TOURISM DEMAND MODELLING AND FORECASTING

Modern Econometric Approaches

Haiyan Song
and
Stephen F. Witt

2000

PERGAMON
An Imprint of Elsevier Science
Amsterdam – Lausanne – New York – Oxford – Shannon – Singapore – Tokyo

ELSEVIER SCIENCE Ltd
The Boulevard, Langford Lane
Kidlington, Oxford OX5 1GB, UK

First edition 2000

Library of Congress Cataloging in Publication Data
Song, Haiyan.
Tourism demand modelling and forecasting: modern econometric approaches/Haiyan Song and Stephen F. Witt.—1st ed.
p. cm.—(Advances in tourism research series)
Includes bibliographical references (p.).
ISBN 0-08-043673-0 (hc.)
1. Tourism—Econometric models. 2. Tourism—Forecasting. I. Witt, Stephen F. II. Title. III. Series.

G155.A1 S593 2000
338.4'791—dc21 00-021690

British Library Cataloguing in Publication Data

Tourism demand modelling and forecasting: modern econometric approaches. – (Advances in tourism research)
1. Tourist trade – Mathematical models. 2. Tourist trade – Forecasting
I. Song, Haiyan II. Witt, Stephen F.
338.4'791

ISBN 0-08-043673-0

⊗The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).
Printed in The Netherlands.

ADVANCES IN TOURISM RESEARCH SERIES

Series Editor: Stephen Page
Massey University at Albany, New Zealand

Published:

RYAN & PAGE
Tourism Management: Towards the New Millennium

Forthcoming titles include:

ASHWORTH & TUNBRIDGE
The Tourist-Historic City: Retrospect and Prospect of Managing the Heritage City

BAUM & LUNDTORP
Seasonality in Tourism: An Exploration of Issues

TEO, CHANG & HO
Interconnected Worlds: Tourism in Southeast Asia

Related Elsevier journals – sample copies available on request

Annals of Tourism Research
Cornell Hotel and Restaurant Administration Quarterly
International Journal of Hospitality Management
International Journal of Intercultural Relations
Tourism Management
World Development

Contents

Preface vii

1 Introduction to Tourism Demand Analysis **1**
1.1 Introduction 1
1.2 Determinants of Tourism Demand 2
1.3 Functional Form and Demand Elasticities 8
1.4 Summary 13

2 Traditional Methodology of Tourism Demand Modelling **15**
2.1 Introduction 15
2.2 Traditional Methodology of Tourism Demand Modelling 16
2.3 Failure of the Traditional Approach to Tourism Demand Modelling 26

3 General-to-Specific Modelling **27**
3.1 Introduction 27
3.2 General-to-Specific Modelling 28
3.3 Tests of Restrictions on Regression Parameters 33
3.4 Diagnostic Checking 34
3.5 Model Selection 42
3.6 Worked Example 43

4 Cointegration **53**
4.1 Spurious Regression 53
4.2 The Concept of Cointegration 55
4.3 Test for Order of Integration 57
4.4 Test for Cointegration 68

5 Error Correction Model **73**
5.1 Cointegration and Error Correction Mechanism 73
5.2 Estimating the ECM 76
5.3 Worked Examples 79

6 Vector Autoregression (VAR) and Cointegration **91**
6.1 Introduction 91
6.2 Basics of VAR Analysis 92
6.3 Impulse Response Analysis 94
6.4 Forecasting with VAR and Error Variance Decomposition 97
6.5 Testing for Granger Causality and Exogeneity 98
6.6 An Example of VAR Modelling 101
6.7 VAR Modelling and Cointegration 111

7 Time Varying Parameter Modelling **123**
7.1 Introduction 123

7.2 Are Tourism Demand Elasticities Constant Over Time? 124
7.3 The TVP Model and Kalman Filter 128
7.4 A Worked Example 131
7.5 Summary and Conclusions 134

8 Panel Data Analysis **137**
8.1 Introduction 137
8.2 Model Specification and Estimation 138
8.3 An Empirical Study of a Tourism Demand Model Based on the Panel Data Approach 142
8.4 Summary and Conclusions 152

9 Evaluation of Forecasting Performance **157**
9.1 Introduction 157
9.2 Measures of Forecasting Accuracy 159
9.3 Forecasting Evaluation of Different Models 161
9.4 Concluding Remarks 164

References **167**

Author index 173

Subject index 175

Preface

The phenomenal growth of both the world-wide tourism industry and academic interest in tourism over the last thirty years has generated great interest in tourism demand modelling and forecasting from both business and academic sectors. However, although econometric modelling and forecasting have experienced considerable changes over the last two decades, these developments have not been incorporated in any of the existing tourism forecasting books. For example, the development of cointegration techniques has enabled researchers to overcome the problems of spurious regression and non-stationarity of time series.

The purpose of this book is to introduce students, researchers and practitioners to the recent advances in econometric methodology within the context of tourism demand analysis, and to illustrate these new developments with actual tourism applications. The book is designed for final-year undergraduate, taught postgraduate and research students in tourism studies, as well as researchers and practitioners who wish to apply the recent advances in econometric modelling and forecasting to tourism demand analysis. It is assumed that the reader has taken an introductory course in statistics and multiple regression analysis and has some knowledge of tourism. The concepts and computations of modern econometric modelling methodologies are introduced at a level that is accessible to non-specialists. The methodologies introduced include general-to-specific modelling, cointegration, vector autoregression, time varying parameter modelling and panel data analysis. In order to help the reader understand the various methodologies, extensive tourism demand examples are provided. The computer packages used for the estimation of the various tourism demand models in the examples are Microfit 4.0 for Windows and Eviews 3.1.

The book starts with an introduction to the fundamentals of tourism demand analysis (Chapter 1). The problems of traditional tourism demand modelling and forecasting, i.e., data mining and spurious regression due to common trends in the time series, are addressed in Chapter 2. Since general-to-specific modelling, including the use of autoregressive distributed lag processes, cointegration analysis and error correction models, has been regarded as a remedy for the problems of data mining (at least to a large extent) and spurious regression, the general-to-specific approach to tourism demand modelling and forecasting is explored in Chapters 3–6. The size of the tourism industry has experienced a phenomenal increase over the last three decades, and also it is known that the demand for international tourism (both inbound and outbound) is sensitive to regime shifts (changes in government policy, political and economic conditions); therefore demand elasticities are unlikely to have remained unchanged over time. The time varying parameter model together with the use of the Kalman filter as an estimation method is a useful tool for examining the effects of regime shifts on tourism demand elasticities, and this is explored in Chapter 7. Often there is a lack of sufficient time series data in terms of the number of observations on tourism demand variables and this can result in estimation and forecasting biases. The panel data approach is introduced in Chapter 8 as a way of overcoming this problem. The last chapter (Chapter 9) evaluates the empirical forecasting performance of the various models and puts forward general conclusions.

The authors have considerable research experience related to modelling and forecasting the demand for international tourism, with over 70 publications in this area. They have also been teaching tourism demand modelling and forecasting for many years.

November 1999

Haiyan Song
Stephen F. Witt

1

Introduction to tourism demand analysis

1.1 Introduction

Tourism demand is the foundation on which all tourism-related business decisions ultimately rest. Companies such as airlines, tour operators, hotels, cruise ship lines, and many recreation facility providers and shop owners are interested in the demand for their products by tourists. The success of many businesses depends largely or totally on the state of tourism demand, and ultimate management failure is quite often due to the failure to meet market demand. Because of the key role of demand as a determinant of business profitability, estimates of expected future demand constitute a very important element in all planning activities. It is clear that *accurate* forecasts of tourism demand are essential for *efficient* planning by tourism-related businesses, particularly given the perishability of the tourism product. For example, it is impossible for an airline to recoup the potential revenue lost by a flight taking off with empty seats. The loss resulting from these unsold seats contrasts with the case of, say, a car manufacturer, where if a car does not sell on a particular day it can still be sold subsequently.

The tourism forecasting methods examined in this book are modern econometric approaches. Here the forecast variable is specifically related to a set of determining forces; future values of the forecast variable are obtained by using forecasts of the determining variables in conjunction with the estimated quantitative relationship between the forecast variable and its determinants.

In Section 1.2 the tourism demand function is examined, and measurement issues relating to demand variables are discussed. The functional form of the demand equation and its relationship to demand elasticities are presented in Section 1.3, and the chapter is summarised in Section 1.4.

1.2 Determinants of tourism demand

Tourism visits can take place for various reasons: holidays, business trips, visits to friends and relatives (VFR), conferences, pilgrimages and so on. Most empirical studies of tourism demand examine either total tourist trips (including all the above-mentioned purposes) or just holiday trips (Witt and Witt, 1995). Since the majority of tourist visits take place for holiday purposes, the determinants of demand are generally taken to be the same as those for holiday trips in those published studies which examine total trips. We shall therefore also focus on the determinants of the demand for holiday tourism.

The term 'tourism demand' may be defined for a particular destination as the quantity of the tourism product (i.e., a combination of tourism goods and services) that consumers are willing to purchase during a specified period under a given set of conditions. The time period may be a month, a quarter or a year. The conditions that relate to the quantity of tourism demanded include tourism prices for the destination (tourists' living costs in the destination and travel costs to the destination), the availability of and tourism prices for competing (substitute) destinations, potential consumers' incomes, advertising expenditure, tastes of consumers in the origin (generating) countries, and other social, cultural, geographic and political factors.

The demand function for the tourism product in destination i by residents of origin j is given by

$$Q_{ij} = f(P_i, P_s, Y_j, T_j, A_{ij}, \varepsilon_{ij}) \tag{1.1}$$

where Q_{ij} is the quantity of the tourism product demanded in destination i by tourists from country j;
P_i is the price of tourism for destination i;
P_s is the price of tourism for substitute destinations;
Y_j is the level of income in origin country j;
T_j is consumer tastes in origin country j;
A_{ij} is advertising expenditure on tourism by destination i in origin country j;

ε_{ij} is the disturbance term that captures all other factors which may influence the quantity of the tourism product demanded in destination *i* by residents of origin country *j*.

In empirical investigations it can be difficult to find exact measures of the determinants of tourism demand due to lack of data availability. The variables used in empirical studies of tourism demand functions are now reviewed, and the problems associated with these measures are also addressed.

Dependent variable

International tourism demand is generally measured in terms of the number of tourist visits from an origin country to a destination country, or in terms of tourist expenditure by visitors from the origin country in the destination country. The number of tourist nights spent by residents of the origin in the destination is an alternative tourism demand measure.

International tourism demand data are collected in various ways. Tourist visits are usually recorded by frontier counts (inbound), registration at accommodation establishments (inbound) or sample surveys (inbound and outbound). A problem with frontier counts is that in certain cases a substantial transit traffic element may be present. Accommodation establishment records exclude day-trippers and tourists staying with friends or relatives or in other forms of unregistered accommodation. Sample surveys may be applied at points of entry/exit to returning residents or departing non-residents, or household surveys may be carried out (outbound), but in both cases often the sample size is relatively small. International tourist expenditure data are usually collected by the bank reporting method or sample surveys. The former method is based on the registration by authorised banks and agencies of the buying and selling of foreign currencies by travellers. There are many problems associated with this method of data collection such as identifying a transaction as a tourism transaction, the non-reporting of relevant transactions and the unreliability of its use for measuring receipts from specific origin countries (the geographic breakdown relates to the denomination of the currency and not the generating country). Sample surveys provide more reliable data on tourist expenditures, but as with visit data the sample size is often relatively small.

Explanatory variables

Population. The level of foreign tourism from a given origin is expected to depend upon the origin population. Although population features as a separate explanatory variable in some tourism demand studies, more often the effect of population is accommodated by modifying the dependent variable to become international tourism demand per capita. However, in many cases the impact of population changes is ignored. The main justification for not having population as a separate variable is that its presence may cause multicollinearity problems, as population tends to be highly correlated with income (see next sub-section). On the other hand, the procedure adopted whereby demand is specified in per capita terms in effect constrains the population elasticity to equal unity (if a power model is under consideration). Although it is theoretically incorrect to exclude population, it is likely that population changes in the generating countries will be small over the short–medium term, and hence the model will be affected only marginally.

Income. In tourism demand functions, origin country income or private consumption is generally included as a key explanatory variable, and usually enters the demand function in per capita form (corresponding to the specification of demand in per capita terms). If (mainly) holiday demand or visits to friends and relatives are under consideration then the appropriate form of the variable is personal disposable income or private consumption. 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 GDP) should be used, or a measure of business activity such as aggregate imports/exports between the origin and destination countries.

Own price. The appropriate measure of tourism prices is difficult to obtain. In the case of tourism there are two elements of price: the cost of travel to the destination, and the cost of living for tourists in the destination.

Although the theoretical justification for including transport cost as a demand determinant does not appear to be disputed, many empirical studies exclude this variable from the demand function on the grounds of potential multicollinearity problems and lack of data availability. However, it is possible to obtain an approximate measure of transport cost using representative air fares between origin and destination for air travel, and representative petrol costs and/or ferry fares for surface travel.

Usually the consumer price index (CPI) in a destination country is taken to be a proxy for the cost of tourism in that country. The problem with using the CPI as the cost of tourism in the destination is that the cost of living for the local residents does not always reflect the cost of living for foreign visitors to that destination, especially in poor countries. However, this procedure is adopted on the grounds of lack of more suitable data, i.e., an index 'defined over the basket of goods purchased by tourists, rather than over the usual typical consumer basket' (Kliman, 1981, p. 490). Potential tourists base their decisions on tourism costs in the destination measured in terms of their local currency, and therefore the destination price variable should be adjusted by the exchange rate between the origin and destination currencies.

Exchange rates are also sometimes used separately to represent tourists' costs of living. Although they usually appear in addition to a consumer price index proxy, they may also be used as the sole representation of tourists' living costs. The justification for including a separate exchange rate variable in international tourism demand functions is that consumers are more aware of exchange rates than destination costs of living for tourists, and hence they are driven to use exchange rate as a proxy variable. However, the use of exchange rate alone in the demand functions can be very misleading because even though the exchange rate in a destination may become more favourable, this could be counterbalanced by a relatively high inflation rate. Empirical results presented by Martin and Witt (1987) suggest that the exchange-rate-adjusted consumer price index is a reasonable proxy for the cost of tourism, but that exchange rate on its own is not an acceptable proxy.

Substitute prices. Mostly, those substitution possibilities allowed for in international tourism demand studies are restricted to tourists' destination living costs. There are two possible ways in which these substitute prices may enter the demand function. First, the tourists' cost-of-living variable may be specified in the form of the destination value relative to the origin value, thus permitting substitution between tourist visits to the foreign destination under consideration and domestic tourism. The justification for this form of relative price index is that domestic tourism is the most important substitute for foreign tourism. Second, substitute prices may enter the demand function in a form that allows for the impact of competing foreign destinations by specifying the tourists' cost-of-living variable as destination value relative to a weighted average value calculated for a set of alternative destinations, or by specifying a separate weighted average substitute destination cost variable.

Just as tourists' living costs in substitute destinations are likely to influence the demand for tourism to a given destination, so travel costs to substitute destinations may also be expected to have an impact. However, although some theoretical attention has been paid to the notion of substitute travel costs in the literature, they do not often feature in tourism demand functions. Substitute travel costs may enter the demand function by specifying the travel cost variable as travel cost to the destination relative to a weighted average value calculated for a set of alternative destinations, or by specifying a separate weighted average substitute travel cost variable.

Tastes. Consumer tastes can have an important influence on tourism demand. They are affected by socio-economic factors such as age, sex, education and marital status, and can change as a result of innovation and advertising, and more fundamentally as a consequence of changing values and priorities, and rising living-standards. However, due to data limitations most empirical studies which allow for consumer tastes use a time trend to represent a steady change in the popularity of a destination country over the period considered as a result of changing tastes. The intention of using a time trend is also to capture the time-dependent effects of all other explanatory variables not explicitly included in the demand equation, such as changes in air service frequencies and demographic changes in the origins. However, the problem of including a time trend in tourism demand functions is its ambiguity. We normally do not know exactly what the trend variable actually captures. In many cases, the inclusion of a time trend also causes problems with the income variable, since these two variables are likely to be highly correlated. Song *et al.* (2000) make an attempt to examine the influence of tastes on tourism demand by using a destination preference index. However, much work is still needed to incorporate appropriately consumer tastes in tourism demand functions.

Marketing. National tourist organisations engage in sales-promotion activities specifically to attempt to persuade potential tourists to visit the country, and these activities may take various forms including media advertising and public relations. Hence, promotional expenditure is expected to play a role in determining the level of international tourism demand. Much tourism-related marketing activity is not, however, specific to a particular destination (e.g., general travel agent and tour operator advertising) and therefore is likely to have little impact on the demand for tourism to that destination. The promotional activities of national tourist organisations are destination specific and are therefore more likely to influence tourism flows to the destination concerned.

Marketing has not featured often in tourism demand models, but a critical review of studies which do include some form of marketing variable appears in Witt and Martin (1987a) and a subsequent review appears in Crouch *et al.* (1992).

Expectations and habit persistence. Tourist expectations and habit persistence (stable behaviour patterns) are usually incorporated in tourism demand models through the use of a lagged dependent variable, i.e., an autoregressive term. Once people have been on holiday to a particular country and liked it, they tend to return to that destination. There is much less uncertainty associated with holidaying again in that country compared with travelling to a previously unvisited foreign country. Furthermore, knowledge about the destination spreads as people talk about their holidays and show photographs, thereby reducing uncertainty for potential visitors to that country. In fact this 'word of mouth' recommendation may well play a more important role in destination selection than does commercial advertising. A type of adaptive expectations or learning process is in operation and as people are, in general, risk averse (especially in the case of large outlays of money such as for holidays), the number of people choosing a given destination in any year depends on the numbers who chose it in previous years.

A further justification for the inclusion of a lagged dependent variable in tourism demand functions, which is not related to expectations or habit persistence, is that there may be constraints on supply. These constraints may take the form of shortages of hotel accommodation, passenger transportation capacity and trained staff, and these often cannot be increased rapidly. Time is also required to build up contacts among tour operators, hotels, airlines and travel agencies. Similarly, once the tourist industry in a country has become highly developed it is unlikely to dwindle rapidly. If a partial adjustment mechanism is postulated to allow for rigidities in supply, this results in the presence of a lagged dependent variable in the tourism demand function. This is now demonstrated.

$$Q_t - Q_{t-1} = \mu(Q_t^* - Q_{t-1}) \qquad 0 < \mu < 1 \tag{1.2}$$

where Q_t is the quantity of the tourism product provided in destination i for residents of country j in period t;
Q_t^* is the quantity of the tourism product demanded in destination i by residents of country j in period t;
μ is the speed of adjustment of the level of tourism provided to the level demanded.

(Subscripts i and j have been omitted for simplicity.) The left-hand side of

Equation (1.2) denotes the change in the level of the tourism product provided between periods $t-1$ and t. The bracketed term on the right-hand side of Equation (1.2) is the difference between the level of demand for the tourism product in period t and the level provided in period $t-1$, that is the change demanded in the level of tourism. Thus Equation (1.2) states that the change in the level of tourism provided is a proportion, μ, of change demanded. If $\mu = 0$, then $Q_t = Q_{t-1}$, and there is no movement of the quantity of tourism provided towards the quantity demanded. If $\mu = 1$, then $Q_t = Q_t^*$, i.e., there is complete adjustment of the quantity of tourism provided to the quantity demanded. As we are specifying a *partial* adjustment process, μ lies strictly between zero and unity – there is some adjustment but it is incomplete.

Equation (1.2) may be rewritten as

$$Q_t = (1 - \mu)Q_{t-1} + \mu Q_t^* \tag{1.3}$$

Now Q_t^* is the quantity of tourism demanded and is a function of the set of explanatory variables such as income, own price and substitute prices. Therefore, the only difference between the explanatory variables present in models (1.1) and (1.3) is that the latter includes the dependent variable with a one-period lag – hence the justification for including a lagged dependent variable in tourism demand functions to accommodate supply constraints.

Qualitative effects. Dummy variables can be included in international tourism demand models to capture the impacts of 'one-off' events. For example, when governments impose foreign currency restrictions on their residents (e.g., the £50 annual limit introduced in the UK during 1966 to late 1969), this is expected to reduce outbound tourism. Similarly, the 1973 and 1979 oil crises are expected to have temporarily reduced international tourism demand; although the impacts of the oil crises on holiday prices and consumer incomes are incorporated in these explanatory variables, a further reduction in international tourism demand is likely on account of the psychological impact of the resultant uncertainties in the world economic situation. Witt and Martin (1987b) discuss a range of 'one-off' events which have been accommodated by dummy variables.

1.3 Functional form and demand elasticities

Equation (1.1) is a theoretical model of tourism demand which is simply a mathematical statement that indicates that there is a relationship between the variables under consideration. However, in exactly what form these variables are related is unknown. In practice, we need to specify the form of the tourism demand function, i.e., the manner in which tourism demand is related to its determinants. The two commonly used demand equations assume either a linear relationship or a power relationship between Q_{ij} and its determinants.

The simplest relationship is the linear relationship, and this is expressed as:

$$Q_{ij} = \alpha_0 + \alpha_1 P_i + \alpha_2 P_s + \alpha_3 Y_j + \alpha_4 T_j + \alpha_5 A_{ij} + \varepsilon_{ij} \tag{1.4}$$

where the variables Q_{ij}, P_i, P_s, Y_j, T_j and A_{ij} are defined as in Equation (1.1); $\alpha_0, \alpha_1, \ldots, \alpha_5$ are the coefficients that need to be estimated empirically; and ε_{ij} is the disturbance term. Linear tourism demand equations are popular for the following two reasons. First, empirical studies have shown that many tourism demand relationships can be approximately represented by a linear relationship over the sample period under consideration (Edwards, 1985; Smeral *et al.*, 1992). Second, the coefficients in the linear model can be estimated relatively easily.

Based on Equation (1.4), we can examine how sensitive tourism demand is to changes in the independent variables. The measure of responsiveness of demand to changes in the independent variables is the demand elasticity, which is measured by:

$$\tilde{\omega}_X = \frac{\Delta Q_{ij}/Q_{ij}}{\Delta X/X} = \frac{\Delta Q_{ij}}{\Delta X} \times \frac{X}{Q_{ij}} \tag{1.5}$$

where X denotes an independent variable and Δ denotes 'change in'. Equation (1.5) shows the percentage change in quantity demanded, Q_{ij}, attributable to a given percentage change in an independent variable, X. This equation measures arc elasticity, namely the elasticity over some finite range of the function. At the limit, where ΔX and ΔQ_{ij} are very small, the point elasticity may be obtained:

$$\tilde{\omega}_X = \frac{\partial Q_{ij}}{\partial X} \times \frac{X}{Q_{ij}} \tag{1.6}$$

The partial derivative sign denotes that we are examining the impact on quantity demanded resulting from a change in X, holding all other factors constant. Applying formula (1.6) to Equation (1.4) allows various demand elasticities to be calculated. For example, the own-price elasticity is calculated as:

$$\tilde{\omega}_X = \frac{\partial Q_{ij}}{\partial P_i} \times \frac{P_i}{Q_{ij}} = \frac{\partial(\alpha_0 + \alpha_1 P_i + \alpha_2 P_s + \alpha_3 Y_j + \alpha_4 T_j + \alpha_5 A_{ij} + \varepsilon_{ij})}{\partial P_i} \times \frac{P_i}{Q_{ij}} = \alpha_1 \times \frac{P_i}{Q_{ij}} \tag{1.7}$$

From Equation (1.7) we can see that the point elasticity is obtained from a linear demand function by multiplying the regression coefficient of the variable X by the value of X/Q_{ij} at that point. Given that the value of X/Q_{ij} varies over the time period, the point elasticities based on the linear demand model therefore also vary over time. The substitute price elasticity, income elasticity, etc., are calculated in the same manner.

The most commonly used functional form in tourism demand analysis is the power model (Witt and Witt, 1995). This can be expressed as:

$$Q_{ij} = AP_i^{\alpha_1} P_s^{\alpha_2} Y_j^{\alpha_3} T_j^{\alpha_4} A_{ij}^{\alpha_5} u_{ij} \tag{1.8}$$

where A, α_0, α_1, ..., α_5 are coefficients; the variables are as defined previously; and u_i is the disturbance term. The popularity of the power function is due to its features discussed below, but also due to its relatively good empirical performance. For example, Witt and Witt (1992) and Lee *et al.* (1996) found that the power model outperformed the linear model in terms of expected coefficient signs and statistical significance of the coefficients.

The power function has three important features. First, Equation (1.8) implies that the marginal effects of each independent variable on tourism demand are not constant, but depend on the value of the variable, as well as on the values of all other variables in the demand function. This can be seen by taking the partial derivative of Equation (1.8) with respect to, say, income:

$$\frac{\partial Q_{ij}}{\partial Y_j} = A\alpha_3 P_i^{\alpha_1} P_s^{\alpha_2} Y_j^{\alpha_3 - 1} T_j^{\alpha_4} A_{ij}^{\alpha_5} u_{ij} \tag{1.9}$$

Equation (1.9) shows that the marginal effect of a change in income on tourism demand depends not only on the level of income but also on all the other explanatory variables. This changing marginal relationship is perhaps more realistic than the constant relationship assumed in the linear model. For example, when consumers' disposable income is low, a given increase in income tends to have a large impact on foreign holiday expenditure; but when disposable income is high, the increase in foreign holiday expenditure resulting from the given increase in income is likely to be smaller. This may be due to external constraints imposed on consumers' leisure time, or because consumer demand shifts to a higher order such as buying new houses.

The second feature of the power tourism demand function is that Equation (1.8) may be transformed into a linear relationship using logarithms, again making estimation relatively easy. The transformed Equation (1.8) is

$$\ln Q_{ij} = \alpha_0 + \alpha_1 \ln P_i + \alpha_2 \ln P_s + \alpha_3 \ln Y_j + \alpha_4 \ln T_j + \alpha_5 \ln A_{ij} + \varepsilon_{ij} \tag{1.10}$$

where $\alpha_0 = \ln A$ and $\varepsilon_{ij} = \ln u_{ij}$.

The third feature of the power demand function is that the estimated coefficients in Equation (1.10) are estimates of demand elasticities (which are constant over time). For example, own-price elasticity is measured by:

$$\tilde{\omega}_{P_i} = \frac{\partial Q_{ij}}{\partial P_i} \times \frac{P_i}{Q_{ij}} \tag{1.11}$$

From Equation (1.8)

$$\tilde{\omega}_{P_i} = A\alpha_1 P_i^{\alpha_1 - 1} P_s^{\alpha_2} Y_j^{\alpha_3} T_j^{\alpha_4} A_{ij}^{\alpha_5} u_{ij} \frac{P_i}{Q_{ij}} \tag{1.12}$$

$$= A\alpha_1 P_i^{\alpha_1 - 1} P_s^{\alpha_2} Y_j^{\alpha_3} T_j^{\alpha_4} A_{ij}^{\alpha_5} u_{ij} \times \frac{P_i}{AP_i^{\alpha_1} P_s^{\alpha_2} Y_j^{\alpha_3} T_j^{\alpha_4} A_{ij}^{\alpha_5} u_{ij}}$$

$$= \alpha_1 \frac{AP_i^{\alpha_1} P_s^{\alpha_2} Y_j^{\alpha_3} T_j^{\alpha_4} A_{ij}^{\alpha_5} u_{ij}}{P_i} \times \frac{P_i}{AP_i^{\alpha_1} P_s^{\alpha_2} Y_j^{\alpha_3} T_j^{\alpha_4} A_{ij}^{\alpha_5} u_{ij}}$$

$$= \alpha_1$$

Hence, the own-price elasticity is the same as the coefficient of the own-price variable, and does not depend on the price : quantity ratio. This constant demand elasticity feature holds for all the variables in Equation (1.8) (and therefore also Equation (1.10)). The constant demand elasticity property is useful in that it is easy to understand from a managerial point of view; it allows policymakers to assess the percentage impact on tourism demand resulting from a 1% change in one of the independent variables, while holding all other explanatory variables constant.

The magnitude of the estimated price elasticity of demand can provide useful information for policymakers. Total tourism revenue may increase, decrease or remain the same as a result of a change in tourism prices, and this depends on the value of the price elasticity. There are three ranges of values that are relevant.

$|\tilde{\omega}_{P_i}| > 1$: If the absolute value of the price elasticity exceeds unity, the demand for tourism is price elastic. An increase in tourism price will result in a more than proportionate decrease in quantity demanded, and as a result total tourism revenue will fall.

$|\tilde{\omega}_{P_i}| = 1$: If the absolute value of the price elasticity equals unity, the demand curve is a rectangular hyperbola. Total tourism revenue will remain constant with a change in tourism price.

