An Error Correction Analysis of Visitor Arrivals to the Bahamas

Charles, Jacky S. and Fullerton, Thomas M., Jr.

University of Texas at El Paso

11 February 2012

Tourism Economics

Volume 18, 2011, Pages 253-259

An Error Correction Analysis of Visitor Arrivals to the Bahamas

Jacky S. Charles
Ministry of Finance, Economic Affairs and National Development,
Economic Planning Section, 1st Floor American Drywall Building
Vide Boutielle, Castries, Saint Lucia, Telephone 758-468-3999

Thomas M. Fullerton, Jr.
Department of Economics & Finance, University of Texas at El Paso, El Paso, TX 79968-0543
Telephone 1-915-747-7747, Facsimile 1-915-747-7747, Email tomf@utep.edu

Note: Authors' Manuscript Version

Abstract

Tourism is the major domestic export for many countries in the Caribbean region. Given this, the variables which influence tourism demand in this region, as well as accurate forecasts, can assist policy makers in their planning efforts and growth strategies. This study utilizes error correction models (ECMs) to analyze tourism demand in the Bahamas. Findings suggest that income and habit persistence/word of mouth advertising are the primary determinants of tourism demand in the Bahamas, while the cost of travel is generally insignificant. To further assess model reliability, forecasts of the ECMs are compared to random walk and random walk with drift benchmarks. The study finds that while the ECMs provide fairly reliable forecasts, their performances are not superior to those provided by random walk benchmarks.

Keywords
Tourism, error correction analysis, forecasts, Bahamas

JEL Categories
M21 Business Economics, O54 Caribbean Country Studies

Acknowledgements

Financial support was provided by El Paso Electric Company, Hunt Communities, Hunt Corporation of El Paso, JPMorgan Chase Bank of El Paso, a UTEP College of Business Administration Faculty Research Grant, and the James Foundation Scholarship Fund. Helpful comments and suggestions were provided by Doyle Smith and Dennis Bixler-Marquez. Econometric research assistance was provided by Teodulo Soto and Carlos Morales.

An Error Correction Analysis of Visitor Arrivals to the Bahamas

Introduction

Tourism has been a major part of the economies of the Caribbean region for well over a century and is the major engine of growth for much of the region. The WTTC Travel and Tourism Economic Research report on the area ranked the Caribbean the first in the world (out of 13 regions) for relative contribution of travel and tourism to the national economy (World Travel and Tourism Council, 2007). Given the importance of this sector to the Caribbean, empirical analysis of tourism can potentially yield helpful information for countries located in this region. Of particular interest are models that may provide accurate forecasts of international inbound tourists to the Caribbean region. Such models can assist planning efforts for the tourism sectors in these countries.

The tourism literature has long recognized the benefits of accurate forecasts (Archer, 1976; Archer 1994; Morley 1991, Song and Witt, 2000). Accurate forecasts (both short term and long term) can help improve planning efforts by both private and public sectors. For the private sector, these forecasts are utilized for determining investments in aircraft, hotels, hotel industry staff, physical facilities, water craft, supplies, and so forth. Governments are interested in tourist arrivals for national budgeting purposes, as a large percentage of tax and fee revenues are generated by the tourism sector. Examples include room, sales, departure, and passenger ticket taxes. Accurate forecasts of tourist arrivals are, therefore, helpful for effective public sector budgeting efforts.

A variety of studies examine international tourist flows to various Caribbean countries. Many of these studies utilize a structural econometric approach for analyzing tourism demand, but do not, generally, employ them for out of sample simulation exercises (Clarke, 1978; Carey, 1991; Metzgen-Quemarez, 1990; Vanegas and Croes, 2000; Vanegas and Croes, 2005, Yoon and Shafer, 1996). Studies which develop forecasting models primarily rely on structural time series models (Greenidge, 2000); univariate and transfer function autoregressive integrated moving average (ARIMA) models; and autoregressive (AR) models (Dharmaratne, 1995; Dalrymple and Greenidge, 1999). A growing number of studies employ error correction models (ECMs) to analyze tourism demand in different markets around the world (Song, Witt, and Jensen, 2003; Ouerfelli, 2008). Comparatively few of the Caribbean studies to date, however, have tried to utilize ECMs for forecasting tourist arrivals (Croes and Rivera, 2010).

Error correction terms provide a means of capturing adjustments in a dependent variable which depend not only on the levels of different explanatory variables, but also on the extent to which an explanatory variable deviates from an equilibrium relationship with the dependent variable (Banerjee, Dolado, Galbraith, and Hendry, 1993). Simply put, the idea behind the error correction mechanism is that a percentage of the disequilibrium from one period is corrected in the next period. The objective of this study is to develop a set of error correction models for tourist arrivals to the Bahamas. Out-of-sample forecast properties of the models are also employed as an additional means for empirical performance verification.

The study is organized as follows. Section 2 provides a brief review of related literature. Section 3 discusses the modeling framework and econometric methodology. Data and empirical results are summarized in the fourth section. The final section provides suggestions for future research.

Literature Review

This section summarizes related contributions to the literature on modeling and forecasting international tourism demand. Prior studies suggest a large number of possible approaches to estimating structural demand models. Schulmeister (1980) identifies exogenous variables such as disposable income and relative prices between destinations as useful information in explaining tourism demand. Frequently, lags of dependent variables such as price, income, and consumption are included in different types of dynamic models (Witt and Martin, 1987; Morley, 1991). The prevalence of lagged dependent variables included as regressors reflects the nature of the tourism industry as one that is heavily influenced by individual habits, persistence, and word of mouth advertising of different destinations (Song, Li, Witt, and Fei, 2010).

The high volatility of international tourism poses serious challenges to forecasting international tourist flows. Another problem that frequently occurs is multicollinearity among income, airfare, and other variables typically utilized in these models. This problem is encountered by many researchers using time series data to estimate tourism demand models. In order to deal with multicollinearity, Fuji and Mak (1980) employ ridge regression. Results from that study indicate that employment of ridge regression to control for mutlicollinearity among the explanatory variables can sometimes help identify the variables that should be retained for simulation, but this approach is not often utilized.

