2010-01-01

Forecasting Tourist Arrivals To The Bahamas Using Error Correction Models

Jacky Sharon Charles
University of Texas at El Paso, jacky_sharon@hotmail.com

FORECASTING TOURIST ARRIVALS TO THE BAHAMAS USING ERROR CORRECTION MODELS

JACKY S. CHARLES

Department of Economics and Finance

APPROVED:

Thomas M. Fullerton, Ph.D., Chair

William Doyle Smith, Ph.D.

Dennis Bixler-Marquez, Ph.D.

Patricia D. Witherspoon, Ph.D.
Dean of the Graduate School

Dedication

To my parents, Michael Charles and Cynthia Ryfer, along with my grandmother, Pricillia Paul for the continuous love and support they have given me.

FORECASTING TOURIST ARRIVALS TO THE BAHAMAS USING ERROR CORRECTION MODELS

by

JACKY S. CHARLES, BSc

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Economics and Finance

THE UNIVERSITY OF TEXAS AT EL PASO

May 2010

Acknowledgements

I would like to begin by expressing my sincerest appreciation and gratitude to my advisor and chair of my thesis committee, Dr Thomas Fullerton for his support, guidance and patience throughout my Masters program, and in particular the writing of this thesis. To the other members of my committee Dr William Doyle Smith and Dr Dennis Bixler-Marquez I am grateful for their advice and encouragement throughout this journey.

To Dr Timothy Roth I am eternally grateful for his encouragement and direction throughout my era as a graduate student at the University of Texas at El Paso. I also received a great deal of support and encouragement from colleagues, friends and family from both at home in Saint Lucia, and here in the United States. I thank them all. To my parents; Michael Charles and Cynthia Ryfer, my uncle and aunt Jerimiah and Patricia Paul, and also my sisters; Michelle Petra Charles, Shirnel Chrystal Charles, Betty Mariah Ryfer and Nikita Nicole Ryfer, I say special thanks, and I would be forever grateful for your perpetual love, support and strength. Last but by no means least; I would like to thank the almighty God with whom all things are possible. Undoubtedly, this would not have been possible without you.

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. This is the first empirical attempt to estimate ECMs for tourism demand in the Caribbean region. 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 assess model reliability, the 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 of the benchmarks.

Keywords: tourism, error correction models, forecasts, random walk, random walk with drift, Bahamas

Table of Contents

List of Tables

Chapter 1: 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 2007 World Travel and Tourism Council (WTTC) Travel and Tourism Economic Research on the Caribbean ranked the Caribbean the first in the world (out of 13 regions) for relative contribution of travel and tourism to the national economy (Caribbean region: Review of Economic Growth and Development, Investigation No.332-496). 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 in proper planning for the tourism sectors on these islands/economies.

The tourism literature recognizes 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, infrastructure, supplies and so forth. Governments are interested in tourist arrivals for national budgeting purposes, as they generate a large percentage of revenues from the tourism sector. The latter include room, sales, departure, and passenger ticket taxes. Accurate forecasts of tourist arrivals, are, therefore, critical 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 explaining tourism demand, but do not 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); Box-Jenkins autoregressive integrated moving average (ARIMA) models; and autoregressive (AR) models (Dharmaratne, 1995; Dalrymple and Greenidge, 1999). Two recent studies employ error correction models (ECMs) to analyze tourism demand in Tunisia (Ouerfelli, 2008) and the United Kingdom (Song, Witt and Jensen, 2003). None of the Caribbean studies to date, however, have tried to utilize ECMs for forecasting tourist arrivals.

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, et al, 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 will then be examined.

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.

Chapter 2: Literature Review

This section summarizes contributions to the literature on forecasting international tourism demand. Prior studies suggest a large number of possible approaches to estimating structural demand models. Sculmeister (1980) identifies exogenous variables such as disposable income, and relative prices between destinations as being essential in explaining tourism demand. Frequently, variables such as price, income, and consumption are lagged in these models (Morley, 1991; Witt and Martin, 1987). These lagged dependent variables are included as independent variables to reflect the nature of the tourism industry being one that is heavily influenced by individuals' habits and word of mouth advertising of destinations.