$|\tilde{\omega}_{P_i}| < 1$: If the absolute value of the price elasticity is less than unity, the demand for tourism is price inelastic. An increase in tourism price will result in a less than proportionate decrease in quantity demanded, and as a result total tourism revenue will rise.

According to economic theory, the total revenue (TR) will continue to increase as long as marginal revenue (MR) is positive. If we know the price elasticity of demand at a point, the value of MR can be calculated. This is illustrated as follows:

$$TR = P_i Q_{ij}$$

$$MR = \frac{dTR}{dQ_{ij}} = P_i + Q_{ij}\frac{dP_i}{dQ_{ij}} = P_i\left(1 + \frac{Q_{ij}}{P_i}\frac{dP_i}{dQ_{ij}}\right) = P_i\left(1 + \frac{1}{\tilde{\omega}_{P_i}}\right)$$

This equation states that the MR relates to the price of tourism and price elasticity. If we know the value of the price elasticity, we can calculate the sign of the MR, and hence we can see whether or not the TR is increasing. For example, if the price elasticity of demand for tourism is estimated to be -1.5, the MR revenue will be $0.333P_i$. Since the price of tourism is always positive, the MR is positive. The managerial implication of this would be that a decrease in the tourism price holding other variables constant will bring about an increase in total tourism revenue. However, if the calculated price elasticity is -0.5, the MR will be $-P_i$, which is negative. This suggests that a decrease in the tourism price holding other variables constant will bring about a less than proportionate increase in tourism demand, and so total tourism revenue will decline.

Knowledge of income elasticities is also important for tourism planners. A low income elasticity of demand implies that the demand for tourism in a particular destination is relatively insensitive to the economic situation in the origin country. However, if the calculated income elasticity exceeds unity, then a rise in income in the origin country will be accompanied by a more than proportionate rise in tourism demand in the destination. Destinations should therefore pay particular attention to forecasting the expected levels of future economic activity in those tourism generating countries with high income elasticities.

Price elasticities for substitute destinations also contain useful information for tourism policymakers. If the substitute price elasticity in the tourism demand function for a particular destination is high, this means that tourism in this destination is very sensitive to price changes in other destinations. Planners and decision-makers in the destination under consideration should therefore keep a close watch on prices in competing destinations in order to ensure that the relative price level does not increase significantly.

1.4 Summary

This chapter has introduced some basic concepts related to tourism demand modelling and forecasting. The demand for the tourism product faced by a destination depends on the price of tourism for that destination, the price of tourism for alternative destinations, potential consumers' incomes, consumer tastes, and the promotional efforts of the destination, as well as other social, cultural, geographic and political factors. Demand elasticities such as own price elasticity, income elasticity and substitute price elasticity measure the percentage change in the quantity of tourism demanded in a destination country as a result of a 1% change in one of the determinant variables, while holding the rest of the determinants constant in the tourism demand function. A decline in the price of tourism in a destination results in an increase in total tourism revenue in that destination if demand is price elastic, a decrease in total tourism revenue if demand is price inelastic, and unchanged total tourism revenue if the price elasticity equals unity.

Understanding demand elasticities is very important for tourism policymakers and planners in the destinations. For tourism demand forecasters, correctly identifying the determinants of tourism demand and appropriately specifying the tourism demand models are crucial for the generation of accurate forecasts of future tourism demand.

2

Traditional methodology of tourism demand modelling

2.1 Introduction

Most of the published studies on causal tourism demand models before the 1990s were classical regressions with ordinary least squares (OLS) as the main estimation procedure. The functional form of most of these models was single-equation, and either linear or power models. The data used in estimating tourism demand models are mainly time series, and since most of these time series, such as tourist expenditure, tourist arrivals, income (measured by personal disposable income or GDP), tourists' living costs in the destination, transport prices, and substitute prices are trended (non-stationary), the estimated tourism demand models have tended to have high R^2 values due to these common trends in the data. We shall see in Chapters 3 and 4 that statistical tests based on regression models with non-stationary variables are unreliable and misleading, and therefore any inference drawn from these models is suspect. Moreover, tourism demand models with non-stationary variables tend to cause the estimated residuals to be autocorrelated, and this invalidates OLS. The problem of autocorrelation in tourism demand models has normally been dealt with by employing the Cochrane–Orcutt (1949) iterative estimation procedure. However, the use of the Cochrane–Orcutt procedure diverts attention from searching for the correctly specified model (autocorrelation normally indicates model misspecification).

The organisation of this Chapter is as follows. Section 2.2 reviews the traditional tourism demand modelling methodology and presents the commonly used criteria for model selection. Section 2.3 discuses the limitations of the traditional approach to tourism demand modelling.

2.2 Traditional methodology of tourism demand modelling

The traditional tourism demand modelling methodology proceeds with the following steps: *formulate hypotheses based on demand theory*; *decide model's functional form*; *collect data*; *estimate the model*; *test hypotheses*; *and generate forecasts or evaluate policies.*

2.2.1 Hypothesis formulation

Although O'Hagan and Harrison (1984), Syriopoulos and Sinclair (1993) and Smeral *et al.* (1992) examine tourism demand based on the complete demand system approach as exemplified by Stone (1954), Theil (1965) and Deaton and Muellbauer (1980), most studies which have appeared in the literature are single-equation models based on the demand theory discussed in Chapter 1. This demand theory suggests that the optimal choice of consumer goods depends on consumers' income and the prices of the goods. In case of the demand for tourism, the choice of a tourism destination is related to the relative price of tourism products in that destination compared to alternative destinations and income in tourism-generating countries. For example, if tourists in a particular origin face a choice of two destinations, the tourists' demand functions given these two destinations can be written as

$$Q_1 = f_1(P_1, P_2, Y) \tag{2.1}$$

$$Q_2 = f_2(P_2, P_1, Y) \tag{2.2}$$

Equations (2.1) and (2.2) are simplified versions of specification (1.1), where the consumer tastes and marketing variables have been excluded. These two equations state that the demand for tourism to destination 1/destination 2 is related to the price of tourism in the two destinations concerned (P_1 and P_2) as well as the income level of potential tourists from the origin country (Y). The following hypotheses related to the above functions can be formulated based on economic theory.

Hypothesis I. The Engel curve suggests that if the price of tourism is held constant, an increase in tourists' income will result in an increase in the demand for tourism to both destinations provided that tourism is a normal or necessary good. Therefore, income in the origin country has a positive effect on demand for tourism to both destinations.

Hypothesis II. If the price of tourism in destination 1 increases while the price of tourism in destination 2 and consumers' income in the origin country remain unchanged, tourists will 'switch' from going to destination 1 to destination 2, and therefore the demand for tourism to destination 1 will decrease. This is known as the substitution effect and the substitution effect always moves in the opposite direction to the price changes.

Hypothesis III. With respect to the demand for tourism to destination 1, the effect of a price change in destination 2 can have either a positive or negative effect. If destination 2 is a substitute for destination 1 the demand for tourism to destination 1 will move in the same direction as to the price change in destination 2. On the other hand, if tourists tend to travel to the two destinations together, i.e., the destinations are complementary to each other, tourism demand to one destination will move in the opposite direction to the change in price of tourism in the other.

According to these hypotheses, Equations (2.1) and (2.2) can now be written as:

$$Q_1 = f_1(\underset{-}{P_1}, \underset{+/-}{P_2}, \underset{+}{Y}) \tag{2.3}$$

$$Q_2 = f_1(\underset{-}{P_2}, \underset{+/-}{P_1}, \underset{+}{Y}) \tag{2.4}$$

where the minus and plus signs represent the negative and positive effects of the explanatory variables on tourism demand. Equations (2.3) and (2.4) are theoretical tourism demand models which can be estimated using data on Q_i, P_i and Y (where $i = 1,2$), and hypotheses I, II and III can be statistically tested (the discussion on hypothesis testing is given in Section 2.2.5).

2.2.2 Model specification

Although consumer demand theory suggests possible relationships between tourism demand variables, it does not say anything about the precise functional form of the demand model. A researcher has to select an appropriate functional form for estimation, and the functional form is normally decided according to the ease of estimation and interpretation. In Chapter 1 we presented two commonly used functional forms in tourism demand analysis – the linear and power functions. The linear functional form is easy to estimate using OLS, while the power model is linear in logarithms (also known as the double log or log–log model) and has the additional advantage that the estimated parameters can be interpreted as demand elasticities. Other types of functional form include the log linear model ('log' on the left-hand side of the equation and linear on the right), and the linear log model (linear on the left-hand side of the equation and 'log' on the right).

Functional form selection involves choosing the particular type of model that best fits the data, with the estimated error term satisfying the following conditions: $E(\varepsilon_i) = 0$, $Var(\varepsilon_i) = \sigma^2$, $Cov(\varepsilon_i, \varepsilon_j) = 0$ and ε_i is normally distributed, i.e., $\varepsilon_i \sim N(0, \sigma^2)$.

2.2.3 Data collection

To estimate a tourism demand model, i.e., to obtain the numerical values of the α parameters in Equations (1.2) and (1.9), we have to use actual data. The data can be obtained from either primary sources or secondary sources. Three types of data are normally used for model estimation: they are *time series*, *cross-sectional* and *pooled (the combination of time series and cross-sectional) data*.

Time series data are observed at regular time intervals, such as monthly, quarterly or annually. Most pre-1990s tourism demand models were estimated using time series data without testing the stationarity properties of the data, and therefore have been criticised for generating spurious regression relationships. The spurious regression problem is discussed in detail in subsequent chapters.

Cross-sectional data are data on one or more variables collected at the same point of time, such as the data collected from the consumer expenditure surveys in the UK and the USA. The use of cross-sectional data has been very popular in modelling demand for non-durable commodities (food, clothing, meat, tobacco, etc.), but published studies on tourism demand analysis using cross-sectional data are rare. Exceptions are Quandt and Young (1969) and Mak *et al.* (1977). A major disadvantage of using cross-sectional data is that the estimated residuals of the demand model tend to suffer from the problem of heteroscedasticity. The weighted least squares (WLS) procedure has to be used to correct for heteroscedasticity and this in turn can cause problems if the weights are incorrectly applied.

Pooled data combine both time series and cross-sectional components. Pooled data (also called panel or longitudinal data) are data in which the information on cross-sectional units (say, households, firms or tourism-receiving/generating countries) is observed over time. One of the advantages of using panel data in tourism demand modelling is that it takes both temporal and cross-unit variations into account. Another advantage of panel data over pure time series or cross-sectional data sets is the large number of observations and the consequent increase in the number of degrees of freedom. This reduces collinearity and improves the efficiency of the estimates.

Carey (1991), Tremblay (1989), Yavas and Bilgin (1996) and Romilly *et al.* (1998) pool cross-sectional and time series data in order to estimate tourism demand elasticities with respect to income, exchange rates, relative prices, transport costs and social explanatory variables for various country groupings. Chapter 8 of this book examines in detail the use of pooled data in modelling tourism demand.

The use of panel data is, however, not unproblematic, since the choice of an appropriate model depends *inter alia* on the degree of homogeneity of the intercept and slope coefficients, and the extent to which any individual cross-section effects are correlated with the explanatory variables. If the homogeneity condition for the regression coefficients across individual units is not satisfied, the panel data approach cannot be used.

Data sources. Tourism demand data are available from various sources. The World Tourism Organization (WTO) publishes figures on tourist arrivals and tourism receipts for most countries broken down by country of origin in the *Yearbook of Tourism Statistics*. The International Monetary Fund (IMF) publishes consumer price index (CPI), exchange rate and gross domestic product (GDP) figures for most countries in the world in *International Financial Statistics* (monthly and annually). Some additional destination country data such as visitor nights and length of stay may be obtained from national tourist organisations. Data on personal disposable income, consumers' expenditure and population size are generally published by national statistical offices, e.g., the Central Statistical Office (CSO) in the UK, Statistics Canada, the Bureau of Economic Analysis (BEA) in the USA and the National Statistical Bureau of China. For a full list of data sources related to tourism see BarOn (1989).

Demand analysis at the firm/household level requires company or survey data. Company data cannot be easily accessed due to problems of confidentiality, and the quality of survey data depends on whether the questionnaires are correctly designed and implemented.

Whatever data source is used, one has to bear in mind that the quality of the data may be problematic for one or more of the following reasons. First, most tourism demand analyses are carried out at the national/regional level. The aggregated data may not reflect correctly individual consumers' behaviour. Second, tourism and economic data are non-experimental in nature and therefore observation errors in the data-compiling process will be carried over to the modelling process. Third, survey data collected by means of questionnaires may only partially reflect the characteristics of the whole population; in particular, when the response rate is low, the estimated demand model is likely to suffer from the problem of small sample bias. Fourth, when analysing inbound or outbound tourism demand for a country, the data needed may be compiled in different countries, and given that data definitions and collection methods differ from country to country, forecasting performance comparisons across models can be difficult.

2.2.4 Model estimation

After obtaining the data on the tourism demand variables, the tourism demand model can then be estimated. In traditional tourism demand analysis, the main method of estimation is OLS, and sometimes this method is augmented by the Cochrane–Orcutt procedure if autocorrelated residuals are identified. Lim (1997) reviewed 100 publications on tourism demand modelling during the period 1961–94, and found that more than 70% of these used OLS. The estimation method discussed in this section therefore concentrates on OLS only.

For illustrative purpose, we specify a four-variable tourism demand model (2.5). A subscript t is attached to every variable to indicate that the model is estimated using time series data.

$$Q_{it} = \alpha_0 + \alpha_1 P_{it} + \alpha_2 P_{st} + \alpha_3 Y_t + \varepsilon_{it} \tag{2.5}$$

The OLS procedure is used to obtain the numerical values of the α parameters in Equation (2.5). In order to obtain valid estimates of the parameters, the following assumptions have to be made:

(1) $E(Q_{it}) = \alpha_0 + \alpha_1 P_{it} + \alpha_2 P_{st} + \alpha_3 Y_t$, i.e., the expected value of Q_{it} depends on the values of the explanatory variables and the unknown α parameters. This is equivalent to $E(\varepsilon_{it}) = 0$.

(2) $Var(Q_{it}) = Var(\varepsilon_{it}) = \sigma^2$. This states that the sample variance of Q_{it} or the variance of the error term remains constant over time. If this assumption does not hold, the model suffers from the problem of heteroscedasticity.

(3) $Cov(Q_{it}, Q_{is}) = Cov(\varepsilon_{it}, \varepsilon_{is}) = 0$. This assumes that any two observations on the dependent variable or the residuals are not correlated. Violation of this assumption results in a model with autocorrelated residuals.

(4) It is also assumed that the values of the dependent variable are normally distributed about their mean, i.e., $Q_{it} \sim N(\alpha_0 + \alpha_1 P_{it} + \alpha_2 P_{st} + \alpha_3 Y_t, \sigma^2)$, which is equivalent to $\varepsilon_{it} \sim N(0, \sigma^2)$.

(5) The values of the explanatory variables are known and there is no linear relationship among the explanatory variables. If this condition is not met, the model suffers from the problem of multicollinearity.

OLS minimises the sum of squared residuals (SSR) of Equation (2.5):

$$SSR = \sum_{t=1}^{T} (Q_{it} - \hat{Q}_{it})^2 = \sum_{t=1}^{T} (Q_{it} - \hat{\alpha}_0 - \hat{\alpha}_1 P_{it} - \hat{\alpha}_2 P_{st} - \hat{\alpha}_3 Y_t)^2 \tag{2.6}$$

where $\hat{Q}_{it}$ is the estimated value of Q_{it}.

To find the OLS estimates of the α parameters that minimise (2.6), the partial derivatives of Equation (2.6) with respect to $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\alpha}_3$, are taken:

$$\frac{\partial SSR}{\partial \hat{\alpha}_0} = 2(T\hat{\alpha}_0 + \hat{\alpha}_1 \Sigma P_{it} + \hat{\alpha}_2 \Sigma P_{st} + \hat{\alpha}_3 \Sigma Y_t - \Sigma Q_{it})$$

$$\frac{\partial SSR}{\partial \hat{\alpha}_1} = 2(\hat{\alpha}_0 \Sigma P_{it} + \hat{\alpha}_1 \Sigma P_{it}^2 + \hat{\alpha}_2 \Sigma P_{st} P_{it} + \hat{\alpha}_3 \Sigma Y_t P_{it} - \Sigma Q_{it} P_{it})$$

$$\frac{\partial SSR}{\partial \hat{\alpha}_2} = 2(\hat{\alpha}_0 \Sigma P_{st} + \hat{\alpha}_1 \Sigma P_{it} P_{st} + \hat{\alpha}_2 \Sigma P_{st}^2 + \hat{\alpha}_3 \Sigma Y_t P_{st} - \Sigma Q_{it} P_{st})$$

$$\frac{\partial SSR}{\partial \hat{\alpha}_3} = 2(\hat{\alpha}_0 \Sigma Y_t + \hat{\alpha}_1 \Sigma P_{it} Y_t + \hat{\alpha}_2 \Sigma P_{st} Y_t + \hat{\alpha}_3 \Sigma Y_t^2 - \Sigma Q_{it} Y_t)$$

Setting the above equations equal to zero, we have

$$\begin{aligned}
&T\hat{\alpha}_0 + \hat{\alpha}_1 \Sigma P_{it} + \hat{\alpha}_2 \Sigma P_{st} + \hat{\alpha}_3 \Sigma Y_t = \Sigma Q_{it} \\
&\hat{\alpha}_0 \Sigma P_{it} + \hat{\alpha}_1 \Sigma P_{it}^2 + \hat{\alpha}_2 \Sigma P_{st} P_{it} + \hat{\alpha}_3 \Sigma Y_t P_{it} = \Sigma Q_{it} P_{it} \\
&\hat{\alpha}_0 \Sigma P_{st} + \hat{\alpha}_1 \Sigma P_{it} P_{st} + \hat{\alpha}_2 \Sigma P_{st}^2 + \hat{\alpha}_3 \Sigma Y_t P_{st} = \Sigma Q_{it} P_{st} \\
&\hat{\alpha}_0 \Sigma Y_t + \hat{\alpha}_1 \Sigma P_{it} Y_t + \hat{\alpha}_2 \Sigma P_{st} Y_t + \hat{\alpha}_3 \Sigma Y_t^2 = \Sigma Q_{it} Y_t
\end{aligned}$$

These equations are called *normal equations*. Although it is possible to solve these normal equations algebraically for $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\alpha}_2$ and $\hat{\alpha}_3$, the notation can be very complex without using matrices. Standard computer programs do all the calculations once the data on the dependent and independent variables have been entered. If the five assumptions made earlier hold, the parameter estimates generated by OLS ($\hat{\alpha}$) are the best linear unbiased estimates (BLUE) of α.

The residuals and estimated values of the dependent variables are obtained as

$$\hat{\varepsilon}_{it} = Q_{it} - \hat{\alpha}_0 - \hat{\alpha}_1 P_{it} - \hat{\alpha}_2 P_{st} - \hat{\alpha}_3 Y_t \tag{2.7}$$

$$\hat{Q}_{it} = \hat{\alpha}_0 + \hat{\alpha}_1 P_{it} + \hat{\alpha}_2 P_{st} + \hat{\alpha}_3 Y_t = Q_{it} - \hat{\varepsilon}_{it} \tag{2.8}$$

and the variance of the residuals $\hat{\sigma}^2_{it}$ is calculated from

$$s^2_{it} = \hat{\sigma}^2_{it} = \frac{\Sigma\,\hat{\varepsilon}^2_{it}}{T-k} \tag{2.9}$$

where T is the number of observations used in the estimation and k is the number of estimated regression coefficients including the constant term.

OLS minimises the sum of squares of the residuals (SSR) and ensures that the estimated regression equation is the best in terms of the model's 'fit' to the data. However, the 'best fit' does not necessarily mean that the model explains the variations in the dependent variable well, and therefore a statistic that measures the model's explanatory power over the data, i.e., the 'goodness of fit', is needed. A frequently used measure of goodness of fit in traditional econometric methodology is the *coefficient of determination*, and it is defined as

$$R^2 = 1 - \frac{\Sigma\,\hat{\varepsilon}^2_{it}}{\Sigma\,(Q_{it} - \overline{Q})^2} \tag{2.10}$$

where $\overline{Q}$ is the mean of Q_{it}.

R^2 measures the proportion of the total variation in Q_{it} that can be explained by the estimated model. Since R^2 is a proportion, the value of R^2 must lie between zero and one. A value of one suggests a perfect fit while a value of zero indicates that the model does not explain any variation in the dependent variable at all.

As we can see from formula (2.10), the R^2 value is determined by the sum of squared residuals, $\Sigma\,\hat{\varepsilon}^2_{it}$, and the quantity $\Sigma\,(Q_{it} - \overline{Q})^2$. The value of $\Sigma\,(Q_{it} - \overline{Q})^2$ is a constant if the sample of Q_{it} is fixed. Therefore, the R^2 value is mainly determined by the value of $\Sigma\,\hat{\varepsilon}^2_{it}$. It is easy to show that adding more explanatory variables to Equation (2.5) will reduce the value of $\Sigma \hat{\varepsilon}_{it}^2$, and hence increase the value of R^2, even if the added variables are not actually related to the tourism demand variable, Q_{it}. This problem is partially overcome by using 'adjusted' R^2, which takes into account the number of explanatory variables present in the model. Adding unimportant or irrelevant variables in the manner of 'data mining' is penalised. The 'adjusted' R^2 is calculated as

$$\overline{R}^2 = 1 - \frac{T-1}{T-k}(1 - R^2) \tag{2.11}$$

Although the addition of a variable leads to a gain in R^2, it also leads to a loss of one degree of freedom. Formula (2.11) is a better measure of the goodness of fit because it allows for the trade-off between an increase in R^2 and the loss of degrees of freedom.

The majority of published empirical investigations in the area of tourism demand modelling and forecasting uses R^2 or $\overline{R}^2$ as the main criterion for model selection. Other criteria used in the model selection process (although not featured often in tourism demand modelling) include Akaike's (1974) information criterion (AIC) and Schwarz's (1978) Bayesian criterion (SBC), and these are calculated as:

$$AIC = \ln \frac{\Sigma \hat{\varepsilon}_{it}^2}{T} + \frac{2k}{T} \tag{2.12}$$

$$SBC = \ln \frac{\Sigma \hat{\varepsilon}_{it}^2}{T} + \frac{k \ln T}{T} \tag{2.13}$$

where $\Sigma \hat{\varepsilon}_{it}^2$ is the estimated sum of squared residuals from Equation (2.6), and k and T are the number of estimated parameters and the total number of observations used in the estimation, respectively.

Ideally, we would like a model to have a high $\overline{R}^2$ and low values for the AIC and SBC. But there is always a possibility that one model is superior to another under one criterion and inferior under another. For example, the SBC penalises model complexity more heavily than does the AIC and this may lead to contradictory conclusions.

2.2.5 Hypothesis testing

Apart from estimating unknown parameters in a tourism demand model, testing hypotheses about the parameters in the model is also an important part of the empirical investigation.

Testing the significance of a single coefficient. From Section 2.2.1, we know that hypotheses are formulated according to economic theory, and hypothesis-testing statistics are used to provide support for the theoretical hypotheses. For example, as far as Equation (2.5) is concerned, demand theory suggests the following relationships: $\alpha_1 < 0$; $\alpha_2 \neq 0$ and $\alpha_3 > 0$. A commonly used statistic to test these hypotheses is the *t statistic* which is calculated from:

$$t = \frac{\hat{\alpha}_i}{SE(\hat{\alpha}_i)} \tag{2.14}$$

where $\hat{\alpha}_i$ is the estimated regression coefficient and $SE(\hat{\alpha}_i)$ is the standard error of the coefficient. The number of degrees of freedom for the t statistic is $T - k$, where T is the number of observations and k is the number of coefficients in the regression model including the constant term.

In the cases of $\alpha_1 < 0$ and $\alpha_3 > 0$, a one-tailed t statistic is used, while for $\alpha_2 \neq 0$ a two-tailed statistic is appropriate if it is not clear whether two destination countries are regarded as substitutes or complements. The decision on whether a one-tailed or two-tailed statistic should be used is based on whether you have a strong belief that the regression coefficient has a positive or a negative value. The sign of the estimated coefficient is easy to observe, but the statistical significance of the coefficient needs to be tested by looking at whether or not the coefficient is significantly different from zero given the observed sign. Suppose we obtain a negative value for the estimated coefficient $\hat{\alpha}_1$, but we want to know whether the coefficient is significantly different from zero. The steps involved in this t test are as follows:

Step 1: Set up the null and alternative hypotheses H_0: $\alpha_1 < 0$ and H_1: $\alpha_1 = 0$;
Step 2: Calculate the t statistic based on Equation (2.14);
Step 3: Compare the calculated t statistic, given the degrees of freedom, with the corresponding one-tailed critical value at a specific level of significance, normally 5%;
Step 4: Reject H_0 accept H_1 if the calculated t is greater than the critical value, otherwise accept H_0 and reject H_1.

For a two-tailed t test, the steps are the same except that the critical values for inference are different.

Testing the joint significance of all coefficients. The t test is applied to examine whether the dependent variable Q_{it} is related to a particular explanatory variable. Another statistic that has been frequently reported in traditional tourism demand analyses is the F statistic, which is used to test the relevance of all the explanatory variables in the demand model. The null and alternative hypotheses now become:

H_0: $\alpha_1 = 0$, $\alpha_2 = 0$, $\alpha_3 = 0$

H_1: at least one of the αs is non-zero

If the null hypothesis is true, the joint effect of the explanatory variables provides no explanation for Q_{it}, and thus the estimated tourism demand model is of no value. If the alternative hypothesis is true, then the model has some explanatory power. However, the alternative hypothesis does not suggest which of the explanatory variables is relevant. Therefore, this test has to be used in conjunction with the t test.

The F test is calculated based on the following formula:

$$F(k-1, T-k) = \frac{(SSR_2 - SSR_1)/(k-1)}{SSR_1/(T-k)} \tag{2.15}$$

where SSR_1 is the sum of squared residuals from Equation (2.6), SSR_2 is the sum of squared residuals from the regression equation $Q_{it} = \alpha_0 + u_{it}$, and $(k-1,\ T-k)$ are the degrees of freedom.

The calculated value of this statistic is compared with the critical value from the F distribution table. If the calculated F value is smaller than the critical value, the null hypothesis is accepted and the alternative rejected. If the calculated F is greater than the critical value, we accept the alternative and reject the null.

2.2.6 Forecasting and policy evaluation

Suppose that an estimated tourism demand model satisfies demand theory and passes all the hypothesis tests. The next step is to use the model for forecasting and policy analysis. Since the forecasting issue will be discussed in Chapter 9, we do not go into any details here with respect to how to generate forecasts using econometric models. However, one thing that is worth mentioning here is that in order to generate forecasts for the dependent variable, Q_{it}, we have to forecast the explanatory variables P_{it}, P_{st} and Y_t first, and the forecasts for these explanatory variables can be obtained either by time series extrapolation or from other forecasting sources. One of the advantages of econometric demand models is that they can be used for policy evaluations. For example, suppose that we estimate a tourism demand model in which, apart from the three explanatory variables given in (2.5), a marketing variable is also found to be positively related to the dependent variable. This allows us to look at how tourism demand in a destination country responds to the changes in marketing expenditure by this country. If one finds that tourism demand is very responsive to marketing expenditure, the destination country should adopt relevant marketing policies to promote its tourism products in order to attract more tourists.

2.3 Failure of the traditional approach to tourism demand modelling

In the above section we summarised the traditional econometric approach to tourism demand modelling. From the discussions we can see that this methodology relies heavily on the five assumptions made in Section 2.2.3 being satisfied. Witt and Witt (1995, p. 458) argue that "in the light of the lack of testing to see whether or not the models are well specified and recent developments in econometric methodology which focus the specification of the dynamic structure of time series, the quality of the empirical results obtained is questionable". The traditional methodology also implicitly assumes that tourism demand data are stationary or at least trended stationary, but it has been shown that most tourism demand data are actually non-stationary, and this can lead to serious problems with statistics such as t, DW, and F, as well as R^2, which are reported in most of the empirical studies. The integrating property of the data and the modelling strategy related to non-stationary variables will be further discussed in Chapter 3.

Another serious problem with traditional tourism demand models is that the forecasting performance of these models has been poor in comparison with alternative specifications; in particular, they cannot even compete with the simplest time series models such as the naïve no-change model (see, for example, Martin and Witt, 1989; Witt and Witt, 1992). One possible explanation for this could be that the traditional tourism demand method does not take into account both the long-run cointegrating relationship and short-run dynamics in the estimation of the models.

