp

1. Introduction

The growth of international tourism has been remarkable over the past decades. International tourist arrivals increased from 435 million in 1990 to 1,087 million in 2013. International tourism receipts reached 1,197 billion USD in 2013 (UNWTO, 2015). In the case of Cambodia, inbound tourists increased from 0.1 million tourists in 1993 to 4.5 million tourists in 2014. The estimated amount of tourism receipts in Cambodia increased from 0.1 billion USD in 1995 to 2.7 billion USD in 2014 (Cambodian Ministry of Tourism, 2014). As a result, the substantial growth of inbound tourists has contributed to the rapid expansion of the tourism industry in Cambodia.

Future tourism demand is critical for businesses in the tourism sector, including hotels, airline companies, and tour agents. Since tourism services are produced to serve tourists, production and consumption must occur simultaneously for most tourism services. Unused tourism services, such as empty airline seats and hotel rooms, cannot be stockpiled and the perishable nature of tourism services stands in stark contrast with manufacturing industries where an inventory adjustment can mitigate the issue of unsold manufactured products. An excessive supply of tourism services reduces the prices of tourism services and the profitability of tourism businesses. On the other hand, the supply of tourism services cannot always adjust instantaneously to meet a sharp increase in tourism demand. In the case of the hotel and lodging industry, the number of guest rooms is fixed in the short run, and long-term investment is often required to increase guest rooms. These features of tourism services make accurate forecasting of tourism demand a critical component for management decision making in the tourism sector.

In this paper, I use a dataset for the years 2006–2015 on the monthly number of visitor arrivals by country of residence and by the point of entry: (1) Phnom Penh international airport, (2) Siem Reap international airport, and (3) cross-border check points through land and waterways.[1] I construct samples for aggregate visitor arrivals and for specific regions of origin: Europe, northeast Asia, and Association of Southeast Asian Nations (ASEAN). I adopt a vector autoregressive (VAR) model to estimate a relationship between current and past visitor arrivals disaggregated by point of entry. The VAR model allows me to exploit up-to-date information on tourist arrivals across entry points for forecasting.

In estimating the VAR model, I conduct *ex post* and *ex ante* forecasting. Samples from January 2006 through December 2013 are used to estimate the VAR model, from

[1] While the exact definition of "visitor" and "tourist" may differ by researcher, I use these words interchangeably in this paper.

which I generate a series of *n*-periods-ahead forecast for tourist arrivals across entry points. Using samples from January 2014 to December 2015 as holdout samples, I compare the forecast and observed number of tourist arrivals in 2014 and 2015. The *ex post* forecasting exercise allows me to examine the forecasting accuracy of the VAR model. In the *ex ante* forecasting, I estimate the VAR model using all samples from January 2006 through December 2015 to forecast the monthly number of tourist arrivals for January 2016 through December 2017.

The findings of my analysis can be summarized as follows. First, the VAR model generally shows a high fit with the data on tourist arrivals by entry point. This result supports the validity of VAR specification for accounting for monthly tourist arrivals in Cambodia. Second, *ex post* forecasting practices provide evidence for the forecasting accuracy of the VAR model. In particular, the forecast number of tourists to Siem Reap and Phnom Penh airports tends to match the observed data more accurately than the number of tourists arriving by land and waterways. Lastly, *ex ante* forecasting shows that tourist arrivals to Siem Reap are forecast to increase steadily during both the low and high seasons in future periods. The number of tourist arrivals to Phnom Penh is also forecast to increase gradually in the forecast period, with relatively small fluctuations. Tourist arrivals by land and waterways are forecast to exhibit a large fluctuation over time, which should be interpreted with caution given the potential for large forecast error.

Because forecasting tourism demand is an important topic, various forecasting methods have been exploited in the prior literature. Athanasopoulos et al. (2011) evaluate the forecasting performances of various methods for tourism data including time series approaches and causal models with explanatory variables.[2] Song and Witt (2006) specify a standard tourism demand function to represent the determinants of tourism consumption such as tourists' income, relative tourism prices, and travel costs. They adopt a VAR model to estimate the structural relationship in tourism demand. In contrast, this paper does not focus on the estimation of structural relationships in tourism demand. The main objective is to exploit readily available data on visitor arrivals to Cambodia and to forecast the number of visitor arrivals based on the most recent information. In this respect, the VAR model is chosen to use the most recent data on visitor arrivals by entry point to forecast future arrivals. Additionally, there is little prior work on tourism demand in Cambodia. To the best of my knowledge, Hor and Thaiprasert (2015) is the only prior study to estimate tourism demand using panel data

[2] For instance, Santos et al., (2009) compare the forecasting performances of time-series models to forecast the number of overnight stays.

models. In this respect, forecasting tourism demand contributes to the understanding of tourism demand in Cambodia.

The rest of this paper is organized as follows. Section 2 provides a data description on tourist arrivals to Cambodia and outlines the empirical framework. Section 3 discusses the estimation results of the empirical model and compares *ex post* forecasts with observed data on tourist arrivals. Section 4 presents the forecast of tourist arrivals in future periods. Section 5 concludes.

2. Empirical Framework

In this section, I describe data on inbound tourists to Cambodia and outline an empirical model for estimating the number of tourist arrivals. Because the objective in this paper is to forecast the number of inbound tourists in future periods, an estimation of the structural relationships in tourism demand is not pursued. The selection of a forecasting method is made mainly on the basis of readily available data needed for up-to-date forecasting as well as the characteristics of tourist arrivals in Cambodia.

2.1. Data on Inbound Tourists in Cambodia

Data on inbound tourists to Cambodia are constructed from the series of *Annual Tourism Statistics Report* compiled by the Statistics and Tourism Information Department in the Ministry of Tourism in Cambodia. Using data sources provided by the Immigration Department in the Ministry of Interior in Cambodia, the report provides information on inbound and outbound tourists in Cambodia. In this paper, I exploit a dataset for the years 2006–2015 on the monthly number of visitor arrivals by country of residence and by point of entry to Cambodia through the following locations: (1) Phnom Penh international airport, (2) Siem Reap international airport, and (3) cross-border check points through land and waterways. Land and waterway entries include 16 border check points for land and 3 check points for international ports. The report provides information on tourist arrivals disaggregated by the purpose of visit, such as vacation, business, and so on. However, the data on tourist arrivals are aggregated across countries and entry points. Additionally, data on tourism expenditures are reported, but not disaggregated across countries and entry points. For these reasons, I focus on the number of arrivals.

