INTERNATIONAL JOURNAL OF TOURISM RESEARCH
Int. J. Tourism Res. **5**, 369–380 (2003)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jtr.440

Tourism Demand Modelling: Some Issues Regarding Unit Roots, Co-integration and Diagnostic Tests†

Paresh Kumar Narayan*
Department of Economics, Box 94, Monash University, 3800 Victoria, Australia

ABSTRACT

This paper investigates the all important issue of diagnostic tests, including unit roots and cointegration, in the tourism demand modelling literature. The origins of this study lie in the apparent lack in the tourism economics literature of detail concerning the diagnostic test aspect. Study of this deficiency has suggested that previous literature on tourism demand modelling may be divided into two categories: the pre-1995 and post-1995 studies. It was found that the pre-1995 and some post-1995 studies have ignored unit root tests and co-integration and, hence, are vulnerable to the so-called 'spurious regression' problem. In highlighting the key diagnostic tests reported by post-1995 studies, this paper contends that there is no need to report the autoregressive conditional heteroskedasticity (ARCH) test, which is applicable only to financial market analysis where the dependent variable is return on an asset. More generally, heteroskedasticity is not seen as a problem in time-series data. However, the reporting of a greater than necessary range of diagnostic tests — some of which do not have any theoretical justification with regard to tourism demand analysis — does not diminish the precision of the results or the model. This paper should appeal to scholars involved in tourism demand modelling. Copyright © 2003 John Wiley & Sons, Ltd.

*Correspondence to: P. K. Narayan, Department of Economics, Box 94, Monash University, 3800 Victoria, Australia.
E-mail: Paresh.Narayan@BusEco.monash.edu.au
†This paper is a chapter from the author's PhD thesis.

Received 1 May 2002; revised 29 April 2003; accepted 6 May 2003

Keywords: tourism demand; modelling; diagnostic tests; co-integration.

INTRODUCTION

There are a number of excellent reviews of the literature on tourism demand modelling (Crouch, 1994a,b; Witt and Witt, 1995; Lim, 1997a,b). Careful examination shows that the literature is essentially divided into two categories. One that has used traditional econometric techniques in modelling tourism demand can be broadly categorised as the pre-1995 literature, whereas the other, which has used modern econometric techniques, can be categorised as the post-1995 literature. The specific concern of this paper is with the unit roots, co-integration and diagnostic tests used (or not used) by these two categories of studies. The justification for this is embedded in the fact that these tests play a crucial part in determining the accuracy of models and results derived from it (see McAleer, 1994, for a survey).

This paper elucidates the issue of unit roots, co-integration and the main diagnostic tests and examines their application in the empiri-

Table 1. Regression results of the determinants of international tourist expenditures in South Korea. (Source: Lee *et al.*, 1996, p. 537)

Origin country	Real per capita income	Relative prices	$\bar{R}^2$	DW
Japan	11.140 (10.922)	−3.290 (−7.945)	0.9093	1.6516
USA	6.506 (4.194)		0.8329	1.2689
Hong Kong	3.507 (6.528)		0.9102	1.3299
UK	6.592 (12.839)		0.9301	1.7991
West Germany	6.954 (6.148)		0.8857	1.8560
Philippines	14.317 (4.766)	−7.014 (−4.495)	0.6365	1.5987
Canada	13.946 (2.927)		0.6785	1.6460

cal tourism demand literature. It simultaneously remarks on the likely implications of these tests on the results obtained from tourism demand models.

The rest of the paper is organised as follows. In the next section, the issue of unit root tests is analysed, its importance is highlighted and its use in the tourism demand literature is investigated. A brief outline of co-integration and error correction models is provided next, and tourism demand studies that have applied co-integration are highlighted. In the subsequent sections the various diagnostic tests that examine the auxiliary assumptions of the models are briefly elucidated and their application in tourism demand studies are examined. The final section concludes.

UNIT ROOT TESTS

The determination of good models for prediction is an important element in much practical econometric research. As economic time-series often display non-stationary characteristics, an aspect of model determination that must be addressed in practical work is how to model the non-stationarity in the data (Philips, 1996, p. 763). The issue of unit roots was introduced by Nelson and Plosser (1982) in their path-breaking paper in which they argued that most macroeconomic series have unit roots. A series that has unit roots is also known as a non-stationary time series (see Appendix). Yule (1926) suggested that regressions based on non-stationary series are known as 'nonsense' regression. Granger and Newbold (1974) termed this problem as 'spurious' regressions. They were critical of the specifications of regression equations in terms of the levels of economic time-series. Granger and Newbold contend that the levels of many economic time-series are non-stationary; hence, regression equations based on levels of time-series often produce high R^2 and display highly autocorrelated residuals (low Durbin–Watson (DW) statistic); hence, the usual *t*- and *F*-tests on the regression parameters may be very misleading (see Phillips, 1986).

The unit root test is imperative and a fundamental first step in econometric modelling. However, the pre-1995 studies on tourism demand modelling have ignored unit root tests, and have failed to differentiate data and account for error correction terms when it is $I(1)$. (Post-1995 tourism demand studies have found tourism demand variables to be $I(1)$, therefore the variables need to be differentiated.) It follows, then, that upon ignoring the requirement for stationarity, the parameter values are unreliable and, in particular, the standard *t*- and *F*-tests give misleading results. Estimation using standard regression techniques, which ignore the non-stationarity of the data, is flawed (Hausman, 1978; Phillips, 1986).