Modern econometric methodologies such as cointegration, error correction models, vector autoregressive models, time varying parameter models and panel data approaches can be employed to overcome the problems associated with the traditional single-equation demand models. These methodologies will be discussed, and the estimation and forecasting results compared, in subsequent chapters.

3

General-to-specific modelling

3.1 Introduction

The traditional approach to tourism demand modelling starts by constructing a simple model that is consistent with demand theory. Such a model is then estimated and tested for statistical significance. The estimated model is expected to have a high R^2, and the coefficients are expected to be both 'correctly' signed and statistically significant (usually at the 5% level). In addition, the residuals from the estimated model should be properly behaved, i.e., they should be normally distributed with zero mean and constant variance.

As an example, consider the following. A simple outbound tourism demand model for country/region j may be specified as:

$$q_{it} = \alpha_0 + \alpha_1 y_t + \alpha_2 p_{it} + u_{it} \tag{3.1}$$

where q_{it} is tourism demand, which is normally measured by total/per capita visits/expenditure to destination i; y_t is a measure of income in the tourism-generating country/region j; and p_{it} is a price variable which represents living costs for tourists in destination i. The subscript j is omitted for simplicity. Following the traditional approach, Equation (3.1) is estimated using OLS. If the results are 'unsatisfactory', either in terms of a low R^2, or 'wrongly' signed/insignificant coefficients, the model is then re-specified by introducing new explanatory variables, using a different functional form, or selecting a different estimation method if the residuals exhibit heteroscedasticity, autocorrelation or lack of normality. For example, the exchange rate between the origin and destination countries, tourists' living costs in substitute destinations, or event dummy variables may be added to the model. Alternatively, a lagged dependent variable could be introduced if it is believed that a partial adjustment process of actual to desired tourism demand is in operation. Although such a procedure starts from a relatively simple specification, the final model, that is both theoretically sound and statistically acceptable, may be very complex. This methodology is called the *specific-to-general approach*.

The specific-to-general modelling approach is often criticised for its excessive data mining, since researchers normally publish only their final models which often appear to be acceptable on both theoretical and statistical grounds, with the intermediate modelling process omitted. With this methodology, the same set of data is simply fitted to a range of potential models and the same statistics are calculated repeatedly until an equation that fits the researcher's a priori beliefs is discovered. It is possible, therefore, that different researchers equipped with the same data set and statistical tools could end up with totally different model specifications. The problems associated with the traditional approach to tourism demand modelling are discussed in Witt and Witt (1995).

3.2 General-to-specific modelling

The *general-to-specific* modelling approach was initiated by Sargan (1964), and subsequently developed by Davidson *et al.* (1978), Hendry and von Ungern-Sternberg (1981) and Mizon and Richard (1986). In contrast to the specific-to-general modelling methodology, the general-to-specific approach starts with a general model which contains as many variables as possible suggested by economic theory. According to this framework, if a dependent variable y_t is determined by k explanatory variables, the data-generating process (DGP) may be written as an autoregressive distributed lag model (ADLM) of the form:

$$y_t = \alpha + \sum_{j=1}^{k} \sum_{i=0}^{p} \beta_{ji} x_{jt-i} + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t \tag{3.2}$$

where p is the lag length, which is determined by the type of data used. As a general guide, $p = 1$ for annual data, $p = 4$ for quarterly data, $p = 6$ for bimonthly data and $p = 12$ for monthly data. However, the lag lengths of

Table 3.1 Variations of the autoregressive distributed lag model.

Model	Restrictions	Equation[1]
1. ADLM	None	$y_t = \beta_0 x_t + \beta_1 x_{t-1} + \phi_1 y_{t-1} + \varepsilon_t$
2. Static	$\beta_1 = \phi_1 = 0$	$y_t = \beta_0 x_t + \varepsilon_t$
3. Autoregressive (AR)	$\beta_0 = \beta_1 = 0$	$y_t = \phi_1 y_{t-1} + \varepsilon_t$
4. Growth rate	$\phi_1 = 1, \beta_0 = -\beta_1$	$\Delta y_t = \beta_0 \Delta x_t + \varepsilon_t$
5. Leading indicator	$\beta_0 = \phi_1 = 0$	$y_t = \beta_1 x_{t-1} + \varepsilon_t$
6. Partial adjustment	$\beta_1 = 0$	$y_t = \beta_0 x_t + \phi_1 y_{t-1} + \varepsilon_t$
7. Common factor	$\beta_1 = -\beta_0 \phi_1$	$y_t = \beta_0 x_t + \varepsilon_t,\ \varepsilon_t = \beta_1 \varepsilon_{t-1} + u_t$
8. Finite distributed lag	$\phi_1 = 0$	$y_t = \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$
9. Dead start	$\beta_0 = 0$	$y_t = \beta_1 x_{t-1} + \phi_1 y_{t-1} + \varepsilon_t$
10. Error correction	None	$\Delta y_t = \beta_0 \Delta x_t + (\beta_1 - 1)(y - Kx)_{t-1} + \varepsilon_t$

Note: [1]The constant term α is omitted for simplicity.
Source: This table is adapted from Hendry (1995, p. 232).

the time series may vary, and they are normally decided by experimentation. ε_t in Equation (3.2) is the error term which is assumed to be normally distributed with zero mean and constant variance, σ^2, i.e., $\varepsilon_t \sim N(0, \sigma^2)$.

For simplicity, let us assume that there are only two variables, y_t and x_t, involved in the general model and that the lag length $p = 1$. Equation (3.2) can now be simplified to give:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \phi_1 y_{t-1} + \varepsilon_t \tag{3.3}$$

Although the subsequent discussions are based on this simplified ADLM, the principles apply generally to the more complicated specifications.

With certain restrictions imposed on the parameters in Equation (3.3), a number of econometric models may be derived, and these models have been widely used in empirical studies. The models are presented in Table 3.1.

Static model

The static model states that the lagged dependent and lagged independent variables do not influence the dependent variable, i.e., the restrictions imposed on the coefficients of the general model are $\beta_1 = \phi_1 = 0$. Many early tourism demand models, such as Gray (1966), Artus (1972), Kwack (1972) and Loeb (1982), were static models in which the current value of tourism demand is related only to the current values of the explanatory variables. However, the error terms in static tourism demand models have generally been found to be highly autocorrelated, and this indicates that the demand relationships are likely to be spurious and that the normal *t* and *F* statistics are invalid. In attempting to solve the problem of spurious correlation, some researchers and practitioners, such as Witt (1980a, 1980b) and Little (1980), began to introduce dynamic effects, such as lagged values of the dependent and independent variables, into tourism demand models. Static regression has recently become popular again in the application of the Engle and Granger (1987) two-stage cointegration analysis to tourism demand modelling and forecasting (see Kim and Song, 1998 and Song *et al.*, 2000). The Engle–Granger two-stage cointegration approach requires that the static long-run equilibrium relationship is estimated first, and the disequilibrium errors are then incorporated into the corresponding short-run error correction model (this approach will be discussed in detail in the following two chapters).

Autoregressive (AR) model

Lagged dependent variables, but no independent variables, are involved in the AR model. The AR(1) process in Table 3.1 is a special case of the Box–Jenkins (1976) integrated autoregressive and moving average (ARIMA) representation. Such univariate time series models have been widely used for ex ante tourism demand forecasting, and they have been shown by researchers, such as Witt and Witt (1992), Kulendran (1996), and Kulendran and King (1997), to be powerful competitors to econometric forecasting. However, univariate time series analysis is not useful for understanding the causal relationships between tourism demand and its determinants. Indeed, the nature of univariate time series extrapolation makes policy evaluation impossible. However, when the data on tourism demand are limited, autoregressive modelling may be a very useful alternative for tourism demand forecasting.

Growth rate model

The growth rate specification requires that the following hypothesis regarding the coefficients in the general model hold: $\phi_1 = 1, \beta_0 = -\beta_1$. This type of model is an early attempt to deal with the problem of nonsensical correlation caused by trended economic variables. However, although the growth rate model overcomes the problem of spurious regression results, the long-run properties of the economic model are lost due to data differencing.

Leading indicator model

In the case of the leading indicator model, the restrictions on the general model are: $\beta_0 = \phi_1 = 0$. The leading indicator model is a useful tool for macroeconomic forecasting. The model does not normally have behavioural underpinning, and the inclusion of leading indicators in the regression model is usually determined by trial and error. Turner *et al.* (1997) employed the leading indicator model to forecast tourism demand to Australia. One of the important conditions for accurate forecasts is that the coefficients of the leading indicators should be constant. If this condition is violated, the leading indicator model tends to produce poor forecasts, especially in periods of rapid economic change when accurate forecasts are crucial.

Partial adjustment model

The restrictions on the coefficients in the general model in the case of the partial adjustment model are: $\beta_1 = 0$. The partial adjustment model has been widely used in modelling macroeconomic activities that involve habit persistence and adaptive expectations processes, such as the consumption function based on the permanent income hypothesis, the demand for durable goods and aggregate investment determination. In tourism demand analysis, the partial adjustment model has also been used extensively in situations which involve habit persistence and the influences of social factors, such as cultural status, personal preferences and expectations. It has also been used to accommodate supply constraints. Tourism demand studies which use the partial adjustment model include Witt (1980a, 1980b), Kliman (1981) and Martin and Witt (1988).

*Common factor (*COMFAC*) model*

In order to meet the requirements of a *COMFAC* model, the coefficients of the general model have to satisfy: $\beta_1 = -\beta_0\phi_1$. The common factor model is also called the autoregressive error model. It is similar to the static model, but with an autoregressive error term. The *COMFAC* model is a statistical model without a well-established economic-theory basis (Hendry, 1995, p. 267). The tourism demand model developed by Lee, Var and Blaine (1996) is a *COMFAC* model.

Finite distributed lag model

Lagged dependent variables do not enter the regression in finite distributed lag models. Although the lag length of the finite distributed lag model presented in Table 3.1 is equal to 1, in practice the model could be specified with current values of the explanatory variables, and as many lagged explanatory variables as the data permit. There is little evidence to suggest that tourism demand can be represented by a finite distributed lag process.

Dead start model

In the dead start model, all current values of the independent variables are assumed to be irrelevant to the dependent variable. The dead start model is a partial-adjustment-type model. The model could arise for two possible reasons in tourism demand modelling. First, the model is a structural model that represents the decision-making process of economic agents. Second, the model derivation may be based on a given hypothesis; e.g., in macroeconomic modelling the Hall (1978) consumption model utilises the permanent income hypothesis together with rational expectations. Since the structure of the dead start model is similar to the partial adjustment process, and also incorporates leading indicators in the estimation, it is likely that this model will be a good contender in the model selection process in tourism demand analysis.

Error correction model

In the subsequent chapters, we shall show that not only does the error-correction model avoid the problem of spurious regression, but also it avoids the problems associated with the use of the simple growth rate model. In particular, the inclusion of the error correction term in Model 9 in Table 3.1 ensures that no information on the levels variables is ignored. The error correction model has been used successfully in many areas of economics since the mid-1980s. However, only recently has this type of model started to appear in the tourism literature. Tourism demand studies which use the error correction model include Kulendran (1996), Kulendran and King (1997), Kulendran and Witt (2001), Kim and Song (1998) and Song *et al.* (2000).

In order to decide which specification of the above models is appropriate in modelling tourism demand, statistical tests on the parameters of the general regression model (3.3) have to be carried out. These tests are known as restriction tests.

3.3 Tests of restrictions on regression parameters

In some cases, the tests of restrictions on parameters in the general model (3.3) are straightforward, while in other cases they are much more complicated. The null hypothesis of the restriction tests is that the restrictions imposed on the coefficients of the general model are true, i.e., these restrictions cannot be rejected at a given level of significance.

If a general tourism demand ADLM version of Equation (3.3) is specified with the dependent variable y_t being a measure of tourism demand and the explanatory variable x_t being, say, income, then the parameter restrictions in cases 6, 8 and 9 in Table 3.1 can be tested easily using the t statistic, while an F statistic would be appropriate for testing the restrictions in cases 2, 3, 4 and 5. For example, in order to test whether tourism demand may be modelled by a static process (case 2 in Table 3.1) against the general ADLM, the F test is carried out as follows:

(1) Estimate both the unrestricted general model (3.3) and the restricted static model using OLS.
(2) Obtain the residual sum of squares, SSR_0, from the general model, and the residual sum of squares, SSR_1, from the restricted (static) model. It should be noted that SSR_0 is normally smaller than SSR_1, since the unrestricted model (which has more explanatory variables) is likely to explain more variations in the dependent variable than the restricted models.
(3) Calculate the F statistic based on

$$F(r, n-k) = \frac{(SSR_1 - SSR_0)/r}{SSR_0/(n-k)} \tag{3.4}$$

where r is the number of restrictions, n is the number of observations, and k is the number of explanatory variables (including the constant term) in the unrestricted equation. If the restrictions are valid, i.e., $\beta_1 = 0$, $\phi_1 = 0$ in this case, SSR_1 would not be significantly larger than SSR_0, and the calculated F statistic should approach zero.
(4) Decide whether the restrictions should be rejected or accepted by comparing the calculated F statistic with $(r, n-k)$ degrees of freedom with the critical value. If the calculated F statistic is greater than the

critical value, then the null hypothesis that the restrictions are true should be rejected. The F test can be carried out easily in most econometric packages (e.g., Eviews 3.1 and Microfit 4.0).

The F test loses its power when the restrictions are more complicated or non-linear, such as in case 7 in Table 3.1. Alternative tests have been developed for both linear and non-linear restrictions. Three of the commonly used tests are the Wald, Likelihood Ratio (LR) and Lagrange Multiplier (LM) tests. (Detailed descriptions of these tests are given in Thomas (1997, pp. 355–60).) The inference procedure for each of these tests is the same as that for the F test; i.e., the calculated statistic based on the sample data is compared with the critical value of the statistic at a certain level of significance, with a larger value leading to rejection of the null hypothesis.

The error correction model is a re-parameterisation of the general ADLM, and it incorporates both short-run and long-run relationships in the modelling process. To test for the existence of an error correction mechanism, the presence of a long-run cointegration relationship needs to be tested for first. These topics are discussed in the next chapter.

3.4 Diagnostic checking

Restriction tests are useful in selecting the model type, but the final model will have had to be subject to rigorous statistical checking in order to determine its statistical acceptability. In practice, the final model used for policy evaluation and forecasting should not exhibit autocorrelation, heteroscedasticity, or structural instability. The final model should also be correctly specified in terms of functional form, and the independent variables should be exogenous. Moreover, the final model should be able to compete with all rival models, i.e., the final model should encompass all other models which have been used to explain the same variables. If any of these conditions is not met, the model is mis-specified and should not be used for policy evaluation and forecasting. Diagnostic statistics for testing model specification are available in most econometric and forecasting packages. The most commonly used diagnostic statistics are now described.

Testing for autocorrelation

(a) The Durbin–Watson (DW) statistic. The DW statistic is the most widely used test for detecting the problem of autocorrelation in the regression

residuals. The statistic is defined as:

$$DW = \frac{\sum_{t=2}^{n} (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2} \tag{3.5}$$

where $\hat{\varepsilon}_t$ is the residual from the estimated regression equation.

The statistic can be used to test for the presence of first-order autocorrelation in the regression. If there is no autocorrelation, the value of the DW statistic should be approximately 2. A DW statistic of 0 would suggest perfect positive autocorrelation, while a DW value of 4 would indicate perfect negative autocorrelation. Although the DW statistic is standard output for all regression programs, it has severe limitations. The first limitation is the inconclusive region which varies according to the size of the sample. The second is its inability to detect higher-order autocorrelation. The third limitation is that the DW statistic is biased towards 2 when a lagged dependent variable is included as an explanatory variable in the model.

(b) The Lagrange Multiplier (LM) test. This is also known as the Breusch–Godfrey test and was developed by Breusch (1978) and Godfrey (1978). The LM statistic is not limited to testing for the presence of first-order autocorrelation, and is still valid even if a lagged dependent variable is present on the right-hand side of the regression model. Due to these two important characteristics, the LM test is far more generally applicable than the DW test. The calculation of the test is based on an auxiliary equation of the form

$$\hat{\varepsilon}_t = \alpha + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_k X_{kt} + \rho_1 \hat{\varepsilon}_{t-1} + \rho_2 \hat{\varepsilon}_{t-2} + \ldots + \rho_p \hat{\varepsilon}_{t-p} + u_t \tag{3.6}$$

where the X_{it}s are explanatory variables, the β_is and ρ_js are parameters, and the $\hat{\varepsilon}_{t-j}$s are the lagged residuals from the estimated regression model. Under the null hypothesis of no autocorrelation, $H_0\!:\rho_1 = \rho_2 = \ldots = \rho_p = 0$, the test statistic is nR^2, where n is the sample size and R^2 is calculated from Equation (3.6). In large samples, the statistic has a χ^2 distribution with p degrees of freedom. If the value of nR^2 exceeds the critical value of χ^2, this suggests the existence of autocorrelation.

Testing for heteroscedasticity

(a) The Goldfeld–Quandt test. The underlying null hypothesis is that the residuals are homoskedastic, and this is tested against the alternative that

the variance of the residuals increases as the values of one of the explanatory variables increase. The test is carried out as follows:

(i) Identify the explanatory variable that is related to the variance of the residuals and re-order the observations of the explanatory variable from the largest value to the smallest one.
(ii) Divide the re-ordered observations into two equal-sized subsamples by omitting *c central* observations.
(iii) Calculate the OLS regression for each of the subsamples and obtain the residual sum of squares from each of these two regressions.

The Goldfeld–Quandt statistic is an F statistic based on the ratio

$$F[1/2(n-c)-k, 1/2(n-c)-k] = \frac{\Sigma\,\hat{\varepsilon}_1^2}{\Sigma\,\hat{\varepsilon}_2^2} \tag{3.7}$$

where $\Sigma\,\hat{\varepsilon}_1^2$ is the residual sum of squares from the subsample containing the larger values of the explanatory variable and $\Sigma\hat{\varepsilon}_2^2$ is the residual sum of squares from the subsample containing the smaller values of the explanatory variable. The numbers in the square brackets are the degrees of freedom. If there is no heteroscedasticity in the residuals, the calculated F statistic should not exceed the critical value at the appropriate level of significance.

A limitation of the Goldfeld–Quandt test is that different choices of the number of central observations may lead to different results.

(b) The White test. This was developed by White (1980). Suppose that we have a multiple regression model with two explanatory variables of the form

$$Y_t = \beta_1 + \beta_2 X_{1t} + \beta_3 X_{2t} + \varepsilon_t \tag{3.8}$$

In order to test whether the residual ε_t has constant variance or not, the following auxiliary equation is estimated

$$\hat{\varepsilon}_t^2 = \alpha_1 + \alpha_2 X_{1t} + \alpha_3 X_{2t} + \alpha_4 X_{1t}^2 + \alpha_5 X_{2t}^2 + \alpha_6 X_{1t} X_{2t} + u_t \tag{3.9}$$

where $\hat{\varepsilon}_t$ is the estimated residual from Equation (3.8). If the regression model has more than two explanatory variables, the $\hat{\varepsilon}_t^2$ should be regressed against all the explanatory variables together with their squares and cross-products.

The test statistic is equal to nR^2, where R^2 is calculated from the OLS estimation of Equation (3.9). The statistic has a χ^2 distribution, with degrees of freedom equal to the number of regressors excluding the intercept. The null hypothesis that there is no heteroscedasticity is rejected if the calculated White statistic exceeds the critical value of χ^2.

(c) Testing for an autoregressive conditional heteroscedasticity (ARCH) process. This test was developed by Engle (1982). Instead of relating $\hat{\varepsilon}_t^2$ to a vector of explanatory variables as in the White test, $\hat{\varepsilon}_t^2$ is assumed to depend on past squared errors, $\hat{\varepsilon}_{t-1}^2, \hat{\varepsilon}_{t-2}^2, \ldots, \hat{\varepsilon}_{t-p}^2$. The ARCH process is autoregressive in the second moment, and the ARCH test is calculated using an auxiliary equation of the form

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \ldots + \alpha_p \hat{\varepsilon}_{t-p}^2 + u_t \tag{3.10}$$

The test statistic has a χ^2 distribution with p degrees of freedom. The most common form of the ARCH test is the first-order autoregressive model in which $p = 1$.

Testing for normality

The Jarque–Bera (J–B) test. The normality test was developed by Jarque and Bera (1980). If the normality assumption regarding the residuals does not hold, t and F statistics are invalid in small samples. Jarque and Bera (1980) show that the normality property may be tested using the statistic

$$n\left[\frac{\mu_3^2}{6\mu_2^3} + \frac{(\mu_4/\mu_2^2 - 3)^2}{24}\right] \tag{3.11}$$

where

$$\mu_2 = \sum_{t=1}^{n} \hat{\varepsilon}_t^2/n, \quad \mu_3 = \sum_{t=1}^{n} \hat{\varepsilon}_t^3/n \quad \text{and} \quad \mu_4 = \sum_{t=1}^{n} \hat{\varepsilon}_t^4/n$$

are the second, third and fourth moments of the residuals, respectively. The J–B statistic has a χ^2 distribution with two degrees of freedom. Under the null hypothesis of normally distributed residuals, the third and fourth moments, μ_3 and μ_4, should take certain specified values. If the moments of the residuals depart sufficiently far from the expected values, i.e., the calculated J–B statistic exceeds the critical χ^2 value, then the hypothesis of normally distributed residuals is rejected.

Testing for mis-specification

The Ramsey RESET test. The Ramsey (1969) RESET test is designed to test for model mis-specification due to either the omission of important explanatory variables or incorrect choice of functional form. Consider a regression model with two explanatory variables as an example. The test involves three steps. The first step is to estimate the proposed regression model using OLS and to retain the estimated values of Y_t from Equation (3.12):

$$\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_{1t} + \hat{\beta}_3 X_{2t} \tag{3.12}$$

and the second step is to estimate the augmented equation of the form

$$Y_t = \beta_1 + \beta_2 X_{1t} + \beta_3 X_{2t} + \alpha_1 \hat{Y}_t^2 + \alpha_2 \hat{Y}_t^3 + \alpha_3 \hat{Y}_t^4 + u_t \tag{3.13}$$

Finally, the significance of the α parameters is tested using a standard restriction test, such as the F or Wald test. In carrying out the functional form test, we introduce, one by one, the estimated dependent variables with various powers into regression model (3.13) with the lowest powered variable first. If only $\hat{Y}_t^2$ is included, the significance of its coefficient may be tested using the t statistic. It should be noted that the RESET test is a general mis-specification test. The null hypothesis is that the model is correctly specified, but there is no specific alternative hypothesis. Therefore, rejection of the null hypothesis is merely an indication that the model is incorrectly specified, but how the model is mis-specified is beyond the concern of this test.

Testing for structural instability

For an econometric model to be able to produce accurate forecasts, the structure of the model should be constant over time, i.e., the values of the parameters of the model that represents the economic relationship should be the same for both the sample period and the forecasting period. This assumption is often violated by regime shifts, such as changes in economic policy, evolution of consumers' tastes, oil crises and political unrest. Another potential source of structural instability is model mis-specification, such as the omission of important explanatory variables and incorrect specification of the functional form. If structural instability is caused by one of the above factors, then the model has to be re-specified and tested until a stable specification is reached. Testing for structural stability may be carried out using two tests developed by Chow (1960), and also a recursive least squares procedure.

(a) The Chow parameter constancy test (or breakpoint test). This test examines whether there is a statistically significant difference between the

OLS regression residuals from two sub-samples, and the statistic is calculated as follows:

$$F_{Chow1} = \frac{(SSR_0 - SSR_1 - SSR_2)/k}{(SSR_1 + SSR_2)/(n_1 + n_2 - 2k)} \sim F(k, n_1 + n_2 - 2k) \tag{3.14}$$

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 sub-samples, and n_1 and n_2 are the numbers of observations in the first and second sub-sample periods, respectively. If the calculated F statistic is larger than the critical value, the null hypothesis of parameter constancy between the two sub-sample periods is rejected. In performing this test, the break point is assumed to be known.

(b) The Chow predictive failure test. In this second Chow test of structural instability, the equation which was estimated using the first n_1 observations is used to forecast the dependent variable for the remaining n_2 data points. If the differences between the predicted values and the actual values are large, this suggests that the structure of the model in the estimation period is not the same as that in the forecast period. The test is calculated from the OLS residuals using the following formula:

$$F_{Chow2} = \frac{(SSR_0 - SSR_1)/n_2}{SSR_1/(n_1 - k)} \sim F(n_2, n_1 - k) \tag{3.15}$$

The null hypothesis of structural stability between the two sub-samples is rejected if the test statistic (3.15) exceeds the critical F value. The difference between the Chow parameter constancy test and the Chow predictive failure test is that the latter does not involve estimating the regression for the second sub-sample. This can sometimes be useful when the second sub-sample is too short, such as when we evaluate the *ex post* forecasting performance of a regression model in which only a small number of observations is reserved for forecasting comparison purposes.

(c) The recursive least squares procedure. A critical feature of the two Chow tests is that we have to know exactly at which point we suspect the structural break takes place, but sometimes this is impossible as the change in structure may not happen suddenly, but rather evolve gradually over time. In such situations the Chow tests are not very useful. An alternative way of looking at structural instability is to employ the recursive least squares procedure.

In recursive least squares, the parameters of the model are first estimated using OLS with a small sub-sample of observations, $t = 1, 2, 3, \ldots, m$, where $m > k$. The sample period is then extended to $t = 1, 2, 3, \ldots, m + 1$ observations and the model is re-estimated. The procedure is repeated until all the sample observations are used up. At each step, the values of the estimated parameters are noted, and these values are plotted against time at the end of the estimation process. Under the assumption of constant structure, the plotted estimates should converge to a constant value. Any significant divergence of the recursive estimates from their mean values suggests structural instability of the model.

Testing for exogeneity

Although a more detailed discussion of exogeneity will take place in Chapter 6 when we examine the vector autoregressive (VAR) modelling approach, a brief introduction to the exogeneity test within a single equation framework is now provided. The test was developed by Wu (1973) and subsequently modified by Hausman (1978).

Consider a regression equation with two explanatory variables

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + u_t \tag{3.16}$$

Suppose we suspect that X_{2t} also depends on the current value of Y_t, i.e., X_{2t} is not exogenous to Y_t. If this is the case, the variable X_{2t} will be correlated with the error term u_t. Testing the null hypothesis that there is no correlation between X_{2t} and u_t is therefore a test of the exogeneity of the variable X_{2t}. The test is carried out as follows. First, regress X_{2t} against a number of instrumental variables (normally lagged Y_t, X_{1t} and X_{2t}):

$$X_{2t} = \alpha_0 + \alpha_i \sum_{i=1}^{p} X_{2t-i} + \gamma_i \sum_{i=1}^{p} X_{1t-i} + \lambda_i \sum_{i=1}^{p} Y_{t-i} + \varepsilon_t \tag{3.17}$$

Then estimate:

$$Y_t = \beta_0' + \beta_1' X_{1t} + \beta_2' X_{2t} + \beta_3' \hat{\varepsilon}_t + u_t \tag{3.18}$$

where $\hat{\varepsilon}_t$ is the estimated error term from Equation (3.17).

The hypothesis $\beta_3' = 0$ in Equation (3.18) is then tested using the t statistic. Rejection of the null hypothesis suggests that the variable X_{2t} is endogenous.

If it is suspected that two or more explanatory variables in a multiple regression model are endogenous, these variables should be regressed individually against the instrumental variables. The estimated residuals are then collected from each of these models and introduced into the proposed multiple regression model. In this case, the appropriate test is whether the variables in question are jointly exogenous, and a standard restriction test such as the F or Wald test can be used.

Encompassing test

The concept of encompassing is associated with Mizon and Richard (1986). Encompassing tests are particularly useful when a researcher has to choose only one model for policy evaluation and forecasting from a number of competing models. Encompassing means that 'a model should not just be able to explain existing data. It should also be able to explain why previous models could or could not do so. In particular, if different researchers using different models have come to different conclusions, a preferred model should be able to explain why this is the case' (Thomas, 1997, p. 362). For simplicity, we shall concentrate on the case where there are only two competing models, M1 and M2. M1 is said to encompass M2 if M1 can explain all the results of M2. For example, suppose a tourism demand model M1 contains a variable that captures some irregular movements in the dependent variable due to, say, policy changes. If an alternative model M2 does not include this variable, structural instability may occur in this second model, and M1 will be able to predict the structural instability of M2. If this is the case, we say that M1 encompasses M2.