Table 1 presents the total number of monthly inbound tourists for years 2006, 2010, and 2015 across entry points. The number of tourist arrivals for Siem Reap airport increased from 580,404 in 2006 to 1,404,174 in 2015. The arrivals for Phnom Penh airport increased from 419,344 in 2006 to 1,058,516 in 2015. Additionally, the arrivals by land and waterways increased from 469,772 in 2006 to 2,273,896 in 2015. For the past decade, the number of inbound tourists increased substantially across all entry points. For each entry point, there is a seasonal fluctuation in tourist arrivals. For instance, the arrivals in Siem Reap for 2015 numbered 166,451 in January and 74,424 in June. Such a fluctuation is also observed for tourist arrivals in Phnom Penh airport and land and waterways. In general, tourist arrivals tend to be greater in the dry season (November to May) and lower in the rainy season (May to October).

---Table 1 here---

Taken together, the available data on inbound tourists, from a forecasting perspective, can be summarized as follows. First, up-to-date information on monthly tourist arrivals are readily available for forecasting. The number of tourist arrivals is disaggregated by three entry modes: Siem Reap, Phnom Penh, and land/waterways. Second, the data on inbound tourists exhibit both upward and seasonality trends across entry points. This suggests that the data on tourist arrivals in one entry mode may contain useful information for the data on tourist arrivals in another entry mode. For instance, tourist arrivals to Phnom Penh airport in past periods may improve forecasting of tourist arrivals to Siem Reap in future periods, and vice versa. Finally, the data on tourist arrivals are disaggregated by tourists' country of residence. While forecasting aggregate tourist arrivals is of primary interest, the data allow me to forecast tourist arrivals from specific regions and/or countries.

2.2. Empirical Model

Based on the previous discussions up to this point, I adopt a vector autoregressive (VAR) model to forecast the number of monthly inbound visitor arrivals disaggregated by point of entry (Phnom Penh airport, Siem Reap airport, and land and waterway check points). To exploit past information on tourists for forecasting, I specify the VAR model for year-month *t* and the number of lags *p* for arrivals as follows:

$$Y_t = \Gamma_0 + \Gamma_1 Y_{t-1} + \Gamma_2 Y_{t-2} + \cdots + \Gamma_p Y_{t-p} + f_q + \gamma t + \varepsilon_t \qquad (1)$$

where Y_t is a vector of variables on tourist arrivals {TRSR, TRPP, TRLW}. *TRSR* is the number of visitor arrivals to Siem Reap airport. *TRPP* is the number of visitor arrivals to Phnom Penh airport. *TRLW* is the number of arrivals through land and waterway checkpoints. f_q is a set of seasonal dummy variables for quarter q. The linear time trend is captured by the time period variable, *t*. Finally, ε_t is the error term.

As discussed previously, the data on tourist arrivals exhibit seasonal fluctuations across months. Quarterly-level dummy variables are included to capture the seasonality effects that are common across years. However, there may remain useful information in past months for forecasting monthly tourist arrivals for each year. For instance, the number of tourist arrivals in a month for this year is likely to be a good predictor of the number of tourist arrivals in the same month for next year. To exploit information on the same month for the previous year, I include 12 lags for each variable in equation (1) to capture the impact of each of the preceding 12 months. To check if 12 months of lags are sufficiently long, I examine Akaike's information criterion (AIC), Schwarz's Bayesian information criterion (SBIC), and the Hannan and Quinn information criterion (HQIC) lag-order selection statistics. On examination, these statistics generally support the lag length of the past 12 months in the sense that the lag-order selection statistics do not strongly indicate the need for longer than 12 months.

I also conduct *ex post* and *ex ante* forecasting practices. For *ex post* forecasting, I use samples from January 2006 through December 2013 to estimate the VAR model. Based on the estimation results, I generate the one-period-ahead forecasts as follows:

$$Y_{t+1} = \hat{\Gamma}_0 + \hat{\Gamma}_1 Y_t + \hat{\Gamma}_2 Y_{t-1} + \cdots + \hat{\Gamma}_{12} Y_{t-11} + \widehat{f_q} + \hat{\gamma} t \tag{2}$$

where a hat indicates the estimated coefficients. In the same way, I generate a series of *n*-periods-ahead forecast for tourist arrivals across entry points. Using samples from January 2014 to December 2015 as holdout samples, I compare the forecast and observed number of tourist arrivals in the years 2014 and 2015. The *ex post* forecasting provides a method to examine the forecasting accuracy of the VAR model for tourist arrivals in each entry mode. Additionally, I conduct *ex ante* forecasting of tourist arrivals, which is the primary objective of this paper. Using all samples from January 2006 through December 2015, I estimate the VAR model for future periods. Based on the estimation results, I forecast the number of tourist arrivals for the 24-month period from January 2016 to December 2017.

It should be emphasized that my empirical model is not intended to estimate a structural relationship in tourism demand on the basis of consumer theory in microeconomics (Candela and Figini, 2012, Chapters 5 and 6). A standard function of tourism demand suggests that tourism consumption depends on tourists' income, relative tourism prices, and travel costs (Song et al., 2009, Chapter 1). In the prior literature, such a structural relationship was estimated using a VAR model (Song and Witt, 2006). In contrast, my VAR approach is distinctive in that current tourist arrivals depends on a vector of past tourist arrivals in all points of entry to Cambodia. As a theoretical justification for my approach, a possible interpretation is that tourists in a current period take into account economic factors in the current period, whereas those in past periods considered similar economic factors in past periods. Given that the number of tourist arrivals is a reflection of tourists' decisions based on these economic factors, the number of tourist arrivals should reflect the economic information considered by tourists. Consequently, the number of past tourist arrivals should deliver crucial economic information to be considered in the decisions by tourists in the current period.[3]

3. Estimation Results

This section presents the estimation results of the VAR model for aggregate tourist arrivals in Cambodia.[4] I also estimate the VAR model for tourist arrivals from specific regions: Europe, northeast Asia, and ASEAN. As the estimation results of the VAR model are extensive, I discuss only diagnostic statistics for these regions.[5] Finally, I present the results of *ex post* forecasting of tourist arrivals.

3.1. VAR Results

I start by discussing the estimation results of the VAR model in equation (1) for aggregate tourist arrivals. The sample periods are from January 2007 to December 2013, and the number of observations is 84. Table 2 shows the result of the equation for number of tourist arrivals to Siem Reap airport. I find that some of the lagged variables for tourists to Siem Reap are significant, indicating that the number of past tourist arrivals to Siem Reap has the significant relationship with the number of tourist arrivals to Siem Reap in a current period. I also find that there are significant coefficients for the lagged tourists to Phnom Penh and by land and waterways. This implies that the number of past tourist arrivals in other entry points has a significant relationship with the number of tourist arrivals to Siem Reap even after exploiting past information on tourists to Siem Reap. Additionally, the quarterly dummy variables have significant coefficients. Compared with the baseline quarter from April through July, the arrivals to Siem Reap are significantly larger by 13,410 for July to August, by 22,854 for September to December, and by 15,668 for January to March. Finally, the linear time trend is not significant.