The high volatility of international tourism poses serious obstacles 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 help identify the variables that should be retained for simulation.

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 seems to provide excellent forecasts in the short term (1-2 years), with almost a perfect fit with the actual data. However, as the number of forecasting years increase the forecasts seem to deviate considerably from the actual data, with larger standard errors. In general this approach is able to generate good forecasts in the short term (1-2years).

Dalrymple and Greenidge (1999) argue that quarterly data are more useful in policy settings. Univariate ARIMA models are used to generate forecasts of quarterly arrivals to Barbados. 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. Greenidge estimates Basic Structural models (BSMs) which exclude the 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 overall the BSM produce better in-sample and 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; Error Correction Models (ECMs), univariate ARIMA models, BSMs, and regression based time series models. The relative performance of each model is found to vary from country to country. It is recognized that the ARIMA models produce more reliable short term forecasts, while the ECMs are more reliable for long term forecasts. The main conclusion from this study is that the error correction models perform poorly in comparison with the time series models.

Kim and Song (1998) use co-integration 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. Forecasting results indicate that the best models tend to be ECM or univariate ARIMA models, depending on the tourist generating market. For the United States (US) and United Kingdom (UK) markets, the

ECM is the most accurate. For the German and Japanese markets, the ECM is outperformed by the ARIMA methodology.

Song, Romilly and Liu (2000) use a general to specific approach to construct United Kingdom (UK) demand for outbound tourism models to twelve destinations. Ex posts 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) model, and a Vector Auto Regression (VAR) model. Results from this study suggest that the ECM model provides the best forecasting performance in comparison 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 and forecasting purposes.

A study carried out by Kulendran and Witt (2001) investigates the performance of ECMs relative to other time series models such as univariate ARIMA models and BSMs. In a large majority of cases the ECM outperforms these other models. Overall the empirical results show that the adoption of cointegration/ECM methods result in more accurate tourism demand forecasts than those generated by univariate time series models and other least squares regression models.

Song and Witt (2003) also employ a general to specific approach using data on inbound tourism to South Korea from its four major generating countries. Their study finds that the ECM has some advantages over other econometric models, and they confirm the applicability of this model to avoiding the problem of spurious regression. In general, the ex-post forecasting performances of the ECMs on average produce accurate results.

Ouerfelli (2008) uses co-integration 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 from one country to another. Findings from this study also show multiple statistically significant long-run relationships. Empirical results indicate that ECMs produce reliable accurate 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.

Chapter 3: Theoretical Model

Error correction models (ECMs) are estimated in order to capture both the 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, the business cycles of these countries 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:

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

$$\text{and } a_1 > 0,\ a_2 < 0.$$

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*.

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:

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

$$\text{and } b_1 > 0,\ b_2 < 0,\ b_3 < 0.$$

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 and long-run forces through the error correction term U_{t-1}, which measures the equilibrium error from the previous period. 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. Writing Equation (1) at time t-1 and solving for U_{t-1} yields the following result:

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

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

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

If there is a long-run equilibrium relationship between tourist arrivals and the explanatory variables in Equation (1), then the 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:

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

where y_t is the relevant time series variable, T is a linear deterministic trend, and ν_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.

Co-integration requires that all variables in the long-run co-integration equation be integrated of order 1 or *I* (1). Engle and Granger (1987) demonstrate that if co-integration is found among a set of variables in Equation (1) then the co-integration 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 as follows:

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

and $c_1>0$, $c_2<0$, $c_3<0$, $c_4>0$, $c_5<0$.

Equation (6) includes the effects of both short-run and long-run forces on changes in tourist arrivals. Changes in tourist 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. The error correction models will be used to produce out-of-sample simulations. To determine model reliability the accuracy of each model's forecasts will be compared to a random walk and a random walk with drift benchmark. To date, none of the previous studies on tourism demand have used these benchmarks to gauge model reliability.