Although a number of encompassing tests are available in the literature, only two of them are introduced here. Suppose we are interested in choosing one model from the following two competing models:

$$\text{M1:} \quad Y_t = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 X_{2t} + \alpha_3 X_{3t} + u_t \tag{3.19}$$

$$\text{M2:} \quad Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 Z_{1t} + \beta_3 Z_{2t} + e_t \tag{3.20}$$

The first test is known as the J-test. This test was developed by Davidson and MacKinnon (1981) and can be carried out as follows. First estimate M1 and M2 using OLS, and save the estimated values of $\hat{Y}_t$ from the models. Next, the following two models are estimated using OLS:

$$Y_t = \alpha'_0 + \alpha'_1 X_{1t} + \alpha'_2 X_{2t} + \alpha'_3 X_{3t} + \alpha'_4 \hat{Y}_{M2t} + u'_t \tag{3.21}$$

$$Y_t = \beta'_0 + \beta'_1 X_{1t} + \beta'_2 Z_{1t} + \beta'_3 Z_{2t} + \beta'_4 \hat{Y}_{M1t} + e'_t \tag{3.22}$$

where $\hat{Y}_{M2t}$ and $\hat{Y}_{M1t}$ are the estimated Y_ts from Equations (3.20) and (3.19), respectively. If the calculated t statistic associated with α'_4 in Equation (3.21) is insignificant, we say that M1 encompasses M2 (or M2 is nested in M1). If the calculated t statistic of β'_4 in Equation (3.22) is insignificant, this suggests that M2 encompasses M1 (or M1 is nested in M2). Although there is a possibility that the results will suggest that both M1 encompasses M2 and vice versa, in practice this is unlikely to happen.

The second test is simply termed Encompassing test. In the case of whether M1 encompasses M2, the test involves computing the F statistic for testing the restrictions $\gamma_1 = \gamma_2 = 0$ in the following equation:

$$Y_t = \alpha_0 + \varphi X_{1t} + \alpha_2 X_{2t} + \alpha_3 X_{3t} + \gamma_1 Z_{1t} + \gamma_2 Z_{2t} + u_t \tag{3.23}$$

where $\varphi = \alpha_1 + \beta_1$, and Z_{1t} and Z_{2t} are variables that cannot be expressed as exact linear combinations of the regressors of M1. Equation (3.23) combines the common variable, X_{1t}, in M2 and M1 as φX_{1t}. Rejection of the null hypothesis suggests that M1 encompasses M2. The test of whether M2 encompasses M1 involves testing the restrictions $\alpha_2 = \alpha_3 = 0$ in Equation (3.23).

3.5 Model selection

In summary, the general-to-specific modelling approach involves the following steps. First, a general demand model that has a large number of explanatory variables, including the lagged dependent and lagged explanatory variables, is constructed in the form of an ADLM. Economic theory suggests the possible variables to be included, and the nature of the data suggests the lag length. Second, the t, F, and Wald (or LR or LM as appropriate) statistics are used to test various restrictions in order to achieve a simple but statistically significant specification. Third, the normal diagnostic tests, such as those for autocorrelation, heteroscedasticity, functional form and structural instability, are carried out to examine whether the final model is statistically acceptable or not. Fourth, the final model can be used for policy evaluation or forecasting.

Thomas (1993) has summarised the various criteria for model selection within the framework of general-to-specific modelling. The criteria are: *consistent with economic theory*, *data coherency*, *parsimony*, *encompassing*, *parameter constancy* and *exogeneity*. The first criterion for model selection is that the final model should be consistent with economic theory. This is very important; e.g., in general we cannot use a demand model for policy evaluation and forecasting if the model has a negative income elasticity. Although such a model may be acceptable according to the diagnostic statistics, it should still be rejected because it invalidates a law of economics.

The data coherency criterion ensures that economic data also have a role to play in the determination of the structure of the final model. It implies that the preferred model should have been subject to rigorous diagnostic checking for mis-specification. The parsimony criterion states that simple specifications are preferred to complex ones. In the case of modelling tourism demand, if two equations have similar powers in terms of explaining the variation in the dependent variable, but one has six explanatory variables while the other has only two, the latter should be chosen as the final model. This is because we gain very little by including more variables in the model, and moreover large numbers of explanatory variables tend to result in inadequate degrees of freedom and imprecise estimation. The encompassing principle requires that the preferred model should be able to encompass all, or at least most, of the models developed by previous researchers in the same field. The encompassing criterion does not necessarily conflict with that of parsimony; the preferred model may be structurally simpler than other models, but still encompasses them. The parameter constancy criterion is particularly important when we use econometric models to forecast. In order to generate accurate forecasts, the parameters of the model should be constant over time. The final criterion for selecting a model is that the explanatory variables should be exogenous, that is they should not be contemporaneously correlated with the error term in the regression.

In modelling tourism demand, the final preferred model should ideally satisfy all of the above criteria. However, this can sometimes be very difficult due to various reasons, such as data limitations, errors in variables and insufficient knowledge of the demand system. Any of these may result in the above criteria being not satisfied. Even if we find a demand model that satisfies all the criteria, it should be borne in mind that the model can still serve only as an approximation to the complex behaviour of tourists, and it is possible that the decision-making process of tourists will change due to changes in expectations, tastes and economic regimes. Therefore, we should always be prepared to revise our model to take account of such changes.

3.6 Worked example

In this section we illustrate the general-to-specific modelling approach using the example of outbound tourism demand from the UK to Spain. We start with the construction of a general demand model. Various restrictions on the parameters are tested to determine which type of model is appropriate, and diagnostic checking is then carried out in the model selection process.

The data

The data are annual and cover the period 1966–94. The following variables are included in the model: the number of holidays in Spain taken by UK residents (VSP), UK real personal disposable income (PDI), the implicit deflator of UK consumer expenditure (which is calculated by dividing total consumer expenditure in current prices by consumer expenditure in constant prices) (PUK), an exchange rate index of the UK pound against the US dollar (EXUK), the consumer price index in Spain (CPISP), an exchange rate index of Spanish pesetas against the US dollar (EXSP), and the UK population (POP). The data on VSP are obtained from *Social Trends* (various issues); the PDI, PUK and POP series are compiled from *Economic Trends* (1996); and the data on EXUK, EXSP and CPISP are obtained from the *IMF International Financial Statistics Yearbook* (1995). The data are given in Table A3.1 in Appendix 3.1.

In modelling UK outbound tourism demand to Spain, the dependent variable VSP and the independent variable PDI are first transformed into per capita terms. In order to examine the influence of tourism prices on tourism demand, a relative cost-of-tourism variable is created using the following formula:

$$RCSP = \frac{(CPISP/EXSP)}{(PUK/EXUK)} \tag{3.24}$$

This is a relative price variable adjusted by exchange rates. The demand equation is assumed to be a double log function. Since annual data are used, a general tourism demand model is specified with one lag for each variable:

$$\begin{aligned} LPVSP_t = {} & \alpha_0 + \alpha_1 LPVSP_{t-1} + \alpha_2 LPPDI_t + \alpha_3 LPPDI_{t-1} \\ & + \alpha_4 LRCSP_t + \alpha_5 LRCSP_{t-1} + \varepsilon_t \end{aligned} \tag{3.25}$$

where *LPVSP* is the logarithm of per capita holiday visits to Spain; *LPPDI* is the logarithm of per capita personal disposable income, and *LRCSP* is the logarithm of living costs in Spain relative to UK living costs adjusted by the exchange rate.

Estimated general model

Equation (3.25) is estimated using OLS based on data for the period 1966–94:

$$\begin{aligned} LPVSP_t = &-\underset{(3.268)}{8.018} + \underset{(0.133)}{0.648}\, LPVSP_{t-1} + \underset{(1.288)}{1.249}\, LPPDI_t - \underset{(1.185)}{0.429}\, LPPDI_{t-1} \\ &- \underset{(0.266)}{0.470}\, LRCSP_t - \underset{(0.284)}{0.440}\, LRCSP_{t-1} \end{aligned} \tag{3.26}$$

$$\begin{array}{llll} R^2 = 0.938 & \sigma = 0.1304 & \text{SSR} = 0.374 & \chi^2_{Auto}(2) = 0.150 \\ \chi^2_{Norm}(2) = 1.705 & \chi^2_{ARCH}(1) = 0.083 & \chi^2_{White}(17) = 16.97 & \\ \chi^2_{RESET}(2) = 4.710 & F_{Forecast}(4, 20) = 1.058 & & \end{array}$$

where σ is the standard error of the regression; SSR is the sum of squared residuals; $\chi^2_{Auto}(2)$ is the Breusch–Godfrey LM test for autocorrelation; $\chi^2_{Norm}(2)$ is the Jarque–Bera normality test; $\chi^2_{ARCH}(1)$ is the Engle test for autoregressive conditional heteroscedasticity; $\chi^2_{White}(17)$ is the White test for heteroscedasticity; $\chi^2_{RESET}(2)$ is the Ramsey test for omitted variables/functional form; and $F_{Forecast}(4, 20)$ is the Chow predictive failure test (when calculating this test, 1991 was chosen as the starting-point for forecasting). The figures in parentheses under the estimated coefficients are standard errors, and this is the case throughout this chapter. The diagnostic statistics show that the general model passes all but one of the tests. The model only fails the RESET test for model mis-specification.

Tests for restrictions

The model reduction process can be started by performing various restriction tests. This process aims to determine a parsimonious specification that can be used to model the demand by UK residents for holidays to Spain. Since the demand model involves two explanatory variables, the restrictions on the coefficients are slightly more complicated than is the case in the general model (3.3). However, the restrictions can still be easily tested using either the *F* test or the Wald (or LR or LM) statistic.

(a) Testing for the *static model*. The null hypothesis is that $\alpha_1 = \alpha_3 = \alpha_5 = 0$, so there are three restrictions on the regression coefficients of model (3.25). In order to test these restrictions we also need to estimate the restricted static model (3.27):

$$LPVSP_t = \alpha_0 + \alpha_2 LPPDI_t + \alpha_4 LRCSP_t + u_t \tag{3.27}$$

OLS estimation of Equation (3.27) gives:

$$LPVSP_t = \underset{(1.887)}{-23.05} + \underset{(0.220)}{2.369}\, LPPDI_t - \underset{(0.339)}{0.861}\, LRCSP_t$$

$$R^2 = 0.819 \qquad \sigma = 0.220 \qquad \text{SSR} = 1.261$$

Since 28 observations are used in estimating the unrestricted general model (3.26) and there are three restrictions involved, the F statistic can be calculated as:

$$F(r, n-k) = \frac{(SSR_1 - SSR_0)/r}{SSR_0/(n-k)} = \frac{(1.126 - 0.374)/3}{0.374/(28-6)} = 14.75$$

The critical value of the F statistic with [3, 22] degrees of freedom at the 5% significance level is 3.05. As the calculated F value exceeds this critical value, the restrictions should not be accepted, i.e., the static model is not an appropriate specification for modelling the demand by UK residents for tourism in Spain.

(b) Testing for the *autoregressive model*. The restrictions in this case are $\alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$. The F statistic is calculated in the same way as in (a), but the degrees of freedom are [4,22], as there are now four restrictions instead of three. The estimated restricted model is

$$LPVSP_t = \underset{(0.173)}{-0.184} + \underset{(0.061)}{0.914}\, LPVSP_{t-1}$$

$$R^2 = 0.896 \qquad \sigma = 0.155 \qquad \text{SSR} = 0.625$$

and the restriction test statistic is

$$F(r, n-k) = \frac{(SSR_1 - SSR_0)/r}{SSR_0/(n-k)} = \frac{(0.625 - 0.374)/4}{0.374/(28-6)} = 4.56$$

Since the critical value of $F[4, 22]$ is 2.84, the autoregressive model cannot be accepted.

(c) Testing for the *growth rate model*. The growth rate model requires that the coefficients in the general model satisfy $\alpha_1 = 1$, $\alpha_2 = -\alpha_3$ and $\alpha_4 = -\alpha_5$; i.e., the demand model may be rewritten as:

$$\Delta LPVSP_t = \alpha_0 + \alpha_2 \Delta LPPDI_t + \alpha_4 \Delta LRCSP_t + u_t \qquad (3.28)$$

OLS estimation of Equation (3.28) yields

$$\Delta LPVSP_t = \underset{(0.043)}{-0.061} - \underset{(1.430)}{0.045}\, \Delta LPPDI_t - \underset{(0.306)}{0.107}\, \Delta LRCSP_t$$

$$R^2 = 0.010 \qquad \sigma = 0.163 \qquad \text{SSR} = 0.668$$

The results show that this restricted model fits the data badly. The income coefficient has the ‘wrong’ sign and all the coefficients are statistically insignificant. These empirical results indicate that the growth rate model cannot be accepted. This is confirmed by the F test, with the calculated $F[3, 22]$ statistic (5.77) exceeding the critical value (3.05).

(d) Testing for the *leading indicator* model. In the leading indicator model, the lagged dependent variable and the current values of independent variables do not enter the equation. The null hypothesis for the tourism demand model is, therefore, $\alpha_2 = \alpha_3 = \alpha_5 = 0$. The estimated leading indicator model is

$$LPVSP_t = \underset{(1.800)}{-22.73} + \underset{(0.210)}{2.333}\, LPPDI_{t-1} - \underset{(0.309)}{1.331}\, LRCSP_{t-1}$$

$$R^2 = 0.832 \qquad \sigma = 0.201 \qquad \text{SSR} = 1.006$$

The calculated value of the $F[3,22]$ statistic is 12.392, which clearly suggests rejection of the leading indicator model.

(e) Testing for the *partial adjustment* process. The restrictions on the coefficients of the general model are $\alpha_4 = \alpha_6 = 0$. The estimated restricted model is

$$LPVSP_t = \underset{(3.001)}{-5.247} + \underset{(0.122)}{0.755}\, LPVSP_{t-1} + \underset{(0.314)}{0.534}\, LPPDI_t - \underset{(0.212)}{0.666}\, LRCSP_t$$

$$R^2 = 0.927 \qquad \sigma = 0.134 \qquad \text{SSR} = 0.435$$

The F statistic is given by

$$F(r, n-k) = \frac{(SSR_1 - SSR_0)/r}{SSR_0/(n-k)} = \frac{(0.435 - 0.374)/2}{0.374/(28-6)} = 1.974$$

This calculated $F[2,22]$ value is lower than the critical value (3.44), which suggests that the goodness of fit does not deteriorate much when the restrictions for a partial adjustment process are imposed. Since this model involves fewer variables than the general model and has more degrees of freedom, the partial adjustment model is preferred to the general demand model. The various diagnostic statistics for the partial adjustment model are as follows:

$$\chi^2_{Auto}(2) = 1.899 \qquad \chi^2_{Norm}(2) = 2.356 \qquad \chi^2_{ARCH}(1) = 0.018$$
$$\chi^2_{White}(10) = 4.647 \qquad \chi^2_{RESET}(2) = 1.394 \qquad F_{Forecast}(4,20) = 1.120$$

The model passes all the diagnostic tests, and in particular the Ramsey RESET statistic now does not suggest a model mis-specification problem in terms of omitted explanatory variables or incorrect functional form. An important implication of rejecting the general model is that the error correction model should also be rejected, since it can be shown that the error correction model is a re-parameterisation of the ADLM process (Chapter 4).

(f) Testing for the *COMFAC* model. The restrictions for the *COMFAC* model in the present example are $\alpha_3 = -\alpha_2\alpha_1$ and $\alpha_5 = -\alpha_4\alpha_1$. Since the restrictions are non-linear, the Wald test is employed. The calculated Wald statistic is $\chi^2(2) = 12.554$,[1] which suggests that the restrictions cannot be accepted, i.e., the *COMFAC* specification is not appropriate.

(g) Testing for the *finite distributed lag model*. In this case, the restriction is $\alpha_1 = 0$ and the estimated restricted model is

$$LPVSP_t = \underset{(1.659)}{-22.910} + \underset{(1.615)}{4.144}\, LPPDI_t - \underset{(1.625)}{1.801}\, LPPDI_{t-1}$$

$$- \underset{(0.376)}{0.534}\, LRCSP_t - \underset{(0.388)}{0.789}\, LRCSP_{t-1}$$

$$R^2 = 0.869 \qquad \sigma = 0.184 \qquad \text{SSR} = 0.779$$

The calculated $F[2, 22]$ value is 23.83, so clearly the restriction $\alpha_1 = 0$ is not valid.

(h) Testing for the *dead start model*. The restrictions in this case are $\alpha_2 = \alpha_4 = 0$. The estimated dead start model is

$$LPVSP_t = \underset{(2.989)}{-7.139} + \underset{(0.119)}{0.677}\, LPVSP_{t-1} + \underset{(0.314)}{0.731}\, LPPDI_{t-1} - \underset{(0.230)}{0.748}\, LRCSP_{t-1}$$

$$R^2 = 0.929 \qquad \sigma = 0.134 \qquad \text{SSR} = 0.438$$

The calculated $F[2, 22]$ statistic is 1.585, which is smaller than the corresponding critical value (3.44), and therefore we cannot reject the restrictions at the 5% level.

The dead start model is preferred to the general demand model. The diagnostics are as follows:

$$\chi^2_{Auto}(2) = 0.016 \quad \chi^2_{Norm}(2) = 1.911 \quad \chi^2_{ARCH}(1) = 0.085 \quad \chi^2_{White}(10) = 12.72$$
$$\chi^2_{RESET}(2) = 2.504 \quad F_{Forecast}(4, 20) = 1.379$$

From the diagnostic statistics we can see that the model is well specified and the coefficients of all the variables have 'correct' signs and are statistically significant at the 5% level. These results imply that there are two specific models that we can choose from – the partial adjustment model and the dead start model. Apart from the fact that both models pass all the diagnostic tests, the values of R^2 and the standard errors of regression for the two models are almost identical. The decision between these two models therefore has to be based on other criteria, such as the encompassing test.

In order to choose which is the better model – the partial adjustment model (M1) or the dead start model (M2) – we carry out the two encompassing tests explained in the previous section. The J-test is based on the following two regressions:

$$LPVSP_t = \alpha'_0 + \alpha'_1 LPVSP_{t-1} + \alpha'_2 LPPDI_t + \alpha'_3 LRCSP_t + \alpha'_4 \overline{LPVSP}_{M2t} + u'_t \tag{3.29}$$

$$LPVSP_t = \beta'_0 + \beta'_1 LPVSP_{t-1} + \beta'_2 LPPDI_{t-1} + \beta'_3 LRCSP_{t-1} + \beta'_4 \overline{LPVSP}_{M1t} + e'_t \tag{3.30}$$

where $\overline{LPVSP}_{M2t}$ is the estimated $LPVSP$ from the dead start model and $\overline{LPVSP}_{M1t}$ is the estimated $LPVSP$ from the partial adjustment model. The statistics concerned are the t ratios for α'_4 and β'_4. If the coefficient α'_4 is not significantly different from zero, we say that the dead start model (M2) is nested in the partial adjustment model (M1), and if the coefficient β'_4 is not significantly different from zero, we say that M1 is nested in M2. After estimating the partial adjustment and the dead start models the estimated values of $LPVSP$ are saved as $\overline{LPVSP}_{M1t}$ and $\overline{LPVSP}_{M2t}$, respectively. These two variables are then included in Equations (3.30) and (3.29), and OLS estimation of these Equations gives:

$$\begin{aligned} LPVSP_t = &-\underset{(1.602)}{4.092} + \underset{(0.065)}{0.787}\, LPVSP_{t-1} + \underset{(0.168)}{0.412}\, LPPDI_t - \underset{(0.116)}{0.431}\, LRCSP_t \\ &+ \underset{(0.113)}{0.891}\, \overline{LPVSP}_{M2t} \end{aligned}$$

$$\begin{aligned} LPVSP_t = &-\underset{(3.578)}{3.658} + \underset{(0.301)}{0.218}\, LPVSP_{t-1} + \underset{(0.371)}{0.376}\, LPPDI_{t-1} - \underset{(0.274)}{0.482}\, LRCSP_{t-1} \\ &+ \underset{(0.382)}{0.631}\, \overline{LPVSP}_{M1t} \end{aligned}$$

The calculated t statistic for α'_4 in Equation (3.29) is

$$\frac{\hat{\alpha}'_4}{SE(\hat{\alpha}'_4)} = \frac{0.891}{0.113} = 7.869$$

so α'_4 is highly significant. This suggests that the partial adjustment model (M1) does not encompass the dead start model (M2). The calculated t

statistic for β_4' in Equation (3.30) is $(0.631/0.382) = 1.647$, so β_4' is not significantly different from zero even at the 10% level. The dead start model (M2) therefore encompasses the partial adjustment model (M1).

The second encompassing test based on Equation (3.23) is exactly the same as testing for the partial adjustment model and dead start model as described in (e) and (h). Therefore, the encompassing statistics are $F[2,22] = 1.974$ and $F[2,22] = 1.585$, respectively. These two statistics suggest that the dead start model does not encompass the partial adjustment model and also that the partial adjustment model does not encompass the dead start model.

Although the J-test suggests that the dead start model is better than the partial adjustment model, this is not confirmed by the second encompassing test. Therefore, further criteria such as the ex post forecasting performance should be used to assist in model selection (the forecasting performance criteria will be fully discussed in Chapter 10).

Using the Wu–Hausman test, the exogeneity of the explanatory variables $LPPDI_t$ and $LRCSP_t$ can be tested. In the following example we explain the Wu–Hausman test using the dead start model. The test procedure is as follows. First, estimate the two auxiliary equations in which $LPPDI_t$ and $LRCSP_t$ are regressed against all the lagged variables in the general model (3.25). The results are given as

$$\begin{gathered} LPPDI_t = \underset{(0.508)}{1.230} + \underset{(0.053)}{0.875}\, LPPDI_{t-1} + \underset{(0.020)}{0.055}\, LPVSP_{t-1} - \underset{(0.039)}{0.013}\, LRCSP_{t-1} \\ R^2 = 0.986 \qquad \sigma = 0.023 \qquad \text{SSR} = 0.012 \end{gathered} \tag{3.31}$$

$$\begin{gathered} LRCSP_t = \underset{(2.450)}{1.398} + \underset{(0.189)}{0.622}\, LRCSP_{t-1} - \underset{(0.258)}{0.140}\, LPPDI_{t-1} + \underset{(0.864)}{0.097}\, LPVSP_{t-1} \\ R^2 = 0.372 \qquad \sigma = 0.109 \qquad \text{SSR} = 0.288 \end{gathered} \tag{3.32}$$

Second, the estimated residuals from these two models are saved as u_{LPPDI_t} and u_{LRCSP_t}, respectively. Third, the terms u_{LPPDI_t} and u_{LRCSP_t} are introduced into the dead start specification, and the resultant model is estimated using OLS:

$$\begin{gathered} LPVSP_t = -\underset{(2.919)}{7.140} + \underset{(0.116)}{0.677}\, LPVSP_{t-1} + \underset{(0.307)}{0.731}\, LPPDI_{t-1} - \underset{(0.225)}{0.748}\, LRCSP_{t-1} \\ + \underset{(1.288)}{1.249} u_{LPPDI_t} - \underset{(0.267)}{0.670} u_{LRCSP_t} \\ R^2 = 0.923 \qquad \sigma = 0.130 \qquad \text{SSR} = 0.374 \end{gathered} \tag{3.33}$$

The Wu–Hausman statistic is calculated by imposing the restriction that the coefficients of u_{LPPDI_t} and u_{LRCSP_t} are jointly equal to zero. The F version of the Wu–Hausman statistic is calculated as

$$F(r, n-k) = \frac{(SSR_1 - SSR_0)/r}{SSR_0/(n-k)} = \frac{(0.438 - 0.374)/2}{0.374/(28-6)} = 1.882$$

where SSR_0 is the sum of squared residuals from Equation (3.33) and SSR_1 is the sum of squared residuals from the estimated dead start model.

The calculated $F[2, 22]$ statistic is 1.882, which is less than the critical value of 3.44, indicating that the coefficients of u_{LPPDI_t} and u_{LRCSP_t} are jointly equal to zero. Therefore, we can conclude that the variables $LPPDI_t$ and $LRCSP_t$ are exogenous to $LPVSP_t$. The same conclusion is reached if the partial adjustment process is used as the base model.

Note

1. In contrast to the F and LR tests in which the statistics are calculated using the information from both the restricted and unrestricted models, the Wald test does not require estimation of the restricted model, and this is particularly useful if the restricted model is difficult to estimate as in the case of the *COMFAC* model. (For a detailed explanation of the Wald test see Stewart and Gill, 1998, pp. 129–33.)

Appendix 3.1

Table A3.1 Worked example: data used in estimation.

Year	*VSP*	*PDI*	*PUK*	*EXUK*	*POP*	*CPISP*	*EXSP*
1966	1.282300	201207.0	0.134250	0.485709	54.64300	0.096155	0.435193
1967	1.271800	204171.0	0.137742	0.563200	54.95900	0.102310	0.505549
1968	1.537000	207772.0	0.144298	0.568364	55.21600	0.107425	0.506419
1969	1.950100	209684.0	0.152272	0.564515	55.46100	0.109823	0.508160
1970	1.830000	217675.0	0.161250	0.566155	55.63200	0.113580	0.505694
1971	2.612600	220344.0	0.175189	0.576441	55.92800	0.122852	0.519910
1972	3.153500	238744.0	0.186568	0.626625	56.07900	0.133083	0.500617
1973	3.060100	254329.0	0.202143	0.703709	56.22300	0.148190	0.498295
1974	2.596600	252360.0	0.236360	0.706515	56.23600	0.170490	0.498295
1975	2.881500	253814.0	0.292056	0.784033	56.22600	0.200623	0.507580
1976	2.151400	253012.0	0.337778	0.924894	56.21600	0.235952	0.575470
1977	2.544000	247695.0	0.387586	0.863692	56.19000	0.293742	0.712845
1978	2.860200	265925.0	0.424315	0.867812	56.17800	0.351930	0.662508
1979	2.861500	281084.0	0.482226	0.802721	56.24000	0.407082	0.632045
1980	2.924900	285411.0	0.560568	0.724715	56.33000	0.470626	0.733154
1981	3.582000	283176.0	0.623576	0.826736	56.35200	0.539365	0.822732
1982	4.451100	281722.0	0.677889	0.925951	56.31800	0.538406	1.004932
1983	5.404900	289204.0	0.710628	0.978113	56.37700	0.604108	1.189962
1984	6.104200	299756.0	0.712308	1.148640	56.50600	0.672368	1.232828
1985	5.427200	309821.0	0.785876	1.030533	56.68500	0.731596	1.228113
1986	6.305600	323622.0	0.817104	1.124233	56.85200	0.795940	1.174585
1987	6.400000	334702.0	0.852381	1.027294	57.00900	0.837343	1.121564
1988	6.720000	354627.0	0.894970	1.007860	57.15800	0.877548	1.107348
1989	5.880000	371676.0	0.947763	1.109298	57.35800	0.937175	1.045840
1990	4.637000	378325.0	1.000000	1.000000	57.56100	1.000000	1.000000
1991	4.428300	377969.0	1.073715	1.036266	57.80800	1.058988	1.003191
1992	4.648000	386804.0	1.124222	1.232419	58.00600	1.121493	1.143178
1993	5.866800	393125.0	1.163626	1.256732	58.19100	1.173048	1.416842
1994	7.216700	396181.0	1.192823	1.266178	58.39500	1.228119	1.394937

4

Cointegration

4.1 Spurious regression

Many tourism demand variables (dependent and explanatory), such as aggregate tourism expenditure, total tourist arrivals, tourism costs in the destination countries and income in the tourism-generating country, are often trended, i.e., the variables are non-stationary. This can be a potential problem for tourism demand analysis. We tend to obtain a high R^2 and significant t statistics for the regression coefficients if the variables in the demand model have a common deterministic trend, but this does not necessarily mean that these variables are actually related, i.e., the regression may be spurious.

A typical spurious correlation is illustrated in the following example. The implicit deflator of consumer expenditure for the UK and the consumer price index (CPI) for Spain are plotted in Figure 4.1. The figure shows that both variables are upward trended.