---Table 2 here---

[3] Another justification is that up-to-date data on these variables are not readily available at the monthly level across countries, making it difficult to estimate a structural demand function for forecasting monthly tourist arrivals in Cambodia.
[4] Appendix Table 1 presents a list of countries in the sample.
[5] Estimation results are available upon request.

Table 3 presents the result of the equation for tourist arrivals to Phnom Penh airport. I find that some lagged variables for the number of tourists to Siem Reap and Phnom Penh airports are significant. For instance, the number of tourists to Phnom Penh in a current period correlates significantly and positively with the one-month lag of tourists to Phnom Penh. It is also significantly and positively associated with the third-month lag of tourists to Siem Reap. On the other hand, the lagged number of past tourist arrivals by land and waterways has no significant correlation with the number of current tourist arrivals to Phnom Penh. The quarterly dummy variables are significant and positive, consistent with the result for the equation of tourists to Siem Reap. The time trend variable is not significant.

---Table 3 here---

Table 4 presents the result of the equation for tourist arrivals by land and waterways. The result shows that the number of tourists by land and waterways in a current period correlates significantly with some of the lagged variables of tourist arrivals to Siem Reap and Phnom Penh. While tourists arriving by land and waterways in past periods do not correlate with Phnom-Penh tourists in a current period, the Phnom-Penh tourists in some past periods have a significant correlation with the tourists arriving by land and waterway in a current period. The arrivals by land and waterways in the 6-month and 12-month lagged periods also have a significant correlation with the current number of tourists. Finally, the quarterly dummy variable is significant and positive only for January to March, whereas the time trend variable is significant and positive. This implies that the seasonality and long-term trends in tourists arriving by land and waterways differ from those for tourists arriving to Siem Reap and Phnom Penh.

---Table 4 here---

Column (1) in Table 5 presents the summary of diagnostic statistics for the VAR model of aggregate tourist arrivals. The R-squared values are 0.981 for the equation of tourists to Siem Reap airport, 0.976 for the equation of tourists to Phnom Penh airport, and 0.994 for the equation of tourists arriving by land and waterways. These results indicate that the VAR model fits significantly well with the data on tourist arrivals across entry points. Comparing root mean squared errors (RMSE) across the equations, the equation for Phnom Penh has the smallest RMSE. Additionally, I present the diagnostic statistics of the VAR model for tourist arrivals from Europe in Column (2), northeast Asia in Column (3), and ASEAN in Column (4).[6] Across alternative samples, the VAR model generally shows a high fit with the data on tourist arrivals across entry modes. Thus, these results provide support for the VAR specification to explain a variation in tourist arrivals in Cambodia.

---Table 5 here---

3.2. Comparison of Forecast with Observed Data

I proceed to examine the forecasting accuracy based on the estimated VAR model for samples from January 2006 to December 2013. Samples from January 2014 to December 2015 are reserved as holdout samples to compare the forecast and observed number of tourist arrivals for these periods. Figure 1 presents the *ex post* forecast and observed data on aggregate tourists. Note that the number of tourist arrivals is shown in log scale.[7] It shows that the forecasts for Siem Reap and Phnom Penh airports closely correspond with observed data, whereas the forecast for land and waterway arrivals tends to have large deviations from the observed data. Since tourist arrivals by land and waterways are larger in number than other entry points, the deviation between forecast and observed data tends to be larger. On the other hand, the forecast number of tourists to Siem Reap appears to be slightly underestimated across periods as compared with the observed data. The forecast number tourists to Phnom Penh also tend to be underestimated in many periods. Taken together, these results suggest that the forecasts for Siem Reap and Phnom Penh arrivals tend to be more accurate than for land and waterway arrivals.

---Figure 1 here---

Next, I turn to examine the *ex post* forecasts for tourist arrivals from the various regions. Figure 2 presents the results for Europe. While the forecast number of tourists arriving by land and waterways are also generated, these forecasts tend to exhibit relatively large deviations from observed data. For this reason, only forecasts for international airports are shown. Consistent with the results for the aggregate sample of tourists, the forecast number of tourists from Europe closely corresponds with the observed data. In

[6] Europe includes 6 countries with the largest inbound tourists from European region: France, Germany, Italy, Russia, Spain, and United Kingdom. Northeast Asia includes China, Japan, South Korea, and Taiwan. ASEAN includes Brunei Darussalam, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, and Vietnam.

[7] The following figures also show the number of tourist arrivals in log scale.

particular, the fluctuations in tourist arrivals are close between the forecast and observed data for Phnom Penh. The forecast for Siem Reap tends to be a slight underestimate compared with the observed data, with the difference more pronounced during the low season. Figure 3 presents the results for northeast Asia. Consistent with the previous results, the forecast number of tourists appears to match the observed data well. In contrast with the results of Europe, the forecast number of tourists tends to be an overestimation relative to observed data for Siem Reap during the high season. Also, deviations in forecast tourists for Phnom Penh tend to increase in periods further out from the projection sample. Finally, Figure 4 shows the results for ASEAN. The forecast number of tourists to Phnom Penh appears to capture observed data accurately, whereas forecast tourists to Siem Reap are not well matched with fluctuations in the observed data. The number of forecast tourists arriving by land and waterways do not match the observed data in the sense that some fluctuations in the observed data are not reflected in the forecast. The forecast number of tourists appears to exhibit a linear time trend that is not apparent in the observed data.

---Figures 2, 3, and 4 here---

Taken together, the *ex post* forecasts provide evidence for the accuracy of this method for forecasting tourist arrivals, which shows several findings. First, the forecast number of tourists to Siem Reap and Phnom Penh airports tends to be more accurate than the forecast for tourists arriving by land and waterways. Across the different regions of tourist residence, the fluctuations in the forecast number of tourists arriving in Siem Reap and Phnom Penh tend to match the fluctuations in the observed data. Second, the VAR model tends to generate less accurate forecasts for tourists arriving by land and waterways. A possible reason is that visitors arriving by land and waterways have a variety of visit purposes, modes of transport, and types of border check points to cross. These characteristics may make it difficult to generate an accurate forecast. Lastly, the deviation between forecast and observed data does not necessarily increase with future time periods. At least when forecasting for a period up to two years in advance, the *ex post* forecasting results provide evidence for the accuracy of the forecasting in both the short-term and long-term periods.

4. Forecasting Inbound Tourists in Future Periods

I turn to conduct *ex ante* forecasting of tourist arrivals. As explained in section 2, I first re-estimate the VAR model for the whole samples from January 2006 through

December 2015, and use the estimation results to generate the forecast number of tourist arrivals in the upcoming 24 months from January 2016 through December 2017. The forecast number of tourist arrivals is generated both for aggregate tourists and tourists from specific regions.

