This is an Accepted Manuscript of a book chapter published by Routledge in Econometric Modelling and Forecasting of Tourism Demand on 27 October 2022, available online: http://www.routledge.com/9781003269366.

Chapter 3

THE AUTOREGRESSIVE DISTRIBUTED LAG MODEL

ANYU LIU: 0000-0003-1459-0163

XINYANG LIU: 0000-0001-6689-798X

Abstract

This chapter introduces the autoregressive distributed lag (ADL) model. The ADL model is incorporated with the general-to-specific (GETS) approach for model specification, estimation, and selection to capture the relationship between tourism demand and its economic determinants. The model is applied to analyse and forecast inbound tourism demand for Thailand from selected source markets. Following a discussion of the forecasting performance of the ADL model, future research directions are proposed.

3.1 Introduction

Tourism demand forecasting has been a popular research topic for approximately five decades. The rapid development of tourism forecasting research has motivated recent review articles, such as that of Song, Qiu, and Park (2019), which uses 211 studies published between 1968 and 2018 to comprehensively summarise the new trends and developments in this area. Tourism demand forecasting models are generally divided into three types: time series, econometrics, and artificial intelligence.

In this chapter, we focus on econometric model. The greatest advantage in using econometric model among these three categories is the ability to include information about explanatory variables in forecasting tourism demand. With a strong foundation in economics, econometrics can offer clear-cut practical implications (Song & Li, 2008). Econometric models are generally classified into single-equation and system demand models. Traditionally, single-equation models include the static regression model, the ADL model (Song, Witt, & Li, 2003; Song & Lin, 2010; Huang, Zhang, & Ding, 2017), the error correction (EC) model (Lee, 2011; Goh, 2012; Vanegas, 2013), and the time-varying parameter (TVP) model (Song & Wong, 2003; Song, Li, Witt, & Athanasopoulos, 2011; Page, Song, & Wu, 2012). Widely used system models include the vector auto-regressive (VAR) model (Wong, Song, & Chon, 2006; Song & Witt, 2006) and the almost ideal demand system (AIDS) model (Li, Song, & Witt, 2004).

Tourism demand is normally regarded as a dynamic process in which tourists' decisions about destinations are affected by time. There are many reasons to capture the time-lag phenomenon. Tourists tend to revisit destinations at which they have had a pleasant experience because there is less uncertainty associated with returning to a familiar destination than with travelling to an unfamiliar one. Word of mouth is also an important factor. People often share their travel experiences with their friends and relatives after a holiday, and word-of-mouth recommendations can significantly influence a potential visitor's selection of destinations for future holidays. In addition, because people are generally risk-averse and prefer outcomes with low uncertainty to those with high uncertainty, tourists often consider a destination's popularity when making their choice. Therefore, sought-after destinations are likely to continue to receive large numbers of tourists in the future.

The ADL model is the most widely used econometric method, and early studies such as those of Hendry (1995) and Pesaran and Shin (1995) have applied it to capture the dynamic pattern in economic variables (Liu, Lin, Li, & Song, 2022). Since being introduced into tourism demand forecasting research by Song and Witt (2003), the ADL model has been shown to have powerful analytical and predictive capabilities.

Including the current value of independent variables and the lagged terms of dependent and independent variables in model specifications, the ADL model is more accurate to capture the relationship between tourism demand and its determinants compared to static regression models and can help to reveal how certain economic factors affect others.

The ADL model views the time dynamics in demand variables as accounting for the intertemporal relationships between tourism demand and various explanatory variables. However, the introduction of more independent variables leads to a challenge of model specification and selection. Song et al. (2003) introduced the general-to-specific (GETS) modelling specification approach into the tourism literature. It starts from a general ADL model and removes insignificant variables sequentially according to certain criteria such as the Akaike information criterion (AIC), corrected Akaike information criterion (AICc), and Schwarz Bayesian information criterion (BIC). Song et al. (2003) found that the specified ADL model selected by GETS performed well according to both economic and statistical criteria. The ADL model incorporated with GETS has also been applied in studies of tourism in various destinations, such as Thailand (Song et al., 2003), Fiji (Narayan, 2004), and mainland China (Song & Fei, 2007). In addition, numerous studies have combined the ADL and EC models. Song and Lin (2010) and Lin, Liu, and Song (2015) further demonstrated that the ADL-EC model can consider not only the long-term relationship between tourism demand and its determinants but also the short-term error correction mechanism in modelled estimates.

The ADL model has been further developed for incorporation into other methods of tourism demand forecasting. Athanasopoulos, Song, and Sun (2018) incorporated the bootstrap aggregation method into the ADL model to forecast tourism demand in six source markets for Australia, and their results showed the superior forecasting performance of bootstrap aggregation in improving the robustness of the ADL model. Song, Liu, Li, and Liu (2021) confirmed the forecasting performance of the ADL model with Bayesian bootstrap aggregation, which showed lower variance in forecasting results compared to its ordinary bootstrap aggregation counterpart. In addition, incorporating spatial dependence and spatial heterogeneity is an effective way to improve the ADL model, as suggested by Jiao, Li, and Chen (2021). By fully reflecting the spatial heterogeneity of European tourism demand forecasting models, the proposed general nesting spatiotemporal model outperformed the benchmark models. An alternative approach to improve the forecasting performance of the ADL model is to use judgemental adjustment. Song, Gao, and Lin (2013) utilised expert adjustments as inputs to combine statistical results with reliable consensus, and they demonstrated improved forecasting results for Hong Kong tourism.

This chapter introduces the ADL model with the incorporation of GETS procedure and its application in forecasting tourism demand for Thailand from four source markets, with R codes included. To showcase how the model can be applied, the key procedures and forecasting practices are also provided.