Dharmaratne (1995) employs a univariate ARIMA model approach utilizing annual time series data for a period of thirty eight (38) years to forecast long stay visitors to Barbados. The ARIMA model is found to provide excellent forecasts in the short term (1-2 years), with a tight fits around the actual data. However, as the number of years simulated increases, the forecasts deviate considerably from the actual data, with increasingly large standard errors. Dalrymple and Greenidge (1999) also employ univariate ARIMA equations to model arrivals to Barbados, but argue that quarterly data are more useful in policy settings. Results from diagnostic tests coupled with in-sample and out-of-sample forecasts confirm the reliability of ARIMA models in producing short term forecasts.

Greenidge (2000) employs a Structural Times Series Model (STM) to explain and forecast tourist arrivals to Barbados using quarterly data. Initially, Basic Structural Models (BSMs) are estimated which exclude all explanatory variables, and only include trend, seasonal, and cyclical components. Also estimated are General Structural models (GSMs) which include these components as well as the explanatory variables. The study finds that the BSM produce, overall, better in-sample and better out-of-sample forecasts than the GSM.

Kulendran and King (1997) forecast arrivals to Australia using quarterly data on inbound tourist flows. Models estimated include ECMs, univariate ARIMA equations, BSMs, and regression based time series models. The relative performance of each model is found to vary among countries of origin for the various traveler groups. In general, the ARIMA models are found to produce more reliable short term forecasts, while the ECMs perform somewhat better for longer term forecasts. ECMs do not fare very well, for most series, in this market.

Kim and Song (1998) use cointegration and error correction techniques to analyze long-run and short-run inbound tourism demand in South Korea. Ex post forecasts with four different time horizons are generated from seven different modeling approaches. Simulation results indicate that the best models tend to be ECM or univariate ARIMA models, depending on the tourist generating market. For the United States (USA) and United Kingdom (UK) source markets, the ECM is the most accurate. For the German and Japanese markets, the ECM is outperformed by the ARIMA methodology. Similar evidence of mixed results for ECMs are also reported in several other studies (Song and Witt, 2000; Kulendran and Witt, 2001; Song and Witt, 2003).

Song, Romilly and Liu (2000) use a general to specific approach to construct UK demand for outbound tourism models to twelve destinations. Ex post forecasts are generated over a period of six years from ECMs, with results obtained compared to those of a naïve model, an autoregressive AR(1) model, an autoregressive moving average (ARMA) equation, and a Vector Auto Regression (VAR) model. Results suggest that the ECM model provides the best forecasting performance realtive to the other models. Diagnostic tests for normality, heteroscedasticity, serial correlation, functional form, and structural stability indicate that the ECM can be used for policy analysis as well as forecasting purposes.

Ouerfelli (2008) uses cointegration analysis and ECMs to estimate long-run tourism demand elasticities and forecast quarterly European tourism demand for a one-year-ahead horizon. The behavior of European tourists varies noticeably from one country to another. Findings from this study also indicate that multiple statistically significant long-run relationships can be documented for tourism flows. Empirical results indicate that ECMs produce relatively accurate short-range forecasts.

A number of studies indicate that ECM analysis offers a viable means for modeling and forecasting international visitor flows. To date, this technique has not been tested using data from the Bahamas. Given that the Bahamas accounts for a significant share of the tourism sector in the Western Hemisphere, this country provides a logical candidate for examining whether ECM analysis also works reliably using tourism data from it.

Theoretical Model

Error correction models (ECMs) are potentially useful because they allow capturing both long-run and short-run dynamics of tourist arrivals to the Bahamas (Engle and Granger, 1987). Because the United States, Canada, and Europe send the most visitors to it, the business cycles of these large economies will likely influence the bulk of tourist arrivals to the Bahamas. The basic arrangement of these models incorporates the hypothesis that both long-run and short-run forces may influence changes in tourist arrival behavior.

The long-run tourism demand model for tourist generating country i may be expressed as:

$$\ln TA_{it} = a_0 + a_1 \ln Y_{it} + a_2 \ln P_t + U_t \quad (1)$$

where TA_{it} is tourism demand, measured by tourist arrivals from origin country i in year t ; i = 1,2,3 represents United States, Canada and Europe respectively; Y_{it} is real income, measured by gross domestic product (GDP) or disposable personal income (PDI) in origin country i in year t; and P_t is the price of oil in year t. The price of oil, of course, influences the price of overseas travel. The coefficient signs in Equation (1) are hypothesized as $a_1>0$ and $a_2<0$.

Equation (1) is used to estimate the long-run impact of percentage fluctuations in income and oil prices on tourist arrivals. Although the equilibrium long-run relationship can be estimated directly using Equation (1), it is also important to consider short-run dynamics since the system may not always be in equilibrium. A simple dynamic model of short-run adjustment can be written as:

$$\Delta \ln TA_{it} = b_0 + b_1 \Delta \ln Y_{it} + b_2 \Delta \ln P_t + b_3 U_{t-1} + V_t \quad (2)$$

where Δ is the first difference operator and U_{t-1} is a random error term. Changes in tourist arrivals are determined by short-run movements in the explanatory variables and by long-run forces through the error correction term U_{t-1}, which measures the equilibrium error from the previous period.