Figure 5 presents the forecast number of total tourists by entry points. Tourist arrivals to Siem Reap exhibit a sharp fluctuation between low and high seasons. Despite this large seasonal variation, the number of tourists is forecast to increase steadily in both the low and high seasons over the forecast period. This result has crucial implications for management decisions in the tourism sector and at Siem Reap airport. As tourism demand is expected to increase in future periods, an expansion of tourism supply is needed to meet this growth in tourist demand. However, a large fluctuation in forecast arrivals poses a challenge for the tourism sector to address an excessive capacity of tourism supply in the low season. Additionally, the number of tourist arrivals to Phnom Penh is forecast to fluctuate in future periods, but also to increase gradually at the end of forecast periods. Lastly, tourist arrivals by land and waterways are forecast to exhibit a large fluctuation over time. While the *ex post* forecasting indicates a potentially large forecast error for tourists arriving by land and waterways, the *ex ante* forecasting indicates an upward trend over time.

---Figure 5 here---

Figure 6 shows the forecasts for arrivals to Siem Reap and Phnom Penh airports from Europe. The forecast number of tourists exhibit sharp fluctuations across months, as seen for observed data in past periods. Interestingly, such variations are not necessarily characterized by the distinction between dry and rainy seasons in Cambodia. In particular, there is a sharp increase in tourist arrivals to Siem Reap for August. Despite the monthly fluctuations, tourist arrivals are forecast to increase through the end of forecast period. Additionally, Figure 7 presents the predicted tourist arrivals to Siem Reap and Phnom Penh airports from northeast Asia. Compared with the result for Europe, tourist arrivals to Siem Reap exhibit clear movements between high and low reasons. Tourist arrivals are predicted to increase gradually over the forecast period. On the other hand, tourist arrivals from northeast Asia to Phnom Penh fluctuate significantly, but do not increase markedly in the forecast period.

---Figures 6 and 7 here---

Figure 8 shows the forecast number of tourist arrivals from ASEAN. Tourist arrivals to Siem Reap and Phnom Penh airports are predicted to increase steadily in the forecast period. Compared with the forecasts for Europe and northeast Asia, fluctuations in tourist arrivals appear to be smaller. On the other hand, tourist arrivals by land and waterways exhibit unclear patterns over time. Because the *ex post* forecasting exercise showed a relatively large forecasting error for tourist arrivals by land and waterways, this result must be interpreted with caution.

---Figure 8 here---

5. Conclusion

The remarkable growth of international tourists has contributed to the expansion of the tourism sector in developing economies. Since the success of tourism businesses critically depends on tourism demand, forecasting future demand is of great interest for management decisions in the tourism industry for businesses such as hotels, airlines, and tour agents. In this paper, I adopt the VAR model to forecast the monthly number of inbound visitor arrivals in Cambodia based on the sample from January 2006 to December 2015. Data on arrivals are disaggregated by Siem Reap airport, Phnom Penh airport, and land and waterway check points. To forecast arrivals from specific regions, I also disaggregate total visitor arrivals by tourists arriving from Europe, northeast Asia, and ASEAN.

The results show that the VAR model fits well with data on tourist arrivals across entry points. Movements in forecast and observed data on tourist arrivals are closely matched in *ex post* forecasting. In particular, the forecasts are more accurate for arrivals to Siem Reap and Phnom Penh airports than for arrivals by land and waterways. By forecasting the monthly number of total tourist arrivals in 2016 and 2017, I find that the arrivals increase steadily in both low and high seasons for Siem Reap, whereas arrivals to Phnom Penh increase gradually with relatively small fluctuations. Tourist arrivals by land and waterways tend to contain a relatively large forecast error.

This paper presents a formal approach based on available past information on inbound tourism demand to forecast future tourist arrivals. The forecasting results should provide useful information for the tourism industry in Cambodia. However, there remain several issues to improve the forecast. While the VAR model indicates a high fit with data, there may be cointegration relationships among tourist variables. In this case, my approach can be extended to estimate the vector error-correction model, which may improve the accuracy of the forecast results. On the other hand, univariate time-series methods can also be used to compare the forecasting performances of alternative empirical models. Additionally, my approach relies on the fact that forecasting future tourist arrivals exploits only past information on tourist arrivals. While unexpected events in future periods may affect inbound tourism demand, it is challenging to precisely take such events into account in *ex ante* forecasting. Because forecasting tourism demand is subject to uncertainty in the environment, forecast results must be interpreted with caution.

References

Cambodian Ministry of Tourism, 2014. *Annual Tourism Statistics Report* 2014. Phnom Penh.

Candela, Guido and Paolo, Figini. 2012. *The Economics of Tourism Destinations*. Heidelberg: Springer.

Athanasopoulos, George, Rob J., Hyndman, Haiyan, Song, and Doris C., Wu. 2011. The tourism forecasting competition. *International Journal of Forecasting*, 27 (3), 822-844.

Hor, Chantha, and Nalitra Thaiprasert. 2015. Analysis of international tourism demand for Cambodia. In Huynh, V.-N., Kreinovich, V., Sriboonchitta, S., Suriya, K. (Eds.), *Econometrics of Risk*, 414-425. Switzerland: Springer International Publishing.

Santos, Carlos, Gualter, Couto, and Pedro M., Pimentel. 2009. Forecasting hotel overnights in the autonomous region of the Azores. In Matis, Á., Nijkamp, P., Sarmento, M. (Eds.), *Advances in Tourism Economics – New Development*, 175-186. Heidelberg: Physica-Verlag Heidelberg

Song, Haiyan, and Stephen F. Witt. 2006. Forecasting international tourist flows to Macau. *Tourism Management*, 27 (2), 214-224.

Song, Haiyan, Stephen F., Witt, and Gang, Li. 2009. *The Advanced Econometrics of Tourism Demand*. New York: Taylor & Francis.

United Nations World Tourism Organization, 2015. UNWTO Tourism Highlights 2015 Edition, Madrid.

Table 1. Total number of monthly inbound tourists in Cambodia

	Siem Reap International Airport			Phnom Penh International Airport			Land and Waterways		
Month	2006	2010	2015	2006	2010	2015	2006	2010	2015
1	59,180	72,158	166,451	34,699	54,303	93,573	43,814	88,788	196,804
2	55,539	69,763	156,325	33,879	51,641	84,137	42,920	88,012	187,493
3	53,086	69,844	133,914	32,664	52,282	94,224	38,840	95,135	186,067
4	40,090	53,513	100,969	33,038	43,591	79,663	36,731	92,077	177,164
5	31,234	42,002	81,250	28,307	40,804	79,019	31,050	77,558	152,450
6	33,122	38,673	74,424	29,747	40,520	79,909	29,713	78,951	161,081
7	37,309	47,933	99,202	34,377	49,229	93,707	36,577	97,444	169,505
8	43,775	57,985	110,874	34,191	46,374	92,631	43,920	93,767	160,045
9	37,541	45,062	86,777	31,251	40,502	78,536	35,755	79,199	146,679
10	48,783	58,031	103,719	36,776	49,083	86,926	39,392	90,721	214,184
11	66,008	78,696	135,438	43,318	59,526	97,402	42,355	103,727	207,997
12	74,737	76,329	154,831	47,097	60,265	98,789	48,705	105,215	314,427
Total	580,404	709,989	1,404,174	419,344	588,120	1,058,516	469,772	1,090,594	2,273,896