3.2 Methods

3.2.1 The ADL model specification

From the perspective of neoclassical economic theory, tourism demand is usually related to potential consumers' income, the price of visiting the destination, and the comparable price for competing destinations. Following Chapter 2, a tourism demand function of a specific destination can be written as

$$\ln y_{i,t} = \beta_{0,i} + \beta_{2,i} \ln income_{i,t} + \beta_{3,i} \ln price_{i,t} + \varepsilon_{i,t}, \tag{3.1}$$

where $\beta_{0,i}$ and $\varepsilon_{i,t}$ are the constant and disturbance terms, respectively. $\beta_{2,i}$ and $\beta_{3,i}$ represent the income and the destination's own price elasticities, respectively. To measure the dynamic features of tourism demand, the static model in Equation (3.1) can be written in an ADL model as

$$\ln y_{i,t} = \beta_0 + \sum_{j=1}^{J} \beta_{1,i,j} \ln y_{i,t-j} + \sum_{k=0}^{K} \beta_{2,i,k} \ln income_{i,t-k} + \sum_{n=0}^{N} \beta_{3,i,n} \ln price_{i,t-n} + Dummies + \varepsilon_{i,t} \quad (3.2)$$

Considering the time lag of tourists' decision-making progress, the ADL model offers more explanatory power, as tourism demand is affected by both the current value of its determinants and the lagged terms of itself and the determinants. The lag length of the time series may vary depending on data frequency. In common practice, one lag is frequently used for annual data, four lags for quarterly data, and twelve lags for monthly data (Song & Witt, 2003). The dummy variables in Equation (3.2) are used to capture the seasonal effect and offset the impact of one-off events such as the severe acute respiratory syndrome (SARS) outbreak in 2003.

3.2.2 Stationarity and cointegration tests

The concept of cointegration is used to test the existence of a long-run equilibrium between a pair of non-stationary variables in the same economic system, such as tourism demand and its determinants (Engle & Granger, 1987).

To prepare for a cointegration test, we begin with unit root tests for all of input variables to identify their stationarity. A stationary time series is a series that has a constant mean, variance, and covariance over time, and is denoted by $I(0)$. A non-stationary time series has unit roots and is usually called an integrated process. The number of unit roots contained in the series equals the times that the series must be differenced before a stationary process is reached. The simplest autoregressive model (AR(1)) is

$$y_t = \lambda_0 + \lambda_1 y_{t-1} + e_t, \tag{3.3}$$

where λ_0 is the intercept, e_t is the white noise, and λ_1 is the parameters of the model. If y_t has one unit root, denoted by $I(1)$, then $\lambda_1 = 1$. Moreover, when the constant intercept $\lambda_0 = 0$, the process is termed a random walk; when the constant intercept $\lambda_0 \neq 0$, the process is termed a random walk with drift.

To determine objectively whether a series contains a unit root process, in most unit root tests, the null hypothesis $H_0: \lambda_1 = 1$ is tested against the alternative hypothesis, $H_1: \lambda_1 < 1$, based on Equation (3.3). In empirical work, the Dickey–Fuller (DF) test, the augmented Dickey–Fuller (ADF) test, and the Phillips–Perron (PP) test are frequently applied.

The DF test assumes that the time series can be modelled by an AR(1) process. Instead of testing $\lambda_1 = 1$ directly, Dickey and Fuller (1979) transformed Equation (3.3) by subtracting y_{t-1} from both sides:

$$\Delta y_t = \psi y_{t-1} + e_t \tag{3.4}$$

$$\Delta y_t = \pi_0 + \pi_1 t + \psi y_{t-1} + e_t \tag{3.5}$$

where Δ is the first difference operator and t represents the time trend variable. Equation (3.4) is used to test a random walk process and Equation (3.5) for a random walk with drift process. Note that the null hypothesis changes to $H_0: \psi = 0$, against the alternative hypothesis $H_1: \psi < 0$.

$$t = \frac{\hat{\psi}}{SE(\hat{\psi})} \tag{3.6}$$

The t value, which is computed by Equation (3.6), can be compared to the critical value for the one-tailed test because $H_1: \psi < 0$ indicates that the rejection region is on the left.

However, an AR(1) process may not model every time series; in these cases, the error will be serially correlated. To avoid the problem of autocorrelation in the residuals, the ADF test includes lagged dependent variables in the model specifications. Two variants of the ADF models are based on the following equations:

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{n} \omega_i \Delta y_{t-i} + e_t, \tag{3.7}$$

$$\Delta y_t = \pi_0 + \pi_1 t + \psi y_{t-1} + \sum_{i=1}^{n} \omega_i \Delta y_{t-i} + e_t, \tag{3.8}$$

where the *n* lagged first differences approximate the autoregressive moving average dynamics of the time series. Equation (3.7) is used to test a random walk process and Equation (3.8) for a random walk with drift process. The unit root test is then carried out under the null hypothesis $H_0: \psi = 0$, against the alternative hypothesis $H_1: \psi < 0$. Therefore, a one-tailed *t*-test is conducted to examine whether we need to reject the null hypothesis when the t value is less than the critical value.

Both the DF and ADF tests assume that the residual in the regression is identically and independently distributed, and they thus are fairly restrictive. Phillips and Perron (1988) generalised the DF test and developed the PP test. However, the model specification of the PP test is rather complex and outside of the scope of this book. Many econometric software suites provide functions to conduct a PP test.

For the cointegration test, a conditional ADL-EC model is used to test the existence of long-term relationships between tourism demand and the explanatory variables:

$$\Delta \ln y_{i,t} = \alpha_0 + \sum_{j=1}^{J} \varphi_{1,i,j} \Delta \ln y_{i,t-j} + \sum_{k=0}^{K} \varphi_{2,i,k} \Delta \ln income_{i,t-k} + \sum_{l=0}^{L} \varphi_{3,i,l} \Delta \ln price_{i,t-l} + \theta_{1,i} \ln y_{i,t-1} + \theta_{2,i} \ln income_{i,t-1} + \theta_{3,i} \ln price_{i,t-1} + Dummies + e_{i,t}, \tag{3.9}$$

where φ is the short-run impact parameter that captures short-run deviations from the equilibrium, θ is the long-run impact parameter that gives the long-run equilibrium and the cointegrating relationships, and α_0 is the constant intercept. To eliminate insignificant terms and reach the optimal lag structure, the lag orders *J*, *K*, and *L* are determined by a number of criteria, such as AIC, AICc, and BIC. The mathematical details of the ADL-EC model can be found in Song and Turner (2006).