Hypothesized parameter signs in Equation (2) are and $b_1>0$, $b_2<0$, and $b_3<0$. The coefficient of U_{t-1} is expected to be negative and significant, implying that the model adjusts toward equilibrium by removing b_3 units of the error observed during the prior period. Re-writing Equation (1) at time t-1 and solving for U_{t-1} yields the following result:

$$U_{t-1} = \ln TA_{it-1} - a_0 - a_1 \ln Y_{it-1} - a_2 \ln P_{t-1} \quad (3)$$

Substitution of Equation (3) into Equation (2) and rearrangement generates the tourist arrivals error correction equation:

$$\Delta \ln TA_{it} = (b_0 - a_0 b_3) + b_1 \Delta \ln Y_{it} + b_2 \Delta \ln P_t + b_3 \ln TA_{it-1} + a_1 b_3 \ln Y_{it-1} + a_2 b_3 \ln P_{t-1} + V_t \quad (4)$$

If there is a long-run equilibrium relationship between tourist arrivals and the explanatory variables in Equation (1), then those variables should be co-integrated. The Engle and Granger (1987) two stage approach is as follows. The first stage is to estimate the parameters of the co-integrating Equation (1) and then test for the existence of unit roots in the estimated error term. In testing for unit roots, the augmented Dickey and Fuller (1979), ADF, procedure can be used. The ADF test is based on the following equation:

$$\Delta y_t = \alpha + \rho y_{t-1} + \beta T + \sum_{i=1}^{p} \gamma_i \Delta y_{t-i} + \nu_t \quad (5)$$

where y_t is the relevant time series variable, T is a linear deterministic trend, and v_t is an error term, which is assumed to have a mean of zero and constant variance (Kim and Song, 1998; Song, Romilly and Liu, 2000). If there are problems of serial correlation and heteroscedasticity when carrying out the ADF test, then the ADF statistic will be invalid, and in this case the Phillips and Perron (1988), PP, test should be employed. The PP test is also based on Equation (5), but assumes that the residuals are serially correlated.

Cointegration requires that all variables in the long-run cointegration equation be integrated of order 1 or *I* (1). Engle and Granger (1987) demonstrate that, if cointegration is found among a set of variables in Equation (1), then the cointegration regression can always be transformed into an ECM of the form in Equation (4). Estimation of this dynamic specification forms the second stage of the procedure.

Equation (4) can be rewritten more simply as follows:

$$\Delta \ln TA_{it} = c_0 + c_1 \Delta \ln Y_{it} + c_2 \Delta \ln P_t + c_3 \ln TA_{it-1} + c_4 \ln Y_{it-1} + c_5 \ln P_{t-1} + V_t \quad (6)$$

with the arithmetic signs of each coefficient expected to be $c_1>0$, $c_2<0$, $c_3<0$, $c_4>0$, $c_5<0$. Equation (6) includes the effects of both short-run and long-run forces on changes in visitor arrivals. Changes in arrivals are expected to be determined by variations in the level of income, oil prices, and the level of tourist arrivals at time t-1. Following parameter estimation, the ECMs are also used to produce out-of-sample simulations. That step is taken as an additional means for examining model reliability. The accuracy of each equation's forecasts is compared to those generated by random walk and random walk with drift benchmarks. To date, relatively few tourism demand studies have employed these benchmarks to gauge model reliability, but they have proven useful in related travel and transportation contexts (Fullerton, 2004; De Leon, Fullerton, and Kelley, 2009).

Data and Empirical Results

Annual data on visitor arrivals to the Bahamas for 1977 – 2007 are used as the dependent variables. The sample period and data frequency are chosen because they provide the most consistent data set available. Because tourists travel to the Bahamas by either air or sea, the dependent variable is represented by both total air and cruise ship arrivals. Also, since not all tourists to the Bahamas stay more than twenty four hours, another measure of arrivals used is stop-over visitors (tourists that stay 24 hours or more). As note above, the visitor data are from the three major tourist generating economies (USA, Canada, and Europe).

If holiday demand or visits to friends and relatives are under consideration, then the appropriate metric for the income variable is either PDI or private consumption, where personal disposable income is defined as the amount of current income that individuals have available for either spending or saving. That definition means PDI is personal income minus personal income taxes and national insurance contributions (Lipsey and Chrystal, 2004). However, because business visits may form an important part of the total, a more general income variable such as national income or GDP may also be appropriate (Song and Witt, 2000).

Given that tourism demand to the Bahamas is largely for holiday purposes, PDI is probably the most appropriate measure of income to utilize. Unfortunately, as a consequence of data constraints, GDP has to be used to measure income for Canada and Europe. PDI estimates are utilized, however, in the case of the USA. The price of jet fuel is also used as an explanatory variable to measure travel cost to the Bahamas. Kim and Song (1998) measure this variable in the form of return airfares. Variable definitions and data sources are provided in Table 1. Data for all variables utilized are reported in Appendix A.

Table 1 About Here

Results from the unit root tests are reported in Table 2. The ADF test is undertaken for 3 variables for each origin country. Outcomes indicate that all of the variables are stationary after the first difference [i.e., all variables are I (1) variables]. Given this, standard regressions in level form may be spurious. Column 1 lists the explanatory variables for each dependent variable in the sample.

Table 2 About Here

Cointegration tests are next carried out to determine if linear combinations of these $I(1)$ variables are stationary or $I(0)$. Results from the co-integration tests are shown in Tables 3 and 4. The Johansen (1988) cointegration test in Table 3 indicates that there is one cointegrating vector at the 5-percent level of significance for stopover visitors from USA, Canada, and Europe. The test also indicates one cointegrating vector for total tourist air arrivals, while there was no vector identified for total cruise ship arrivals.

Table 3 About Here

In Table 4, the Engle and Granger (1987) procedure also indicates that the residuals from the cointegrating regressions are likely to be $I(0)$. Those results are also confirmed by the ADF test statistics. These outcomes suggest that the variables in each of these long-run regressions are cointegrated. Given the results of the co-integrating regressions, the corresponding error correction models can be estimated, by incorporating the lagged error terms from the co-integrating models. Although cruise ship arrivals do not satisfy the cointegration test, an ECM is still estimated for comparison purposes. The lack of cointegration in the cruise ship arrivals model is possibly because of the omission of one or more explanatory variables that are specific to demand by these tourists. Also, in a few models there are estimated parameters which have algebraic signs opposite of those hypothesized. Given these concerns, care should be taken with respect to the interpretation of the econometric output obtained below.

Table 4 About Here

The coefficients of the corresponding ECMs in Table 5 are the long-run and short-run demand elasticities. The income coefficient is significant in the models for Canada, Europe and total tourist air arrivals. The short-run income elasticity for Canadian stopovers is inelastic, while those for European stopovers and for total air arrivals are elastic. In the case of European stopover tourists, the long-run income coefficient is inelastic. Collectively, the results suggest that income fluctuations play important roles in determining tourist flows to the Bahamas (Witt and Martin, 1987; Morley, 1991). This appears to especially hold true for air arrivals and travelers from Europe.