Tables 1 and 2 below show some results extracted from tourism demand studies that have ignored unit root tests. Apart from the relatively large parameter estimates and *t*-statistics (particularly with regard to the income variable) the results also contain a high adjusted R^2 and untoward DW statistics — symptomatic of the 'nonsense' or 'spurious' regression phenomenon.

Moreover, the results obtained from data-series in the absence of the unit root test cannot be construed as long-run parameter estimates. This is because if only two variables are used,

Table 2. Regression results of the determinants of visitor arrivals to Singapore. (Source: Gunadhi & Boey, 1986, p. 245)

Origin	Income	Price	$\bar{R}^2$	DW
World	4.694 (9.360)	−0.9176 (−3.92)	0.98	1.35
Australia	5.454 (16.37)		0.98	1.88
UK	7.304 (12.45)		0.91	1.91
Indonesia	0.8070 (3.77)	−1.127 (−22.60)	0.99	1.77

both need to be integrated of the same order. If the number of variables is greater than two, the order of integration of the dependent variable cannot be higher than the order of integration of any of the explanatory variables (Charemza and Deadman, 1997). (The order of integration is not a concern if the estimation of equations utilises the autoregressive distributed lag (ARDL) method of co-integration recently developed by Pesaran and Shin (1995). The ARDL approach to co-integration does not require knowledge of whether the variables under consideration are $I(1)$ or $I(0)$. Put differently, the method avoids the requirements of pre-testing of the order of integration, which is necessary in other co-integration methods (see Pesaran *et al.*, 1996). For an application of the ARDL model to tourism demand, see Narayan (2004).) Once the integration properties of the variables are established then a test for co-integagaration (which is discussed in the next section) is essential for only the presence of co-integration among variables allows the estimation of the long-run parameters.

CO-INTEGRATION AND ERROR CORRECTION MODELS

Error correction models (ECM) are associated with Sargan (1964), Hendry and Anderson (1977) and Davidson *et al.* (1978). The essence of an ECM is to capture adjustments in a dependent variable that depend on the extent to which an explanatory variable deviates from an equilibrium relationship with the dependent variable (Banerjee *et al.*, 1993). The ECMs provide a way of combining both levels and difference of variables, hence capturing the dynamics of both short-run (changes) and long-run (levels) adjustments.

The notion of co-integration was first introduced by Granger (1981) and Granger and Weiss (1983). It was further extended and formalised by Engle and Granger (1987). Co-integration describes the existence of an equilibrium or stationary relationship among two or more time-series, each of which is individually non-stationary. The advantage of the co-integration approach is that it allows integration of the long-run and short-run relationships between variables within a unified framework. In addition, the presence of co-integration rules out the spurious regression problem. Since the seminal work of Engle and Granger (1987), research on co-integration techniques has multiplied, with a focus on determining the number of linearly independent co-integration vectors, or the co-integrating rank, in a general vector autoregressive process. The research on cointegration essentially has taken two routes: single equation-based tests and systems of equation-based tests. The former follows the work of Engle and Granger (1987), Phillips and Ouliaris (1990), and Hansen (1992) and Park (1990), whereas the latter has roots in the work of Johansen (1988, 1991), Johansen and Juselius (1990) and Stock and Watson (1988), amongst others.

Given that co-integration gained popularity only in the late 1980s, it is not surprising to note that none of the tourism demand studies applied co-integration until the mid-1990s. However, rather surprisingly, three of the post-1995 studies have ignored the issue of co-integration analysis in modelling tourism demand (see Table 4). Most studies have applied the Johansen (1988) and Johansen and Juselius (1990) approach to co-integration (Kulendran, 1996; Seddighi and Shearing, 1997; Lathiras and Siriopoulos, 1998; Kulendran and Wilson, 2000; Lim and McAleer, 2001; Daniel and Ramos, 2002; Narayan, 2002, 2003), two have applied the Engle and Granger (1997) approach to co-integration (Kim and Song, 1998; Song *et al.*, 2000), whereas Narayan (2004) has applied the bounds test approach to co-integration within an autoregressive distributed lag framework developed by Pesaran and Shin (1995) and Pesaran *et al.* (1996).

It follows, then, that the ignorance of unit root tests and the absence of co-integration makes it impossible to place much reliance on the majority of the pre-1995 empirical studies on tourism demand modelling, particularly with respect to policy decisions. Thus, one needs to apply caution in interpreting these results. However, it is clear that the parameter estimates of the price and income variables in most of the pre-1995 studies have appeared with the correct sign — consistent with economic theory; but the magnitude of the parameter estimates and their level of significance are indeed doubtful.

DIAGNOSTIC TESTS

Although the proliferation of applied empirical work on the determinants of tourism demand is welcomed, a common deficiency in much of the work is the absence of statistical diagnostic testing. In the common stochastic specification of econometric models, the error terms are assumed to be normally distributed with mean zero, to have constant variance and to be serially correlated. Diagnostic tests are used to test these auxiliary assumptions of the models. A valid statistical model is one with underlying model assumptions that are not violated, such that the actual distributions of common test statistics will not differ from those expected. If this does not hold, then inferences will be invalid and estimated elasticities and tests of economic theory will have no statistical validity. In the light of this, the econometric literature has attained a general consensus: that the final econometric model needs to undergo rigorous statistical checking in order to ascertain its statistical acceptability.