The bounds test is an advanced method proposed by Pesaran, Shin, and Smith (2001) that has been widely used in tourism demand forecasting to examine the long-term relationships between tourism demand and explanatory variables (Song & Lin, 2010). Compared to the traditional cointegration test, the bounds test provides two flexibilities: 1) it is applicable irrespective of the integration order, which means that the variables can be a mixture of $I(0)$ and $I(1)$, and 2) variables considered in the ADL model can have different lag terms.

The bounds test conducts an *F*-test with the null hypothesis that no cointegration relationships exist between the variables (H_0: $\theta_{1,i} = \theta_{2,i} = \delta\theta_{3,i} = 0$ for Equation (3.9)) against the alternative hypothesis that cointegration exists. Based on the situation in which the model variables are an arbitrary mix of $I(0)$ and $I(1)$, the lower bound assumes that all of the variables are $I(0)$, and the assumption of the upper bound is that all of the variables are $I(1)$. By definition, if the computed *F*-statistic is less than the critical value of the lower bound, we cannot reject the null hypothesis. If the *F*-statistic exceeds the upper bound, cointegration may exist. If the *F*-statistic falls between the two boundaries, the test result is inconclusive. The critical values of the lower and upper bounds are provided by Pesaran et al. (2001) in Tables CI (pp. 300-301) and CII (pp. 303-304) in their study.

For further checking, a *t*-test should also be conducted to identify the existence of cointegration (Pesaran et al., 2001). The *t*-test considers the null hypothesis $H_0: \theta_{1,i} = 0$, which means that there is no cointegration relationship in the lag terms of tourism demand. Therefore, the *t*-test would confirm the existence of cointegration by rejecting the null hypothesis if the computed *t*-statistic is greater than the upper critical value, and it would indicate that all of the variables are stationary if the computed *t*-statistic is less than the lower critical value.

3.2.3 Diagnostic tests

Before we generate a forecast, the estimated model must be checked with numerous statistical tests to ensure that it is properly specified. For the final model, the regression residual is expected to be normality and not to contain autocorrelation or heteroscedasticity. In addition, the model should choose the function form correctly and avoid omitting important explanatory variables. Here we list the most used diagnostic tests.

Autocorrelation test

The Durbin–Watson statistic is a standard test to determine the existence of autocorrelation (Durbin & Watson, 1950). The statistic is

$$d = \frac{\sum_{i=2}^{n}(\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=1}^{n}\varepsilon_i^2}, \tag{3.10}$$

where ε_i represents the residuals and n is the sample size.

The statistic d ranges from 0 to 4, where a value of 0 means that the data are perfectly positively autocorrelated and a value of 4 means that the data are perfectly negatively autocorrelated. If the value is around 2, there is no autocorrelation in the regression residuals. However, the DW statistic, constructed by Equation (3.10), is only able to detect first-order autocorrelation. The advanced method known as the Breusch–Godfrey (BG) test is a more general test, as it does not allow broader autocorrelation order, and it includes the lagged terms of dependent variables.

Heteroscedasticity test

Heteroscedasticity, which occurs when the modelling errors do not have the same variance, is a major concern in regression analysis, as it can result in biased standard errors. To make it simple but without the loss of universality, a White test (White, 1980) posits two explanatory variables consisting of a multiple regression model as follows:

$$y_t = \rho_1 + \rho_2 x_{1t} + \rho_3 x_{2t} + \tau_t \tag{3.11}$$

Then, the following auxiliary equation is used to test whether the error term τ_t is homoscedastic or heteroscedastic:

$$\hat{\tau}_t^2 = \gamma_1 + \gamma_2 x_{1t} + \gamma_3 x_{2t} + \gamma_4 x_{1t}^2 + \gamma_5 x_{2t}^2 + \gamma_6 x_{1t} x_{2t} + \epsilon_t \tag{3.12}$$

where $\hat{\tau}_t^2$ is the estimated residual from Equation (3.11). The equations can also be expanded to test heteroscedasticity when the regression model has more than two explanatory variables. The null hypothesis of the White test is that the variances of the modelling error terms are equal, $H_0 = \sigma_t^2 = \sigma^2$, and the alternative hypothesis is that the variances are not equal. The test statistic nR^2, where R^2 is from the auxiliary regression in Equation (3.12), has a χ^2 distribution with degrees of freedom equal to the number of regressors. If the calculated statistic is greater than the critical χ^2

value at the specific level of significance, the null hypothesis is rejected. In econometrics modelling practice, the Breusch–Pagan (BP) test is an effective method for testing heteroscedasticity. Derived from the Lagrange multiplier test, the BP test was developed by Breusch and Pagan in 1979, and subsequent studies have developed theoretical extensions of this test.

Testing for normality

The normality test is used to determine how well a data set is modelled by a normal distribution and how likely a random variable is to be normally distributed. In terms of model selection, the normality test is useful in measuring the goodness of fit of a model to the data. Developed by Shapiro and Wilk (1965), the Shapiro–Wilk normality test is

$$W = \frac{(\sum_{i=1}^{n} a_i r_{(i)})^2}{\sum_{i=1}^{n}(r_i - \bar{r})^2} \tag{3.13}$$

where r_i is the ordered random samples or the model's residual. For $r_{(i)}$, the subscript indices that are enclosed within parentheses represent the *i*th smallest number in the sample. a_i is the constants computed from the variance, covariance, and means of the sample from a normal distribution. The null hypothesis of the Shapiro–Wilk normality test is that the tested sample is normally distributed. If the small value of *W* rejects the null hypothesis, the sample is not normally distributed.