Table 5 About Here

Four of the five short-run jet fuel price parameters in Table 5 fail to satisfy the 5-percent significance criterion. In the equation for Canadian stopovers, the short-run fuel price is significant, but positive. Four of the five long-run jet fuel price coefficients also fail to satisfy the significance threshold. The estimated fuel price parameter for stopover visitors from Europe is, however, significant and has a plausible magnitude associated with it. Taken as whole, the results suggest that changes in the price of jet fuel and, by extension, the cost of transportation do not exercise very much influence over the volume of vacationers that visit the Bahamas.

The one period lags of the dependent variables, not differenced, are the most consistently significant regressors in Table 5. It suggests that habit persistence and/or word of mouth recommendation are the major driving forces for holiday tourism demand to the Bahamas. That result has also been documented for other vacation destinations using alternative methodologies (Witt 1980). This result is interesting and warrants further research for other markets in the Caribbean and elsewhere.

As an additional step to assess model reliability, out-of-sample simulations for the ECMs are compared to random walk (RW) and random walk with drift (RWD) benchmarks. The ECMs are re-estimated using sub-sets of available history and used to generate two year dynamic simulations. The metrics used in this study for the evaluation of forecasting performance are the root mean square error (RMSE), and the Theil inequality coefficient, *U* (Pindyck and Rubinfeld, 1998). U-statistics can take values between 0 and 1. The ideal distribution of the inequality coefficient second moment proportions (bias proportion, variance proportion and covariance proportion) is 0, 0, and 1, respectively. Information on to calculate each measure is reported in Appendix B.

Table 6 summarizes predictive accuracy results for each of the tourism demand series. The most striking result is that none of the ECM forecasts prove as accurate as either of the random walk benchmarks. Although simulation bias hampers the performances for stopover visitors from Europe and for total air arrivals, the fact that none of the ECM forecasts obtain U-statistics that are lower than either benchmark is cause for concern and goes beyond the mixed evidence reported for other regions with this technique (Song and Witt, 2000; Kulendran and Witt, 2001; Song and Witt, 2003). In spite of that, it should also be noted, however, that the Theil inequality coefficients reported in Column 3 of Table 6 all point to good forecasting performances for each of the ECMs. Along those lines, the data in Column 5 further indicate that all of the ECMs except that for cruise ship arrivals do good jobs in terms of simulating the variability of tourism flows to the Bahamas.

Table 6 About Here

Annual data are utilized in this study. Before concluding that the ECM approach is not a good candidate for analyzing tourism demand in the Bahamas, additional testing using higher frequency quarterly and/or monthly data is in order. Experimentation with other model specifications than those utilized above may also prove helpful. Beyond those steps, the results reported herein also provide a fairly strong indication that other modeling approaches such as those provided by structural econometric and ARIMA time series frameworks should be considered for this important holiday destination (Dharmaratne, 1995; Kulendran and King, 1997; Dalrymple and Greenidge, 1999; Greenidge, 2000). The outcomes shown in Table 4 also indicate that a structural econometric approach is likely to meet with success.

The relative accuracy performance of the random walk benchmarks shown in Table 6 indicates that recent historical evidence regarding tourism flows should probably be carefully considered as part of any type of planning exercises. These results are interesting because previous studies on tourism demand have not used these benchmarks to assess predictive accuracy. Given that, it is difficult to ascertain if these outcomes are unique to the Bahamas. Further research employing different sample data sets would be required in order to confirm these outcomes. Their inclusion in future work of this nature is potentially informative.

Conclusion

In this study, error correction models are estimated and used to forecast inbound tourism demand in the Bahamas from its major tourist generating regions. The models are estimated for total air and cruise ship arrivals, as well as for stopover visitors, to the Bahamas. Time series test procedures suggest that co-integration/long-run equilibrium relationships exist for 4 of the 5 models identified.

The empirical results provide some useful insights concerning tourism demand in the Bahamas. Income is found to be significant in the models for Canada and Europe, while insignificant in the model for the United States. Further, income is found to be a primary determinant of tourist arrivals by air, while insignificant in determining the number of cruise ship arrivals. The short-run and long-run income elasticities suggest that the demand for tourism in the Bahamas tends to be relatively income inelastic in the short-run, but elastic in the long-run. Income variations in the major tourist generating countries influence the numbers of tourist air arrivals to the Bahamas, but further research is needed to determine what factors influence cruise ship visitors.

Fuel price elasticities in the error correction equations are generally found to be insignificant in the decision making process of tourists who travel to the Bahamas by both air and sea. That suggests that tourist arrivals are affected more by business cycle fluctuations than by the cost of travel to the Bahamas. The cointegration equations indicate, however, that fuel prices do exercise noticeable impacts on most visitor categories. Because the Bahamas are very close to the United States, it would not be surprising to discover that transport costs exert greater influence on tourism flows to less accessible locations elsewhere in the Caribbean.

The lagged dependent variable is significant in all models and suggests that habit persistence and word of mouth recommendation are primary determinants of the demand for the

Bahamas as a holiday destination. Accordingly, it is important that the tourism product the Bahamas provides is of high quality and offers a good vacation experience. In this context, future efforts should consider attempting to examine the effectiveness of different marketing, advertising, and promotional campaigns.

Out-of-sample simulations generated from the error correction models are compared with random walk and random walk with drift benchmarks. These results show that while the ECMs provide a fairly reliable means of forecasting tourist arrivals to the Bahamas, they are not more accurate than random walk and random walk with drift extrapolations over the course of the sample simulation period. Also, because bias is a problem with the error correction model forecasts for European stopovers and for total tourist air arrivals, care should be exercised with respect to using these out-of-sample forecasts. Further research appears warranted with respect to model specifications, data frequencies, and alternative estimation techniques.

References

B.H. Archer, 1976, *Demand Forecasting in Tourism,* Bangor, UK: University of Wales Press.