Chapter 4: Data and Empirical Results

In this study annual data from 1977 – 2007 for tourist arrivals to the Bahamas is used as the dependent variable. This period and frequency is chosen because it provides the most consistent data set available. Since 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 these tourist arrivals to the Bahamas are expected to stay more than twenty four hours, another appropriate measure of tourist arrivals used is stop-over visitors (tourists staying 24 hours or more) for the major tourist generating countries (United States, Canada and Europe).

If (mainly) holiday demand or visits to friends and relatives are under consideration, then the appropriate form of the income variable is personal disposable income or private consumption, where personal disposable income is defined as the amount of current income that individuals have available for spending and saving. That means that it is personal income minus personal income taxes and national insurance contributions (Lipsey and Chrystal, 2004). However, if attention focuses on business visits (or they form an important part of the total), then a more general income variable such as national income or gross domestic product should be used (Song and Witt, 2000). This study therefore utilizes personal disposable income (PDI) or gross domestic product (GDP) as explanatory variables. Given that tourism demand to the Bahamas is largely for holiday purposes, personal disposable income is clearly the most appropriate measure of income. Unfortunately, due to data unavailability, GDP has to be used to measure income in all cases other than the US. 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.

Table 1: Variable Definitions and Data Sources.

Variable	Definition and Source
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*.
All tourist arrival data are obtained from the Bahamas Ministry of Tourism website *(www.tourismtoday.com)*.	
GDP	Natural logarithm of real gross domestic product at time *t* in country *i* in constant United States dollars, using 2000 as the base year. 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).*
PDI	Natural logarithm of real disposable personal income at time *t* in United States Dollars, using 2005 as the base year. Data are obtained from the United States Bureau of Economic Analysis website *(www.bea.gov).*
P	Natural logarithm of jet fuel prices at time *t*. Data are from the United States Energy Information Administration website *(www.eia.gov).*
v_t	Stochastic disturbances with zero means, constant variances, and serial independence.
Δ	First difference operator, also known as backshift or lag operator.
Sample	Annual frequency data, 1977-2007.
ws	Indicates parameter sign opposite of that hypothesized.

To begin empirical analysis, the order of integration of all variables is examined by conducting unit root tests on each of the variables. As shown in Table 2, the augmented Dickey and Fuller (ADF) unit root test is undertaken for 3 variables for each origin country. The results show that all variables become stationary after the first difference [i.e., all variables are *I* (1) variables]. Given this, standard regressions in level form may be spurious.

Having identified the common trends between the tourism demand variable and the income and price variables, co-integration tests are carried out to determine if a linear combination of these *I* (1) variables are stationary or *I* (0). The results for co-integration are shown in tables 3 and 4. The Johansen co-integration test (Table 3) indicates that there is one co-integrating vector at the 5% level of significance for stop-overs from the United States, Canada and Europe. The test also indicates one co-integrating vector for total tourist air arrivals, while there was no vector identified for total cruise ship arrivals.

The Engle and Granger procedure (Table 4) also indicates that the residuals from these co-integrating regressions are likely to be *I* (0)s, and this is confirmed by the augmented Dickey and Fuller (ADF) test statistic. These outcomes suggest that the variables in each of these long-run regressions are co-integrated. 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 co-integration test, an ECM is still estimated for comparison purposes. The lack of co-integration 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 had 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 2: Results of ADF Test for Unit Roots

		Levels						First Differences		
		Dependent Variable						Dependent Variable		
Variable	lnSUS	lnSCA	lnSEU	lnTA	lnCS	lnSUS	lnSCA	lnSEU	lnTA	lnCS
lnTA				-3.470 (1)					-4.679 (1)**	
lnCS					-1.068 (0)					-5.466 (0)**
lnSUS	-1.611 (1)					-5.631 (0)**				
lnSCA		-1.473 (6)					-4.940(0)**			
lnSEU			-0.788 (5)					-4.042 (4)*		
lnGDP		-2.590(1)	-2.610 (1)	-3.805 (1)	-3.805 (1)		-3.750 (1)*	-2.493 (0)*	-4.120 (1)*	-4.120 (1)*
lnPDI	-2.028 (0)					-5.429 (0)**				
lnp	-0.531 (2)	-0.439 (2)	-2.023 (0)	-0.531 (2)	-0.531 (2)	-4.376 (1)*	-4.454 (1)*	-4.096 (1)*	-4.376 (1)*	-4.376 (1)*