Although there is no reason for us to believe that inflation in the UK determines inflation in Spain or vice versa, the correlation between these two variables is very high since both series exhibit a similar trend over the same period. This can be shown by examining the OLS estimates of the regression of the log of the Spanish consumer price index on the log of the implicit deflator of UK consumer expenditure (the equivalent CPI for the UK). The apparent correlation between the two price variables as measured by the extremely high value of R^2 (0.99) in Equation (4.1) is

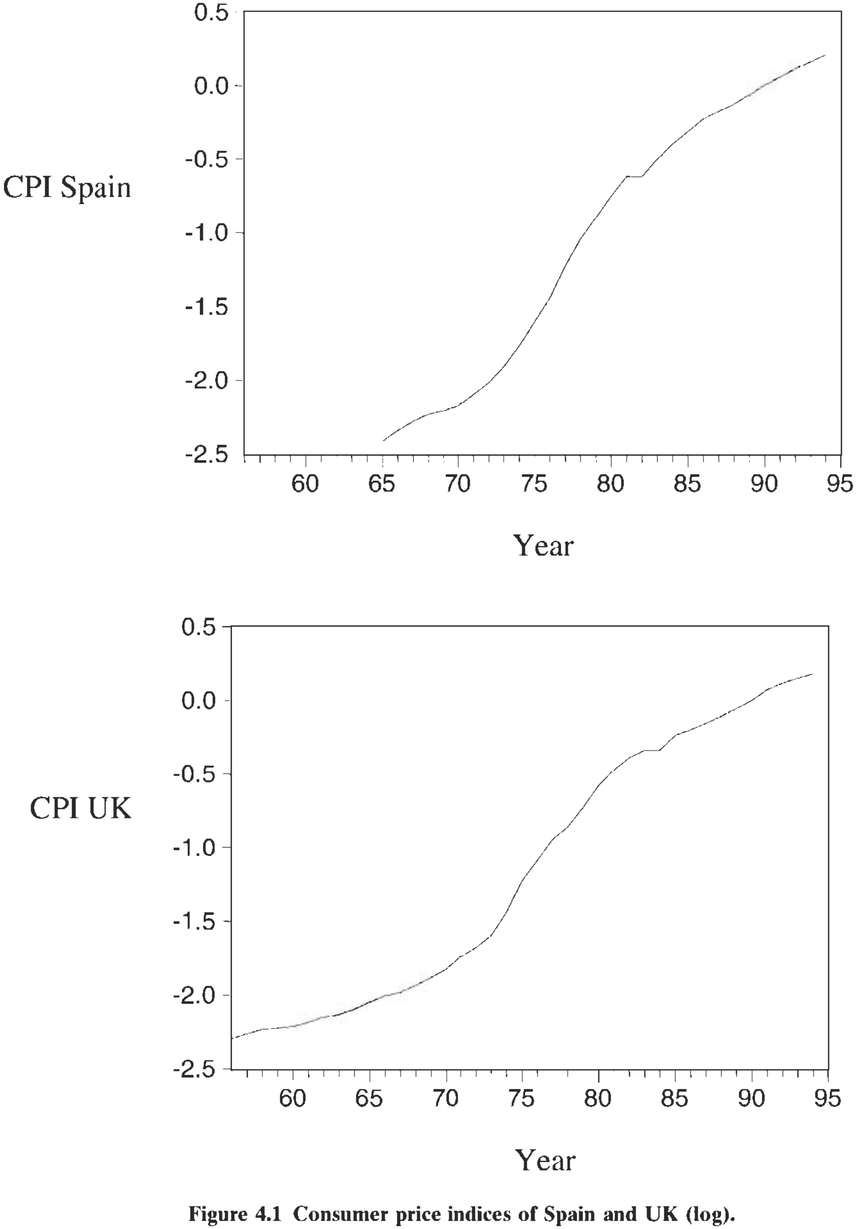

Figure 4.1 Consumer price indices of Spain and UK (log).

clearly spurious.

$$\log(CPI)_{Spain} = \underset{(0.016)}{0.041} + \underset{(0.014)}{1.171} \log(CPI)_{UK} \tag{4.1}$$

$$R^2 = 0.990 \qquad \mathrm{DW} = 0.565 \qquad \chi^2_{Auto}(2) = 16.42^{***}\ \chi^2_{Norm}(2) = 4.712^{*}$$
$$\chi^2_{ARCH}(1) = 4.700^{**}\ \chi^2_{White}(2) = 3.675\ \chi^2_{RESET}(2) = 22.82^{***}$$

where figures in parentheses are standard errors and *, ** and *** denote significance at 10%, 5% and 1% levels, respectively.

Even if there is a weak linkage between Spanish inflation and the UK price level due to the spill-over effect caused by imports and exports between these two countries, the estimated standard errors (or t statistics) for the regression coefficients and the R^2 values are likely to be very much exaggerated. Moreover, all diagnostic statistics apart from the White heteroscedasticity test statistic are violated, implying that the model is unacceptable statistically (the Chow breakpoint test was also performed using 1973, 1979 and 1982 as breakpoints; model stability was rejected in all cases).

Until recently very little attention was paid to the spurious regression problem by tourism forecasters and practitioners. Early attempts to solve the problem used differenced variables in the regression model. However, by doing this, important information concerning the long-run equilibrium relationship among the economic variables is lost. The cointegration technique developed by Engle and Granger (1987), coupled with the error correction mechanism, has proved to be a useful tool for solving the problem of spurious correlation when non-stationary time series are used in econometric modelling. This chapter is devoted to examining the concept of and tests for cointegration.

4.2 The concept of cointegration

According to Engle and Granger (1987), if a pair of non-stationary economic variables x_t and y_t belongs to the same economic system, such as tourism demand and income, there should be an *attractor* or *cointegration relationship* that prevents these two time series from drifting away from each other; i.e., there exists a *force of equilibrium* that keeps the two variables, x_t and y_t, moving together in the long run. More formally, if x_t and y_t move together in the long run, they can be modelled by a long-run equilibrium model of the form:

$$y_t = \beta_0 + \beta_1 x_t \tag{4.2}$$

The disequilibrium error of Equation (4.2) is:

$$\varepsilon_t = y_t - \beta_0 - \beta_1 x_t \tag{4.3}$$

Engle and Granger (1987) state that if the long-run equilibrium relationship (4.2) exists, the disequilibrium error (4.3) should 'rarely drift far from zero'. This means that if the long-run equilibrium regression (4.2) is estimated using OLS, the residuals of the model should follow a stationary process and fluctuate around the value of zero over time, i.e., the two non-stationary variables, x_t and y_t, are cointegrated.

A non-stationary time series is also called an integrated process. The order of integration for a series is determined by the number of times the series must be differenced to achieve a stationary process. Therefore, a stationary variable is a series that is integrated of order zero, or $I(0)$ for short. A series is said to be integrated of order 1, or $I(1)$ if it becomes stationary after taking first differences. In general, if a time series needs to be differenced d times before achieving stationarity, then the series is said to be integrated of order d, or $I(d)$.

Cointegration may be defined as follows: x_t and y_t are said to be cointegrated of order d, b, denoted by $(x_t, y_t) \sim CI(d, b)$, if x_t and y_t are integrated of order d and there exists a vector of parameters (β_0, β_1) such that the linear combination $y_t = \beta_0 + \beta_1 x_t$ is integrated of order $(d - b)$, where $b > 0$. The vector (β_0, β_1) is called the cointegration vector. Although this definition refers to the two-variable case, it can be extended to a *k-variable* cointegration system.

The following points need to be made regarding this definition. First, all variables in the cointegration regression must be integrated of the same order (although some of the recent cointegration literature allows for different orders of integration). This pre-condition for a cointegration relationship does not necessarily mean that all economic variables that have the same integration order are cointegrated. For example, the consumer price indices in the UK and Spain may be integrated of the same order, but they are clearly not cointegrated. Second, in a two-variable cointegration equation, the cointegration vector (β_0, β_1) is unique. However, if there are k variables in the system, there may be as many as $k - 1$ cointegration relationships. In Chapter 6 we will look at multiple cointegration relationships within a vector autoregressive (VAR) modelling framework. Third, most of the empirical literature on cointegration concentrates on the model in which all variables are integrated of order one, i.e., the variables are $CI(1, 1)$. Our subsequent discussions of cointegration also follow this tradition.

The cointegration relationship may be tested using the OLS residuals, $\hat{\varepsilon}_t$, from Equation (4.2). The aim is to test whether $\hat{\varepsilon}_t$ is a stationary process or not. If the $\hat{\varepsilon}_t$ term from the OLS estimation of (4.2) is found to be stationary, we will say that x_t and y_t are cointegrated. On the other hand if the residuals are found to be nonstationary, the conclusion would be that the two variables concerned are not cointegrated. The discussion on how to test for cointegration will be given in Section 4.5. Since the cointegration relationship requires that the variables in the cointegration regression are integrated of order one, we need to test for the integration order of each individual time series in the long-run equilibrium model before the cointegration test can be carried out. The tests used for identifying the order of integration are called the Dickey and Fuller (1981) tests and they are presented in the following section.

4.3 Test for order of integration

4.3.1 Stationary and non-stationary time series

We begin with the definition of a stationary series. *A time series is said to be stationary if its mean, variance and covariance remain constant over time*, that is:

$$E(y_t) = \mu \tag{4.4}$$

$$E[(y_t - \mu)^2] = Var(y_t) = \sigma^2 \tag{4.5}$$

$$E[(y_t - \mu)(y_{t-p} - \mu)] = Cov(y_t - y_{t-p}) = \Omega_p \tag{4.6}$$

Equations (4.4) and (4.5) state that the time series, y_t, has a constant mean, μ, and a constant variance, σ^2, respectively, while Equation (4.6) indicates that the covariance between any two values from the time series depends only on the time interval, p, between those two values and not on time itself. If a time series violates any of the above conditions, it is regarded as a non-stationary series. Figure 4.2 shows the log of per capita holiday visits by UK residents to Spain ($LPVSP_t$) and the same variable expressed in first differences ($\Delta LPVSP_t$). It is clear that $LPVSP_t$ is a non-stationary variable since its mean increases over time, while $\Delta LPVSP_t$ fluctuates around a mean value of 0.06, suggesting that it is likely to be a stationary variable. (This will be confirmed by the standard test statistics in Section 4.3.7.)

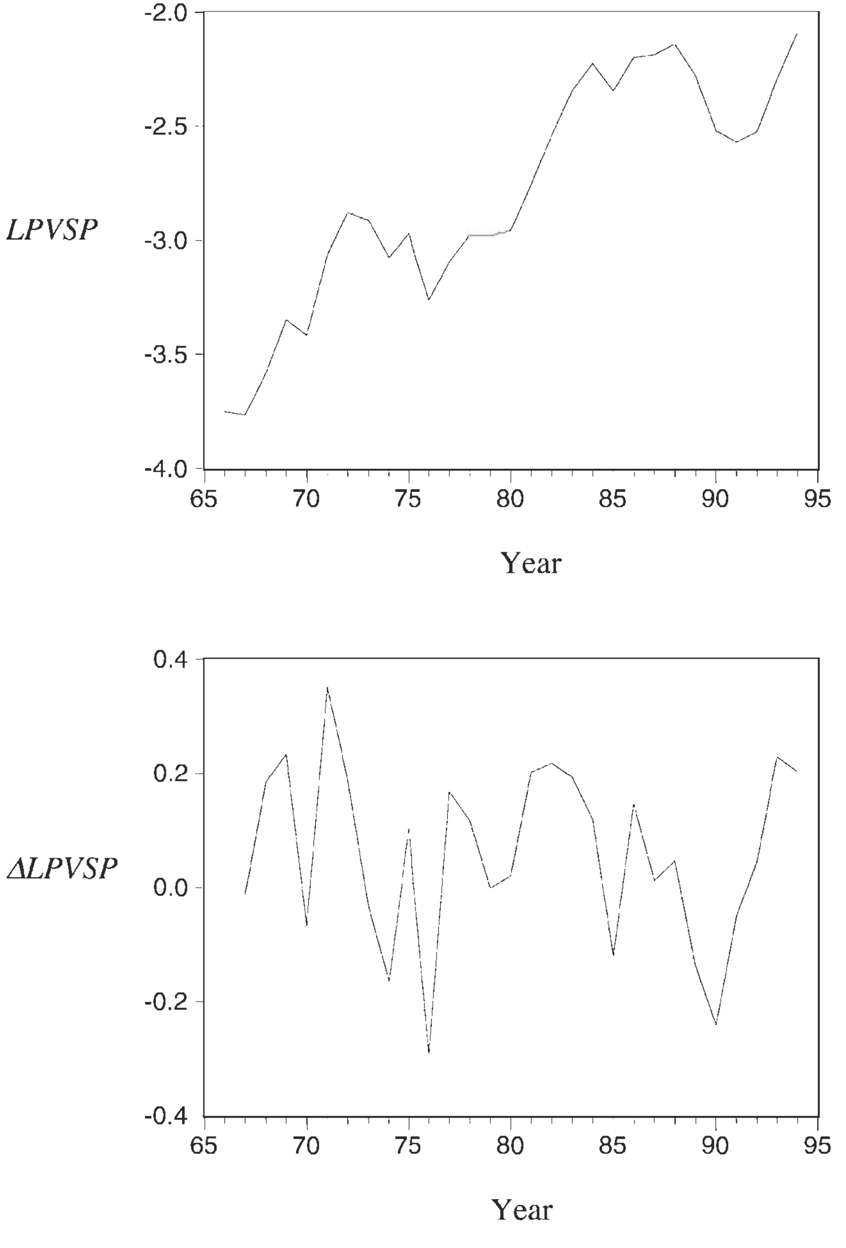

Figure 4.2 Comparison of stationary and non-stationary time series.

4.3.2 *Non-stationary series and unit roots*

A non-stationary time series is also known as a series that has unit roots. The number of unit roots contained in the series is equal to the number of times the series must be differenced in order to achieve a stationary series. For example, if the time series y_t becomes stationary after taking first differences, we say that y_t has one unit root, denoted as $y_t \sim I(1)$, or is integrated of order one. If y_t needs to be differenced twice in order to become a stationary process, then it has two unit roots, i.e., y_t is an $I(2)$ series. Therefore an $I(\mathrm{d})$ series means that the series has d unit roots.

A non-stationary series can always be represented by an autoregressive process of order $p(\mathrm{AR}(p))$:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_p y_{t-p} + e_t \tag{4.7}$$

where e_t is an error term which is assumed to have a zero mean and constant variance. The condition for y_t to be a non-stationary series is: $\beta_1 + \beta_2 + \ldots + \beta_p = 1$. If the order of the AR process is $p = 1$, equation (4.7) becomes:

$$y_t = \beta_0 + \beta_1 y_{t-1} + e_t \tag{4.8}$$

and if y_t is a non-stationary series, $\beta_1 = 1$, i.e.,

$$y_t = \beta_0 + y_{t-1} + e_t \tag{4.9}$$

Equation (4.9) is also called a random walk with drift (RWD). If the constant term equals zero in (4.9), the process is termed a random walk (RW). To test for stationarity is therefore to test either that $\beta_1 + \beta_2 + \ldots + \beta_p = 1$ or $\beta_1 = 1$ in Equations (4.7) and (4.8), respectively, and this can be accomplished by using the Dickey–Fuller (DF) and the augmented Dickey–Fuller (ADF) tests (Dickey and Fuller, 1981).

4.3.3 *The DF and ADF tests for unit roots*

DF test. Let us assume that the time series can be modelled by an AR(1) process as in Equation (4.8). Our objective is to test the null hypothesis, H_0: $\beta_1 = 1$, against the alternative, H_1: $\beta_1 < 1$. An obvious way of testing for $\beta_1 = 1$ is to look at the t ratio:

$$t = \frac{\hat{\beta}_1 - 1}{SE(\hat{\beta}_1)} \tag{4.10}$$

However, instead of testing $\beta_1 = 1$ directly, Dickey and Fuller (1979) transformed (4.8) into (4.11):

$$\Delta y_t = \beta_0 + \phi y_{t-1} + e_t \tag{4.11}$$

This is done by subtracting y_{t-1} from both sides of Equation (4.8). In Equation (4.11) $\Delta y_t = y_t - y_{t-1}$ and $\phi = \beta_1 - 1$. Note that testing for $H_0{:}\beta_1 = 1$ against $H_1{:}\beta_1 < 1$ in (4.8) is exactly the same as testing $H_0{:}\phi = 0$ against $H_1{:}\phi < 0$ in (4.11). The commonly used test for $\phi = 0$ is of course the t ratio: $\hat{\phi}/[SE(\hat{\phi})]$. However, under the null hypothesis of non-stationarity, the t ratio has a non-standard distribution, and therefore the conventional critical values for the t statistic are not applicable. Dickey and Fuller obtained critical values based on Monte Carlo simulations, and a selection of these critical values is presented in Table 4.1 under the column τ_μ. If the calculated t value is lower than the critical value of τ_μ, the null hypothesis that the series is non-stationary should be rejected. Acceptance of the null hypothesis means that the time series has at least one unit root. The test is therefore repeated on the differenced series of y_t if the levels variable is found to be non-stationary. The DF regression now becomes:

$$\Delta^2 y_t = \beta_0 + \phi \Delta y_{t-1} \tag{4.12}$$

where $\Delta^2 y = \Delta y_t - \Delta y_{t-1}$.

If $\phi = 0$ in Equation (4.12) is rejected, then the time series y_t is said to be $I(1)$, or has one unit root. However, if we still cannot reject the null, then the DF test must be carried out again, now with the dependent variable expressed in third difference form and the explanatory variable in second difference form. This process is repeated until the null hypothesis of non-stationarity is rejected.

The DF test explained so far is based on the assumption that the time series can be represented by an AR(1) process as in Equation (4.8). There are other variants of the DF tests, and these are based on the following two equations:

$$\Delta y_t = \phi y_{t-1} + e_t \tag{4.13}$$

$$\Delta y_t = \beta_0 + \lambda T + \phi y_{t-1} + e_t \tag{4.14}$$

where T is a time trend variable. The null hypothesis is still $\phi = 0$, and the critical values for the DF tests are given in Table 4.1 under the columns τ and τ_τ, corresponding to Equation (4.13) and Equation (4.14), respectively. Regression (4.13) is valid only when the overall mean of the series is zero, while regression (4.14) relates to a series that has a deterministic trend. Equation (4.14) encompasses both (4.11) and (4.13). The DF test therefore normally starts from the general specification of (4.14), and restrictions may then be imposed on the coefficients β_0 and λ to decide whether Equation (4.11) or Equation (4.13) is preferred.

ADF test. If a simple AR(1) DF regression is used when in fact the time series y_t follows an AR(p) process, then the error term will be autocorrelated and this will invalidate the DF statistics. To avoid the problem of autocorrelation in the residuals, the DF regressions may be augmented by including lagged dependent variables. Corresponding to the three DF regressions (4.11), (4.13) and (4.14), the three ADF models are, respectively:

$$\Delta y_t = \beta_0 + \phi y_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta y_{t-i} + e_t \tag{4.15}$$

$$\Delta y_t = \phi y_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta y_{t-i} + e_t \tag{4.16}$$

$$\Delta y_t = \beta_0 + \lambda T + \phi y_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta y_{t-i} + e_t \tag{4.17}$$

To test the null hypothesis, we again calculate the DF t statistic, $\hat{\phi}/[SE(\hat{\phi})]$, which is compared with the corresponding critical values in Table 4.1. It is important to select the appropriate lag length for the dependent variable. Too few lags may result in over-rejection of the null hypothesis when it is true, while too many lags may reduce the power of the test due to loss of degrees of freedom. The criteria for selecting the lag length are $\overline{R}^2$, the Akaike information criterion (AIC) and the Schwarz Bayesian criterion (SBC). The preferred lag length in the ADF regression should maximise $\overline{R}^2$ and minimise both the AIC and SBC.

Although the idea of testing for unit roots is simple, the test procedure is not a straightforward process. In the following subsection, the procedure is illustrated using examples.

4.3.4 Phillips–Perron tests

The distribution theory supporting the Dickey–Fuller tests assumes that the error term in the DF (ADF) regression is identical and independently distributed (IID). In order to use the Dickey–Fuller tests, there should therefore be no autocorrelation or heteroscedasticity present in the estimated residuals, but these assumptions are rather restrictive. Phillips and Perron (1988) have developed a generalisation of the Dickey–Fuller

Table 4.1. Critical values for the DF tests.

	Significance level								
	Critical value of τ			Critical value of τ_{μ}			Critical value of τ_{τ}		
Sample size	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
25	−2.66	−1.95	−1.60	−3.75	−3.00	−2.63	−4.38	−3.60	−3.24
50	−2.62	−1.95	−1.61	−3.58	−2.93	−2.60	−4.15	−3.50	−3.18
100	−2.60	−1.95	−1.61	−3.51	−2.89	−2.58	−4.04	−3.45	−3.15

Source: Fuller (1976).

Notes: If the DF or ADF regression does not have constant and trend terms, τ is applicable. If the DF or ADF regression has a constant term but no trend, τ_{μ} is relevant. If both trend and constant terms are present, τ_{τ} should be used.

procedure that relaxes these assumptions. The expressions of the Phillips–Perron tests are extremely complex and are beyond the scope of this book. However, many econometric software packages provide standard print-outs of the Phillips–Perron statistics, and the critical values are exactly the same as those of the Dickey–Fuller tests.

4.3.5 The unit root testing procedure

The unit root test starts from the general specification of Equation (4.17) since this encompasses all other testing models, and a test down procedure is then followed to decide on the final DF (ADF) regression from which the DF or ADF statistic is calculated.

Step 1. Estimate the general AR(p) model (4.17) using OLS or the Phillips–Perron procedure. The lag length for the dependent variable is determined by $\bar{R}^2$, AIC and SC.

Step 2. Test the joint hypothesis (restrictions) that $\lambda = \phi = 0$ using the F statistic. However, since the distribution of the calculated F statistic based on the ADF regression is not standard, the critical values of the conventional F statistic cannot be used. Dickey and Fuller (1981) derived the critical values of the F statistic (Φ_3 in Dickey and Fuller, 1981, p. 1063) based on Monte Carlo simulations. Selected critical values of Φ_3 are presented in Table 4.2.

If the joint hypothesis $H_0 : \lambda = \phi = 0$ is accepted in accordance with the critical values in Table 4.2, then go to Step 4.

Table 4.2. Critical values of the Dickey–Fuller Φ_3 statistic.

	Significance level		
Sample size: n	0.01	0.05	0.10
25	10.61	7.42	5.91
50	9.31	6.73	5.61
100	8.73	6.49	5.47
500	8.34	6.30	5.36

Source: Dickey and Fuller (1981).

Step 3. If the null hypothesis in Step 2 is rejected, this suggests that the trend variable is statistically significant and the standard t statistic, $\hat{\phi}/[SE(\hat{\phi})]$, can then be used to test that $\phi = 0$ instead of using the DF τ_τ statistic. If the standard t statistic indicates that $\phi = 0$, we conclude that the series is a non-stationary process with a deterministic trend; otherwise we conclude that the series is a trended stationary series.

Step 4. Since the joint null hypothesis $H_0\colon \lambda = \phi = 0$ is not rejected in Step 2, this means that the series is a non-stationary series without trend, but with possible drift. This conclusion may be confirmed by looking at the calculated t statistic in Equation (4.17) and comparing it with the corresponding critical value τ_τ in Table 4.1. The test is usually terminated at this stage, since the main aim of the DF and ADF tests is to find out whether or not a series is stationary. However, occasionally, we may wish to test whether a non-stationary series is a random walk or random walk with drift. In this case, further tests may be required. Interested readers are referred to Dickey and Fuller's Φ_2 and Φ_1 statistics (Dickey and Fuller, 1981).

4.3.6 Seasonal unit roots

If seasonal time series are used in tourism demand modelling, testing for seasonal unit roots and seasonal cointegration is necessary. Seasonal unit root tests are much more complicated than the simple unit root tests, since seasonal time series tend to have different unit roots such as quarterly, semi-annual and annual. If any of these unit roots are present in the time series, then the relevant differencing strategy is needed to achieve stationarity; e.g., in the case of a quarterly time series, annual unit roots require annual differencing ($\Delta_4 y_t = y_t - y_{t-4}$) while semi-annual unit roots need semi-annual differencing ($\Delta_2 y_t = y_t - y_{t-2}$). There are a number of seasonal unit root tests available, such as Dickey *et al.* (1984), Osborn *et al.* (1988) and Hylleberg, Engle, Granger and Yoo (HEGY) (1990), but here we present only the HEGY seasonal unit root test. Our discussion focuses on quarterly seasonal data.

The simplest HEGY seasonal unit root test uses the following regression:

$$\Delta_4 y_t = \sum_{i=1}^{4} \pi_i Y_{it-1} + e_t \tag{4.18}$$

where the Y_{it}s are new series generated from the following:

$$\begin{aligned} Y_{1t} &= y_t + y_{t-1} + y_{t-2} + y_{t-3} \\ Y_{2t} &= -y_t + y_{t-1} - y_{t-2} + y_{t-3} \\ Y_{3t} &= -y_t + y_{t-2} \\ Y_{4t} &= -y_{t-1} + y_{t-3} \end{aligned}$$

The HEGY regression may be estimated by OLS. The null hypothesis of seasonal unit roots implies that: $\pi_1 = \pi_2 = \pi_3 = \pi_4 = 0$. Alternatively, we say that the series does not have any unit roots or it is stationary if the null hypothesis is rejected. However, if some of the πs are zero while others are not, the interpretations will be different. In particular, if only $\pi_1 = 0$, then the series has normal unit roots only as discussed above. If $\pi_2 = 0$, this suggests that there are semi-annual seasonal unit roots, and $\pi_3 = \pi_4 = 0$ indicates that annual unit roots are present. Testing for $\pi_1 = 0$ and $\pi_2 = 0$ may be achieved by looking at the t statistics, while $\pi_3 = \pi_4 = 0$ can be tested using the F statistic. However, the t and F statistics from the HEGY regression are non-standard. The critical values have to be generated from Monte Carlo simulations, and they can be found in HEGY (1990).

The HEGY test may be generalised by including a constant term and/or deterministic seasonalilty such that:

$$\Delta_4 y_t = \pi_0 + \sum_{i=1}^{4} \pi_i Y_{it-1} + e_t \tag{4.19}$$

$$\Delta_4 y_t = \sum_{i=1}^{4} \pi_i Y_{it-1} + \sum_{i=1}^{4} b_i Q_{it} + e_t \tag{4.20}$$

where the Q_{it}s are seasonal dummies. The critical values for the HEGY tests where there is either an intercept or deterministic seasonal components are different from those of the simplest HEGY test relevant to Equation (4.18), and they can also be found in HEGY(1990).

Table 4.3. Empirical results for general ADF regression (4.17).

	Lag length				
Statistics	5	4	3	2	1
ADF	−2.661	−3.570	−3.031	−2.879	−2.649
$\bar{R}^2$	0.171	0.286	0.162	0.165	−0.159
AIC	−3.676	−3.736	−3.329	−3.657	−3.699
SC	−3.281	−3.393	−3.336	−3.415	−3.507

The above HEGY regressions can also be augmented by lagged dependent variables to ensure that the residuals from the HEGY regressions do not exhibit autocorrelation. As in the case of the ADF test, the inclusion of the lagged dependent variables does not affect the distribution of the statistics, and the critical values are the same as those for the model without the augmented terms.

4.3.7 Examples of unit root tests

Although most econometric packages have built-in options to test for unit roots, in this subsection we give a step-by-step guide to testing for unit roots. We use the three tourism demand series introduced in Chapter 3 for illustrative purposes. The variables that we wish to test for unit roots are: the log of per capita holiday visits to Spain by UK residents, $LPVSP_t$; the log of UK per capita personal disposable income, $LPPDI_t$; and the log of living costs in Spain relative to UK living costs, adjusted by the corresponding exchange rate, $LRCSP_t$. We present a step-by-step unit root testing procedure for the variable $LPVSP_t$, but for the variables $LPPDI_t$ and $LRCSP_t$ we report the results only, with the intermediate steps omitted.

In Step 1, the general ADF regression corresponding to Equation (4.17) is constructed:

$$\Delta LPVSP_t = \beta_0 + \lambda T + \phi LPVSP_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta LPVSP_t + e_t \tag{4.21}$$

There are 28 observations available for the estimation of Equation (4.21). We use OLS to estimate the equation, first with five lagged dependent variables, and then we reduce by one lag at a time in order to see what is the appropriate lag length according to $\bar{R}^2$, AIC and SC. The results are reported in Table 4.3.

Since the ADF regression with four lagged dependent variables maximises $\overline{R}^2$ and minimises the AIC (although lag one minimises the SC), we set the lag length equal to 4 when we conduct the ADF test.