Table 2. VAR estimation result for aggregate tourists to Siem Reap airport

	Tourists to Siem Reap			Tourists to Phnom Penh			Tourists by Land/waterway		
	Coefficient		t statistic	Coefficient		t statistic	Coefficient		t statistic
Lag 1	0.90	***	7.39	-0.21		-0.95	0.0003		0.00
Lag 2	-0.29	*	-1.82	-0.47	*	-1.76	0.14		1.26
Lag 3	0.54	**	3.11	0.09		0.34	-0.19	*	-1.96
Lag 4	-0.99	***	-5.37	0.49	*	1.87	0.25	**	2.41
Lag 5	0.86	***	5.17	0.10		0.37	-0.21	*	-1.94
Lag 6	-0.41	**	-2.33	0.01		0.04	0.28	**	2.47
Lag 7	0.34	*	1.81	-0.20		-0.8	-0.09		-0.75
Lag 8	-0.71	***	-4.04	0.38		1.48	-0.14		-1.18
Lag 9	0.66	***	3.94	-0.51	*	-1.92	-0.07		-0.60
Lag 10	0.02		0.11	-0.43		-1.41	0.03		0.29
Lag 11	0.32	*	1.83	-0.49		-1.49	0.08		0.78
Lag 12	-0.01		-0.06	-0.003		-0.01	0.02		0.21
Dummy for Jul.-Aug.	13,410	***	3.03						
Dummy for Sep.-Dec.	22,854	***	4.36						
Dummy for Jan.-Mar.	15,668	***	3.91						
Trend	148.4		0.82						
Constant	18,588	**	2.16						

Notes: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 3. VAR estimation result for aggregate tourists to Phnom Penh airport

	Tourists to Siem Reap			Tourists to Phnom Penh			Tourists by Land/waterway	
	Coefficient		t statistic	Coefficient		t statistic	Coefficient	t statistic
Lag 1	0.03		0.44	0.38	***	3.45	-0.08	-1.65
Lag 2	-0.05		-0.63	-0.03		-0.22	0.01	0.20
Lag 3	0.23	***	2.74	-0.03		-0.27	0.04	0.75
Lag 4	-0.17	*	-1.96	0.07		0.57	-0.004	-0.09
Lag 5	0.10		1.23	0.20		1.59	0.02	0.36
Lag 6	-0.04		-0.42	0.11		0.87	-0.05	-0.89
Lag 7	0.13		1.39	-0.24	**	-2.04	0.01	0.21
Lag 8	-0.15	*	-1.82	0.16		1.25	-0.06	-1.03
Lag 9	0.06		0.77	0.01		0.04	0.00	0.03
Lag 10	0.14		1.61	-0.54	***	-3.70	0.01	0.24
Lag 11	0.05		0.59	0.09		0.54	0.05	0.99
Lag 12	-0.04		-0.44	0.04		0.30	0.07	1.47
Dummy for Jul.-Aug.	8,242	***	3.85					
Dummy for Sep.-Dec.	13,858	***	5.48					
Dummy for Jan.-Mar.	9,154	***	4.72					
Trend	121.8		1.39					
Constant	8,494	**	2.04					

Notes: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 4. VAR estimation result for aggregate tourists arriving by land and waterways

	Tourists to Siem Reap		Tourists to Phnom Penh		Tourists by Land/waterway	
	Coefficient	t statistic	Coefficient	t statistic	Coefficient	t statistic
Lag 1	-0.17	-1.38	0.61 **	2.64	0.04	0.35
Lag 2	0.54 ***	3.27	-0.88 ***	-3.23	-0.15	-1.30
Lag 3	0.11	0.61	-0.61 **	-2.19	0.03	0.29
Lag 4	-0.15	-0.80	0.10	0.36	-0.10	-0.98
Lag 5	0.27	1.57	0.41	1.51	0.10	0.94
Lag 6	0.15	0.85	-0.96 ***	-3.48	0.33 ***	2.85
Lag 7	0.05	0.28	-0.20	-0.80	0.07	0.54
Lag 8	-0.25	-1.40	0.62 **	2.35	0.13	1.08
Lag 9	0.12	0.73	0.12	0.43	-0.08	-0.71
Lag 10	-0.15	-0.80	-0.42	-1.36	0.14	1.27
Lag 11	0.30 *	1.72	0.08	0.23	0.08	0.74
Lag 12	0.34 *	2.01	-0.08	-0.28	-0.22 **	-2.24
Dummy for Jul.-Aug.	-5008.6	-1.11				
Dummy for Sep.-Dec.	-3609.8	-0.68				
Dummy for Jan.-Mar.	8791.0 **	2.15				
Trend	1126.9 ***	6.07				
Constant	-3136.3	-0.36				

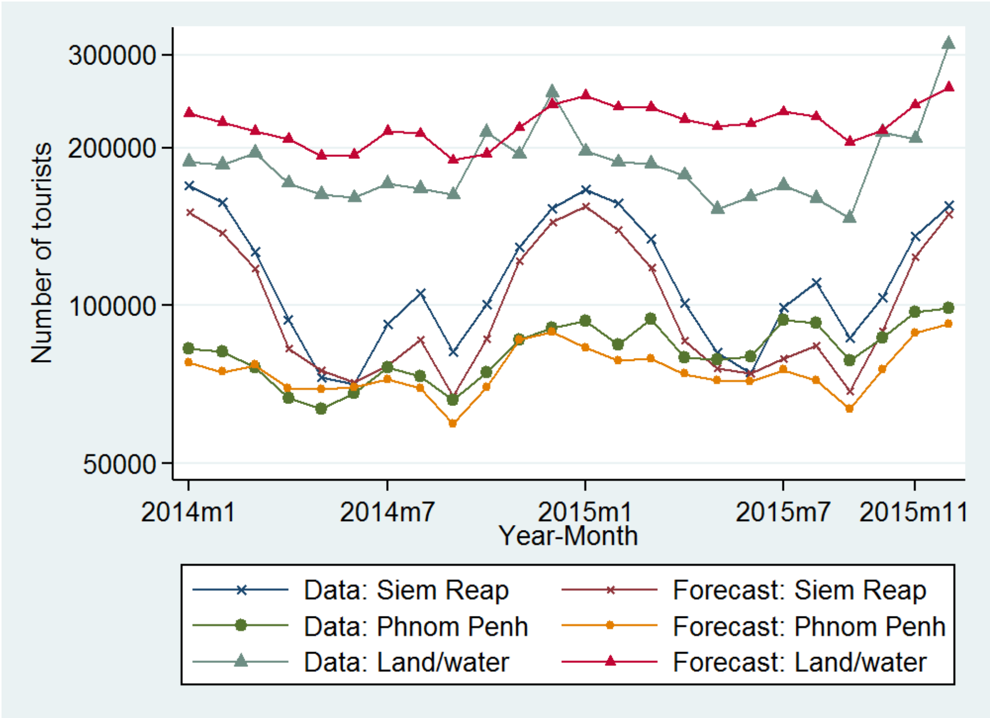
Notes: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 5. VAR estimation results of monthly inbound tourists to Cambodia