B.H. Archer, 1994, "Demand Forecasting and Estimation," Chapter 10 in *Travel, Tourism and Hospitality Research: A Handbook for Managers and Researchers*, Edited by J.R.B. Ritchie and C.R. Goeldner, New York, NY: John Wiley & Sons.

A. Banerjee, J.Dolado, J.W. Galbraith, and D. F. Hendry, 1993, *Co-integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford, UK: Oxford University Press.

J. Blackwell, 1970, "Tourist Traffic and the Demand for Accommodation: Some Projections," *Economic & Social Review* 1, 323-343.

K. Carey, 1991, "Estimation of the Caribbean Tourism Demand: Issues in Measurement and Methodology," *Atlantic Economic Journal* 19, 32-40.

C.D. Clarke, 1978, *An Analysis of the Determinants of Demand for Tourism in Barbados*, Ph.D. Dissertation, Fordham University.

R.S. Cline, 1975, "Measuring Travel Volumes and Itineraries and Forecasting Future Travel Growth to Individual Pacific Destinations," Chapter 9 in *Management Science Applications to Leisure Time Operations*, Edited by S.P. Ladany, New York, NY: North Holland.

R. Croes and M.A. Rivera, 2010, "Testing the Empirical Link between Tourism and Competitiveness: Evidence from Puerto Rico," *Tourism Economics* 16, 217-234.

K. Dalrymple and K. Greenidge, 1999, "Forecasting Arrivals to Barbados," *Annals of Tourism Research* 26, 188-190.

M. De Leon, T.M. Fullerton, Jr., and B.W. Kelley, 2009, "Tolls, Exchange Rates, and Borderplex International Bridge Traffic," *International Journal of Transport Economics* 36, 223-259.

G.S. Dharmaratne, 1995, "Forecasting Tourist Arrivals in Barbados," *Annals of Tourism Research* 22, 804-818.

D.A. Dickey and W.A. Fuller, 1979, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association* 74, 427-431.

R.F. Engle and C.W.J. Granger, 1987, "Cointegration and Error Correction Representation, Estimation, and Testing," *Econometrica* 55, 251-276.

R.F. Engle and H. White, 1999, *Co-integration, Causality and Forecasting*, Oxford, UK: Oxford University Press.

T.M. Fullerton, Jr., 2004, "Borderplex Bridge and Air Econometric Forecast Accuracy," *Journal of Transportation & Statistics* 7, 7-21.

C.W.J. Granger and P. Newbold, 1974, "Spurious Regressions in Econometrics," *Journal of Econometrics* 2, 111-120.

C.W.J. Granger, 1981, "Some Properties of Time Series Data and their use in Econometric Model Specification," *Journal of Econometrics* 16, 121-130.

K. Greenidge, 2000, "Forecasting Tourism Demand, an STM Approach," *Annals of Tourism Research* 28, 98-112.

S. Johansen, 1988, "A Statistical Analysis of Co-integration Vectors," *Journal of Economic Dynamics and Control* 12, 231-254.

S. Johansen and K. Juselius, 1990, "Maximum Likelihood Estimation and Inference on Co-integration with Application to the Demand for Money," *Oxford Bulletin of Economics and Statistics* 52, 169-210.

S. Kim and H. Song, 1998, "Analysis of Inbound Tourism Demand in South Korea: A Cointegration and Error Correction Approach, *Tourism Analysis* 3, 25-41.

N. Kulendran and M.L. King, 1997, "Forecasting International Tourist Flows using Error Correction and Time Series Models," *International Journal of Forecasting* 13, 319-327.

N. Kulendran and S.F. Witt, 2001, "Cointegration versus Least Squares Regression," *Annals of Tourism Research* 28, 291-311.

C. Lim and M. McAleer, 2001, "Forecasting Tourist Arrivals," *Annals of Tourism Research* 28, 965-977.

R.G. Lipsey and K.A. Chrystal, 2004, *Economics*, Oxford, UK: Oxford University Press.

J. Mackinnon, 1991, "Critical Values for Cointegration Tests," Chapter 13 in *Long-Run Economic Relationships: Readings in Cointegration*, Edited by R.F. Engle and C.W.J. Granger, Oxford, UK: Oxford University Press.

G.S. Maddala, I. Kim, 1998, *Unit Roots, Co-integration, and Structural Change*, Cambridge, UK: Cambridge University Press.

Y. Metzgen-Quemarez, 1990, *Estimating the Demand for International Tourist Service: The US and the Caribbean*, Ph.D. Dissertation, Princeton University.

C. Morley, 1991, "Modeling International Tourism Demand: Model Specification and Structure," *Journal of Travel Research* 30 (1), 40-44.

C. Morley, 1997, "An Evaluation of the use of Ordinary Least Squares for Estimating Tourism Demand Models, *Journal of Travel Research* 35 (4), 69-73.

C. Ouerfelli, 2008, "Co-integration Analysis of Quarterly European Tourism Demand in Tunisia," *Tourism Management* 29, 127-137.

P.C.B. Phillips, 1986, "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics* 33, 311-340.

P.C.B. Phillips and P. Perron, 1988, "Testing for a Unit Root in Time Series Regression," *Biometrika* 75, 335-346.

R.S. Pindyck and D.L. Rubinfeld, 1998, *Econometric Models and Econometric Forecasts*, Boston, MA: Irwin/McGraw-Hill.

G. Riddington, 1993, "Time Varying Coefficient Models and their Forecasting Performance," *Omega: International Journal of Management Sciences* 21, 571-583.

S. Schulmeister, 1980, *Tourism and the Business Cycle: Econometric Models for the Purpose of Analysis and Forecasting of Short-Term Changes in the Demand for Tourism*, Vienna, AT: Austrian Institute for Economic Research.

H.Y. Song, G. Li, S.F. Witt, and B.G. Fei, 2010, "Tourism Demand Modelling and Forecasting: How should Demand be Measured?" *Tourism Economics* 16, 63-81.

H.Y. Song, P. Romilly and X. Liu, 2000, "An Empirical Study of Outbound Tourism Demand in the UK," *Applied Economics* 32, 611-624.