Test for misspecification

The Ramsey RESET test is used to examine whether the omission of important explanatory variables or the non-linear format of an equation causes model misspecification (Ramsey, 1969). Taking Equation (3.11) as the example again, the fitted values of y_t is:

$$\hat{y}_t = \hat{\rho}_1 + \hat{\rho}_2 x_{1t} + \hat{\rho}_3 x_{2t} \tag{3.14}$$

The second step is to test whether higher-powered dependent variables have explanatory power for y_t:

$$y_t = \mu_1 + \mu_2 X_{1t} + \mu_3 X_{2t} + \vartheta_1 \hat{Y}_t^2 + \vartheta_2 \hat{Y}_t^3 + \vartheta_3 \hat{Y}_t^4 + \xi_t \tag{3.15}$$

Finally, the Wald test is used to identify the significance of ϑ_1 through ϑ_3. The null hypothesis is that the model is correctly specified and that all ϑ coefficients are zero. If the null hypothesis is rejected because at least one powered fitted value can further explain the dependent variable, then the model suffers from misspecification. However, the RESET test is a general misspecification test, which means that rejection of the null hypothesis only identifies the misspecified condition of the model, not how the model is misspecified.

Once the model passed the cointegration and most diagnostic tests, the income and own price elasticities in Equation (3.2) can be derived as $\frac{\sum_{k=0}^{K} \beta_{2,i,k}}{1-\sum_{j=1}^{J} \beta_{1,i,j}}$ and $\frac{\sum_{n=0}^{N} \beta_{3,i,n}}{1-\sum_{j=1}^{J} \beta_{1,i,j}}$, respectively. Then the model can be used to predict the tourism demand and the forecasting results should be evaluated by forecasting accuracy measured such as MAPE, RMSE and MASE as introduced in Chapter 2.

3.3 Application

In this chapter, the ADL model with the incorporation of GETS procedure is used to generate *ex ante* forecasts of visitor arrivals from mainland China, Malaysia, Russia and the USA to Thailand.

3.3.1 Unit root test results

Before modelling and forecasting, unit root tests are conducted for all of the log-transformed variables to avoid spurious regressions. Table 3.1 demonstrates the stationarity status of all of the variables based on the results of the ADF, PP, and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests. The null hypotheses of ADF and PP tests are that there is a unit root present in the tested series, and the null hypothesis of the KPSS test is that no such unit root is present.

An identified integration order is set depending on the majority result of the three unit root tests. In the ADF and PP tests, in the Malaysian model, the null hypothesis of a unit root is rejected for the log-transformed visitor arrivals variable. Taking the first differences, all of the variables are stationary series in the three unit root tests except for the income variable in the Chinese mainland model. The results of the ADF and KPSS tests suggest that a higher order of integration exists. The above results show that most of the variables in the models are either $I(0)$ or $I(1)$, which justifies the use of the following bounds test procedure.

[Insert Table 3.1 about here]

3.3.2 Bounds test results

After the model specification and selection by GETS procedure, we conduct the bounds test to check whether there is cointegration between the remaining independent variables and visitor arrivals. With the null hypothesis that there is no cointegration between the variables, the bounds test outputs the F-statistic and t-statistic together. These are then compared with the threshold interval at the 10% and 1% significance levels (see Table 3.2). Because the final models for each source country have different numbers of independent variables, the corresponding threshold intervals are also provided for reference.

The results in Table 3.2 show the F-statistics all lie on the right side of the upper bound at the 1% significance level. The computed t-statistic supports the existence of a long-run relationship between tourism demand and its determinants in Malaysia and the USA but fails to reject the null hypothesis with the bounds test for mainland China and Russia, which may relate to the fact that those two markets are largely influenced by domestic policy and international politics. The fluctuating outbound tourism series and the limited sample size also provide relatively weak confirmation of cointegration. Therefore, we need to interpret the modelling and forecasting results with caution of the two destinations. As the results show that the variables have long-term relationships across these tourism demand models, the demand elasticities can be reasonably obtained.

[Insert Table 3.2 about here]

3.3.3 Diagnostic test results

The estimation results and diagnostic tests are included in Table 3.3 and 3.4, respectively. All four models have high goodness of fit, as suggested by the high values of the adjusted R^2, which exceed 0.95 for mainland China, Russia, and the USA. These results show that most of the variations in visitor arrivals from the relevant markets over the 2000Q1-2016Q4 period can be explained by the estimated models. In addition, all of the F-statistics are significant at the 1% significance level. The diagnostic statistics in Table 3.4 show that all of the models are well constructed except for the RESET tests. All of the models reject the null hypothesis of the RESET test that the correct specification is linear, probably because income and own prices are limited in their ability identify tourism demand. Other explanatory variables such as alternative price could be included with caution. The Chinese mainland model fails the BG test because of the substantial volatility in the Chinese outbound market over the sample period. The USA model fails the Shapiro–Wilk test but passes the others.

Overall, the diagnostic testing results indicate that these four models are reasonably valid and reliable and can undergo further analysis.

[Insert Table 3.3 and 3.4 here]

3.3.4 Demand elasticities

The signs for the income variables are positive for all four markets, which is consistent with theory (Table 3.4). The income elasticities for all of the models are greater than one, which suggests that demand for tourism in Thailand from these source markets is income elastic. These results imply that these visitors are sensitive to changes in income and that travel to Thailand has the attributes of a luxury product, probably because leisure travel makes up a large proportion of Thailand's tourism market. In addition, the magnitudes of the estimated elasticities vary across markets. It is notable that the income elasticities of mainland China and Malaysia are smaller than those of Russia and the USA, perhaps due to the relatively short distances between Thailand and these short-haul markets. Income is a less significant influencing factor for mainland China and Malaysia.

According to the law of demand, all price elasticities are expected to be negative. The computed price elasticity for the USA is less than one, revealing that American visitors are relatively less sensitive to changes in the prices of tourism products and services in Thailand. Visitors from mainland China and Russia are more sensitive to price elasticity, which suggests that price campaigns would be an effective way to attract visitors from these two source markets. The sign of the price elasticity of tourism demand from Malaysia is positive but not significant.

3.3.5 Tourism demand forecasts

The quarterly forecasts of visitor arrivals in Thailand from the selected short- and long-haul markets are generated based on the above estimated equations. One- to 12-

step-ahead forecasts are generated in a rolling window from 2017Q1 to 2019Q4. To eliminate potential forecasting outliers, we only consider the forecasting performances of one to eight steps ahead because they contain adequate forecast points. For example, the forecasting period from 2017Q1 to 2019Q4 allows five eight-step-ahead forecasts using the rolling window. The forecasting performance is compared with the three most-used time series models: the seasonal autoregressive integrated moving average (ARIMA), exponential smoothing (ETS), and seasonal naïve (SNAIVE) models.