	(1)	(2)	(3)	(4)
Sample	Total	Europe	Northeast Asia	ASEAN
Sample period	2007m1 - 2013m12	2007m1 - 2013m12	2007m1 - 2013m12	2007m1 - 2013m12
No. of observations	84	84	84	84
Log likelihood	-2337.1	-1949.5	-2185.1	-2072.5
Akaike information criterion	58.6	49.3	55.0	52.3
Hannan-Quinn information criterion	60.0	50.8	56.4	53.7
Schwarz-Bayesian information criterion	62.1	52.9	58.5	55.8
Equation of tourists to Siem Reap airport				
Root mean squared error	4777.0	872.8	4243.5	600.0
R-squared	0.9811	0.9692	0.9702	0.9362
Equation of tourists to Phnom Penh airport				
Root mean squared error	2307.9	527.1	1391.3	888.0
R-squared	0.9766	0.9710	0.9529	0.9413
Equation of tourists by land and waterway				
Root mean squared error	4870.7	1231.7	1329.2	4644.4
R-squared	0.9946	0.9843	0.8259	0.9922

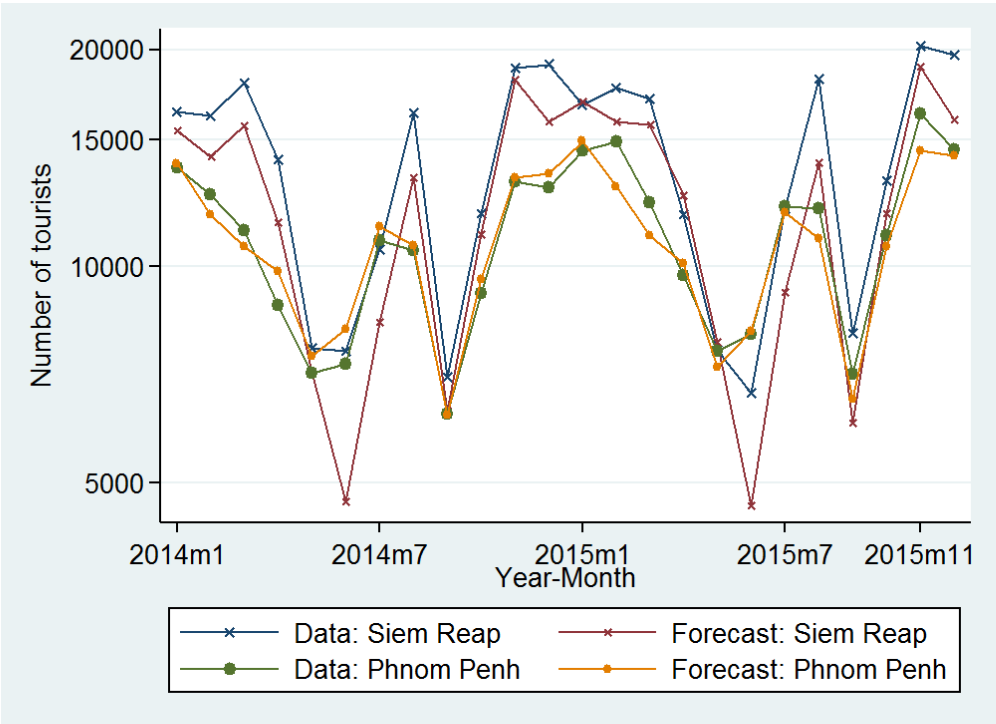
Notes: Europe includes France, Germany, Italy, Russia, Spain, and United Kingdom; northeast Asia includes China, Japan, South Korea, and Taiwan; ASEAN includes Brunei Darussalam, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, and Vietnam.

Figure 1. Ex post forecast and observed data on aggregate tourists



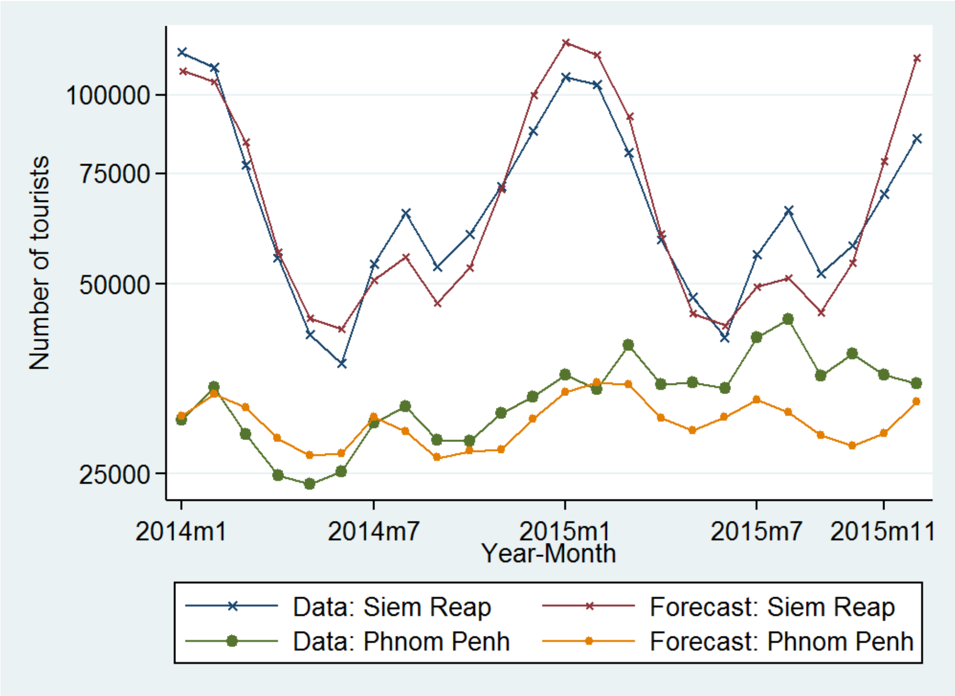
Notes: The number of tourists is shown in log scale.