H.Y. Song and S.F. Witt, 2000, *Tourism Demand Modeling and Forecasting: Modern Econometric Approaches*, Amsterdam, NL: Pergamon.

H.Y. Song and S.F. Witt, 2003, “Tourism Forecasting: The General to Specific Approach,” *Journal of Travel Research* 42 (1), 65-74.

H.Y. Song, S.F. Witt and T.C. Jensen, 2003, “Tourism Forecasting: Accuracy of Alternative Econometric Models,” *International Journal of Forecasting* 19, 123-141.

H.Y. Song and K.F. Wong, 2003, “Tourism Demand Modeling: A Time Varying Parameter Approach,” *Journal of Travel Research* 42 (1), 57-64.

C. Ouerfelli, 2008, “Co-integration Analysis of Quarterly European Tourism Demand in Tunisia,” *Tourism Management* 29, 127-137.

Sr. M. Vanegas and R.R. Croes, 2000, “Evaluation of Demand: US Tourists to Aruba,” *Annals of Tourism Research* 27, 946-963.

Sr. M. Vanegas and R.R. Croes, 2004, “An Econometric Study of Tourist Arrivals in Aruba and its Implications,” *Tourism Management* 26, 879-890.

S.F. Witt, 1980, “An Abstract Mode – Abstract (Destination) Node Model of Foreign Holiday Demand,” *Applied Economics* 12, 163-180.

S.F. Witt and C.A. Martin, 1987, “Econometric Models for Forecasting International Tourism Demand,” *Journal of Travel Research* 15 (3), 23-30.

C.A. Witt and S.F. Witt, 1990, “Appraising an Econometric Forecasting Model,” *Journal of Travel Research* 28 (3), 30-34.

World Travel and Tourism Council, 2007, *Caribbean Region: Review of Economic Growth and Development*, WTTC Investigation Number 332-496.

J. Yoon and E.L. Shafer, 1996, “Models of U.S. Travel Patterns for the Bahamas,” *Journal of Travel Research* 35 (1), 50-60.

Table 1
Variable Definitions

Variable	**Definition**
SUS	Natural logarithm of stop over visitors from the United States at time *t*.
SCA	Natural logarithm of stop over visitors from Canada at time *t*.
SEU	Natural logarithm of stop over visitors from Europe at time *t*.
TA	Natural logarithm of tourist air arrivals at time *t*.
CS	Natural logarithm of cruise ship arrivals at time *t*.
GDP	Natural logarithm of real gross domestic product at time *t* in country *i*.
PDI	Natural logarithm of USA real disposable personal income at time *t*.
P	Natural logarithm of jet fuel prices at time *t*.

Notes:
All visitor data are from the Bahamas Ministry of Tourism website (*www.tourismtoday.com*).
All real GDP data are in constant USA dollars, using 2000 as the base year.
GDP data are from the April 2009 IMF *International Financial Statistics CD-ROM*, and from the World Bank's Development Indicators (WDI) online database (*www.worldbank.org*).
USA PDI data are in constant USA dollars, using 2005 as the base year.
USA PDI data are from the USA Bureau of Economic Analysis website (*www.bea.gov*).
Jet fuel price data are from the USA Energy Information Administration website (*www.eia.gov*).
Sample period, 1977-2007.
Data frequency, annual.

Table 2
ADF Unit Root Test Results

Explanatory Variable	**Dependent Variable (Levels)** lnSUS	lnSCA	lnSEU	lnTA	lnCS
lnTA				-3.470 (1)	
lnCS					-1.068 (0)
lnSUS	-1.611 (1)				
lnSCA		-1.473 (6)			
lnSEU			-0.788 (5)		
lnGDP		-2.590(1)	-2.610 (1)	-3.805 (1)	-3.805 (1)
lnPDI	-2.028 (0)				
lnP	-0.531 (2)	-0.439 (2)	-2.023 (0)	-0.531 (2)	-0.531 (2)

Explanatory Variable	**Dependent Variable (First Differences)** lnSUS	lnSCA	lnSEU	lnTA	lnCS
lnTA				-4.679 (1)***	
lnCS				-5.466 (0)***	
lnSUS	-5.631 (0)***				
lnSCA		-4.940(0)***			
lnSEU			-4.042 (4)**		
lnGDP		-3.750 (1)**	-2.493 (0)**	-4.120 (1)**	-4.120 (1)**
lnPDI	-5.429 (0)***				
lnP	-4.376 (1)**	-4.454 (1)**	-4.096 (1)**	-4.376 (1)**	-4.376 (1)**

Notes:
Numbers in parentheses are the number of lags used for the ADF test.
The numbers of lags are using the Akaike Information Criterion.
* Denotes significance at the 10-percent level.
** Denotes significance at the 5-percent level.
*** Denotes significance at the 1-percent level.

Table 3
Johansen Test for Cointegration

Dependent Variable	Hypothesized Number of Co-integrating Vectors	Trace Statistic	Max-Eigen Statistic
lnSUS	None	36.012**	18.773
	At most 1	17.240	11.070
	At most 2	6.171	6.171
lnSCA	None	43.711**	37.971**
	At most 1	5.740	5.682
	At most 2	0.058	0.058
lnSEU	None	30.307**	17.113
At most 1		13.195	11.864
At most 2		1.331	1.331
lnTA	None	34.798**	28.274
	At most 1	6.523	6.243
	At most 2	0.280	0.280
lnCS	None	25.541	15.995
	At most 1	9.546	9.009
	At most 2	0.537	0.537

Notes:
* Denotes significance at the 10-percent level.
** Denotes significance at the 5-percent level.
*** Denotes significance at the 1-percent level.