In general, the ADL model performs well across most forecasting horizons in terms of the rankings of MAPE, RMSE, and MASE (details of the forecasting performance can be found in Tables 3.5-3.7). When the forecasting horizon is extended, the forecast accuracy of the benchmark models deteriorates, but the ADL model outperforms the benchmark models. The empirical results suggest that the ADL model offers a highly consistent, superior forecasting performance.

[Insert Table 3.5-3.7 here]

3.4 Conclusion and future directions

This chapter introduces the ADL model with the incorporation of GETS procedure and its application to forecasting demand for tourism in Thailand. The advantage of the ADL model lies in its ability to introduce into a demand model the economic determinants that are revealed as demand elasticities in a log-transformed demand function. Capturing the dynamics of dependent and independent variables contributes to the accurate prediction of future demand, particularly over the longer term. The empirical application in this chapter illustrates the model's superior forecasting performance in most cases and shows that its performance remains highly stable over different forecasting horizons and with different forecast error measures.

Based on the general form of the ADL model, this method can be further improved for incorporation into other techniques for forecasting tourism demand. The ADL model has shown significant improvements in forecasting tourism demand, as studies have expanded it with the EC model and time series bootstrap aggregation methods; however, developments remain limited (Song & Lin, 2010; Athanasopoulos et al., 2018; Song et al., 2021). Integrations of the ADL model with other advanced forecasting approaches such as boosting, staking, and mixed-data sampling methods has yet to applied in the tourism demand forecasting context. In addition, as a powerful tool in identifying causal relationships, the ADL model can be applied to investigate new aspects of the tourist industry and verify their general applicability for forecasting tourism demand.

Self-study questions

(1) What are the advantages of the ADL model?

(2) What are the limitations of the ADL model?

(3) What are the practical implications for income/price elasticities larger than one (or less than one)?

References

Athanasopoulos, G., Song, H., & Sun, J. A. (2018). Bagging in tourism demand modeling and forecasting. *Journal of Travel Research*, *57*(1), 52-68. https://doi.org/10.1177/0047287516682871

Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression: I. *Biometrika*, *37*(3/4), 409-428. https://doi.org/10.2307/2332391

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 251-276. https://doi.org/1913236

Goh, C. (2012). Exploring impact of climate on tourism demand. *Annals of Tourism Research*, *39*(4), 1859-1883. https://doi.org/10.1016/j.annals.2012.05.027

Hardy, M. A. (1993). Regression with dummy variables. *Sage University Paper Series on Quantitative Applications in the Social Sciences, series No. 07-093*. *Newbury Park, CA: Sage.*

Hendry, D. F. (1995). Dynamic econometrics. *Oxford University Press.*

Huang, X., Zhang, L., & Ding, Y. (2017). The Baidu Index: Uses in predicting tourism flows–A case study of the Forbidden City. *Tourism Management*, *58*, 301-306. https://doi.org/10.1016/j.tourman.2016.03.015

Jiao, X., Chen, J. L., & Li, G. (2021). Forecasting tourism demand: Developing a general nesting spatiotemporal model. *Annals of Tourism Research*, 90, https://doi.org/10.1016/j.annals.2021.103277 .

Lee, K. N. (2011). Forecasting long-haul tourism demand for Hong Kong using error correction models. *Applied Economics*, *43*(5), 527-549. https://doi.org/10.1080/00036840802599743

Li, G., Song, H., & Witt, S. F. (2004). Modeling tourism demand: A dynamic linear AIDS approach. *Journal of Travel Research*, *43*(2), 141-150. https://doi.org/10.1177/0047287504268235

Lin, V. S., Liu, A., & Song, H. (2015). Modeling and forecasting Chinese outbound tourism: An econometric approach. *Journal of Travel and Tourism Marketing*, *32*(1-2), 34-49. https://doi.org/10.1080/10548408.2014.986011

Liu, A., Lin, V. S., Li, G., & Song, H. (2022). Ex ante tourism forecasting assessment. *Journal of Travel Research*, *61*(1), 64-75. https://doi.org/10.1177/0047287520974456

Narayan, P. (2004). Reformulating critical values for the bounds F-statistics approach to cointegration: An application to the tourism demand model for Fiji. (Vol. 2, No. 04). Australia Monash University.

Page, S., Song, H., & Wu, D. C. (2012). Assessing the impacts of the global economic crisis and swine flu on inbound tourism demand in the United Kingdom. *Journal of Travel Research*, *51*(2), 142-153. https://doi.org/10.1177/0047287511400754

Peng, B., Song, H., & Crouch, G. I. (2014). A meta-analysis of international tourism demand forecasting and implications for practice. *Tourism Management*, *45*, 181-193. https://doi.org/10.1016/j.tourman.2014.04.005

Pesaran, M. H., & Shin, Y. (1995). An autoregressive distributed lag modelling approach to cointegration analysis. *University of Cambridge, working paper, 9514 (9514)*. https://doi.org/10.1017/CCOL0521633230.011

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, *16*(3), 289-326. https://doi.org/10.1002/jae.616

Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335-346. https://doi.org/10.1093/biomet/75.2.335

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society: Series B (Methodological)*, *31*(2), 350-371. https://doi.org/10.1111/j.2517-6161.1969.tb00796.x

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3/4), 591-611. https://doi.org/10.2307/2333709

Song, H., & Fei, B. (2007). Modelling and forecasting international tourist arrivals to mainland China, *China Tourism Research*, 3(1),3-40.