There are numerous forms and types of diagnostic tests, each performing a different function. In what follows, these tests are reviewed, their importance in applied empirical work on tourism demand is highlighted and tourism demand studies that have reported or not reported diagnostic tests are identified.

THE TOURISM ECONOMICS LITERATURE: WHAT IS KNOWN REGARDING DISGNOSTIC TESTS?

Most diagnostic tests in econometric software programs have been available only since the 1980s (Lim, 1997b). Hence, it is not surprising that studies before the 1980s have not reported any diagnostic tests, apart from the Durbin and Watson (1950) test which checks for first-order autocorrelation. The DW test is defined as:

$$DW = \sum_{t=2}^{n}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2 \Big/ \sum_{t=1}^{n}\hat{\varepsilon}_t^2$$

where $\hat{\varepsilon}_t$ is the residual from the estimated regression equation. If there is no autocorrelation, the value of the DW statistics should be approximately 2.

Between 1961 and 79, for instance, only slightly over 50% of the studies reported DW test. Most surprisingly, between 1990–2002, despite the well documented state of diagnostic tests, 31% of the studies failed to report any of these tests (Table 3).

TESTING FOR AUTOCORRELATION

It is generally accepted that most economic data consist of time-series and there are often correlations in model error terms corresponding to successive time periods. This is the problem of autocorrelation and the reason might be omitted variables, mis-specification of the dynamic process, etc. (Autocorrelation may be caused by the omission of important variables from the regression model when they are correlated with the dependent variable, the inclusion of variables in an incorrect functional form or inadequately modelled seasonality. Errors in the independent variables, ignoring the simultaneous equation system, inaccurate modelling of dynamic systems, changing regression parameters, heteroskedasticity, temporal aggregation or other transformations of

Table 3. Diagnostic tests in studies on tourism demand modelling, 1961–2002. (Source: *Lim, 1997a, Table 6, p. 76; **author's own research)

Years	Only DW	Others[a]	None	Total
1961–79*	16	0	14	30
1980–89*	25	9	16	50
1990–94*	6	4	10	20
1995–2002**	2	14	2	18
Total	49	27	42	118

[a]Others include: the Wallis test for fourth-order serial correlation in the error term; the Breusch–Pagan chi-square test for heteroskedasticity; Goldfeld–Quandt test for homoskedastic residuals; the Breusch–Godfrey (this is the LM test for autocorrelation); the Jarque–Bera normality test; the Engle test for autoregressive conditional heteroskedasticity; the Ramsey test for omitted-variables/functional form; the White test for heteroskedasticity; and the Chow predictive failure and breakpoint tests.

independent variables, such as their deflation or use of proportions, may also be causes (Praetz, 1987, p. 129).) Tests for serial correlation are important because tourism demand model equations are often used for forecasting. Serial correlation in residuals leads to misspecified dynamics, which produce inaccurate forecasts.

The DW statistic is the most commonly used to test for detecting the problem of autocorrelation in the regression residuals. In the empirical tourism demand literature, many studies have reported the DW test statistic. Lim (1997a, p. 77), in a survey of empirical studies published over the period 1961–94, found that 'where a diagnostic test has been reported, it has invariably been only the DW statistic'. The DW test, however, has some limitations; for instance, it cannot detect higher order autocorrelation, and it is biased towards 2 when a lagged dependent variable is used as an explanatory variable in the model. It is common in the tourism demand literature to see the lagged dependent variable (either measured in terms of tourism numbers or expenditure) used as an explanatory variable to capture the so-called 'word of mouth effect'.

To remedy these deficiencies, particularly when a lagged dependent variable is included in the model, the Lagrange multiplier (LM) test (also known as the Breusch–Godfrey test — after Godfrey, 1978) is used. The null hypothesis is that there is no autocorrelation; this is tested against the alternative that there is autocorrelation. (The calculation of the test is based on an auxiliary equation of the form,

$$\hat{\varepsilon}_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_k X_{kt} + \rho_1 \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + \ldots + \rho_\rho \hat{\varepsilon}_{t-\rho} + \mu_t$$

where X_{it} are explanatory variables, β_i are parameters and $\hat{\varepsilon}_{t-j}$ are the lagged residuals from the estimated regression model. Under the null hypothesis of no autocorrelation, H_0: $\rho_1 = \rho_2 = \ldots = \rho_p = 0$.) Hence, it is imperative to report the LM test statistic, which, although absent in the pre-1995 studies, also has been ignored by some post-1995 studies (Vogt and Wittayakorn, 1998; Vanegas and Croes, 2000).

TESTING FOR NORMALITY

The normality test is known as the Jarque and Bera (J–B) (1980) test. The null hypothesis of normally distributed residuals is tested against the alternative that the residuals are not normally distributed. The test statistic is computed as:

$$\xi_{\text{LM}} = N\left[\frac{1}{6}\left(\frac{1}{N}\sum_{i=1}^{N}\hat{\varepsilon}_t^3/\hat{\sigma}^3\right)^2 + \frac{1}{24}\left(\frac{1}{N}\sum_{i=1}^{N}\hat{\varepsilon}_i^4/\hat{\sigma}^4 - 3\right)^2\right]$$

where N is the number of observations. The above equation is a weighted average of the squared sample movements corresponding to skewness and excess kurtosis, respectively. Under the null hypothesis, it is asymptotically distributed as a chi-squared with two degrees of freedom; see Godfrey (1988, pp. 143–145) for more details.