( ) the numbers in parentheses are the number of lags used for the test determined by AIC.
**Denotes significance at the 5% level.*
***Denotes significance at the 1% level.*

Table 3 : Johansen test for co-integration

Dependent Variable	Hypothesized Number of Co-integrating Vectors	Trace Statistic	Max-Eigen Statistic
lnSUS	None	36.012^{A}	18.773
	At most 1	17.240	11.070
	At most 2	6.171	6.171
lnSCA	None	43.711^{A}	37.971^{A}
	At most 1	5.740	5.682
	At most 2	0.058	0.058
lnSEU	None	30.307^{A}	17.113
	At most 1	13.195	11.864
	At most 2	1.331	1.331
lnTA	None	34.798^{A}	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

^{A}*Denotes significance at the 5% level.*

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 with the magnitude ranging from 0.036 to 1.751. In the case of Europe and total tourist air arrivals, the income coefficient is greater than one in the long-run, while in the short-run the magnitude of this coefficient lies closer to zero.

Table 4 : Co-integrating equations

Estimated Equation				ADF test Statistic	No. of lags
$\ln SUS =$	$0.737 \ln PDI_{1t}$	$- \; 0.101 \ln P_t$	$- \; 7.065$		
	(7.072)^B	*(-1.860)*^{C}	*(-6.562)*^{B}	(-3.639)^B	0
$\ln SCA =$	$^{WS}- 0.970 \ln GDP_{2t}$	$- \; {}^{WS}0.028 \ln P_t$	$+ \; 7.143$		
	(-5.092)^B	(0.359)	(3.750)^B	(-2.933)^B	1
$\ln SEU =$	$1.123 \ln GDP_{3t}$	$- \; 0.033 \ln P_t$	$- \; 23.553$		
	(2.489)^A	*(-2.077)*^{A}	(-2.748)^B	(-2.636)^B	1
$\ln TA =$	$0.285 \ln GDP_{4t}$	$- \; 0.001 \ln P_t$	$- \; 8.203$		
	(5.223)^B	(-0.272)	(-5.009)^B	(-2.052)^A	0
$\ln CS =$	$2.053 \ln GDP_{5t}$	$- \; 0.208 \ln P_t$	$- \; 60.228$		
	(12.582)^B	*(-1.830)*^{C}	*(-12.271)*^{B}	(-2.612)^A	0

^{A}*Denotes significance at the 5% level.*
^{B}*Denotes significance at the 1% level.*
^{C}*Denotes significance at the 10% level.*

In the equation for Canada the income elasticity is less than one in both periods, although one of the estimated parameters exhibits a negative sign. Income is not significant in the equations for stop-overs from the United States and tourists who travel to the Bahamas by sea. These results indicate that while income generally affects the number of tourist air arrivals it is not significant in determining the number of cruise ship arrivals to the Bahamas. This result is interesting and warrants further research.

Table 5 : Estimated error correction models

Estimated Equation

$$\Delta lnSUS = -\,0.801 + 0.420\Delta lnPDI_{1t} - 0.022\Delta lnP_t - 0.322lnSUS_{1t-1} + 0.094lnPDI_{1t-1} - 0.043lnP_{t-1} + v_t$$

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

R-square	*0.410*	*F-Statistic*	*3.343*
R- square (adj)	*0.288*	*S.E.*	*0.060*
Log likelihood	*45.120*	DW	1.766

$$\Delta lnSCA = 3.471 + 0.603\Delta lnGDP_{2t} + {}^{WS}0.036\Delta lnP_t - 0.423lnSCA_{2t-1} - {}^{WS}0.448lnGDP_{2t-1} - 0.054lnP_{t-1} + v_t$$