Table 4
Cointegrating Equations

Estimated Equation	ADF Statistic	Lags
$\ln SUS_t = 0.737 \ln PDI_{1t} - 0.101 \ln P_t - 7.065$ (7.072)*** (-1.860)* (-6.562)***	(-3.639)***	0
$\ln SCA_t = {}^{ws}- 0.970 \ln GDP_{2t} - {}^{ws}0.028 \ln P_t + 7.143$ (-5.092)*** (0.359) (3.750)***	(-2.933)***	1
$\ln SEU_t = 1.123 \ln GDP_{3t} - 0.033 \ln P_t - 23.553$ (2.489)** (-2.077)** (-2.748)***	(-2.636)***	1
$\ln TA_t = 0.285 \ln GDP_{4t} - 0.001 \ln P_t - 8.203$ (5.223)*** (-0.272) (-5.009)***	(-2.052)**	0
$\ln CS_t = 2.053 \ln GDP_{5t} - 0.208 \ln P_t - 60.228$ (12.582)*** (-1.830)* (-12.271)***	(-2.612)**	0

Notes:
* Denotes significance at the 10-percent level.
** Denotes significance at the 5-percent level.
*** Denotes significance at the 1-percent level.
ws Denotes coefficient sign opposite of that hypothesized.

Table 5
Estimated Error Correction Models

USA Stopover Visitors

$$\Delta \ln SUS_t = -0.801 + 0.420\Delta \ln PDI_{1t} - 0.022\Delta \ln P_t - 0.322\ln SUS_{1t-1} + 0.094\ln PDI_{1t-1} - 0.043\ln P_{t-1}$$

(-0.749) (0.475) (-0.372) (-2.928)*** (0.875) (-1.243)

R-squared	0.410	F-statistic	3.343
R- squared (adjusted)	0.288	Standard Error	0.060
Log likelihood	45.120	DW statistic	1.766

Canada Stopover Visitors

$$\Delta \ln SCA_t = 3.471 + 0.603\Delta \ln GDP_{2t} + {}^{WS}0.036\Delta \ln P_t - 0.423\ln SCA_{2t-1} - {}^{WS}0.448\ln GDP_{2t-1} - 0.054\ln P_{t-1}$$

(1.997)** (1.742)** (2.820)*** (-3.400)*** (-2.290)** (-1.085)

R-squared	0.478	F-Statistic	4.210
R-squared (adjusted)	0.364	Standard Error	0.084
Log likelihood	34.058	DW statistic	1.823

European Stopover Visitors

$$\Delta \ln SEU_t = -7.147 + 1.751\Delta \ln GDP_{3t} - 0.116\Delta \ln P_t - 0.293\ln SCA_{3t-1} + 0.360\ln GDP_{3t-1} - 0.299\ln P_{t-1}$$

(-1.380)* (1.991)* (-0.732) (-3.052)*** (1.338)*** (-3.278)***

R-squared	0.414	F-statistic	3.390
R-squared (adjusted)	0.292	Standard Error	0.150
Log likelihood	17.638	DW statistic	1.125

Total Air Arrival Visitors

$$\Delta \ln TA_t = -1.529 + 1.1891\Delta \ln GDP_{4t} + {}^{WS}0.062\Delta \ln P_t - 0.328\ln TA_{4t-1} + 0.054\ln GDP_{4t-1} - 0.008\ln P_{t-1}$$

(-0.948) (2.328)** (1.299) (-2.703)*** (0.992) (-0.307)

R-squared	0.465	F-statistic	4.354
R-squared (adjusted)	0.358	Standard Error	0.051
Log likelihood	51.611	DW statistic	2.158

Total Cruise Ship Arrivals

$$\Delta \ln CS_t = -3.742 - {}^{WS}1.038\Delta \ln GDP_{5t} - 0.021\Delta \ln P_t - 0.113\ln CS_{5t-1} + 0.130\ln GDP_{5t-1} + {}^{WS}0.009\ln P_{t-1}$$

(-0.557) (-0.960) (-0.189) (-1.175) (0.570) (0.148)

R-squared	0.189	F-statistic	1.342
R-squared (adjusted)	0.049	Standard Error	0.108
Log likelihood	31.445	DW statistic	2.121

Notes:
* Denotes significance at the 10-percent level.
** Denotes significance at the 5-percent level.
*** Denotes significance at the 1-percent level.
ws Denotes coefficient sign opposite of that hypothesized.

Table 6
Out-of-Sample Simulation Predictive Accuracy

Series & Forecast	RMSE	Theil U Coefficient	Bias Proportion	Variance Proportion	Covariance Proportion
USA Stopovers					
ECM	0.073	0.028	0.271	0.000	0.729
RW	0.049	0.019	0.000	0.015	0.986
RWD	0.045	0.017	0.071	0.000	0.929
Canada Stopovers					
ECM	0.017	0.112	0.275	0.013	0.713
RW	0.000	0.000	0.022	0.067	0.916
RWD	0.008	0.050	0.040	0.006	0.954
Europe Stopovers					
ECM	0.024	0.154	0.632	0.269	0.118
RW	0.000	0.000	0.236	0.352	0.435
RWD	0.014	0.080	0.021	0.417	0.590
Tourist Air Arrivals					
ECM	0.065	0.023	0.694	0.046	0.263
RW	0.000	0.000	0.026	0.002	0.972
RWD	0.054	0.018	0.036	0.229	0.750
Tourist Cruise Ship Arrivals					
ECM	0.515	0.079	0.178	0.458	0.394
RW	0.000	0.000	0.228	0.238	0.550
RWD	0.317	0.050	0.001	0.171	0.839

Notes:
15 data points are used to calculate each inequality coefficient.
Sample subset estimation periods are 1977-1999; 1977-2000; 1977-2001; 1977-2002; 1977-2003; 1977-2004; 1977-2005; 1977-2006.
Out-of-sample simulation periods are 2000-2001; 2001-2002; 2002-2003; 2003-2004; 2004-2005; 2005-2006; 2006-2007; 2007.
Random walk forecasts are calculated as last available historical observations.
Random walk with drift forecasts are calculated as last available historical observation plus last observed percentage changes.