The second step is to test the joint hypothesis that $\lambda = \phi = 0$ in Equation (4.17). First, let us examine the estimated ADF regression:

$$\Delta LPVSP_t = \underset{(-3.425)}{-3.134} + \underset{(3.169)}{0.047T} - \underset{(-3.571)}{0.866 LPVSP_{t-1}} + \underset{(2.600)}{0.568 \Delta LPVSP_{t-1}}$$

$$+ \underset{(2.447)}{0.571 \Delta LPVSP_{t-2}} + \underset{(1.795)}{0.432 \Delta LPVSP_{t-3}} + \underset{(1.414)}{0.333 \Delta LPVSP_{t-4}} \tag{4.22}$$

$$\overline{R}^2 = 0.286, \quad \text{AIC} = -3.736 \quad \text{SC} = -3.393 \quad \text{SSR} = 0.319$$

The figures in parentheses are t ratios. The ADF statistic in this case is -3.571. To test for $\lambda = \phi = 0$, the F test for restrictions discussed in Chapter 3, Section 3, may be used. In calculating the F statistic we need to estimate both the unrestricted model (4.17) and the restricted model (4.23) below, and obtain the SSRs from both equations.

$$\Delta LPVSP_t = \beta_0 + \sum_{i=1}^{p-1} \gamma_i \Delta LPVSP_{t-i} + e_t \tag{4.23}$$

Here we do not report the estimated results for Equation (4.23), but the sum of squared residuals from (4.22) and the estimated model (4.23) are 0.3194 and 0.5688, respectively. The F statistic is therefore

$$F(r, n-k) = \frac{(SSR_1 - SSR_0)/r}{SSR_0/(n-k)} = \frac{(0.5688 - 0.3194)/2}{0.3194/(24-7)} = 6.636$$

The critical value of the ADF F statistic (Φ_3) with 25 observations in Table 4.3 is 7.24. As the calculated F value is lower than the critical value, this suggests that we should accept the joint hypothesis that $\lambda = \phi = 0$. This result means that the time series $LPVSP_t$ is non-stationary with a possible drift.

Since the restriction that $\lambda = \phi = 0$ cannot be rejected, the next step is to compare the calculated ADF statistic, -3.571, with the corresponding critical value of τ_τ in Table 4.1. We can see that the critical values of τ_τ with 25 degrees of freedom at the 1% and 5% significance levels are -4.38 and -3.60, respectively. The calculated ADF statistic is higher than these two values, indicating that the series is likely to be non-stationary.

Since the variable $LPVSP_t$ is non-stationary in levels, we need to perform

Table 4.4 Empirical results for ADF regression (4.24).

	Lag length					
Statistics	5	4	3	2	1	0
ADF	−2.054	−2.570	−2.403	−2.359	−2.852	−4.380
$\overline{R}^2$	0.267	0.318	0.309	0.351	0.359	0.398
AIC	−3.259	−3.376	−3.260	−3.314	−3.401	−3.507
SC	−2.863	−3.031	−2.965	−3.070	−3.207	−3.363

the ADF test based on the differenced series, and the ADF regression now becomes:

$$\Delta^2 LPVSP_t = \beta_0 + \lambda T + \phi \Delta LPVSP_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta^2 LPVSP_{t-i} + e_t \qquad (4.24)$$

where $\Delta^2 LPVSP_t = \Delta LPVSP_t - \Delta LPVSP_{t-1}$.

Equation (4.24) is estimated with different numbers of lagged dependent variables, and the values of $\overline{R}^2$ and AIC and SC are reported in Table 4.4.

According to the results given in Table 4.4, we can see that the ADF model with no lagged dependent variable, i.e., the DF regression, should be used, and the estimated model is presented below:

$$\Delta^2 LPVSP_t = \underset{(1.321)}{0.095} - \underset{(-0.617)}{0.025T} - \underset{(-4.382)}{0.906\,\Delta LPVSP_{t-1}} \qquad (4.25)$$

$$\overline{R}^2 = 0.398, \qquad AIC = -3.507 \qquad SC = -3.363 \qquad SSR = 0.648$$

Now the joint hypothesis that $\lambda = \phi = 0$ is tested again using the F statistic. After estimating the restricted model with only the constant term in the ADF regression, the value 1.167 is obtained for the SSR. Since we already know the SSR from the unrestricted DF regression (4.25), the F statistic can be calculated as:

$$F(r, n-k) = \frac{(SSR_1 - SSR_0)/r}{SSR_0/(n-k)} = \frac{(1.167 - 0.648)/2}{0.648/(27-3)} = 9.611$$

Comparing this $F(2,24)$ statistic with the corresponding critical value of the Dickey–Fuller F statistic (Φ_3) in Table 4.2, we can see that the null hypothesis $\lambda = \phi = 0$ is rejected at the 5% significance level. This suggests that the variable $\Delta LPVSP_t$ is a stationary series without trend but with a possible drift, and this is confirmed by the ADF value of −4.382 which is

Table 4.5. ADF test results for $LPPDI_t$, $LRCSP_t$ and $LPVSP_t$.

Variable	ADF statistic	Critical value 0.01	Critical value 0.05	Lag length	T and β_0[1]
$LPPDI_t$	−3.630	−4.38	−3.60	1	T & β_0
$\Delta LPPDI_t$	−4.195	−3.75	−3.00	1	β_0
$LRCSP_t$	−2.226	−3.75	−3.00	4	β_0
$\Delta LRCSP_t$	−7.212	−2.66	−1.95	0	–
$LPVSP_t$	−3.571	−4.38	−3.60	4	T & β_0
$\Delta LPVSP_t$	−4.395	−3.75	−3.00	0	β_0

Note: 1. The tests for the presence of trend and constant terms in the ADF regression are carried out using the ADF, Φ_2 and Φ_1 statistics (Dickey and Fuller, 1981).

lower than the critical value of the ADF test as given in Table 4.1. If we look at the estimated parameters in Equation (4.25) we can see that the trend variable is insignificant. If the variable T is dropped from the ADF regression, the ADF statistic becomes −4.395, and this value is compared with the critical value under the column τ_μ in Table 4.1. We can see that the null hypothesis of non-stationarity is clearly rejected, and therefore we conclude that the variable $\Delta LPVSP_t$ is a stationary series.

The ADF tests are performed in the same way for the other two variables $LPPDI_t$ and $LRCSP_t$. The results are summarised in Table 4.5. Although the test on $LPPDI_t$ shows that the calculated ADF statistic (just) leads to rejection of the null hypothesis of non-stationarity at the 5% level, we are reluctant to conclude that the per capita disposable income variable is a trend stationary variable, due to the small sample size. We can see from the ADF test results that all three variables appear to be integrated of order one.

4.4 Test for cointegration

Once the order of integration of a time series has been identified, we can test for cointegration. The approach followed in this section is that developed by Engle and Granger (1987).

As has already been mentioned in Section 4.2, the cointegration relationship requires that all variables in the cointegration regression should be integrated of order one (although stationary series, such as event dummies

Table 4.6. Critical values for cointegration ADF test.

	Sample size/Significance level							
	25		50		100		∞	
m^1	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
2	−3.59	−3.22	−3.46	−3.13	−3.39	−3.09	−3.90	−3.05
3	−4.10	−3.71	−3.92	−3.58	−3.83	−3.51	−3.74	−3.45
4	−4.56	−4.15	−4.31	−3.98	−4.21	−3.89	−4.10	−3.81
5	−5.41	−4.96	−5.05	−4.69	−4.88	−4.56	−4.70	−4.42

Note: 1. m is the number of $I(1)$ variables in the long-run static model.
Source: Calculated from MacKinnon (1991).

and a time trend, could also be included). The variables that are integrated of higher orders therefore need to be differenced in order to reduce the order of integration. For example, if a variable is an $I(2)$ series, then the first difference of the variable will be an $I(1)$ series.

Suppose that two variables, x_t and y_t, are found to be $I(1)$ series. The cointegration relationship can be tested by first running the following long-run static regression:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \tag{4.26}$$

The ADF statistic is then employed to test whether the estimated residual, $\hat{\varepsilon}_t$, is stationary or not based on:

$$\Delta\hat{\varepsilon} = \phi^* \hat{\varepsilon}_{t-1} + \sum_{i=1}^{p} \gamma_i^* \Delta\hat{\varepsilon}_{t-i} + u_t \tag{4.27}$$

Note that neither constant nor trend terms should be included in Equation (4.27). The null hypothesis is that the estimated error term in the long-run static model (4.26) is a non-stationary process, i.e., $\hat{\varepsilon}_t$ has unit roots. The test statistic is the t value of the estimated coefficient, ϕ^*. However, the critical values for the Dickey–Fuller tests presented in Table 4.1 cannot be used for the following reasons. First, the OLS estimator tends to minimise the sum of squared residuals from Equation (4.26), which makes the error term $\hat{\varepsilon}_t$ appear to be stationary even if the variables are not cointegrated. Second, the distribution of the t statistic under the null of non-cointegration is affected by the number of variables included in the long-run static model. Therefore, the use of the ADF critical values in Table 4.1 tends to over-reject the null hypothesis of non-cointegration. Using

Monte Carlo simulations, MacKinnon (1991) re-calculated the critical values of the ADF regression for differing numbers of $I(1)$ variables included in the long-run static model. Table 4.6 presents some of the critical values calculated from the MacKinnon (1991) simulation.

If the calculated t statistic of the coefficient ϕ^* in the ADF regression (4.27) is lower than the critical value in Table 4.6, the null hypothesis of non-cointegration is rejected.

The example considered here is the model of demand for tourism to Spain by UK residents. Suppose we want to test whether a long-run equilibrium (cointegration) relationship exists between the three tourism demand variables: $LPVSP_t$, $LPPDI_t$ and $LRCSP_t$. The estimated long-run static model would be:

$$LPVSP_t = -23.010 + 2.369\,LPPDI_t - 0.861\,LRCSP_t \tag{4.28}$$
$$R^2 = 0.82 \qquad \text{DW} = 0.66$$

Notice that the standard errors for the estimated cointegration parameters are omitted. This is because the variables in the static model are non-stationary. Although the OLS estimators are still consistent under the assumption of cointegration, the standard errors are non-standard, and therefore the estimated standard errors cannot be used in significance testing. The diagnostic statistics, apart from the DW statistic, are also not reported, since they are not relevant in the estimation of the long-run static model. The DW statistic can be used as an informal criterion for cointegration; Engle and Granger (1987) called it the Cointegration Durbin-Watson (CIDW). If the value of the CIDW is significantly greater than zero (normally larger than the R^2 from the long-run static model), then the error term from the static model is likely to be a stationary process.

The residuals from the above estimated long-run regression are now tested for stationarity. The ADF regressions with different lag lengths are estimated, and the best model is chosen based on $\bar{R}^2$, AIC and SC. The results are presented below:

$$\Delta\hat{\varepsilon}_t = \underset{(-3.453)}{-0.842\,\hat{\varepsilon}_{t-1}} + \underset{(1.473)}{0.337\,\Delta\hat{\varepsilon}_{t-1}} + \underset{(2.239)}{0.488\,\Delta\hat{\varepsilon}_{t-2}} + \underset{(2.612)}{0.548\,\Delta\hat{\varepsilon}_{t-3}} + \underset{(0.955)}{0.197\,\Delta\hat{\varepsilon}_{t-4}} \tag{4.29}$$
$$\bar{R}^2 = 0.33 \qquad \text{AIC} = -3.639 \qquad \text{SC} = -3.394$$

where the figures in parentheses are t values.

As we can see from Equation (4.29), the calculated ADF t statistic is -3.453[1]. If we compare this value with the critical value where $m = 3$ and the sample size is 25 in Table 4.6, we can see that this value is greater than the corresponding critical values at the 5% and also 10% levels. We therefore conclude that there is no cointegration relationship between $LPCSP_t$, $LPPDI_t$ and $LRCSP_t$. This is also supported by the CIDW statistic. The failure to find a cointegration relationship between these variables may well be an indication that some important variables are missing. This is also consistent with the early results in Chapter 3, where we found that the demand for tourism to Spain can best be modelled by a dead start or partial adjustment process.

To summarise, the procedure for testing for a cointegration relationship involves the following two steps:

(1) Test for the order of integration of all the variables in the long-run static model using the ADF test. The cointegration regression requires that all the explanatory variables are integrated of order one, so variables that have a higher order of integration should be differenced in order to yield $I(1)$ series. Since most tourism demand variables are likely to be integrated of order one, we will find in most cases that the long-run cointegration relationship relates only to $I(1)$ variables.
(2) Use OLS to estimate the long-run static regression of the form:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \beta_m x_{mt} + \varepsilon_t \tag{4.30}$$

The estimated residuals from Equation (4.30) are then tested for stationarity using the ADF statistic. The calculated t ratio of the coefficient ϕ^* with n degrees of freedom in (4.27) is compared with the MacKinnon critical values presented in Table 4.6. If the calculated t value is lower than the corresponding MacKinnon critical value, the null hypothesis of non-cointegration is rejected. Moreover, the Cointegration Durbin-Watson (CIDW) statistic can also be used as a rough guide for cointegration.

Note

1. The corresponding Phillips–Perron test statistics is −2.954.

5

Error correction model

5.1 Cointegration and error correction mechanism

Engle and Granger (1987) show that cointegrated variables can always be transformed into an error correction mechanism (ECM) and vice versa. This bi-directional transformation is often called the 'Granger Representation Theorem' and implies that there is some adjustment process that prevents economic variables from drifting too far away from their long-run equilibrium time path. The cointegration and error correction models are very useful in situations where both long-run equilibrium and short-run disequilibrium behaviour are of interest. In tourism demand analysis, the long-run equilibrium behaviour of tourists is expected to be a major concern of policy-makers and planners whilst the short-run dynamics are likely to provide useful information for short-term business forecasting and managerial decisions.

This chapter explores how cointegration and error correction are linked, and the ways in which error correction models are estimated within a single equation framework.

5.1.1 From ADLM to error correction model

We started in Chapter 3 with the general ADLM of the form

$$y_t = \alpha + \sum_{j=1}^{k} \sum_{i=0}^{p} \beta_{ji} x_{jt-i} + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t \tag{5.1}$$

Equation (5.1) can be re-parameterised into an ECM of the form

$$\Delta y_t = (\textit{current and lagged } \Delta x_{jt}\text{s, } \textit{lagged } \Delta y_t\text{s})$$
$$-(1-\phi_1)\left[y_{t-1} - \sum_{j=1}^{k} \xi x_{jt-1}\right] + \varepsilon_t \tag{5.2}$$

We shall demonstrate this in the case of an ADLM(1,1) model, but the derivation can be extended to a general ADLM(p, q) process. The ADLM(1,1) model takes the form

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \phi_1 y_{t-1} + \varepsilon_t \tag{5.3}$$

Subtracting y_{t-1} from both sides of Equation (5.3) yields:

$$\Delta y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} - (1-\phi_1) y_{t-1} + \varepsilon_t$$

or

$$\Delta y_t = \alpha + \beta_0 \Delta x_t + (\beta_0 + \beta_1) x_{t-1} - (1-\phi_1) y_{t-1} + \varepsilon_t. \tag{5.4}$$

Equation (5.4) can be further re-parameterised to give

$$\Delta y_t = \beta_0 \Delta x_t - (1-\phi_1)[y_{t-1} - k_0 - k_1 x_{t-1}] + \varepsilon_t \tag{5.5}$$

where $k_0 = \alpha/(1-\phi_1)$, $\quad k_1 = (\beta_0 + \beta_1)/(1-\phi_1)$.

The parameter β_0 is called the impact parameter, $(1-\phi_1)$ is the feed-back effect, k_0 and k_1 are the long-run response coefficients, and the combination of the terms in the square brackets is called the error correction mechanism. Since the coefficient ϕ_1 is less than 1 and greater than 0, the coefficient of the error correction term, $-(1-\phi_1)$, is greater than -1 and less than 0. This implies that the system will adjust itself towards equilibrium by removing $(1-\phi_1)$ of a unit from the error made in the previous period. Although Equations (5.3) and (5.5) are different in their functional forms, they actually represent the same data-generating process.

Equation (5.5) has the following advantages over Equation (5.3). First, Equation (5.5) reflects both the long-run and short-run effects in a single model. The specification indicates that changes in y_t depend on changes in x_t and the disequilibrium error in the previous period.

Second, Equation (5.5) overcomes the problem of spurious correlation by employing differenced variables. It can be easily shown that the term $[y_{t-1} - k_0 - k_1 x_{t-1}]$ is a stationary process if y_t and x_t are cointegrated. Therefore, it is unlikely that the residuals in (5.5) will be correlated. Moreover, the use of (5.5) avoids the problems associated with the growth rate model in which only differenced data are used.

Third, the ECM fits in well with the general-to-specific methodology.

Since the ECM is another way of writing the general ADLM, acceptance of the ADLM is equivalent to acceptance of the error correction model.

Fourth, the estimation of Equation (5.5) reduces the problem of data mining, since in the model reduction process one is permitted to eliminate the differenced variables according to statistical significance. However, the elimination of lagged levels variables is not permitted since they represent the cointegration relationship. For example, suppose that a cointegration relationship is found between the variables y_t, x_t and z_t. In the estimation of the ECM in which both differenced and lagged levels forms of y, x and z are involved, the researcher is free to eliminate Δy, Δx and/or Δz, but the lagged levels variables, y, x and z should always appear in the final ECM.

Finally, estimation of the general ADLM (5.3) which involves a large number of explanatory variables tends to suffer from the problem of multicollinearity, i.e., several of the explanatory variables are likely to be highly correlated which will result in abnormally large standard errors and hence the calculated t statistics cannot be used as a reliable criterion for hypothesis testing. However, the corresponding variables in the ECM are less likely to be highly correlated. In fact, Engle and Granger (1987) show that the explanatory variables in the ECM are almost orthogonal (i.e., the correlation is almost zero). This is a desirable property, as the t statistics provide a reliable guide for the elimination of differenced variables. Consequently it is easier for a researcher to arrive at a sufficiently parsimonious final preferred model using the testing-down procedure of the general-to-specific methodology.

5.1.2 From ECM to cointegration regression

Engle and Granger (1987) demonstrate that if a pair of economic variables is cointegrated, they can always be represented by an ECM and vice versa. This can be shown by the following transformation. The long-run steady state suggests that $y_t = y_{t-1}$ and $x_t = x_{t-1}$, i.e., $\Delta y_t = \Delta x_t = 0$. Therefore, the ECM (5.5) becomes:

$$0 = -(1 - \phi_1)[y_t - k_0 - k_1 x_t]$$

i.e.,

$$y_t = k_0 + k_1 x_t \tag{5.6}$$

Equation (5.6) is the long-run cointegration regression with k_0 and k_1 being the long-run cointegration coefficients. Now $k_0 = \alpha/(1 - \phi_1)$ and $k_1 = (\beta_0 + \beta_1)/(1 - \phi_1)$, and therefore the long-run cointegration coefficients (vector) can be obtained from the estimates of the general ADLM model (5.1).

5.2 Estimating the ECM

Various procedures for estimating the single equation ECM have been proposed. In this section three estimation methods are explained: the Engle–Granger two-stage, the Wickens–Breusch one-stage and the ADLM procedures.

5.2.1 Engle–Granger two-stage approach

In their seminal paper Engle and Granger (1987) suggest testing for the long-run equilibrium relationship between a set of economic variables and modelling their short-run dynamics via an ECM in a two-step procedure. The first step is to test for a cointegration relationship between the two $I(1)$ variables, y_t and x_t, based on the static long-run equilibrium regression:

$$y_t = k_0 + k_1 x_t + \varepsilon_t \tag{5.7}$$

If the estimated residual term is a stationary process, y_t and x_t are said to be cointegrated. According to Stock (1987) if y_t and x_t are cointegrated, OLS estimates of the cointegration vector (k_0, k_1) will be 'super-consistent'.

After confirming the acceptance of a cointegration relationship, the second step is to estimate the ECM

$$\Delta y_t = \sum_{i=0}^{p} \beta_i \Delta x_{t-i} + \sum_{j=1}^{p} \phi_j \Delta y_{t-j} - \lambda \hat{\varepsilon}_{t-1} + u_t \tag{5.8}$$

where $\hat{\varepsilon}_{t-1} = y_{t-1} - \hat{k}_0 - \hat{k}_1 x_{t-1}$ is the OLS residual from (5.7).

As we can see, the error correction equation (5.8) consists of differenced variables with appropriate lags, and the lag structure is determined by experimentation. In this stage the short-run parameters are obtained. According to Engle and Granger (1987), the estimates of the short-run parameters are consistent and efficient. The estimated standard errors of the parameters in the second stage are the true standard errors, and therefore the model can be used for forecasting and policy evaluation.

One of the concerns about the Engle and Granger two-stage approach is its first step. As we already know that the variables in the cointegration regression are $I(1)$ variables, the estimated standard errors of the cointegration coefficients are not standard normal. That is why many researchers do not even bother to report them in their empirical studies. Moreover, the Engle–Granger two-stage approach does not prove that the cointegration regression is really a long-run one. This is an assumption and cannot be tested statistically. Researchers therefore have to have a good justification that the variables in the cointegration regression represent the long-run equilibrium relationship, and normally the justification is the relevant economic theory. Another problem associated with the Engle–Granger method is that it tends to produce biased estimates of the long-run coefficients in small samples.

5.2.2 Wickens–Breusch one-stage approach

It has been shown by Wickens and Breusch (1988) that although the estimates of the short-run ECM of the Engle–Granger two-stage approach are consistent and efficient in large samples, they are biased in small samples. An alternative estimation method which overcomes this problem has been suggested by Wickens and Breusch, and involves estimating both the long-run and short-run parameters in a single step using OLS. The estimation is based on:

$$\Delta y_t = \alpha + \sum_{i=0}^{p-1} \beta_i \Delta x_{t-i} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \lambda_1 y_{t-1} + \lambda_2 x_{t-1} + u_t \tag{5.9}$$

Again the lag lengths of the differenced variables are determined by the statistical significance of the estimated coefficients in (5.9). After estimating (5.9), the long-run cointegration parameters may be derived from:

$$y_t = -\frac{\hat{\alpha}}{\hat{\lambda}_1} - \frac{\hat{\lambda}_2}{\hat{\lambda}_1} x_t$$

or

$$y_t = k_0^* + k_1^* x_t \tag{5.10}$$

where $k_0^* = -\hat{\alpha}/\hat{\lambda}_1$ and $k_2^* = -\hat{\lambda}_2/\hat{\lambda}_1$.

Wickens and Breusch (1988) show that the OLS estimates of both the long-run and short-run parameters in Equation (5.9) are consistent, efficient and unbiased.

5.2.3 ADLM approach

This method of estimating the ECM was developed by Pesaran and Shin (1995), and involves estimating both the long-run cointegration relationship and the short-run ECM based on a general ADLM. Our discussion of this estimation approach is based on Equation (5.3), but the principle is the same for the more general form of the ADLM as given in Equation (5.1).

The ADLM approach first estimates Equation (5.3) with extended lag structure using OLS.

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \ldots + \beta_p x_{t-p} + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_t \tag{5.11}$$

The optimal lag structure of the ADLM from which the ECM is derived is determined by the following criteria: $\overline{R}^2$, AIC and SBC as described in Chapter 2. The long-run cointegration regression is $y_t = \hat{k}_0 + \hat{k}_1 x_t$, where

$$\hat{k}_1 = \frac{\hat{\beta}_0 + \hat{\beta}_1 + \ldots + \hat{\beta}_p}{1 - \hat{\phi}_1 - \hat{\phi}_2 - \ldots - \hat{\phi}_p} \tag{5.12}$$

and

$$\hat{k}_0 = \frac{\hat{\alpha}}{1 - \hat{\phi}_1 - \hat{\phi}_2 - \ldots - \hat{\phi}_p} \tag{5.13}$$

The corresponding short-run ECM may be derived from Equation (5.11). Since

$$\Delta y_t = y_t - y_{t-1}$$
$$y_t = \Delta y_t + y_{t-1}$$

Also

$$x_t = \Delta x_t + x_{t-1}$$

Substituting these relationships into Equation (5.11) and rearranging the terms gives the following ECM representation:

$$\Delta y_t = \sum_{i=0}^{p-1} \beta_i^* \Delta x_{t-i} + \sum_{i=1}^{p-1} \phi_i^* \Delta y_{t-i} - \lambda EC_{t-1} + \varepsilon_t \tag{5.14}$$

where $\lambda = 1 - \hat{\phi}_1 - \hat{\phi}_2 - \ldots - \hat{\phi}_{p-1}$, $EC_{t-1} = y_{t-1} - \alpha - \beta_1 x_{t-1}$, which is the error-correction term, and the coefficients β_i^* and ϕ_j^* are calculated from:

$$\begin{aligned}
\beta_0^* &= \hat{\beta}_p + \hat{\beta}_{p-1} + \ldots + \hat{\beta}_2 + \hat{\beta}_1 \\
\beta_1^* &= \hat{\beta}_p + \hat{\beta}_{p-1} + \ldots + \hat{\beta}_2 \\
&\vdots \\
\beta_{p-1}^* &= \hat{\beta}_p
\end{aligned}$$

and

$$\begin{aligned}
\phi_1^* &= \hat{\phi}_p + \hat{\phi}_{p-1} + \ldots + \hat{\phi}_3 + \hat{\phi}_2 \\
\phi_2^* &= \hat{\phi}_p + \hat{\phi}_{p-1} + \ldots + \hat{\phi}_3 \\
&\vdots \\
\phi_{p-1}^* &= \hat{\phi}_p.
\end{aligned}$$

The $\hat{\beta}_i$s and $\hat{\phi}_j$s are the estimated parameters from the ADLM (5.11). The standard errors of these estimated coefficients can be calculated using the formula provided by Bewley (1979).

As noted by Pesaran and Shin (1995), however, the ADLM approach also tends to suffer from small-sample bias, although the consistency improves with increasing sample size.

5.3 Worked examples

In this section we use inbound tourism demand data for South Korea to illustrate the error correction modelling procedure. The demand for tourism to Korea by two major tourism-generating countries, the UK and USA, is examined. The total number of tourist arrivals by country ranging from 1962 to 1994 (33 observations) is used as the dependent variable and the data are obtained from the Korea National Tourism Corporation (KNTC) Annual Statistical Report jointly published by the KNTC and the Ministry of Culture and Sports. Since the tourist arrivals variable includes both business and leisure travellers, the gross domestic product (GDP) of the tourism-generating country is used as the income variable, rather than personal disposable income. To reflect the influence of business activities on tourism demand, a trade volume variable measured by the sum of total imports and exports between Korea and the tourist-generating country is also included in the model. The tourism price variable is measured by the relative consumer price index (CPI) of Korea to that of the tourism-generating country, adjusted by the corresponding exchange rate (see Equation 5.16). The data are provided in Table A5.1 in Appendix 5.1.

5.3.1. Testing for cointegration

The proposed long-run demand model is of the form

$$\ln TA_{it} = \alpha_0 + \alpha_1 \ln GDP_{it} + \alpha_2 \ln TV_{it} + \alpha_3 \ln LRCPI_{it} + u_t \tag{5.15}$$

where TA_{it} is total tourist arrivals in South Korea from country i ($i = 1$ represents UK and $i = 2$ represents USA) in year t,

GDP_{it} is the index of the real gross domestic product of origin i in year t,

TV_{it} is real trade volume measured by total imports and exports between South Korea and country i in year t,

$RCPI_{it}$ is the tourism price variable calculated from:

$$RCPI_{it} = \frac{CPI_{Korea,t}/EX_{Koreali,t}}{CPI_{it}} \tag{5.16}$$

This takes into account the effects of both relative inflation and the exchange rate on the demand for tourism to Korea.

The null hypothesis is that the variables in Equation (5.15) are cointegrated. The integration orders (or the numbers of unit roots) of the variables are examined prior to the cointegration test. In testing for unit roots, the Dickey–Fuller and Phillips–Perron tests discussed in Chapter 4 are used. The results of these tests show that all the series (in logarithms) are likely to be $I(1)$ with the possibility that USA GDP is a trended stationary process. Since a trended stationary series still needs to be differenced to get rid of the trend, the inclusion of this variable in the cointegration regression is not unjustified.

The long-run cointegration regressions are estimated using OLS as:

$$\text{UK:}\quad \ln TA_{1t} = -4.257 + 2.097\ln GDP_{1t} + 0.473\ln TV_{1t} - 0.213\ln LRCPI_{1t} \qquad R^2 = 0.995 \qquad \text{CIDW} = 1.516 \tag{5.17}$$

$$\text{USA:}\quad \ln TA_{2t} = -6.158 + 1.655\ln GDP_{2t} + 0.576\ln TV_{2t} - 0.861\ln LRCPI_{2t} \qquad R^2 = 0.971 \qquad \text{CIDW} = 0.490 \tag{5.18}$$

As a rough guide, if there is a cointegration relationship between a set of variables, the CIDW statistic in the cointegration regression should be greater than its R-squared value. According to this criterion the variables in the UK model are likely to be cointegrated whilst those in the USA model are not. This can be further confirmed by the Dickey–Fuller and Phillips–Perron tests on the OLS residuals from Equations (5.17) and (5.18). The results are presented in Table 5.1.