If the regression residuals are not normally distributed, the t and F statistics are invalid, particularly in small samples. For this reason, the J–B test is essential. In the pre-1995 literature the J–B test is absent, however, many post-1995 studies also have ignored this important test (including Lee *et al.*, 1996; Kulendran and King, 1997; Vanegas and Croes, 2000; Lim and McAleer, 2001; see Table 4).

Table 4. Diagnostic tests reported in the post-1995 studies[a] on tourism demand modelling. (Source: author's own research)

Reference	DW	UR	JB	RESET	Chow 1	Chow 2	CUSUMSQ/CUSUM	LM	White	ARCH	Others	C
Akis (1998)	✓							✓				
Daniel and Ramos (2002)	✓	✓										
Icoz *et al*. (1998)	✓											
Kim and Song (1998)	✓	✓	✓	✓	✓			✓	✓	✓		
Kulendran (1996)	✓	✓	✓	✓	✓			✓		✓	✓	✓
Kulendran and King (1997)	✓	✓						✓				
Kulendran and Wilson (2000)	✓	✓										
Kulendran and Witt (2001)	✓	✓		✓	✓			✓			✓	
Lathiras and Siriopoulos (1998)	✓	✓	✓	✓	✓		✓	✓				
Lee *et al*. (1996)	✓											
Levantis and Gani (2000)	✓		✓	✓						✓		
Lim and McAleer (2001)		✓			✓	✓		✓			✓	
Narayan (2002a)	✓		✓	✓			✓	✓	✓		✓	
Narayan (2003)	✓		✓	✓			✓	✓	✓		✓	
Narayan (2004)	✓		✓	✓			✓	✓	✓		✓	
Payne and Mervar (2002)	✓		✓	✓			✓	✓	✓			
Qu and Lam (1997)	✓		✓						✓		✓	✓
Seddighi and Shearing (1997)	✓	✓		✓			✓					
Song *et al*. (2000)	✓	✓	✓	✓	✓			✓	✓	✓		
Vanegas and Croes (2000)												
Vogt and Wittayakorn (1998)		✓										✓

[a] All studies have used the standard R^2 and the F-test; DW is the Durbin–Watson test for first-order serial correlation; UR is the unit root test; JB is the Jarque–Bera test for normality; RESET is the Ramsey test for omitted variables/functional form; Chow 1 is the predictive failure test; Chow 2 is the breakpoint test; CUSUMSQ/CUSUM is the recursive least square procedure for structural instability; LM is the Breusch–Godfrey test for autocorrelation; White is a test for heteroskedasticity; ARCH is the Engle test for autoregressive conditional heteroscedasticity; 'Others' include the Wallis test to test for fourth-order serial correlation in the error term, Breusch–Pagan chi-square test for heteroskedasticity and the Goldfeld–Quandt test for homoskedastic residuals; C stands for those studies that did not use the co-integration technique.

TESTING FOR MIS-SPECIFICATION

The Ramsey (1969) regression specification error test (RESET) test is designed to test for model mis-specification owing to either the omission of important explanatory variables or incorrect choice of functional form. The null hypothesis is that the model is correctly specified, but there is no specific alternative hypothesis. The first stage of the system-wise RESET test is performed by calculating the least squares' predictions from the primary regression (Equation 1), $\hat{Y}_t = (X(X'X)^{-1}X')Y$. These predictions are then used in the following auxiliary regression, $Y_t = X_tB + \hat{Y}_t^2\Phi_1^* + \hat{Y}_t^3\Phi_2^* + \ldots + \hat{Y}_t^{K+1}\Phi_K^* + \delta_t$. The RESET test is now performed by testing the hypothesis that $\Phi_1^* = \ldots = \Phi_K^* = 0$ (for more details, see Godfrey, 1988, pp. 106–107).

This means that in the event the null hypothesis is rejected, it indicates that the model is incorrectly specified (Ramsey, 1974; and Ramsey and Schmidt, 1976). Although the RESET test is non-existent in the pre-1995 research on tourism demand modelling, some post-1995 studies also have not reported this test (including Lee *et al.*, 1996; Kulendran and King, 1997; Vogt and Wittayakorn, 1998; Vanegas and Croes, 2000; see Table 4). Those studies that have performed the RESET test indicated the acceptance of the null hypothesis, implying that the model is correctly specified (Kulendran, 1996; Lathiras and Siriopoulos, 1998; Kulendran and Witt, 2001).

TESTING FOR STRUCTURAL INSTABILITY

There is a general consensus within the econometrics literature that in order for an econometric model to produce accurate forecasts, the structure of the model should be constant over time. Put differently, the values of the parameters of the model should be the same for both the sample and the forecasting periods. There are essentially three methods of carrying out the structural stability test. The most commonly used is the Chow (1960) parameter constancy test (or breakpoint test)

$$F_{\text{CHOW}} = (\text{SSR}_0 - \text{SSR}_1 - \text{SSR}_2)/k/(\text{SSR}_1 + \text{SSR}_2)/(n_1 + n_2 - 2k)$$

where SSR_0 is the residual sum of squares for the whole sample period, SSR_1 and SSR_2 are the residual sums of squares for the two subsamples, and n_1 and n_2 are the number of observations in the first and second subsamples, respectively. If the calculated F statistic is larger than the critical value, the null hypothesis of parameter constancy between the two subsample periods is rejected (Chow, 1960).