(1.997)^A *(1.742)*^{A} *(2.820)*^{B} *(-3.400)*^{B} *(-2.290)*^{A} *(-1.085)*

R-square	*0.478*	*F-Statistic*	*4.210*
R-square (adj)	*0.364*	*S.E.*	*0.084*
Log likelihood	*34.058*	DW	*1.823*

$$\Delta lnSEU = -\,7.147 + 1.751\Delta lnGDP_{3t} - 0.116\Delta lnP_t - 0.293lnSCA_{3t-1} + 0.360lnGDP_{3t-1} - 0.299lnP_{t-1} + v_t$$

(-1.380)^C *(1.991)*^{C} *(-0.732)* *(-3.052)*^{B} *(1.338)*^{B} *(-3.278)*^{B}

R-square	*0.414*	*F-Statistic*	*3.390*
R-square (adj)	*0.292*	*S.E.*	*0.150*
Log likelihood	*17.638*	DW	*1.125*

$$\Delta lnTA = -\,1.529 + 1.1891\Delta lnGDP_{4t} + {}^{WS}0.062\Delta lnP_t - 0.328lnTA_{4t-1} + 0.054lnGDP_{4t-1} - 0.008lnP_{t-1} + v_t$$

(-0.948) (2.328)^A *(1.299) (-2.703)*^{B} *(0.992) (-0.307)*

R-square	*0.465*	*F-Statistic*	*4.354*
R-square (adj)	*0.358*	*S.E.*	*0.051*
Log likelihood	*51.611*	DW	*2.158*

$$\Delta lnCS = -\,3.742 - {}^{WS}1.038\Delta lnGDP_{5t} - 0.021\Delta lnP_t - 0.113lnCS_{5t-1} + 0.130lnGDP_{5t-1} + {}^{WS}0.009lnP_{t-1} + v_t$$

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

R-square	*0.189*	*F-Statistic*	*1.342*
R-square (adj)	*0.049*	*S.E.*	*0.108*
Log likelihood	*31.445*	DW	*2.121*

^{A}*Denotes significance at the 5% level;* ^{B}*Denotes significance at 1% level;* ^{C}*Denotes significance at 10%level.*

The own price elasticity is found to be significant in the equation for Canada although the estimated parameter sign is incorrect. This positive own price elasticity in this model is an interesting result and merits further investigation. The own price elasticity is found to be significant with the expected sign only in the model for Europe, with a magnitude that is significantly less than one. This suggests that although the cost of travel has some minor effects on European residents when they make decisions to travel to the Bahamas, it has no role to play in the decision making processes of tourists from the United States, and cruise ship visitors in general. The results from these ECMs show that the most important determinant of tourism demand to the Bahamas is the lagged dependent variable. This suggests that habit persistence and/or word of mouth recommendation are the major driving forces for holiday tourism demand (Witt 1980).

In order to determine the reliability of these models in producing out-of-sample simulations, the ECMs are re-estimated 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. See the appendix for specific information on the calculation of those measures.

Table 6 summarizes predictive accuracy results for each of the tourism demand series. The ECMs for stop-overs from the United States and Canada seem to offer the best forecasting performance in comparison to the other error correction models. For these models the distributions of the inequality coefficient second moment proportions are much closer to the ideal 0, 0, 1 distribution than any of the other ECMs.

Table 6: Tourism demand series predictive accuracy

Series	RMSE	Theil U	Bias Proportion	Variance Proportion	Covariance Proportion
Stop-overs United States					
LTF	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
Stop-overs Canada					
LTF	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
Stop-overs Europe					
LTF	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					
LTF	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					
LTF	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 used to calculate each inequality coefficient. Out-of-sample simulation periods: 2000-2001; 2001-2002; 2002-2003; 2003-2004; 2004-2005; 2005-2006; 2006-2007; 2007. Random walk forecast is calculated as last available historical observation. Random walk with drift forecast calculated as last available historical observation plus drift.

Also, with the exception of the ECM for cruise ship arrivals, the variance proportions of the majority of these error correction models respective Theil inequality coefficients are close to zero. This implies that these models do a fairly good job in terms of simulating the variability of tourism demand.