Figure 2. Ex post forecast and observed data on tourists from Europe



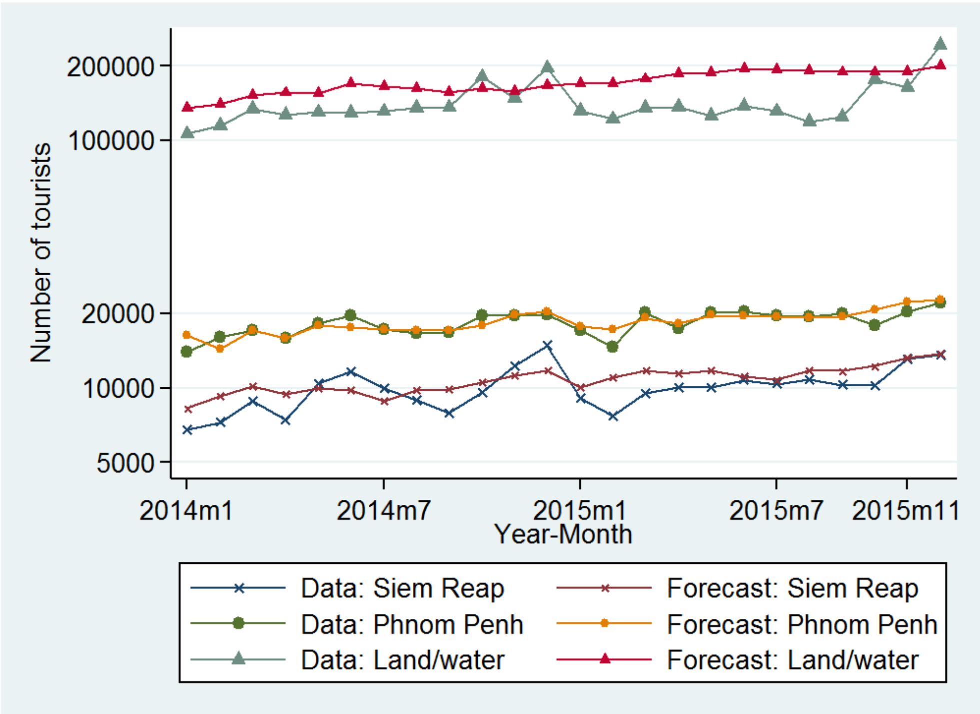
Notes: The number of tourists is shown in log scale.

Figure 3. Ex post forecast and observed data on tourists from northeast Asia



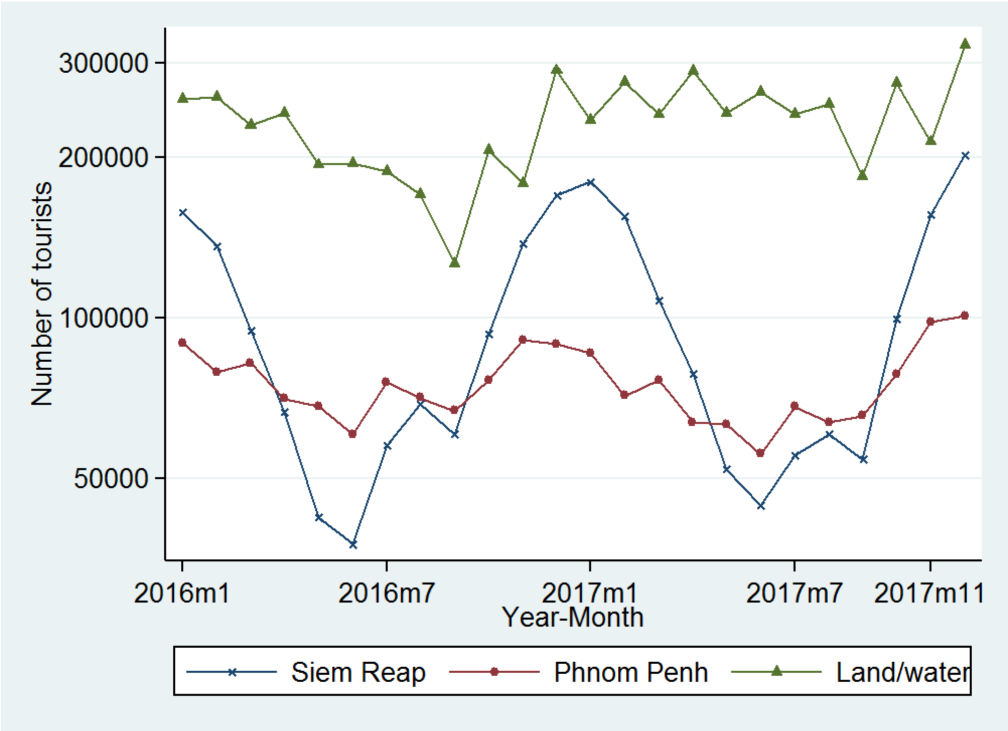
Notes: The number of tourists is shown in log scale.

Figure 4. Ex post forecast and observed data on tourists from ASEAN



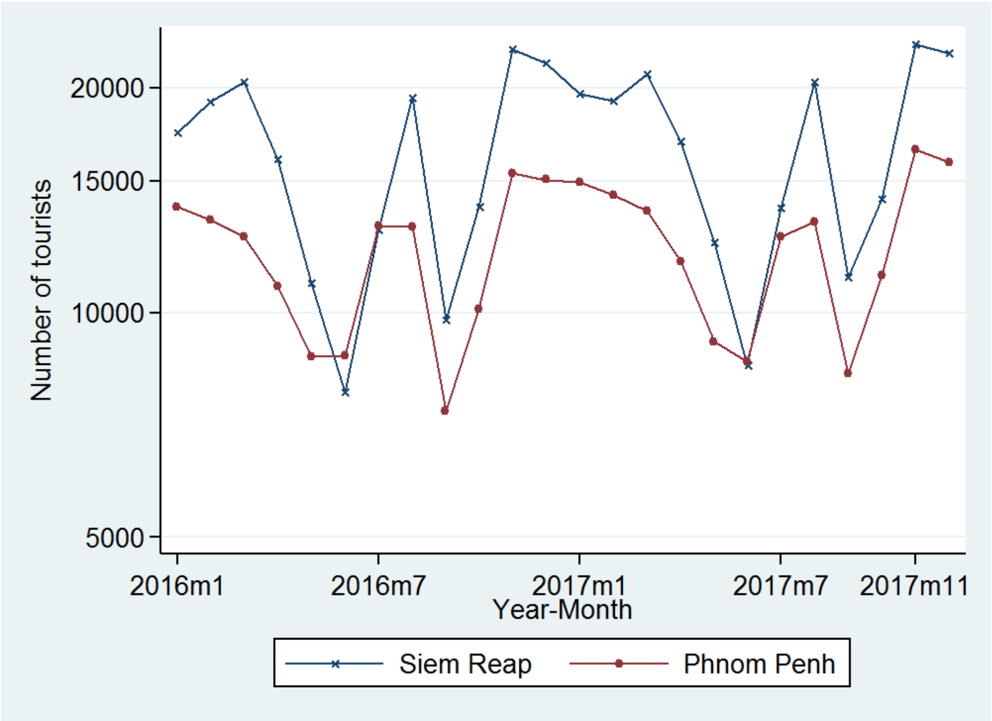
Notes: The number of tourists is shown in log scale.