This test examines whether there is a statistically significant difference between the ordinary least squares regression residuals from the two subsamples. The null hypothesis is parameter constancy between the two subsample periods. In performing this test, the breakpoint is assumed to be known.

Second, is the Chow predictive failure test, also originating from Chow's (1960) seminal contribution

$$F_{\text{CHOW}} = (\text{SSR}_0 - \text{SSR}_1)/n_2/\text{SSR}_1/(n_1 - k)$$

where SSR_0 is the residual sum of squares for the whole sample period, SSR_1 and SSR_2 are the residual sums of squares for the two subsamples, and n_1 and n_2 are the number of observations in the first and second subsamples, respectively. If the calculated F statistic is larger than the critical value, the null hypothesis of structural stability between the two subsamples is rejected. The key difference between this and the parameter constancy test is that the former does not involve estimating the regression for the second subsample (for details, see Song and Witt, 2000, p. 39).

Third is the recursive least squares procedure, which gained popularity owing to the apparent shortcoming of the Chow tests — for which one needs to know exactly the point at which the structural break takes place. This is unrealistic because the change in structure may evolve gradually over time and not rapidly. In such situations the Chow tests are not appropriate; hence, the recursive least square procedure is seen as a good alternative.

The recursive least square technique is used to examine the stability of the regression coefficients using the cumulative sum (CUSUM) and cumulative sum of squares (CUSUMSQ) of the recursive residual tests for structural stability (Brown *et al.*, 1975). The CUSUM test is useful for detecting systematic changes in the regression coefficients, whereas the CUSUMSQ test is useful in situations where the departures from the constancy of the regression coefficients is rather abrupt and sudden. The null hypothesis is that the coefficient vector is the same in every period, whereas the alternative is simply that it is not.

There are some post-1995 studies on tourism demand modelling (Kulendran, 1996; Kim and Song, 1998; Lathiras and Siriopoulos, 1998; Song *et al.*, 2000; Kulendran and Witt, 2001; Lim and McAleer, 2001) that test for structural stability, but none of the pre-1995 studies have reported the Chow tests. On the other hand, only a few studies (Seddighi and Shearing, 1997; Lathiras and Siriopoulos, 1998; Payne and Mervar, 2002) have reported the recursive least squares test for structural stability. Given the importance of structural stability tests, it is imperative that they be given priority if one is to construct accurate econometric models.

TESTING FOR HETEROSKEDASTICITY

Researchers have observed that heteroskedasticity is usually found in cross-sectional data and not in time-series data (Engle and Bollerslev, 1986; Gujarati, 1992). With respect to the autoregressive conditional heteroskedasticity (ARCH) test, Engle (2001, p. 158) noted: 'The ARCH models are used in financial applications where the dependent variable is the return on an asset or protfolio and the variance of the return represents the risk level of those returns'. (According to Engle (2001, p. 158), the goal of such models is to provide a volatility measure — like a standard deviation — that can be used in financial decisions concerning risk analysis, portfolio selection and derivative pricing.)

It should be noted that the bulk of the empirical research on tourism demand has involved times-series data where the dependent variable generally has been either tourist arrivals or tourist expenditure. (With the exception of Mak *et al.* (1977) and Yavas and Bilgin (1996), cross sectional data, and Romilly *et al.* (1998), Carey (1991) and Jud and Joseph (1974), panel data.) Against this background, there is little justification, if any, for the importance of ARCH tests in tourism demand modelling. In spite of poor theoretical justification for the use of ARCH, some of the post-1995 studies on tourism demand modelling have reported the test (Kulendran, 1996; Kim and Song, 1998; Song *et al.*, 2000). An examination of the studies that have reported the ARCH test reveals that except for Song *et al.* (2000) none of the other studies have found any problems with the ARCH test. Song *et al.* (2000), who estimated UK demand for tourism to 12 destinations found that two out of the 12 models had problems with the ARCH test. It is not clear whether the use of the Engle and Granger (1987) approach, as applied by Song *et al.* (2000), to co-integration is responsible for this untoward result. The limitations of the Engle and Granger (1987) approach in small sample sizes and particularly when more than one explanatory variable is used are well documented (see Pattichis, 1999).

There are two other types of tests for heteroskedasticity.