Song, H., Gao, B. Z., & Lin, V. S. (2013). Combining statistical and judgmental forecasts via a web-based tourism demand forecasting system. *International Journal of Forecasting*, *29*(2), 295-310. https://doi.org/10.1016/j.ijforecast.2011.12.003

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. *Tourism Management*, *29*(2), 203-220. https://doi.org/10.1016/j.tourman.2007.07.016

Song, H., Li, G., Witt, S. F., & Athanasopoulos, G. (2011). Forecasting tourist arrivals using time-varying parameter structural time series models. *International Journal of Forecasting*, *27*(3), 855-869. https://doi.org/10.1016/j.ijforecast.2010.06.001

Song, H., & Lin, S. (2010). Impacts of the financial and economic crisis on tourism in Asia. *Journal of Travel Research*, *49*(1), 16-30. https://doi.org/10.1177%2F0047287509353190

Song, H., Liu, A., Li, G., & Liu, X. (2021). Bayesian bootstrap aggregation for tourism demand forecasting. *International Journal of Tourism Research*, *23*(5), 914-927. https://doi.org/10.1002/jtr.2453

Song, H., Qiu, R. T., & Park, J. (2019). A review of research on tourism demand forecasting: Launching the Annals of Tourism Research Curated Collection on tourism demand forecasting. *Annals of Tourism Research*, *75*, 338-362. https://doi.org/10.1016/j.annals.2018.12.001

Song, H., & Turner, L. (2006). Tourism demand forecasting. *International Handbook on the Economics of Tourism*, 89-114.

Song, H., & Witt, S. F. (2003). Tourism forecasting: The general-to-specific approach. *Journal of Travel Research*, *42*(1), 65-74. https://doi.org/10.1177%2F0047287503253939

Song, H., & Witt, S. F. (2006). Forecasting international tourist flows to Macau. *Tourism Management*, *27*(2), 214-224. https://doi.org/10.1016/j.tourman.2004.09.004

Song, H., Witt, S. F., & Li, G. (2003). Modelling and forecasting the demand for Thai tourism. *Tourism Economics*, *9*(4), 363-387. https://doi.org/10.5367%2F000000003322663186

Song, H., & Wong, K. K. (2003). Tourism demand modeling: A time-varying parameter approach. *Journal of Travel Research*, *42*(1), 57-64. https://doi.org/10.1177%2F0047287503253908

UNWTO (2019). UNWTO Tourism Highlights 2016 Edition. Madrid: UNWTO.

Vanegas Sr, M. (2013). Co-integration and error correction estimation to forecast tourism. *Journal of Travel & Tourism Marketing*, *30*(6), 523-537. https://doi.org/10.1080/10548408.2013.810992

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 817-838. https://doi.org/10.2307/1912934

Witt, S. F., & Witt, C. A. (1995). Forecasting tourism demand: A review of empirical research. *International Journal of Forecasting*, *11*(3), 447-475. https://doi.org/10.1016/0169-2070(95)00591-7

Wong, K. K., Song, H., & Chon, K. S. (2006). Bayesian models for tourism demand forecasting. *Tourism Management*, *27*(5), 773-780. https://doi.org/10.1016/j.tourman.2005.05.017

Author bios

Anyu Liu, PhD, is an Assistant Professor in the School of Hotel and Tourism Management at The Hong Kong Polytechnic University. His research interests are in tourism demand modelling and forecasting, tourism's economic impact and big data analytics. Email: an-yu.liu@polyu.edu.hk.

Xinyang Liu is a Ph.D. research student at the University of Surrey, UK. His research interests include tourism demand modelling and forecasting with a focus on the application and development of time series forecasting techniques. Email: xinyang.liu@surrey.ac.uk