The critical value for non-cointegration with four $I(1)$ variables in the long-run model with 25 degrees of freedom is −4.56 at the 5% level and −4.15 at the 10% level. The results in Table 5.1 show that the tourism demand variables are cointegrated in the UK model, but not in the USA equation.

Table 5.1 Dickey–Fuller (DF) and Phillips–Perron (PP) test results.

Tests	UK	USA
DF	−4.733(0)	−3.490(3)
PP	−4.707(3)	−3.188(3)

Figures in parentheses are the lag lengths used in the tests.

The long-run equations (5.17 and 5.18) show that all the demand elasticities are correctly signed and significant at the 5% level. The income elasticities are positive and greater than 1 suggesting that travelling to South Korea is a luxury for tourists from the UK and USA. However, since the variables in the USA long-run model are not cointegrated, the estimated long-run elasticities are biased.

The significance of the trade volume variable is not surprising since a relatively large proportion (about 20%) of tourist arrivals in South Korea is related to business travel. According to Lee, Var and Blaine (1996), business travellers dominated Korean inbound tourism in the 1960s, and they paved the way for the surge of pleasure tourists to South Korea since the early 1970s.

The relative price elasticity in the UK and USA long-run models is less than 1. This indicates that tourism demand is price inelastic.

5.3.2 Error correction models

This section presents the Engle and Granger two-stage, the Wickens and Breusch one-stage and the ADLM procedures for ECM estimation. Although a cointegration relationship does not exist in the USA model, we still try to model the short-run dynamics using the ECM. This will allow us to compare the consequences of ECMs with and without cointegrated variables.

Engle–Granger two-stage approach. As mentioned earlier, if a set of variables is found to be cointegrated, the cointegration regression can always be transformed into an error correction model of the form:

$$\Delta \ln TA_{it} = \beta_0 + \beta_1 \Delta \ln GDP_{it} + \beta_2 \Delta \ln TV_{it} + \beta_3 \Delta \ln RCPI_{it} + \delta \hat{u}_{t-1} + \varepsilon_t \qquad (5.19)$$

where $\hat{u}_{t-1}$ is the estimated error term obtained from the cointegration regression Equation (5.15) and Δ is the first difference operator. The coefficient of $\hat{u}_{t-1}$ is expected to be negative and significant. In estimating Equation (5.19) lagged dependent and independent variables are also included and a 'test-down' procedure is employed repeatedly until the most parsimonious specification is achieved.

The initial estimates of the general ECMs for the UK and USA are given in Equations (5.20) and (5.21):

UK:

$$
\begin{aligned}
\Delta \ln TA_{1t} = {} & \underset{(1.736)}{0.048} - \underset{(\ 0.349)}{0.072}\,\Delta \ln TA_{1t-1} + \underset{(1.645)}{0.718}\,\Delta \ln GDP_{1t} + \underset{(0.220)}{0.126}\,\Delta \ln GDP_{1t-1} \\
& + \underset{(4.837)}{0.340}\,\Delta \ln TV_{it} + \underset{(0.993)}{0.074}\,\Delta \ln TV_{1t-1} + \underset{(0.417)}{0.055}\,\Delta \ln RCPI_{1t} \\
& - \underset{(2.201)}{0.280}\,\Delta \ln RCPI_{1t-1} - \underset{(-1.891)}{0.384}\hat{u}_{1t-1}
\end{aligned}
\tag{5.20}
$$

$$R^2 = 0.608 \qquad \sigma = 0.082$$

USA:

$$
\begin{aligned}
\Delta \ln TA_{2t} = {} & \underset{(-0.411)}{-0.017} + \underset{(1.450)}{0.224}\,\Delta \ln TA_{2t-1} - \underset{(-0.419)}{0.431}\,\Delta \ln GDP_{1t} + \underset{(0.488)}{0.561}\,\Delta \ln GDP_{1t-1} \\
& + \underset{(3.168)}{0.858}\,\Delta \ln TV_{it} + \underset{(0.368)}{0.092}\,\Delta \ln TV_{1t-1} - \underset{(-1.827)}{0.665}\,\Delta \ln RCPI_{1t} \\
& - \underset{(-0.067)}{0.026}\,\Delta \ln RCPI_{1t-1} - \underset{(-3.056)}{0.561}\hat{u}_{2t-1}
\end{aligned}
\tag{5.21}
$$

$$R^2 = 0.610 \qquad \sigma = 0.105$$

The values in parentheses are t ratios. The insignificant variables judged by the t values are eliminated one by one starting with the most insignificant ones until all remaining coefficients are statistically significant. The final ECMs for the UK and USA are given below:

UK:

$$
\begin{aligned}
\Delta \ln TA_{1t} = {} & \underset{(2.595)}{0.052} + \underset{(1.887)}{0.752}\,\Delta \ln GDP_{1t} + \underset{(5.553)}{0.341}\,\Delta \ln TV_{it} \\
& - \underset{(-2.176)}{0.239}\,\Delta \ln RCPI_{1t-1} - \underset{(-2.976)}{0.468}\hat{u}_{1t-1}
\end{aligned}
\tag{5.22}
$$

$$
\begin{gathered}
R^2 = 0.584 \qquad \sigma = 0.078 \qquad \chi^2_{Auto}(2) = 1.469 \\
\chi^2_{Norm}(2) = 1.028 \qquad \chi^2_{ARCH}(1) = 0.049 \qquad \chi^2_{White}(9) = 9.88 \\
\chi^2_{RESET}(2) = 1.388 \qquad F_{Forecast}(5{,}23) = 0.260
\end{gathered}
$$

USA:

$$
\begin{aligned}
\Delta \ln TA_{2t} = {} & \underset{(-0.357)}{-0.013} + \underset{(1.530)}{1.356}\,\Delta \ln GDP_{2t-1} + \underset{(4.195)}{0.823}\,\Delta \ln TV_{2t} \\
& - \underset{(-2.142)}{0.666}\,\Delta \ln RCPI_{2t} - \underset{(-3.155)}{0.440}\hat{u}_{2t-1}
\end{aligned}
\tag{5.23}
$$

$$
\begin{gathered}
R^2 = 0.554 \qquad \sigma = 0.104 \qquad \chi^2_{Auto}(2) = 3.340 \\
\chi^2_{Norm}(2) = 8.165^{**} \qquad \chi^2_{ARCH}(1) = 0.588 \qquad \chi^2_{White}(9) = 19.87^{**} \\
\chi^2_{RESET}(2) = 23.30^{***} \qquad F_{Forecast}(5{,}23) = 0.885
\end{gathered}
$$

** and *** denote that the statistics are significant at the 5% and 1% levels, respectively. The diagnostic statistics are explained in Chapter 3.

The estimation results given in Equations (5.22) and (5.23) show that the UK short-run ECM passes all the diagnostic statistics whilst the USA equivalent fails the normality, homoscedasticity and the functional form tests. This is not surprising since no cointegration relationship was found in the USA long-run model based on the Engle–Granger procedure. In practice, if a cointegration relationship does not exist between a set of economic variables, the ECM specification should not be used. The solution to this problem is either to try an alternative modelling strategy or to re-specify the long-run model in terms of a different functional form and/or inclusion of new explanatory variables based on economic theory.

The estimated coefficients in the short-run error correction model are short-run demand elasticities. Since the USA ECM is not statistically acceptable, there is no point in trying to interpret the economic meanings of the coefficients. However, the estimated demand elasticities in the UK ECM are unbiased, efficient and consistent, and therefore economic interpretations of the estimated coefficients are possible. The short-run income and trade elasticities in Equation (5.22) are lower than those in the long-run model, as expected on the basis of economic theory, but the short-run price elasticity of the demand for Korean tourism by UK residents remains more or less the same as that in the long-run model. This suggests that the relative cost of living for UK tourists in Korea has a similar influence on both long-run and short-run decision-making for UK residents. The coefficient of the error correction term in the UK ECM is correctly signed and significant at the 1% level as expected.

Wickens–Breusch one-stage approach. The general tourism demand models are first estimated using the Wickens–Breusch method based on the same data set, and the results are given below:

UK:

$$\begin{aligned}
\Delta \ln TA_{1t} = {} & \underset{(-0.939)}{-1.339} - \underset{(-0.689)}{0.146}\Delta \ln TA_{1t-1} + \underset{(1.230)}{0.638}\Delta \ln GDP_{1t} - \underset{(-0.220)}{0.021}\Delta \ln GDP_{1t-1} \\
& + \underset{(3.935)}{0.297}\Delta \ln TV_{it} + \underset{(1.012)}{0.080}\Delta \ln TV_{1t-1} + \underset{(0.702)}{0.115}\Delta \ln RCPI_{1t} \\
& - \underset{(-1.965)}{0.290}\Delta \ln RCPI_{1t-1} - \underset{(-1.630)}{0.342}\ln TA_{1t-1} + \underset{(1.414)}{0.752}\ln GDP_{1t-1} \\
& + \underset{(1.369)}{0.143}\ln TV_{1t-1} - \underset{(-0.506)}{0.058}\ln RCPI_{1t-1}
\end{aligned}$$

$$R^2 = 0.656 \qquad \sigma = 0.083$$

USA:

$$\begin{aligned}\Delta \ln TA_{2t} = & \underset{(\ 2.150)}{-4.908} - \underset{(\ 0.164)}{0.028}\Delta \ln TA_{2t-1} - \underset{(\ 0.932)}{0.957}\Delta \ln GDP_{2t} + \underset{(0.100)}{0.116}\Delta \ln GDP_{2t-1} \\ & + \underset{(2.274)}{0.634}\Delta \ln TV_{2t} - \underset{(-0.165)}{0.045}\Delta \ln TV_{2t-1} - \underset{(-0.951)}{0.353}\Delta \ln RCPI_{2t} \\ & + \underset{(0.652)}{0.262}\Delta \ln RCPI_{2t-1} - \underset{(-3.151)}{0.561}\ln TA_{2t-1} + \underset{(1.313)}{0.084}\ln GDP_{2t-1} \\ & + \underset{(1.701)}{0.263}\ln TV_{2t-1} - \underset{(-2.810)}{0.821}\ln RCPI_{2t-1}\end{aligned}$$

$$R^2 = 0.727 \qquad \sigma = 0.095$$

Since some of the coefficients in the models are insignificant, a model-reduction process is needed. Following the general-to-specific methodology the insignificant variables are eliminated based on the t statistic values, and the final ECMs become:

UK:

$$\begin{aligned}\Delta \ln TA_{1t} = & \underset{(-0.837)}{-0.858} + \underset{(4.156)}{0.281}\Delta \ln TV_{it} - \underset{(-2.669)}{0.443}\ln TA_{1t-1} + \underset{(1.488)}{0.604}\ln GDP_{1t-1} \\ & + \underset{(2.376)}{0.208}\ln TV_{1t-1} - \underset{(-2.305)}{0.178}\ln RCPI_{1t-1}\end{aligned} \tag{5.24}$$

$$R^2 = 0.541 \qquad \sigma = 0.082 \qquad \chi^2_{Auto}(2) = 0.717$$
$$\chi^2_{Norm}(2) = 0.974 \qquad \chi^2_{ARCH}(1) = 0.155 \qquad \chi^2_{White}(15) = 15.890$$
$$\chi^2_{RESET}(2) = 4.356 \qquad F_{Forecast}(5{,}23) = 0.137$$

USA:

$$\begin{aligned}\Delta \ln TA_{2t} = & \underset{(-2.510)}{-4.171} + \underset{(2.209)}{0.450}\Delta \ln TV_{2t} + \underset{(1.600)}{0.135}\Delta \ln RCPI_{2t-1} - \underset{(-3.791)}{0.531}\ln TA_{2t-1} \\ & + \underset{(1.875)}{0.836}\ln GDP_{2t-1} + \underset{(1.875)}{0.235}\ln TV_{2t-1} - \underset{(-3.210)}{0.696}\ln RCPI_{2t-1}\end{aligned} \tag{5.25}$$

$$R^2 = 0.703 \qquad \sigma = 0.088 \qquad \chi^2_{Auto}(2) = 0.393$$
$$\chi^2_{Norm}(2) = 0.003 \qquad \chi^2_{ARCH}(1) = 2.652 \qquad \chi^2_{White}(14) = 22.560^{**}$$
$$\chi^2_{RESET}(2) = 22.188^{***} \qquad F_{Forecast}(5{,}23) = 0.467$$

Again the UK model is highly satisfactory in terms of the diagnostic statistics whereas the USA model suffers from the same problems as those in the Engle–Granger two-stage model. Although the coefficient of the variable $\Delta \ln RCPI_{2t-1}$ is not significant at the 5% level, we keep it in the final specification since the elimination of this variable drastically reduces the explanatory power of the model. To obtain the long-run parameters, Equations (5.24) and (5.25) can be transformed to give:

UK:

$$\Delta \ln TA_{1t} = 0.304\Delta \ln TV_{it} - 0.443(\ln TA_{1t-1} + 1.937 - 1.363\ln GDP_{1t-1} - 0.469\ln TV_{1t-1} + 0.402\ln RCPI_{1t-1}) \quad (5.26)$$

USA:

$$\Delta \ln TA_{2t} = 0.450\Delta \ln TV_{2t} + 0.135\Delta \ln RCPI_{2t-1} - 0.530(\ln TA_{2t-1} + 7.870 - 1.577\ln GDP_{2t-1} - 0.443\ln TV_{2t-1} + 1.313\ln RCPI_{2t-1}) \quad (5.27)$$

Theoretically, the Engle–Granger method and the Wickens–Breusch approach should produce similar results for large samples if the cointegration relationship exists in the levels variables. However, since the lagged levels variables in the parentheses in the UK model (5.24) represent the cointegration relationship and the combination of these levels variables tends to be stationary, the OLS estimates of the long-run parameters in this method are more reliable. In contrast, OLS estimation in the first stage of the Engle–Granger method tends to be biased and these biases are likely to be carried over to the second stage of the estimation. For the USA model, since no cointegration relationship has been found, the Wickens–Breusch ECM still suffers from similar problems to those of the corresponding Engle–Granger ECM.

The long-run relationships for the UK and USA can therefore be derived as:

UK:

$$\ln TA_{1t-1} = -1.937 + 1.363\ln GDP_{1t-1} + 0.469\ln TV_{1t-1} - 0.402\ln RCPI_{1t-1} \quad (5.28)$$

USA:

$$\ln TA_{2t-1} = -7.870 + 1.577\ln GDP_{2t-1} + 0.443\ln TV_{2t-1} - 1.313\ln RCPI_{2t-1} \quad (5.29)$$

Comparing Equations (5.28) and (5.29) with (5.17) and (5.18), respectively, we see that the long-run parameters estimated from the Engle–Granger method are not the same as those from the Wickens–Breusch approach for both the UK and USA models.

To conclude, the Engle–Granger and the Wickens–Breusch cointegration-ECM approaches should produce similar results for large samples. In the case of small samples, the Wickens–Breusch method is preferred due to the consistent and unbiased nature of the estimation.

ADLM method. Since the data are annual, we introduce one lag for each of the variables in the general ADLM[1], that is:

$$\ln TA_{it} = \alpha + \beta_{11}\ln GDP_{it} + \beta_{12}\ln GPD_{it-1} + \beta_{21}\ln TV_{it} + \beta_{22}\ln TV_{it-1} + \beta_{31}\ln RCPI_{it} + \beta_{32}\ln RCPI_{it-1} + u_t \quad (5.30)$$

The estimates of the ADLM (5.30) using the UK and USA data are given below[2]:

UK:

$$\begin{aligned}\ln TA_{1t} = & \underset{(-1.089)}{-1.321} + \underset{(3.116)}{0.544}\ln TA_{1t-1} + \underset{(1.217)}{0.582}\ln GDP_{1t} + \underset{(0.342)}{0.194}\ln GPD_{1t-1}\\ & + \underset{(4.017)}{0.281}\ln TV_{1t} - \underset{(-0.956)}{0.072}\ln TV_{1t-1} - \underset{(-0.027)}{0.004}\ln RCPI_{1t} \\ & - \underset{(-1.089)}{0.144}\ln RCPI_{1t-1}\end{aligned} \tag{5.31}$$

$R^2 = 0.997$ $\quad\sigma = 0.082$ $\quad\chi^2_{Auto}(1) = 0.022$ $\quad\chi^2_{Norm}(2) = 0.921$
$\chi^2_{Hetro}(1) = 1.912$[3] $\quad\chi^2_{RESET}(1) = 0.875$

USA:

$$\begin{aligned}\ln TA_{2t} = & \underset{(-3.290)}{-5.140} + \underset{(6.238)}{0.617}\ln TA_{2t-1} - \underset{(-0.772)}{0.644}\ln GDP_{2t} + \underset{(1.641)}{1.555}\ln GPD_{2t-1}\\ & + \underset{(1.957)}{0.440}\ln TV_{2t} - \underset{(-1.614)}{0.334}\ln TV_{2t-1} - \underset{(-1.434)}{0.428}\ln RCPI_{2t} \\ & - \underset{(-0.836)}{0.289}\ln RCPI_{2t-1}\end{aligned} \tag{5.32}$$

$R^2 = 0.994$ $\quad\sigma = 0.091$ $\quad\chi^2_{Auto}(1) = 0.894$ $\quad\chi^2_{Norm}(2) = 7.920^{***}$
$\chi^2_{Hetro}(1) = 4.279^{**}$ $\quad\chi^2_{RESET}(1) = 7.783^{***}$

The ADLM method uses the full lag structure to derive the long-run and short-run parameters, and the insignificant parameters are therefore not eliminated. However, Pesaran and Shin (1995) show that the optimal lag structure can be determined by a number of criteria such as the Akaike information criterion (AIC), Schwarz Bayesian criterion (SBC), and the Hannan–Quinn criterion (HQC). For illustrative purposes we impose the restriction that lag length equals one.

The results show that the UK model is well specified according to the diagnostics, but the USA model still appears to have problems with regard to functional form, heteroscedasticity and normality. Based on the formulae (5.12) and (5.13), we can derive the long-run static parameters. Assuming there are no dynamics in the long run, Equation (5.31) becomes:

$$\begin{aligned}(1 - 0.544)\ln TA_{1t} = & -1.321 + (0.582 + 0.194)\ln GPD_{1t}\\ & + (0.281 - 0.072)\ln TV_{1t}\\ & - (0.004 + 0.144)\ln RCPI_{1t}\end{aligned}$$

If both sides of the above equation are divided by $(1 - 0.544)$, we get:

$$\ln TA_{1t} = -2.897 + 1.702 \ln GDP_{1t} + 0.456 \ln TV_{1t} - 0.322 \ln RCPI_{1t} \quad (5.33)$$

This is the long-run model for the UK. Using the same procedure, the USA long-run model can be derived as follows:

$$\ln TA_{2t} = -13.425 + 2.380 \ln GDP_{1t} + 0.277 \ln TV_{1t} - 1.871 \ln RCPI_{1t} \quad (5.34)$$

Although the standard errors (or t ratios) of the coefficients in (5.33) and (5.34) are omitted, they can be calculated according to the method proposed by Bewley (1979).

The corresponding ECMs are estimated based on the estimated long-run regressions (5.33) and (5.34).

UK:

$$\begin{aligned}\Delta \ln TA_{1t} = \underset{(-1.089)}{-1.321} + \underset{(1.216)}{0.582}\Delta \ln GDP_{1t} + \underset{(4.017)}{0.281}\Delta \ln TV_{it} \\ - \underset{(-0.026)}{0.004}\Delta \ln RCPI_{1t-1} - \underset{(-2.612)}{0.456}\hat{EC}_{1t-1}\end{aligned} \quad (5.35)$$

where

$$\hat{EC}_{1t-1} = \ln \hat{TA}_{1t-1} + 2.897 - 1.702 \ln GDP_{1t-1} - 0.456 \ln TV_{1t-1} + 0.322 \ln RCPI_{1t-1}$$

USA:

$$\begin{aligned}\Delta \ln TA_{2t} = \underset{(-3.290)}{-5.140} - \underset{(-0.772)}{0.644}\Delta \ln GDP_{2t} + \underset{(1.957)}{0.440}\Delta \ln TV_{2t} \\ - \underset{(-1.433)}{0.428}\Delta \ln RCPI_{2t-1} - \underset{(-3.871)}{0.383}\hat{EC}_{2t-1}\end{aligned} \quad (5.36)$$

where

$$\hat{EC}_{2t-1} = \ln \hat{TA}_{2t-1} + 13.425 - 2.379 \ln GDP_{2t-1} - 0.277 \ln TV_{2t-1} + 1.871 \ln RCPI_{2t-1}$$

Apart from the R-squared values, the other diagnostic statistics for (5.35) and (5.36) are the same as those in (5.31) and (5.32).

As can be seen, the ADLM estimation method is also a two-stage process. The advantage of this method over the Engle–Granger two-stage approach is that in the first stage the error term in the ADLM tends to be independent and identically distributed, and therefore the diagnostic statistics and the standard errors of the coefficients are standard and the normal critical values can be used for inference.

As to which method performs best, there is no clear-cut answer to this question. But if the sample is sufficiently large and there is only one cointegration relationship between a number of variables, then all three

Table 5.2 Forecast errors generated from the three UK ECMs.

Model	MAE	RMSE
Engle–Granger	0.04378	0.04690
Wickens–Breusch	0.05928	0.06167
ADLM	0.03221	0.03633

methods should produce similar results. However, in the case of multi-cointegration relationships, a more advanced approach such as the Johansen (1988) method should be used. This approach will be discussed in the next chapter.

Another way of assessing the validity of these estimation methods is to look at the ex post forecasting performance. In the next sub-section we will re-estimate the ECMs using the data from 1962 to 1990 and reserve the final four observations for forecasting comparison purposes. Since the USA ECMs are badly specified for all three methods, we will not carry out the forecasting exercise for the USA.

Forecasting performance of ECMs. The forecasts from 1991 to 1994 are generated from all three ECMs. The forecasting errors are calculated based on the differences between the actual and forecast values. In order to see which model performs best, two measures of forecasting performance are used: the mean absolute forecasting error (MAE) and the root mean square forecasting error (RMSE). The results are presented in Table 5.2:

The results in Table (5.2) show that the ADLM approach produces the smallest forecasting errors.

Notes

1. When quarterly or monthly data are used, the lag structure will be changed accordingly – four lags for quarterly data and 12 for monthly data.
2. The estimation of the cointegration and error correction models based on the ADLM approach is carried out using Microfit 4.0. Therefore, some diagnostic statistics reported may differ from those given by the Eviews program.
3. The heteroscedasticity test reported here is based on the regression of the squared residuals on the squared fitted values.

Appendix 5.1

Table A5.1 Worked examples: data used in estimation.

Year	UKTA	USTA	UKGDP	USGDP	UKTV	USTV	UKCPI	USCPI	KCPI	UKEXR[1]	KEXR[2]
1962	602	7328	48.89	42.83	28.623	841.67	10.7	23.1	3.8	2.80	130.00
1963	705	10178	51.02	44.47	24.727	1113.80	10.9	23.4	4.5	2.80	130.00
1964	737	11530	53.69	46.54	35.145	861.23	11.3	23.7	5.8	2.80	213.85
1965	828	14152	56.05	50.71	17.266	877.69	11.8	24.1	6.6	2.80	266.40
1966	1052	30226	57.11	53.76	25.436	1218.47	12.2	24.9	7.5	2.80	271.34
1967	1522	39274	58.42	55.16	45.993	1542.16	12.5	25.5	8.3	2.77	270.52
1968	1924	41823	60.80	57.43	76.610	2331.86	13.1	26.6	9.3	2.40	276.65
1969	2564	49606	62.05	58.99	139.70	2824.90	13.8	28.1	10.4	2.40	288.16
1970	2680	55352	63.47	58.96	144.54	3076.25	14.7	29.7	12.1	2.40	310.56
1971	3029	58003	64.73	60.79	214.36	3689.30	16.1	31.0	13.7	2.43	347.15
1972	3671	63578	67.00	63.70	299.91	4121.35	17.2	32.0	15.3	2.50	392.89
1973	4980	77537	71.93	67.01	371.83	5744.44	18.8	34.0	15.8	2.45	398.32
1974	5345	80621	70.71	66.59	428.48	6941.30	21.8	37.8	19.6	2.34	404.47
1975	6446	97422	70.20	66.05	566.20	6794.04	27.1	41.2	24.6	2.22	484.00
1976	8899	102199	72.15	69.31	808.37	8470.53	31.6	43.6	28.4	1.81	484.00
1977	9970	113710	73.85	72.44	810.83	9991.04	36.6	46.4	31.3	1.75	484.00
1978	12566	118039	76.40	75.93	1004.98	11827.24	39.6	49.9	35.8	1.92	484.00
1979	13395	127355	78.54	77.84	1537.67	13282.13	44.9	55.6	42.3	2.12	484.00
1980	12414	121404	76.84	77.43	1134.54	12307.89	53.0	63.1	54.5	2.33	607.43
1981	14874	130402	75.85	78.79	1308.42	13924.08	59.3	69.6	66.1	2.03	681.03
1982	16140	151249	77.16	77.09	1659.30	14236.05	64.4	73.9	70.8	1.75	731.08
1983	18598	176488	80.00	80.09	1688.86	16693.46	67.4	76.2	73.3	1.52	775.75
1984	19213	212986	81.86	85.05	1699.55	19512.33	70.7	79.5	75.0	1.34	805.98
1985	21414	239423	94.93	87.75	1604.28	19552.42	75.0	82.4	76.8	1.30	870.02
1986	23481	284571	88.57	90.30	1726.22	23744.78	77.6	83.9	78.9	1.47	881.45
1987	24606	326330	92.83	93.08	2541.86	30704.75	80.8	87.0	81.3	1.64	822.57
1988	33276	347281	97.48	96.74	3129.35	37156.52	84.7	90.5	87.1	1.78	731.47
1989	34423	317133	99.60	99.19	2774.33	36902.69	91.3	94.9	92.1	1.64	671.46
1990	36054	325388	100.00	100.00	2976.00	36392.00	100.0	100.0	100.0	1.78	707.76
1991	35848	315828	98.04	98.84	3320.36	37437.13	105.9	104.2	109.3	1.77	733.35
1992	36284	333805	97.52	102.09	3159.72	36088.30	109.8	107.4	116.1	1.77	780.65
1993	35923	325366	99.72	105.27	2993.16	35255.13	111.5	110.6	121.7	1.50	802.67
1994	40999	332428	103.56	109.57	3325.29	40667.95	114.3	113.4	129.3	1.53	803.45

Notes: [1] Exchange rate in terms of US dollar per pound sterling.
[2] Exchange rate in terms of Korean won per US dollar.

6

Vector autoregression (VAR) and cointegration

6.1 Introduction

Our main focus so far in this book has been on the single equation tourism demand model in which an endogenous tourism demand variable is related to a number of exogenous variables. The single equation approach depends heavily on the assumption that the explanatory variables are exogenous. If this assumption is invalid, a researcher would have to model the economic relationships using a system of (or simultaneous) equations method. The popularity of simultaneous equations approaches dates back to the 1950s and 1960s within the context of structural macroeconomic models which were used for policy simulation and forecasting. In estimating these structural models, restrictions were often imposed in order to obtain identified equations. Sims (1980) argued that many of the restrictions imposed on the parameters in the structural equations were 'incredible' relative to the data-generating process, and hence he suggested that it would be better to use models that do not depend on the imposition of incorrect prior information. Following this argument, Sims developed a vector autoregressive (VAR) model in which all the variables apart from the deterministic variables such as trend, intercept and dummy variables are modelled purely as dynamic processes, i.e., the VAR model treats all variables as endogenous.

More importantly, the VAR technique has been closely associated with some of the recent developments in multivariate cointegration analysis, such as the Johansen (1988) cointegration method. Although there has been increasing interest in using the VAR technique in macroeconomic modelling and forecasting, relatively little effort has been made to use this method for tourism demand analysis. Exceptions are Kulendran and King (1997) and Kulendran and Witt (2001) who model the demand for Australian inbound tourism and UK outbound tourism, respectively, using the cointegrated VAR system.

In this chapter we introduce the VAR modelling approach and its use in causality testing, impulse response analysis, forecasting and cointegration analysis. Although the aim of this book is to use the least complicated mathematical expressions possible, from this chapter onwards it is necessary to use matrices to cope with the complexity of the techniques involved.