(1) One is the Goldfeld and Quandt (1965) test; the underlying null hypothesis is that the residuals are homoskedastic. This is tested against the alternative that the variance of the residuals increases as the value of one of the explanatory variables increases. The test examines whether there is heteroskedasticity in the residuals. The Goldfeld-Quandt (1965) test statistic is based on the ratio: $\Sigma\hat{\varepsilon}_1^2/\Sigma\hat{\varepsilon}_2^2$ is distributed as central $F(r,r)$, $r = (n - c)/2 - k$, if the errors are homoskedastic. If there are no heteroskedasticity in the residuals, the calculated F-statistic should not exceed the critical value at the appropriate level of significance.
(2) The other is the White (1980) test; it tests whether the residual has constant variance or not. The null hypothesis is that there is no heteroskedasticity. Say we have a multiple regression model with two explanatory variables. In order to test whether the residual ε_t has constant variance or not, the following auxiliary equation is estimated $\hat{\varepsilon}_t^2 = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{1t}^2 + \beta_4 X_{2t}^2 + \beta_5 X_{it} X_{2t} + \mu_t$, where $\hat{\varepsilon}_t$ is the estimated residual from the initial equation and is regressed against all its explanatory variables together with their squares and cross-products. The test statistic is equal to nR^2, where R^2 is obtained

from the estimation of the initial equation. The statistic has a χ^2 distribution, with degrees of freedom equal to the number of regressors excluding the intercept.

Numerous post-1995 studies have reported these tests (including Kim and Song, 1998; Song *et al.*, 2000; Lim and McAleer, 2001). It should be noted that none of the studies reviewed here found any problems with respect to the White (1980) test. Although clearly there is little, if any, reason to suspect heteroskedasticity in time-series data used in tourism demand modelling, reporting these tests does not influence or distort either the results or the model.

CONCLUSION

This paper follows on from a number of excellent reviews of the empirical tourism demand modelling literature. There is an apparent lack of analysis of the unit root, co-integration and diagnostic test aspects; these are regarded as being crucial to the precision of econometric models and results. Given the plethora of econometric based research on tourism demand, there is an urgent need to look into this issue, as has been done here for the first time.

Based on previous surveys of tourism demand literature, this study classified the research into two categories: the pre-1995 research, which used traditional econometric techniques, and the post-1995 research, based on modern econometric modelling techniques.

The paper highlighted that, apart from the Durbin–Watson (1950) test, the pre-1995 and some of the post-1995 literature on tourism demand modelling have ignored the unit root and diagnostic tests. Similarly, there has been an absence of tests for co-integration. As a corollary, there are indications of the so-called 'spurious regression' problem, casting doubt on the precision of results and the overall validity of tourism demand models. Some studies have also not accounted for (i) autocorrelation when a lagged dependent variable is used as an explanatory variable, (ii) model specification, (iii) normality and (iv) structural stability. On the other hand, some of the post-1995 literature has reported all the key diagnostic tests, including heteroskedasticity which the econometric literature regards as not as important in time-series data except in the case of financial market analysis, and cross-section data. Put differently, the widely reported ARCH test in tourism demand models has little relevance as it pertains mainly to describing conditional variances of asset price fluctuations, and in modelling associated time-varying risk premia. The ARCH test, together with the inclusion of other forms of heteroskedasticity test, such as the White (1980) and the Goldfeld and Quandt (1965) tests, however, does not diminish the accuracy of the results, but this is an issue that future studies should take into consideration when reporting diagnostic tests.

In summary, the improved econometric techniques applied in most of the post-1995 literature have produced a higher degree of reliability than those studies that have ignored, particularly, the unit root and co-integration tests. Hence, the tourism demand literature must be viewed with caution, a view consonant with that of Witt and Witt (1995) and Lim (1997a).

ACKNOWLEDGMENT

Constructive comments and suggestions by Professors Peter Dixon, Heather Anderson, and Farshid Vahid, Drs Biman Prasad and Brian Chapman, and anonymous reviewers on earlier versions of this article are gratefully acknowledged.

APPENDIX

Several tests for a unit root(s) in a single time-series are available (e.g. Fuller, 1976; Dickey and Fuller, 1979, 1981; Phillips, 1987; Phillips and Perron, 1988; Pantula, 1989). The most commonly used test is that of Dickey and Fuller (1979) — known as the Augmented Dickey–Fuller (ADF) test — taking the following form

$$\Delta y_t = \alpha_0 + \lambda T + \beta y_{t-1} + \sum_{i=4}^{p-1} \phi \Delta y_{t-i} + \varepsilon_t$$

the ADF auxiliary regression tests for a unit root in y_t; T denotes the deterministic time trend (many researchers include T in the test regression to allow for a deterministic trend under the alternative hypothesis); Δy_{t-i} are the lagged first differences to accommodate serial correlation in the errors, ε_t; and α, λ, β and ϕ are the parameters to be estimated. Often the time trend is included in the auxiliary regression equation if the reported ADF t-statistics, with and without a deterministic trend, are substantially different from each other. The null hypothesis is that $\beta = 0$ and the alternative hypothesis is that $\beta < 0$. Appropriate number of lags can be selected for the dependent variable based on the Akaike information criterion (AIC).

REFERENCES

Akis S. 1998. A compact econometric model of tourism demand for Turkey. *Tourism Management* **19**: 99–102.

Banerjee A, Dolado JJ, Galbraith JW, Hendry DF. 1993. *Cointegration, Error Correction, and the Econometric Analysis of Non-stationary Data*. Oxford University Press: Oxford.

Brown RL, Durbin J, Evans JM. 1975. Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society, Series B* **37**: 149–192.

Carey K. 1991. Estimation of Caribbean tourism demand: issues in measurement and methodology. *Atlantic Economic Journal* **19**: 32–39.

Charemza WW, Deadman DF. 1997. *New Directions in Econometric Practice: General to Specific Modelling, Cointegration and Vector Autoregression*. Edward Elgar: Cheltenham.