Appendix: R code

```
### The following codes shows how to estimate an ARDL model by GETS and
generate the tourism demand forecast
### with the package ARDL in R.
### Author: Dr. Anyu Liu and Mr. Xinyang Liu
### Last updated: 12-04-2022
### R version: R 4.0.5

### Clear memory
rm(list=ls())
### install and load ARDL package
#install.packages("ARDL")
#install.packages("dynlm")
#install.packages("aod")
#install.packages("tseries")
#install.packages("lmtest")
#install.packages("forecast")
library(ARDL)
library(dynlm)
library(aod)
library(tseries)
library(lmtest)
library(forecast)
library(Metrics)

###Set the path that the data and function file is stored
setwd("~")
source('ardl_forecast_function.R')

###Input the raw data
data<-read.csv("THA_MAL.csv")

###Generate the relative price and take natural log to non-dummy variables
#data$RP=(data$CPI_D/data$EX_D)/(data$CPI_O/data$EX_O)
data$RP=(data$CPI_D/data$EX_D)/(data$CPI_O/data$EX_O)
data$lnarr=log(data$Arr)
data$lngdp=log(data$GDP)
data$lnrp=log(data$RP)
data1=data[c('lnarr','lngdp','lnrp','Q1','Q2','Q3')]###Modify and specify the dummy
variable names based on the data
data1<-ts(data1,start=c(2000,1),end=c(2019,4),frequency=4)
###Unit root tests
###Level
adf.test(data$lnarr)
adf.test(data$lngdp)
adf.test(data$lnrp)
pp.test(data$lnarr)
```

```
pp.test(data$lngdp)
pp.test(data$lnrp)
kpss.test(data$lnarr)
kpss.test(data$lngdp)
kpss.test(data$lnrp)
###1st order difference
adf.test(diff(ts(data$lnarr),differences=1))
adf.test(diff(ts(data$lngdp),differences=1))
adf.test(diff(ts(data$lnrp),differences=1))
pp.test(diff(ts(data$lnarr),differences=1))
pp.test(diff(ts(data$lngdp),differences=1))
pp.test(diff(ts(data$lnrp),differences=1))
kpss.test(diff(ts(data$lnarr),differences=1))
kpss.test(diff(ts(data$lngdp),differences=1))
kpss.test(diff(ts(data$lnrp),differences=1))

###Generate the training data set
datam<-ts(data1[1:68,],start=c(2000,1),end=c(2016,4),frequency=4)
###Find the best ADL specification based on AIC by GETS
models<-auto_ardl(lnarr ~ lngdp + lnrp + Q1 + Q2 + Q3, data = datam,
                        max_order = 4, fixed_order = c(-1,-1,-1,0,0,0)) ###Modify the
formula based on the inclusion of dummy variables, the number of "0" in brakets
equals to 3+ number of dummies
models$top_orders
ardl_fitted=models$best_model
summary(models$best_model)
###Estimate the ADL-EC model with the best model specification
model_l<-models$best_model
model_ecm<-uecm(model_l)
summary(model_ecm)

###Bounds test
fbd<-bounds_f_test(model_l, case = 2)
tbd<- bounds_t_test(model_l, case = 3, alpha = 0.01)

###Diagnostic tests
res<-model_l$residuals #save the residual
bgtest(model_l)#performs the Breusch-Godfrey test for higher-order serial correlation
bptest(model_l)#Performs the Breusch-Pagan test against heteroskedasticity
resettest(model_l)#Ramsey's RESET test for functional form
shapiro.test(res)#Shaprio test for normality
###Calculate Elasticities
EL<-multipliers(model_l)

###Forecast Generation
###Generate the testing data set
# datat <- ts(data1[69:80,],start=c(2017,1),end=c(2019,4),frequency = 4)
N = nrow(data)
forecasts.matrix = matrix(NA, 12, 8)
LnY0.matrix = matrix(NA, 12, 8)
```

```
colnames(forecasts.matrix) = paste0("h=", 1:8)
rownames(forecasts.matrix) = paste0(rep(2017:2019, each = 4), paste0("Q", 1:4))

# ADL model
# Rolling window: sampling period ending from N-12 to N-1
for (i in 12:1) {
   endi = N - i
   mat = tail(head(data1, endi),68)
   # Fitting ADL model
   fit.adl <- auto_ardl(lnarr ~ lngdp + lnrp + Q1 + Q2 + Q3, data = mat,
                                        max_order=4, fixed_order=c(-1,-1,-
1,0,0,0))$best_model
   # Forecasting h=i
   fc.adl <- ardl_forecast(fit.adl, data1, i)
   # Arranging forecasts and original values into 2 matrices with the same format
   for (j in 1:8) {
      if ((12 - i + j) > 12)
         break
      forecasts.matrix[(12 - i + j), j] = fc.adl[j]
      LnY0.matrix[(12 - i + j), j] = data1[(endi + j), 1]
   }
}

# Computing residuals matrix in original scale
residuals.matrix = exp(forecasts.matrix) - exp(LnY0.matrix)

# Computing MAPE
MAPE = colMeans(abs(residuals.matrix) * 100 / exp(LnY0.matrix), na.rm = TRUE)

# Computing RMSE
RMSE = sqrt(colMeans((residuals.matrix) ^ 2, na.rm = TRUE))

# Computing MASE
MASE.denominator = mean(abs(diff(exp(data1[,1]), 4)))
MASE = colMeans(abs(residuals.matrix) / MASE.denominator , na.rm = TRUE)

#ardl forecasts function
ardl_forecast = function(model, data_full, horizon) {
   data_train <- model$data
   nrow <- dim(data_full)[1]
   ncol <- dim(data_full)[2]
   outro <- data_full[(nrow - horizon + 1): nrow, 2: ncol]
   #construct forecast matrix
   fore_array <- rep(0, horizon)
   for (i in 1: horizon) {
      if (i == 1) {
         updated_matrix <- rbind(data_train, c(fore_array[1:i], matrix(outro,
horizon,)[1:i,]))
      } else {
         updated_matrix <- rbind(data_train, cbind(fore_array[1:i], outro[1:i,]))
```

```
        }
        #with the help of ARDL fitting function, a complete forecast (both current value and lag value are included) matrix is construct
        data_structure <- ardl(as.formula(paste(model$parsed_formula$y_part$var,
paste(model$parsed_formula$x_part$var, collapse = "+"), sep = "~")),
                                    data = updated_matrix, order = model$order)$model

        nrow_structure <- dim(data_structure)[1]
        ncol_structure <- dim(data_structure)[2]
        outro_input <- data_structure[nrow_structure, 2:ncol_structure]
        coef <- model$coefficients[2: ncol_structure]
        coef[is.na(coef)] <- 0

        fore <- as.matrix(outro_input) %*% coef + model$coefficients[1]
        fore_array[i] <- fore
    }
    return(fore_array)
}
```

Table 3.1 Unit root test results using the ADF, PP, and KPSS tests.

		Level						First Difference						Integration Order
		ADF		PP		KPSS		ADF		PP		KPSS		
Mainland China	$\ln y$	-2.218		-19.773	*	1.860	***	-5.817	***	-97.661	***	0.141		$I(1)$
	$\ln income$	0.465		1.711		2.087	***	-2.673		-61.249	***	1.074	***	higher order
	$\ln price$	-2.364		-8.554		0.403	*	-4.298	***	-62.419	***	0.095		$I(1)$
Malaysia	$\ln y$	-4.366	***	-53.668	***	2.062	***	-6.585	***	-75.967	***	0.052		$I(0)$
	$\ln income$	-2.786		-16.182		2.045	***	-5.489	***	-57.582	***	0.086		$I(1)$
	$\ln price$	-2.612		-21.982	**	1.848	***	-5.830	***	-74.142	***	0.266		$I(1)$
Russia	$\ln y$	-1.331		-45.007	***	1.951	***	-3.388	*	-42.207	***	0.191		$I(1)$
	$\ln income$	-1.822		-18.126	*	1.860	***	-3.462	*	-63.076	***	0.215		$I(1)$
	$\ln price$	-1.825		-7.316		0.503	**	-4.580	***	-93.88	***	0.641	**	$I(1)$
US	$\ln y$	-2.233		-49.092	***	1.817	***	-5.267	***	-49.299	***	0.212		$I(1)$
	$\ln income$	-1.561		-3.393		1.991	***	-3.058		-62.368	***	0.206		$I(1)$
	$\ln price$	-1.579		-7.271		1.478	***	-4.606	***	-49.616	***	0.127		$I(1)$

Note: *, **, and *** denote a rejection of the null hypothesis at the 0.1, 0.05, and 0.01 significance levels, respectively.