Appendix A

Table A1: Tourist Air Arrival, Cruise Ship Arrival, and Stopover Historical Data

Year	Tourist Air Arrivals	Cruise Ship Arrivals	USA Stopovers	Canada Stopovers	Europe Stopovers
1977	982,220	399,190	658,690	141,880	64,290
1978	1,181,580	525,370	819,960	143,250	86,740
1979	1,252,280	537,150	851,590	134,710	101,880
1980	1,262,330	642,230	884,030	129,780	114,070
1981	1,105,560	657,760	791,540	109,210	77,750
1982	1,121,070	826,680	910,770	82,730	57,280
1983	1,220,480	1,003,620	1,051,560	86,680	43,910
1984	1,321,330	1,003,920	1,083,240	85,350	40,700
1985	1,385,260	1,246,710	1,205,275	91,700	36,890
1986	1,378,600	1,628,700	1,223,620	72,190	46,450
1987	1,455,921	1,625,449	1,299,215	80,525	67,950
1988	1,448,679	1,709,412	1,274,365	84,330	85,135
1989	1,490,006	1,908,305	1,351,750	94,300	91,320
1990	1,516,396	2,112,123	1,321,930	96,755	96,625
1991	1,303,318	2,318,900	1,176,690	90,120	112,045
1992	1,227,703	2,461,840	1,128,025	97,640	122,140
1993	1,327,319	2,354,941	1,209,550	96,570	133,085
1994	1,332,280	2,114,096	1,254,210	99,025	109,730
1995	1,317,078	1,922,077	1,328,925	85,600	114,950
1996	1,368,038	2,047,820	1,341,300	85,760	127,620
1997	1,368,107	2,078,256	1,310,420	91,330	130,365
1998	1,304,851	2,042,814	1,250,026	83,086	117,954
1999	1,438,887	2,209,404	1,293,235	87,973	125,485
2000	1,481,492	2,722,342	1,294,295	82,840	104,610
2001	1,428,209	2,754,547	1,308,163	79,715	94,047
2002	1,402,894	3,003,077	1,310,140	68,592	79,564
2003	1,428,973	3,165,069	1,305,335	63,148	93,170
2004	1,450,313	3,553,654	1,360,912	68,462	83,590
2005	1,514,532	3,264,885	1,380,083	75,643	85,277
2006	1,491,633	3,238,974	1,365,104	84,639	82,209
2007	1,487,278	3,114,060	1,263,678	100,340	87,170

Notes:
Tourist arrivals and stop-over data are reported in units.

Table A2: USA, Canada, and Europe Income and Jet Fuel Price Historical Data

Year	USA Personal Disposable Income	Canada GDP	Europe GDP	Jet Fuel Price
1977	1,429,661	190,491	51,556,696	2.59
1978	1,602,026	211,103	57,428,798	2.87
1979	1,784,013	236,932	62,544,263	3.90
1980	1,994,796	264,193	71,985,417	6.36
1981	2,227,807	297,909	79,164,594	7.57
1982	2,403,912	306,339	84,471,793	7.23
1983	2,590,456	329,112	89,050,007	6.53
1984	2,879,581	359,664	94,797,990	6.25
1985	3,066,230	388,568	100,551,779	5.91
1986	3,246,952	406,778	105,783,216	3.92
1987	3,421,907	436,680	113,525,689	4.03
1988	3,712,352	475,264	123,384,600	3.80
1989	3,977,160	505,946	133,909,763	4.39
1990	4,239,944	527,070	148,174,487	5.68
1991	4,428,298	531,295	157,298,469	4.83
1992	4,725,797	551,559	162,597,261	4.52
1993	4,912,783	577,127	166,946,001	4.29
1994	5,177,168	616,696	176,056,148	3.95
1995	5,451,187	643,198	190,460,289	4.00
1996	5,753,335	669,488	200,637,764	4.82
1997	6,069,178	711,879	215,112,070	4.53
1998	6,493,891	743,878	227,635,933	3.35
1999	6,799,637	798,732	249,641,876	4.01
2000	7,323,689	860,456	270,999,730	6.64
2001	7,645,115	893,113	289,151,626	5.72
2002	8,005,414	938,984	304,307,120	5.33
2003	8,369,784	979,516	320,138,184	6.46
2004	8,882,065	1,032,144	339,584,121	8.93
2005	9,269,389	1,094,984	364,486,768	12.86
2006	9,905,432	1,160,718	380,206,312	14.80
2007	10,390,289	1,245,065	404,064,465	16.35

Notes:
Income data are reported in millions of nominal dollars.
Jet fuel price data are reported in nominal dollars per million British thermal units (Btu).

Appendix B

Equation (A1) shows how the RMSE is calculated. Y^s is the forecasted value of Y_t; Y^a is the actual value of Y_t, and T is the number of periods.

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T} (Y_t^s - Y_t^a)^2} \qquad \text{(A1)}$$

Equation (A2) shows how the Theil inequality coefficient U is computed.

$$U = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T} (Y_t^s - Y_t^a)^2}}{\sqrt{\sum_{t=1}^{t}(Y_t^s)^2} + \sqrt{\sum_{t=1}^{t}(Y_t^a)^2}} \qquad \text{(A2)}$$

Equations (A3), (A4), and (A5), respectively, show the decomposition of the Theil inequality coefficient. These equations show the computation of the bias, variance and covariance proportions respectively.

$$U^M = \frac{(\bar{Y}_t^s - \bar{Y}_t^a)^2}{(1/T)\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2} \qquad \text{(A3)}$$

$$U^s = \frac{(\sigma_s - \sigma_a)^2}{(1/T)\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2} \qquad \text{(A4)}$$

$$U^C = \frac{2(1-\rho)\sigma_s\,\sigma_a}{(1/T)\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2} \qquad \text{(A5)}$$

The optimal distribution of the second moment inequality proportions is $U^M = 0,\ U^s = 0 \text{ and } U^C = 1$ (Pindyck and Rubinfeld, 1998).

The 13 countries used to compute EU GDP include:

1. UK
2. France
3. Italy
4. Germany
5. Switzerland
6. Spain
7. Netherlands
8. Sweden
9. Ireland
10. Austria
11. Belgium
12. Norway
13. Denmark

Source of nominal GDP is the April 2009 IMF-IFS CD-ROM.

Source for PPP converter: *www.nationmaster.com*

Note that these 13 countries in Europe account for **above 90%** of tourists to the Bahamas.