Chow G. 1960. Tests for equality between sets of coefficients in two linear regressions. *Econometrica* **28**: 591–605.

Crouch GI. 1994a. The study of international tourism demand: a survey of practice. *Journal of Travel Research* **32**: 41–57.

Crouch GI. 1994b. The study of international tourism demand: a survey of findings. *Journal of Travel Research* **32**: 12–23.

Daniel ACM, Ramos FFR. 2002. Modelling inbound international tourism demand to Portugal. *International Journal of Tourism Research* **4**(3): 193–209.

Davidson JEH, Hemdry DF, Srba F, Yeo S. 1978. Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom. *Economic Journal* **88**: 661–692.

Dickey DA, Fuller WA. 1979. Distributions of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* **74**: 427–431.

Dickey DA, Fuller WA. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* **49**: 1057–1072.

Durbin J, Watson G. 1950. Testing for serial correlation in least squares regression — I. *Biometrika* **37**: 409–428.

Engle RF. 2001. GARCH 101: the use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives* **15**: 157–168.

Engle RF, Bollerslev T. 1986. Modelling the persistence of conditional variance. *Econometric Reviews* **5**: 1–50.

Engle RF, Granger CWJ. 1987. Cointegration and error correction: representation, estimation and testing. *Econometrica* **55**: 251–276.

Fuller WA. 1976. *Introduction to Statistical Time Series*. Wiley: New York.

Godfrey LG. 1978. Testing for multiplicative heteroscedasticity. *Journal of Econometrics* **8**: 227–236.

Godfrey LG. 1988. *Mis-specification Tests in Econometrics. The Lagrange Multiplier Principle and Other Approaches*. Cambridge University Press: Cambridge.

Goldfeld SM, Quandt RE. 1965. Some tests for homoskedasticity. *Journal of American Statistical Association* **60**: 539–547.

Granger CWJ. 1981. Some properties of time series data and their use in econometric model specification. *Journal of Econometrics* **16**: 121–130.

Granger CWJ, Weiss AA. 1983. *Cointegrated Variables and Error Correction Models*. Discussion Paper 88–13a, University of California: San Diego.

Granger CWJ, Newbold P. 1974. Spurious regressions in econometrics. *Journal of Econometrics* **35**: 143–159.

Gujarati D. 1992. *Essentials of Econometrics*. McGraw-Hill: New York.

Gunadhi H, Boey CK. 1986. Demand elasticities of tourism in Singapore. *Tourism Management* **17**: 239–253.

Hansen BE. 1992. Tests for parameter instability in regression with I(1) processes. *Journal of Business and Economic Statistics* **10**: 321–335.

Hausman JA. 1978. Specification tests in econometrics. *Econometrica* **46**: 1251–1271.

Hendry DF, Anderson GJ. 1977. Testing dynamic specification in small simultaneous models: an application to a model of building society behaviour in the UK. In *Frontiers of Quantitative Economics*, Intrilligator MD (ed.). North-Holland: Amsterdam; 361–383.

Icoz O, Var T, Kozak M. 1998. Tourism demand in Turkey. *Annals of Tourism Research* **25**: 236–240.

Jarque CM, Bera AK. 1980. Efficient test for normality, homoscedasticity and serial indepen-