Figure 5. Ex ante forecast of aggregate tourists



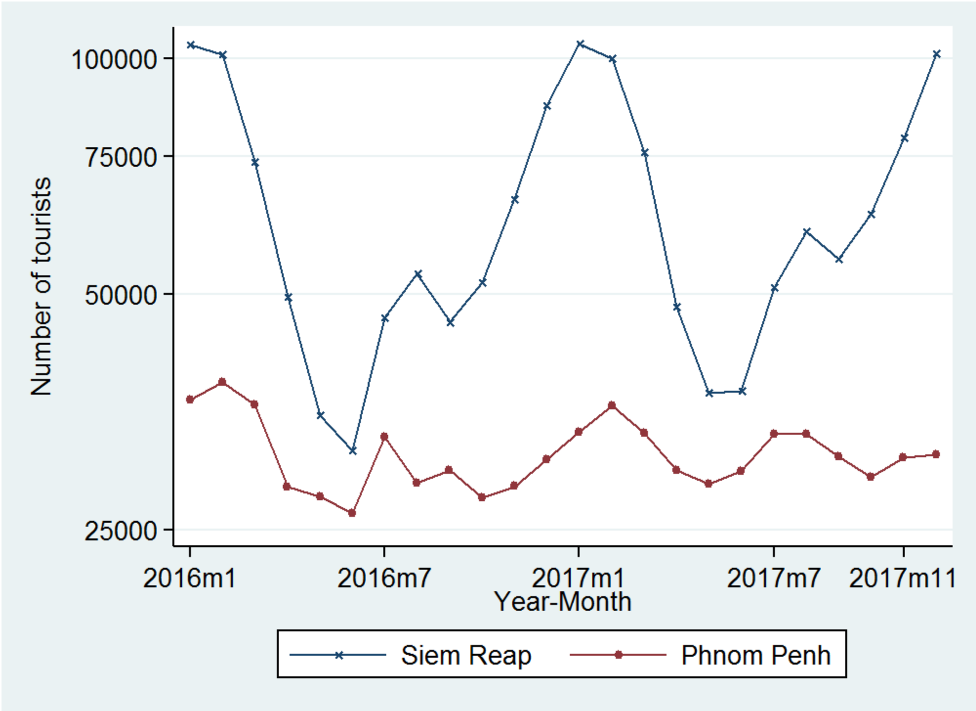
Notes: The number of tourists is shown in log scale.

Figure 6. Ex ante forecast of tourists from Europe



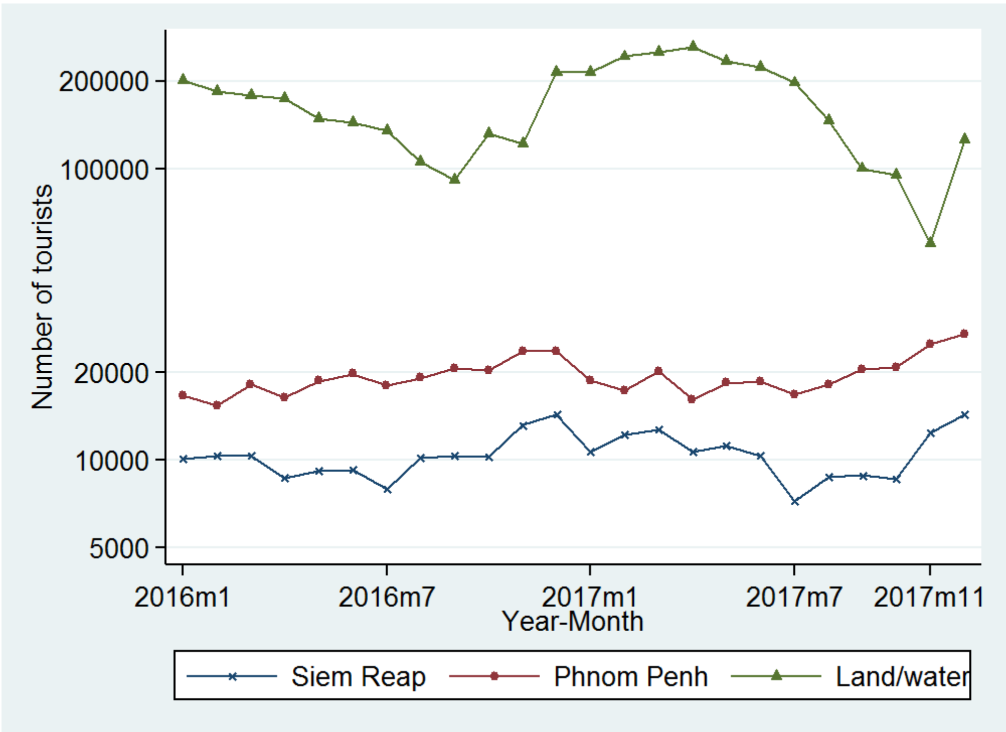
Notes: The number of tourists is shown in log scale.

Figure 7. Ex ante forecast of tourists from northeast Asia



Notes: The number of tourists is shown in log scale.

Figure 8. Ex ante forecast of tourists from ASEAN



Notes: The number of tourists is shown in log scale.

Appendix Table 1. List of sample countries

Afghanistan	Germany	Macau	Slovenia
Argentina	Ghana	Malaysia	South Africa
Australia	Greece	Malta	South Korea
Austria	Honduras	Mexico	Spain
Bangladesh	Hong Kong	Monaco	Sri Lanka
Belgium	Hungary	Mongolia	Sudan
Bhutan	Iceland	Morocco	Sweden
Brazil	India	Myanmar	Switzerland
Brunei Darussalam	Indonesia	Nepal	Syria
Bulgaria	Iran	Netherlands	Taiwan, China
Cameroon	Iraq	New Zealand	Tanzania
Canada	Ireland	Nigeria	Thailand
Chile	Israel	Norway	Tunisia
China	Italy	Oman	Turkey
Colombia	Jamaica	Pakistan	Uganda
Congo	Japan	Peru	Ukraine
Costa Rica	Jordan	Philippines	United Arab Emirates
Croatia	Kazakhstan	Poland	United Kingdom
Czech Republic	Kenya	Portugal	United States
Denmark	Kuwait	Qatar	Uruguay
Ecuador	Kyrgyzstan	Romania	Uzbekistan
Egypt	Laos	Russia	Venezuela
Estonia	Latvia	Saudi Arabia	Vietnam
Fiji	Liberia	Serbia	Yemen
Finland	Lithuania	Singapore	Zimbabwe
France	Luxembourg	Slovakia	