However, examination of the second moment prediction error proportions suggests that the ECMs for Europe and tourist air arrivals tend to be biased. The magnitude of these bias proportions in the ECMs for Europe and tourist air arrivals, and also the root mean square error and variance proportion in the model for cruise ship arrivals, may be as a result of possible misspecification, possibly caused by either the omission of important explanatory variables or by inappropriate use of the proxy variables. Despite these shortcomings, generally the error correction approach offers a reliable means of forecasting tourism demand in the Bahamas, and it is the expectation that the ECMs can be improved with further research and model re-specification.

The error correction or linear transfer function (LTF) models errors are also compared against prediction errors of random walk (RW) and random walk with drift (RWD) benchmarks. Forecasts results for the RW and RWD methodology are also reported in Table 6. Similar to the LTFs, the second moments of the prediction errors are also used to calculate Theil inequality coefficients. Table 6 shows that U-statistics close to zero are obtained for both the random walk and random walk with drift extrapolations, but these results are similar to their LTF counterparts. Outcomes in Table 6 also indicate that though the LTFs for the United States and Canada exhibit the best prediction performance of any of the error correction models, the predictive performance of these models is not superior to their random walk and random walk with drift counterparts.

Also, for the three other series modeled (Europe, tourist air and cruise ship arrivals) both the RW and RWD predictions are both fairly more precise relative to the LTF predictions.

In overall terms, the random walk forecasts point to relatively better simulation accuracy than both LTF and RWD extrapolations. 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 are therefore required in order to confirm these outcomes.

Chapter 5: Conclusion

In this study, error correction models are estimated and used to forecast inbound tourism demand in the Bahamas from its major tourist generating countries. The models are estimated for total air and cruise ship arrivals to the Bahamas. Prior to estimating the models, ADF unit root tests are carried out, which confirm that all variables in this study are I (1). The Johansen (1988) test for co-integration and the Engle and Granger (1987) procedure suggest that co-integration/long-run equilibrium relationships exist for 4 of the 5 models identified in this study.

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 will drive the numbers of tourist air arrivals to the Bahamas, but further research is needed to determine what factors influence cruise ship visitors.

The fuel price elasticity, while significant in the model for Europe with only marginal effects, is generally insignificant in the decision making process of tourists who travel to the Bahamas by both air and sea. This suggests that tourist arrivals are affected by business cycles rather than the cost of travel to the Bahamas. 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 holiday experience. It is also imperative that economic policy makers allocate sufficient resources to enhance the image of the Bahamas through marketing and promotional campaigns.

Out-of-sample simulations generated from the error correction models are compared with random walk and random walk with drift benchmarks. The empirical 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 period. Also, because bias is a problem with the error correction model forecasts for Europe and tourist air arrivals, care should be exercised with respect to using these out-of-sample forecasts. Further research is warranted to improve those model specifications.

Generally, the random walk models have the best overall forecasting performance. Replication of the methodologies used in this study for other tourist destinations in the Caribbean will help determine if these results are specific to the Bahamas or more general in nature. In particular, none of the other destination markets in this region are as close as the Bahamas to the United States. It would not be surprising, therefore, to discover that transport costs influence tourism flows to less accessible locations.

References

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

B.H. Archer, 1994, Demand forecasting and estimation, *Travel, Tourism and Hospitality Research: A Handbook for Managers and Researchers*. J.R.B. Ritchie and C.R. Goeldner (ed.), pp. 105-114, New York: John Wiley.

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 University Press, Oxford.

J. Blackwell, 1970, Tourist traffic and the demand for accommodation: some projections, *Economic and Social Review* 1(3), 323-343.

K. Carey, 1991, Estimation of the Caribbean tourism demand: Issues in measurement and methodology, *Atlantic Economic Journal* 19(3), 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, *In: Management Science Applications to Leisure Time Operations.* S.P. Ladany (ed.), pp. 134-145, North Holland, New York.

K. Dalrymple and K. Greenidge, 1999, Forecasting arrivals to Barbados, *Annals of Tourism Research* 26(1), 188-190.

G.S. Dharmaratne, 1995, Forecasting tourist arrivals in Barbados. *Annals of Tourism Research* 22(4), 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 University Press, Oxford.

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* 121-130.