dence of regression residuals. *Economic Letters* **6**: 255–259.
Johansen S. 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* **12**: 231–254.
Johansen S. 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* **59**: 1551–1580.
Johansen S, Juselius K. 1990. Maximum likelihood estimation and inference on cointegration — with applications to the demand for money. *Oxford Bulletin of Economics and Statistics* **52**: 169–210.
Jud GD, Joseph H. 1974. International demand for Latin American tourism. *Growth and Change* **5**: 25–31.
Kim S, Song H. 1998. Analysis of tourism demand in South Korea: a cointegration and error correction approach. *Tourism Analysis* **3**: 25–41.
Kulendran N. 1996. Modelling quarterly tourist flows to Australia using cointegration analysis. *Tourism Economics* **2**: 203–222.
Kulendran N, King ML. 1997. Forecasting international quarterly tourist flows using error correction and time series models. *International Journal of Forecasting* **13**: 319–327.
Kulendran N, Wilson K. 2000. Modelling business tourism. *Tourism Economics* **6**(1): 47–59.
Kulendran N, Witt SF. 2001. Cointegration versus least squares regression. *Annals of Tourism Research* **28**: 291–311.
Lathiras P, Siriopoulos C. 1998. The demand for tourism to Greece: a cointegration approach. *Tourism Economics* **4**: 171–185.
Lee C, Var T, Blain TW. 1996. Determinants of inbound tourist expenditures. *Annals of Tourism Research* **23**: 527–542.
Levantis T, Gani A. 2000. Tourism demand and the nuisance of crime. *International Journal of Social Economics* **27**: 959–967.
Lim C. 1997a. An econometric classification and review of international tourism demand models. *Tourism Economics* **3**: 69–81.
Lim C. 1997b. Review of international tourism demand models. *Annals of Tourism Research* **24**: 835–849.
Lim C, McAleer M. 2001. Cointegration analysis of quarterly tourism demand by Hong Kong and Singapore for Australia. *Applied Economics* **33**: 1599–1619.
McAleer M. 1994. Sherlock Holmes and the search for truth: a diagnostic tale. *Journal of Economic Surveys* **8**(4): 317–370.
Mak J, Moncur J, Yonamine D. 1977. Determinants of visitor expenditure and visitor lengths of stay: a cross-section analysis of US visitors to Hawaii. *Journal of Travel Research* **15**: 5–8.
Narayan PK. 2002. A tourism demand model for Fiji, 1970–2000. *Pacific Economic Bulletin* **17**(2): 103–116.
Narayan PK. 2003. Determinants of tourist expenditure in Fiji: a cointegration approach. *Pacific Tourism Review*.
Narayan PK. 2004. Fiji's tourism exports: an ARDL approach to cointegration. *Tourism Economics* **10**(1).
Nelson CR, Plosser CI. 1982. Trends and random walk in macroeconomic time series. *Journal of Monetary Economics* **10**: 139–162.
Pantula SG. 1989. Testing for unit roots in time series data. *Econometric Theory* **5**: 256–271.
Park JY. 1990. Testing for unit roots and cointegration by variable addition. *Advances in Econometrics* **8**: 107–133.
Pattichis CA. 1999. Price and income elasticities of disaggregated import demands: results from UECMs and an application. *Applied Economics* **31**: 1061–1071.
Payne JE, Mervar A. 2002. A note in modelling tourism revenues in Croatia. *Tourism Economics* **8**(1): 103–109.
Pesaran HM, Shin Y. 1995. *Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis*. DAE Working Paper Series No. 9514, Department of Applied Economics, University of Cambridge.
Pesaran HM, Shin Y, Smith R. 1996. *Testing the Existence of a Long-run Relationship*. DAE Working Paper Series No. 9622, Department of Applied Economics, University of Cambridge.
Phillips PCB. 1986. Understanding spurious regressions in econometrics. *Journal of Econometrics* **33**: 311–340.
Phillips PCB. 1987. Time series regression with a unit root. *Econometrica* **55**: 277–302.
Phillips PCB. 1996. Econometric model determination. *Econometrica* **64**(4): 763–812.
Phillips PCB, Ouliaris S. 1990. Asymptotic properties of residual based tests for cointegration. *Econometrica* **58**: 165–193.
Phillips PCB, Perron P. 1988. Testing for a unit root in time series regression. *Biometrika* **75**: 335–359.
Praetz P. 1987. Some Aspects of Mis-specification in the linear model. In *Specification Analysis in the Linear Model*, King ML, Giles DEA (eds). Routledge and Kegan Paul: London.
Qu H, Lam S. 1997. A travel demand model for mainland Chinese tourists to Hong Kong. *Tourism Management* **18**(8): 593–597.
Ramsey JB. 1969. Tests for specification errors in classical linear least squares regression analysis. *Journal of the Royal Statistical Society* **31**: 350–371.
Ramsey JB. 1974. Classical model selection through specification error tests. In *Frontiers in Economet-*

rics, Zarembka P (ed.). Academic Press: New York.
Ramsey JB, Schmidt P. 1976. Some further evidence on the use of OLS and BLUS residuals in specification error tests. *Journal of the American Statistical Association* **71**: 389–390.
Romilly P, Liu X, Song M. 1998. Economic and social determinants of international tourism spending: a panel data analysis. *Tourism Analysis* **3**: 3–16.
Sargan JD. 1964. Wages and prices in the UK: a study in econometric methodology. In *Econometric Analysis for National Economic Planning*, Hart PE, Whitaker JK (eds). Butterworth, London. (Reprinted 1984 in *Econometrics and Quantitative Economics*, Hendry DF, Wallis KF (eds). Basil Blackwell: Oxford.
Seddighi HD, Shearing DF. 1997. The demand for tourism in North East England with special reference to Northumbria: an empirical analysis. *Tourism Management* **18**: 499–511.
Song H, Witt S. 2000. *Tourism Demand Modelling and Forecasting: Modern Econometric Approaches*. Elsevier Science Limited: Oxford.
Song H, Romilly P, Liu X. 2000. An empirical study of outbound tourism demand in the UK. *Applied Economics* **32**: 611–624.
Stock JH, Watson MH. 1988. Testing for common trends. *Journal of the American Statistical Association* **83**: 1097–1107.
Vanegas M, Croes RR. 2000. Evaluation of demand: US tourists to Aruba. *Annals of Tourism Research* **27**: 946–963.
Vogt MG, Wittayakorn C. 1998. Determinants of the demand for Thailand's exports of tourism. *Applied Economics* **30**: 711–715.
White H. 1980. A heteroscedasticity-consistent covariance matrix and a direct test for heteroscedasticity. *Econometrica* **48**: 421–448.
Witt DF, Witt CA. 1995. Forecasting tourism demand: a review of empirical research. *International Journal of Forecasting* **11**: 447–475.
Yavas BF, Bilgin Z. 1996. Estimation of tourism demand in Turkey: a pooled cross-section time series analysis. *Tourism Analysis* **1**: 19–27.
Yule GU. 1926. Why do we sometimes get nonsense correlations between time series? a study in sampling and the nature of time series. *Journal of the Royal Statistical Society* **89**: 1–64.
Zarembka P. 1974. *Frontiers in Econometrics*. Academic Press: New York.