Table 3.2 Bounds test results

	k	*F*-statistic		*t*-statistic	
Mainland China	9	13.121 ***		-3.871	
Malaysia	5	8.753 ***		-3.948 *	
Russia	7	8.073 ***		-0.119	
US	9	99.838 ***		-9.726 ***	
		F-statistic		*t*-statistic	
Bounds test interval		Lower-bound *I*(0)	Upper-bound *I*(1)	Lower-bound *I*(0)	Upper-bound *I*(1)
		10% Significance level		10% Significance level	
k=5		2.26	3.35	-2.57	-3.86
k=7		2.03	3.13	-2.57	-4.23
k=9		1.88	2.99	-2.57	-4.56
		1% Significance level		1% Significance level	
k=5		3.41	4.68	-3.43	-4.79
k=7		2.96	4.26	-3.43	-5.19
k=9		2.65	3.97	-3.42	-5.54

Note: 1. k is the number of variables left in the estimated model; 2. *, **, and *** denote a rejection of the null hypothesis at the 0.1, 0.05, and 0.01 significance levels, respectively.

Table 3.3 ADL estimates in the final state and diagnostics

	Mainland China		Malaysia		Russia		US	
$\ln y(-1)$	0.493	***	0.586	***	0.840	***	0.237	***
$\ln y(-2)$	0.022				-0.478	***		
$\ln y(-3)$	0.161	*			0.629	***		
$\ln income$	0.539	***	0.483	**	1.778	***	1.699	***
$\ln income(-1)$					-1.893	***		
$\ln price$	-0.827		0.007		-1.149	***	-0.199	***
$\ln price(-1)$	0.241				0.885	**		
$\ln price(-2)$	0.911							
$\ln price(-3)$	-1.954	*						
Q1	0.297	***	-0.219	***	0.095		-0.061	**
Q2	0.037		-0.166	***	-0.982	***	-0.267	***
Q3	0.195	**	-0.131	***	-0.899	***	-0.285	***
SARS	-1.496	***			-0.882	***	-0.383	***
PC_05	-0.791	***						
FC_08	-0.758	***			0.010		-0.125	***
PC_10	-0.857	***						
SC_10							-0.174	***
SC_14							-0.154	***
Intercept	1.624	***	3.385	***	1.179		1.525	**

Notes: *, **, and *** denote significance at the 0.1, 0.05, and 0.01 levels, respectively.

Table 3.4 Demand elasticities and diagnostic test results

	Mainland China		Malaysia		Russia		US	
Income elasticity	1.66		1.17		2.93		2.23	
Own price elasticity	-5.02		-		-4.16		-0.26	
R^2	0.963		0.893		0.981		0.960	
Adjusted R^2	0.952		0.882		0.976		0.953	
F-statistic	85.86	***	83.37	***	181	***	135.9	***
Diagnostic test								
BG test	6.881	***	0.0560		2.182		0.087	
BP test	13.617		3.921		8.123		9.8	
RESET test	7.282	***	4.905	**	4.92	***	2.961	*
Shapiro–Wilk test	0.957	**	0.958	**	0.963	*	0.98	

Note: *, **, and *** denote a rejection of the null hypothesis at the 0.1, 0.05, and 0.01 significance levels, respectively.

Table 3.5 Forecasting performance evaluation measured by MAPE (%)

	Forecasting horizons				
	1	2	3	4	8
Mainland China					
SNAIVE	13.194	13.683	15.039	15.606	14.882
ETS	12.201	16.127	15.163	13.312	15.984
SARIMA	14.753	13.945	16.658	17.642	21.033
ADL	7.166	11.400	11.122	10.900	7.383
Malaysia					
SNAIVE	7.859	7.559	7.349	8.114	18.818
ETS	7.973	8.113	8.193	8.448	5.081
SARIMA	8.339	8.032	8.422	7.148	8.964
ADL	8.401	8.743	8.318	7.099	12.550
Russia					
SNAIVE	9.800	8.259	7.474	7.783	13.722
ETS	9.956	12.781	16.648	13.287	26.609
SARIMA	12.221	16.511	15.800	19.900	46.633
ADL	12.307	12.278	12.900	10.960	11.817
US					
SNAIVE	5.703	5.563	5.098	4.839	9.991
ETS	3.956	3.919	4.345	6.008	11.263
SARIMA	4.553	7.014	8.247	9.483	14.186
ADL	5.268	5.935	5.854	6.102	5.379

Table 3.6 Forecasting performance evaluation measured by RMSE

	Forecasting horizons				
	1	2	3	4	8
Mainland China					
SNAIVE	435305	451019	473032	490637	465626
ETS	361836	454310	454610	469960	496274
SARIMA	431710	440885	532880	515603	596629
ADL	239047	346058	342582	332689	233116
Malaysia					
SNAIVE	114911	117014	119366	125817	227894
ETS	86084	96150	108836	92312	64159
SARIMA	98563	100908	116794	107835	131294
ADL	96799	111594	111151	107898	156260
Russia					
SNAIVE	60723	49666	50755	53441	72672
ETS	40284	35219	69387	61388	92175
SARIMA	44255	51003	58842	75123	150857
ADL	78401	53962	76022	55736	65652
US					
SNAIVE	17100	16758	15602	15539	32290
ETS	12734	13027	16326	19637	36930
SARIMA	17164	22540	26103	30561	45171
ADL	16907	18062	18546	19978	20142

Table 3.7 Forecasting performance evaluation measured by MASE

	Forecasting horizons				
	1	2	3	4	8
Mainland China					
SNAIVE	1.624	1.691	1.858	1.926	1.790
ETS	1.472	1.916	1.885	1.678	2.002
SARIMA	1.777	1.631	1.963	2.117	2.503
ADL	0.902	1.410	1.418	1.398	0.940
Malaysia					
SNAIVE	1.033	1.024	1.010	1.117	2.720
ETS	0.974	1.016	1.069	1.076	0.727
SARIMA	1.052	1.049	1.121	0.983	1.326
ADL	1.056	1.115	1.103	0.977	1.833
Russia					
SNAIVE	0.830	0.656	0.644	0.698	1.239
ETS	0.657	0.633	1.191	0.912	1.725
SARIMA	0.802	0.959	0.968	1.237	3.059
ADL	1.044	0.836	1.081	0.895	0.943
US					
SNAIVE	0.962	0.936	0.875	0.863	1.862
ETS	0.646	0.652	0.768	1.078	2.110
SARIMA	0.773	1.144	1.409	1.684	2.624
ADL	0.873	0.977	0.989	1.065	0.955