K. Greenidge, 2000, Forecasting tourism demand, an STM approach, *Annals of Tourism Research* 28(1), 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(2), 291-311.

C. Lim and M. McAleer, 2001, Forecasting tourist arrivals, *Annals of Tourism Research* 28(4), 965-977.

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

J. Mackinnon, 1991, Critical values for co-integration tests, in *Long-run Economic Relationships-Readings in Co-integration* (Eds) R. Engle and C.W.J. Granger, Oxford University Press, Oxford.

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

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, *Journal of the Royal Statistical Society, Series A (General) 143(1), 85-86.*

H. Song, P. Romilly and X. Liu, 1997, An empirical study of outbound tourism demand in the UK, *The 17th International Symposium on Forecasting*, Bridgetown, Barbados.

H. Song, P. Romilly and X. Liu, 2000, An empirical study of outbound tourism demand in the UK. *Applied Economics* 32, 611-624.

H. Song and S.F. Witt, 2000, *Tourism demand modeling and forecasting, modern econometric approaches*, Amsterdam, The Netherlands: Pergamon.

H. Song and S.F. Witt, 2003, Tourism forecasting: The general to specific approach, *Journal of Travel Research 42*, 65-74.

H. Song, S.F. Witt and T.C. Jensen, 2003, Tourism forecasting: accuracy of alternative econometric models, *International Journal of Forecasting* 19, 123-141.

H. 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(4), 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 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, investigation no. 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.

Appendix A

Equation A (1) 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.

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

Equation A (2) shows how the Theil inequality coefficient U is computed.

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

Equations A(3), A(4), and A(5) respectively, show the decomposition of the Theil inequality coefficient. These equations show the computation of the bias, variance and covariance proportions respectively.

$$U^M = \frac{(\bar{Y}_t^s - \bar{Y}_t^a)^2}{(1/T)\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2} \qquad \text{A (3)}$$

$$U^s = \frac{(\sigma_s - \sigma_a)^2}{(1/T)\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2} \qquad \text{A (4)}$$

$$U^C = \frac{2(1-\rho)\sigma_s\,\sigma_a}{(1/T)\sum_{t=1}^{T}(Y_t^s - Y_t^a)^2} \qquad \text{A (5)}$$

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

Appendix B

Table B1: Historical data for tourist air arrivals, cruise ship arrivals and stop-overs.

Year	Tourist Air Arrivals	Cruise Ship Arrivals	Stop-overs United States	Stop-overs Canada	Stop-overs Europe
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:
1. Tourist arrivals and stop-over data are reported in units.

Table B2: Historical income data for United States, Canada, and Europe and jet fuel prices

Year	Personal disposable income United States	Gross domestic product - Canada	Gross domestic product - Europe	Jet Fuel Prices
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	

Notes:
1. Income and price data are reported in millions of nominal dollars.

Curriculum Vita

Jacky Charles was born in Castries, Saint Lucia, July 26, 1980 to Michael Charles and Cynthia Ryfer. The second of five siblings, she graduated from the Sir Arthur Lewis Community College in Castries, Saint Lucia in May 1999. In August 2003, she went on to pursue a Bachelor of Science degree in Economics and Management with the University of the West Indies, St Augustine Campus, in Trinidad and Tobago, where she obtained her Bachelor's degree with honors in July 2006. After obtaining her degree she worked as a Budget Analyst and Economist with the Ministry of Finance, Economic Affairs, Economic Planning and National Development in Castries, Saint Lucia. She also worked as a part time Lecturer with the National Research and Development Foundation, also in Saint Lucia. She began her pursuit of a graduate degree in Economics with the University of Texas at El Paso in El Paso, Texas in August 2008. Whilst attending graduate school she worked as an Economics and Mathematics tutor with the Miner Athlete Academic Center in El Paso, Texas and later on as a Graduate Assistant with the Financial Services Department at the University. As of fall 2008, she has been recognized as a James Foundation Scholar for her academic record at the University.

Permanent address: Ministry of Finance, Economic Affairs, Economic Planning and National Development
Castries, Saint Lucia

This thesis was typed by Jacky S. Charles.