6.2 Basics of VAR analysis

In order to introduce the general VAR specification, we start with two AR(1) processes:

$$y_{1t} = \pi_{11} y_{1t-1} + u_{1t}; \qquad u_{1t} \sim IID(0, \sigma_{11})$$
$$y_{2t} = \pi_{22} y_{2t-1} + u_{2t}; \qquad u_{2t} \sim IID(0, \sigma_{22})$$

The above equations may also be written in vector format:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \pi_{11} & 0 \\ 0 & \pi_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

or in matrix form:

$$Y_t = \Pi_1 Y_{t-1} + U_t; \qquad U_t \sim IID(0, \Sigma) \tag{6.1}$$

There are restrictions on the Π_1 matrix in Equation (6.1) that force it to be diagonal. If no restrictions were imposed on the matrix Π_1 we would obtain an unrestricted VAR model of order 1, or VAR(1), which represents the following equations:

$$y_{1t} = \pi_{11} y_{1t-1} + \pi_{12} y_{2t-1} + u_{1t}; \qquad u_{1t} \sim IID(0, \sigma_{11})$$
$$y_{2t} = \pi_{21} y_{1t-1} + \pi_{22} y_{2t-1} + u_{2t}; \qquad u_{2t} \sim IID(0, \sigma_{22})$$

Equation (6.1) can be extended to a general VAR(p) model:

$$Y_t = \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \ldots + \Pi_p Y_{t-p} + U_t; \qquad U_t \sim IID(0, \Sigma) \tag{6.2}$$

An intercept vector can also be added to Equation (6.2). The above equation is known as Sims' general (unrestricted) VAR model, in which the current values of all the variables are regressed against all the lagged values of the same set of variables in the system. U_t is a vector of regression errors that are assumed to be contemporaneously correlated but not autocorrelated. This indicates that each equation in the system can be individually estimated by OLS. Although the use of the seemingly unrelated regression estimator (SURE) is an alternative, it does not add much efficiency to the estimation procedure.

If a VAR model has m equations, there will be $m + pm^2$ coefficients that need to be estimated (including the constant term in each equation). This suggests that an unrestricted VAR model is likely to be over-parameterised. A practical problem with estimating a VAR(p) model is that we would like to include as much information as possible for the purposes of forecasting and policy analysis, but degrees of freedom will quickly run out as more variables are introduced. Therefore, the process of lag length selection is very important for the specification of a VAR model. If the lag length p is too small, the model cannot represent correctly the DGP; if, on the other hand, p is too large, lack of degrees of freedom can be a problem and OLS estimation is biased. To determine the lag length of the VAR model, we normally begin with the longest possible lag length permitted by the sample. For example, if we have about 150 observations on five quarterly series (i.e., there are five equations in the system), we may start with a lag length of 12 quarters based on the *a priori* assumption that a period of three years is sufficiently long to capture the dynamics of the system. There are 305 parameters that need to be estimated for this initial specification and this will probably lead to an over-parameterised VAR. We can then re-estimate the system with a reduced lag structure, say 11 lags, and see whether the shortening of the lag length significantly reduces the model's explanatory power.

A likelihood ratio statistic can be used to test whether the reduction in the model's explanatory power is significant:

$$LR = T(\ln|\Sigma_{11}| - \ln|\Sigma_{12}|) \tag{6.3}$$

where T is the sample size and $\ln|\Sigma_{12}|$ and $\ln|\Sigma_{11}|$ are the natural logarithms of the determinants of the variance/covariance matrices in the initial and reduced VARs, respectively. The LR statistic has an asymptotic χ^2 distribution with degrees of freedom equal to the number of restrictions in the system. In the example given above, we have five equations and the reduction of one lag for each equation means that we will lose 25 parameters, i.e., the number of restrictions for the statistic, RL, is 25. A small value for the calculated χ^2 (25) statistic compared with the critical value suggests that the null hypothesis should be accepted. If this is the case, we can then test whether ten lags are appropriate by calculating the LR statistic:

$$LR = T(\ln|\Sigma_{10}| - \ln|\Sigma_{11}|) \tag{6.4}$$

This process can be repeated until the smallest lag length, normally one lag, is reached.

The LR statistic represented by Equation (6.3) or Equation (6.4) is useful for large sample sizes. However, in small samples, the use of this statistic tends to over-reject the null hypothesis. The degrees-of-freedom-adjusted LR statistic for small samples is calculated as follows

$$LR^* = (T - q - 2 - mp)(\ln|\Sigma_r| - \ln|\Sigma_u|) \tag{6.5}$$

where LR^* is the adjusted LR statistic, q is the number of deterministic variables such as intercept, trend and dummies, p is the order of the unrestricted VAR model and m is the total number of equations in the VAR model. The statistic has a χ^2 distribution with appropriate degrees of freedom.

Alternative criteria that can be used in the determination of the lag length of the VAR model are AIC and SBC which are calculated from

$$AIC = -\frac{Tm}{2}(1 + \ln 2\pi) - \frac{T}{2}\ln|\Sigma| - (m^2 p + mq + 2m) \tag{6.6}$$

$$SBC = -\frac{Tm}{2}(1 + \ln 2\pi) - \frac{T}{2}\ln|\Sigma| - \frac{1}{2}(m^2 p + mq + 2m)\ln(T) \tag{6.7}$$

where $|\Sigma|$ is the determinant of the variance/covariance matrix of the estimated errors and $\pi = 3.1416$. Unlike the single equation case discussed in Chapter 3 where the lowest values of AIC and SBC are favourable, in VAR modelling the highest AIC and SBC are preferred. This occurs because the criteria underlying Equations (6.6) and (6.7) are constructed in a slightly different way.

6.3 Impulse response analysis

One of the advantages of VAR modelling is that the model can be used for policy simulation via impulse response analysis. In tourism demand analysis, we may want to know how 'shocks' to the price of tourism in the destination country and/or 'shocks' to the level of income in the tourism-generating country would affect the demand for tourism in the destination.

In order to investigate how impulse response analysis can be used for policy evaluation, let us consider a simple VAR(1) model:

$$Y_t = \Pi_1 Y_{t-1} + U_t \tag{6.8a}$$

or

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \tag{6.8b}$$

Lagging the above equation by one period and substituting for Y_{t-1} in Equation (6.8a) yields:

$$\begin{aligned} Y_t &= \Pi_1(\Pi_1 Y_{t-2} + U_{t-1}) + U_t \\ &= \Pi_1^2 Y_{t-2} + U_t + \Pi_1 U_{t-1} \end{aligned}$$

If this process is repeated n times we end up with:

$$Y_t = \sum_{i=0}^{n} \Pi_1^i U_{t-i} + \Pi_1^{n+1} Y_{t-n+1} \tag{6.9}$$

If the time series is stationary, that is $0 < |\Pi_1| < 1$, then $\lim_{n\to\infty} \Pi_1^n = 0$ (this is called the stability condition), and Equation (6.9) becomes

$$Y_t = \sum_{i=0}^{\infty} \Pi_1^i U_{t-i} \tag{6.10}$$

Equation (6.10) is called a *vector moving average* (VMA) process since the vector of the dependent variables is represented by an infinite sum of lagged random errors weighted by an exponentially diminishing coefficient.

In the case of a VAR(1) system with only two variables, Equation (6.10) becomes:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \sum_{i=0}^{\infty} \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix}^i \begin{bmatrix} u_{1t-i} \\ u_{2t-i} \end{bmatrix} \tag{6.11}$$

Equation (6.11) suggests that the variables y_{1t} and y_{2t} can be expressed by sequences of 'shocks' to the system. The formula indicates what would be the impacts of unitary changes in the error terms (shocks) on the dependent variables, which is what policy-makers are interested in. However, due to the fact that the error terms u_{1t} and u_{2t} are correlated, the impact of each shock cannot actually be separated from the two equations. To solve this problem, we need to transform Equation (6.11) into a form in which the error terms u_{1t} and u_{2t} are not correlated, and this process is known as orthogonalisation.

According to the assumptions made regarding the error terms in Equation (6.8b) we know that $E(u_{1t}) = 0$, $E(u_{2t}) = 0$, $Var(u_{1t}) = E(u_{1t}^2) = \sigma_{11}$, $Var(u_{2t}) = E(u_{2t}^2) = \sigma_{22}$, and $Cov(u_{1t}, u_{2t}) = \sigma_{12}$. One way of orthogonalising this VAR(1) model is to subtract the result of multiplying the first row of Equation (6.8b) by a constant term $\delta = \sigma_{12}/\sigma_{11}$ from the second row in Equation (6.8b). The resulting model is

$$\begin{bmatrix} y_{1t} \\ y_{2t} - \delta y_{1t} \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi'_{21} & \pi'_{22} \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u'_{2t} \end{bmatrix} \tag{6.12}$$

where $\pi'_{21} = (\pi_{21} - \delta\pi_{11})$, $\pi'_{22} = (\pi_{22} - \delta\pi_{12})$ and $u'_{2t} = (u_{2t} - \delta u_{1t})$. The error terms in Equation (6.12) are no longer correlated since:

$$\begin{aligned} Cov(u_{1t}, u'_{2t}) &= E(u_{1t}, u'_{2t}) = E[(u_{1t} * u_{2t}) - (\sigma_{12}/\sigma_{11}) E(u_{1t}^2)] \\ &= \sigma_{12} - \sigma_{12} = 0 \end{aligned}$$

Now substituting $u'_{2t-i} + \delta u_{1t-i}$ for u_{2t-i} in Equation (6.11), we get:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \sum_{i=0}^{\infty} \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix}^i \begin{bmatrix} 1 & 0 \\ \delta & 1 \end{bmatrix} \begin{bmatrix} u_{1t-i} \\ u'_{2t-i} \end{bmatrix} \tag{6.13}$$

If we denote $\phi_i = [\Pi_1^i] \begin{bmatrix} 1 & 0 \\ \delta & 1 \end{bmatrix}$, Equation (6.13) becomes:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} u_{1t-i} \\ u'_{2t-i} \end{bmatrix} \tag{6.14}$$

or, more compactly,

$$Y_t = \sum_{i=0}^{\infty} \phi_i U'_{t-i} \tag{6.15}$$

where $U'_t = \begin{bmatrix} u_{1t} \\ u'_{2t} \end{bmatrix}$ and the ϕ_i matrices are termed the *impulse response functions* since they represent the behaviour of the time series y_{1t} and y_{2t} in response to the shocks U'_t. When $i = 0$, the four elements in the ϕ_i matrix, $\phi_{11}(0)$, $\phi_{12}(0)$, $\phi_{21}(0)$ and $\phi_{22}(0)$, are called the impact multipliers since they represent the instantaneous impacts of unit changes in u_{1t} and u'_{2t} on the series y_{2t} and y_{1t}. For example, $\phi_{12}(0)$ is the instantaneous impact of a one-unit change in u'_{2t} on y_{1t}. Similarly, $\phi_{21}(1)$ and $\phi_{22}(1)$ represent the one-period responses of unit changes in u_{1t-1} and u'_{2t-1} on y_{2t}, respectively.

The impulse response functions can also be used to look at the accumulated effects of unit changes in u_{1t} and u'_{2t} on the future values y_{1t+n} and y_{2t+n}. For example, $\phi_{12}(n)$ is the effect of u'_{2t} on y_{1t+n}, and the cumulative sum of the effects of u'_{2t} after n periods on y_{1t} would be $\Sigma_{i=0}^{n}\phi_{12}(i)$. The impulse response functions can be plotted against i to give a visual indication of how the series y_{1t} and y_{2t} would respond to various shocks.

6.4 Forecasting with VAR and error variance decomposition

In the single equation approach, in order to forecast the dependent variable we need to forecast the exogenous variables first. This however can be very difficult indeed due to limitations of the data and the understanding of the data-generating processes for the exogenous variables. But there is no need to worry about the forecasts of the exogenous variables in the VAR framework since all explanatory variables are pre-determined. For example, if we want to make n-period-ahead forecasts for the variable Y_t given the estimated coefficient matrix Π_1 in Equation (6.8a) using the information up to period t, the forecasts and their errors can be updated in the following ways.

For one-period-ahead forecasts, since $Y_{t+1} = \Pi_1 Y_t + U_t$, the forecasts can be generated by taking the conditional expectation of Y_{t+1}, that is $\hat{Y}_{t+1} = E_t(Y_{t+1}) = \Pi_1 Y_t$. The forecasting error is $Y_{t+1} - \hat{Y}_{t+1} = U_{t+1}$. In the same way, two-periods-ahead forecasts can be generated from $\hat{Y}_{t+2} = \Pi_1^2 Y_t$, and the forecasting error now becomes $U_{t+2} + \Pi_1 U_{t+1}$. More generally, the n-period-ahead forecast is:

$$\hat{Y}_{t+n} = \Pi_1^n Y_t \tag{6.16}$$

and the corresponding forecasting error is:

$$U_{t+n} + \Pi_1 U_{t+n-1} + \Pi_1^2 U_{t+n-2} + \ldots + \Pi_1^{n-1} U_{t+1} \tag{6.17}$$

In fact, the forecasts and forecasting errors can also be generated from the VMA process of Equation (6.15). From Equation (6.15), we have:

$$Y_{t+n} = \sum_{i=0}^{\infty} \phi_i U'_{t+n-i} \tag{6.18}$$

Based on Equation (6.18) the one-period-ahead forecast $\hat{Y}_{t+1} = E(\Sigma_{i=0}^{\infty}\phi_i U'_{t+1-i})$ and the corresponding one-period-ahead forecast error is: $Y_{t+1} - \hat{Y}_{t+1} = \phi_0 U'_{t+1}$. Following the same procedure, we could obtain the n-period-ahead forecast error:

$$Y_{t+n} - \hat{Y}_{t+n} = \sum_{i=0}^{n-1} \phi_i U'_{t+n-i} \tag{6.19}$$

As far as the variable y_{1t} is concerned, the n-period-ahead forecasting error is measured by

$$\begin{aligned} y_{1t+n} - \hat{y}_{1t+n} = {} & \phi_{11}(0)u_{1t+n} + \phi_{11}(1)u_{1t+n-1} + \ldots + \phi_{11}(n-1)u_{1t+1} \\ & + \phi_{12}(0)u'_{2t+n} + \phi_{12}(1)u'_{2t+n-1} + \ldots + \phi_{12}(n-1)u'_{2t+1} \end{aligned}$$

and the error variance of the n-period-ahead forecast is therefore

$$\begin{aligned} \mathrm{var}(y_{1t+n} - \hat{y}_{1t+n}) = {} & \sigma^2_{y_{1t}}[\phi^2_{11}(0) + \phi^2_{11}(1) + \ldots + \phi^2_{11}(n-1)] \\ & + \sigma^2_{y_{2t}}[\phi^2_{12}(0) + \phi^2_{12}(1) + \ldots + \phi^2_{12}(n-1)] \end{aligned} \tag{6.20}$$

Equation (6.20) shows that the forecasting variance increases as the forecasting horizon lengthens.

Equation (6.20) can also be decomposed to show the shares of the n-period-ahead forecasting error variance resulting from separate shocks in u_{1t} and u'_{2t}. For example, the shares of the n-period-ahead forecasting error variance in the series y_{1t} caused by u_{1t} and u'_{2t} are

$$\sigma^2_{y_{1t}}[\phi^2_{11}(0) + \phi^2_{11}(1) + \ldots + \phi^2_{11}(n-1)]/\mathrm{Var}(y_{1t+n} - \hat{y}_{1t+n})$$

and

$$\sigma^2_{y_{2t}}[\phi^2_{12}(0) + \phi^2_{12}(1) + \ldots + \phi^2_{12}(n-1)]/\mathrm{Var}(y_{1t+n} - \hat{y}_{1t+n})$$

respectively. The forecasting error variance decomposition shows the share of the variation in a time series that is due to it own ‘shocks’ and also the shocks from the other variables. In the two-variable case above, if u'_{2t} cannot explain the forecasting error variance of y_{1t} at all forecasting horizons, then y_{1t} is said to be an exogenous series. On the other hand if u'_{2t} explains all the forecasting error variance of y_{1t}, we say that y_{1t} is completely endogenous. In this respect, the forecasting error variance decomposition can be used as a tool to test for exogeneity of a time series within a VAR framework.

6.5 Testing for Granger causality and exogeneity

6.5.1 The concept of Granger causality

In practice, a researcher may wish to know whether the past values of a time series y_{2t} would help to forecast the current and future values of another series y_{1t}. This problem is known as Granger causality, and is defined as:

y_{2t} fails to Granger cause y_{1t} if for all $n>0$ the mean squared error (MSE) of a forecast of y_{1t+n} based on $(y_{1t}, y_{1t-1}, y_{1t-2}, \ldots)$ is the same as the MSE of a forecast of y_{1t+n} that uses both $(y_{1t}, y_{1t-1}, y_{1t-2}, \ldots)$ and $(y_{2t}, y_{2t-1}, y_{2t-2}, \ldots)$.

Or, more formally, y_{2t} *fails to Granger cause* y_{1t} *if*

$$\begin{aligned} &MSE[E(\hat{y}_{1t}|y_{1t}, y_{1t-1}, y_{1t-2}, \ldots)] \\ &\quad = MSE[E(\hat{y}_{1t}|y_{1t}, y_{1t-1}, y_{1t-2}, \ldots, y_{2t}, y_{2t-1}, y_{2t-2}, \ldots) \end{aligned} \tag{6.21}$$

where $\hat{y}_{1t+n}$ is the n-period-ahead forecast for y_{1t} and '|' stands for 'conditional on'.

We can also say that y_{1t} is *exogenous* with respect to y_{2t} if Equation (6.21) holds. As can be seen from the definition, the concept of Granger causality is not the same as that used in everyday life. The causality here refers only to the predictability of the time series. The rationale behind this concept is given by Granger (1969). He believes that if an event y_2 is the cause of another event y_1, then the former should precede the latter. Although this argument has its philosophic base, quite often we may find this concept difficult to implement with regard to aggregate economic time series, since many of these time series are compiled in a discontinuous way in which the detailed dynamics are 'averaged' out. For example, in tourism demand analysis many researchers use annual data, but the response of tourists' choices of holiday destinations to changes in, say, exchange rates may be more frequent than an annual span.

6.5.2 Tests for Granger causality

In the bivariate VAR model case, testing for Granger causality is straightforward. The normal tests for restrictions discussed in Chapter 3, such as the F (or LR) tests, can be used. The tests can be based on the auxiliary equation proposed by either Granger (1969) or Sims (1972).

The Granger test is based on

$$y_{1t} = \alpha + \sum_{i=1}^{n} \beta_i y_{1t-i} + \sum_{i=1}^{n} \gamma_i y_{2t-i} + u_t \tag{6.22}$$

The two variables in Equation (6.22) are normally assumed to be stationary. The restrictions imposed are $\gamma_1 = \gamma_2 = \ldots = \gamma_n = 0$. If these restrictions are accepted according to the F or LR statistic, we conclude that y_{2t} does not

Granger cause y_{1t}. In estimating Equation (6.22), deterministic variables such as dummies and trend can also be included if necessary.

The Sims (1972) test is based on

$$y_{1t} = \alpha + \sum_{i=1}^{n} \beta_i y_{1t-i} + \sum_{i=-k}^{m} \gamma_i y_{2t-i} + u_t \qquad i \neq 0 \tag{6.23}$$

In Equation (6.23) both the lagged and lead values of y_{2t} are included, but not the current value of y_{2t}. The restrictions now imposed are $\gamma_{-m} = \gamma_{-m+1} = \ldots = \gamma_{-1} = 0$. If these restrictions are accepted, this suggests that future values of y_2 cannot be used to forecast the current value of y_1, and therefore y_1 does not Granger cause y_2. However, if the restrictions are rejected the conclusion would be that y_1 does Granger cause y_2.

Since either Equation (6.22) or Equation (6.23) is only one equation within a bivariate VAR system, each of the tests has to be carried out twice with the dependent variable changed. The final conclusion has to be drawn on the basis of all of the test statistics.

The above tests are related to a two-variable VAR model. If there are more than two variables, the *block Granger causality* test (sometimes also referred as the *block exogeneity* test) should be used. The block Granger causality test examines whether a lagged variable would Granger cause other variables in the system. For example, suppose we have three variables, y_1, y_2 and y_3, and we want to look at whether y_3 Granger causes y_1 and/or y_2. What the test does is to impose the restrictions that all of the coefficients of the lagged y_3 variables in the system are zero. These restrictions can be tested using the *LR* statistic of Equation (6.5). The test is carried out as follows: first, estimate the y_{1t} and y_{2t} equations with n lagged values of y_{1t}, y_{2t} and y_{3t} as regressors and calculate $|\Sigma_u|$; then re-estimate the two equations excluding the lagged y_{3t} values and calculate $|\Sigma_r|$; and finally calculate the *LR* statistic using:

$$LR = (T - c)(\ln|\Sigma_r| - \ln|\Sigma_u|)$$

We already know that this statistic is distributed as χ^2 with degrees of freedom equal to the total number of restrictions. T is the total number of observations for each variable, and c is the total number of parameters in the unrestricted equations. In this case, the degrees of freedom are equal to $2n$ since n lagged y_{3t}s are excluded from two equations, and c equals $2(3n + 1)$ since there are two unrestricted equations and each equation contains a constant term. If the calculated *LR* is greater than the corresponding critical value $\chi^2(2n)$, the restrictions imposed should be rejected and this suggests that y_{3t} does Granger cause y_{1t} and y_{2t}.

The Granger block causality test can also be used to test whether several variables jointly Granger cause a single variable (see Section 6.6.2).

6.6 An example of VAR modelling

The data set on inbound tourism demand to South Korea used in Chapter 5 is used again to illustrate the procedure of VAR modelling. Two VAR models are created based on the demand for tourism by UK and USA residents, respectively. The variables included in each VAR model are tourist arrivals in Korea from the origin country, $\ln TA_{it}$, the GDP of the origin country, $\ln GDP_{it}$, the trade volume between Korea and the origin country, $\ln TV_{it}$, and the relative price variable, $\ln RCPI_{it}$.

In terms of the integration order of the variables, normally the variables included in the model should be $I(0)$. However, Sims (1980) and Doan *et al.* (1984) suggest that the variables in the VAR system should not be differenced since the aim of the VAR model is to determine the inter-relationships between economic variables, and not just the estimates of the parameters. The use of differenced variables in a VAR model results in the loss of useful information related to the co-movement of the time series. Moreover, according to Geweke (1984, pp. 1139–40), various hypothesis tests, such as the Granger causality tests, are still valid if the data are transformed into logarithms, as is the case with our data. Therefore, in this section we will use the levels variables (in logarithmic form) in the construction and analysis of the VAR model.

6.6.1 Order selection of VAR model

Since annual data are used, we initially specify a VAR(4) model for each of the two tourism-generating countries as:

$$Y_t = \Pi_0 + \sum_{i=1}^{4} \Pi_i Y_{t-i} \tag{6.24}$$

where $Y_t = (\ln TA_{it}, \ln GDP_{it}, \ln TV_{it}, \ln RCPI_{it})'$, Π_0 is a 1×4 vector and the Π_is are 4×4 matrices. The *LR* statistic of Equation (6.5) is calculated in a pairwise fashion, i.e., order four against order three, order three against order two, and so on. The *AIC* and *SBC* are also calculated using Equations (6.6) and (6.7) for each of the orders from one to four. Tables 6.1 and 6.2

Table 6.1. Choice criteria for selecting the order of the VAR model (UK).

Based on 29 observations from 1966 to 1994. Order of VAR = 4
List of variables included in the unrestricted VAR:
LTAUK LGDPUK LTVUK LRCPIUK
List of deterministic and/or exogenous variables:
INPT

Order	LL	AIC	SBC	LR test	Adjusted LR test
4	176.2106	108.2106	61.7226	—	—
3	162.9013	110.9013	75.3516	CHSQ(16) = 26.6186[0.046]	11.0146[0.809]
2	149.4026	113.4026	88.7913	CHSQ(32) = 53.6161[0.010]	22.1860[0.902]
1	137.8001	117.8001	104.1271	CHSQ(48) = 76.8211[0.005]	31.7881[0.966]
0	20.1782	16.1782	13.4436	CHSQ(64) = 312.0648[0.000]	129.1303[0.000]

LTAUK = $\ln TA_{1t}$, LGDPUK = $\ln GDP_{1t}$, LTVUK = $\ln TV_{1t}$ and LRCPIUK = $\ln RCPI_{1t}$.

Table 6.2 Choice criteria for selecting the order of the VAR model (USA).

Based on 29 observations from 1966 to 1994. Order of VAR = 4
List of variables included in the unrestricted VAR:
LTAUS LGDPUS LTVUS LRCPIUS
List of deterministic and/or exogenous variables:
INPT

Order	LL	AIC	SBC	LR test	Adjusted LR test
4	256.4243	188.4243	141.9362	—	—
3	237.3447	185.3447	149.7950	CHSQ(16) = 38.1590[0.001]	15.7899[0.468]
2	321.9148	185.9148	161.3035	CHSQ(32) = 69.0189[0.000]	28.5595[0.641]
1	205.2028	185.2028	171.5299	CHSQ(48) = 102.4429[0.000]	42.3902[0.701]
0	78.1155	74.1155	71.3809	CHSQ(64) = 356.6175[0.000]	147.5658[0.000]

LTAUS = $\ln TA_{2t}$, LGDPUS = $\ln GDP_{2t}$, LTVUS = $\ln TV_{2t}$ and LRCPIUS = $\ln RCPI_{2t}$.

are the computer printouts (Microfit 4.0) of the test results for the VAR order selection.

The results in Tables 6.1 and 6.2 show that, as expected, the values of log-likelihood increase with increases in the lag length. In the case of the UK VAR model, the *AIC* and *SBC* indicate order one and this is confirmed by the *LR** (adjusted LR) statistic. For the USA VAR model, the *SBC* and *LR** suggest order 1 whilst the *AIC* points to order 4. However, based on all three criteria we prefer a VAR(1) specification for both the UK and USA models.

Table 6.3. LR test of block Granger non-causality in the VAR model (UK).

Based on 29 observations from 1966 to 1994. Order of VAR = 4
List of variables included in the unrestricted VAR:
LTAUK LGDPUK LTVUK LRCPIUK
List of deterministic and/or exogenous variables:
INPT
Maximised value of log-likelihood = 176.2106

List of variable(s) assumed to be 'non-causal' under the null hypothesis:
LGDPUK LTVUK LRCPIUK
Maximised value of log-likelihood = 161.2347

LR test of block non-causality, CHSQ(12) = 29.9519[0.003]

The above statistic is for testing the null hypothesis that the coefficients of the lagged values of:
LGDPUK LTVUK LRCPIUK
in the block of equations explaining the variable(s):
LTAUK
are zero. The maximum order of the lag(s) is 4.

6.6.2 Block Granger causality test

Since our prime interest is in whether $\ln GDP_{it}$, $\ln TV_{it}$ and $\ln RCPI_{it}$ cause $\ln TA_{it}$, the block Granger causality test is conducted to see whether the lagged values of $\ln GDP_{it}$, $\ln TV_{it}$ and $\ln RCPI_{it}$ help to predict the tourism demand variable $\ln TA_{it}$. Tables 6.3 and 6.4 present the computer printout results of this test for the two tourism-generating countries, UK and USA. The null hypothesis is that $\ln GDP_{it}$, $\ln TV_{it}$ and $\ln RCPI_{it}$ do not jointly Granger cause $\ln TA_{it}$.

The first section of the tables presents the variables included in the unrestricted VAR(4) model. The maximum log-likelihood value of the estimated VAR(4) model is also reported here. The second section of the tables gives a list of the variables that we have assumed to be useful in predicting $\ln TA_{it}$. The maximised value of the log-likelihood of the estimated VAR(4) model without these variables is also given. The next section of the tables presents the LR statistic of the block Granger causality test with some explanation about the statistic. The LR statistics in both Tables 6.3 and 6.4 indicate that the null hypothesis of non-Granger block causality is rejected, which means that the variables $\ln GDP_{it}$, $\ln TV_{it}$ and $\ln RCPI_{it}$ do jointly Granger cause $\ln TA_{it}$, as expected. The reader may also wish to conduct the Granger causality test to see whether there are

Table 6.4. LR test of block Granger non-causality in the VAR model (USA).

Based on 29 observations from 1966 to 1994. Order of VAR = 4			
List of variables included in the unrestricted VAR:			
LTAUS	LGDPUS	LTVUS	LRCPIUS
List of deterministic and/or exogenous variables:			
INPT			
Maximised value of log-likelihood = 256.4243			
List of variable(s) assumed to be 'non-causal' under the null hypothesis:			
LGDPUS	LTVUS	LRCPIUS	
Maximised value of log-likelihood = 220.2853			
LR test of block non-causality, CHSQ(12) = 72.2778[0.000]			
The above statistic is for testing the null hypothesis that the coefficients of the lagged values of:			
LGDPUS	LTVUS	LRCPIUS	
in the block of equations explaining the variable(s):			
LTAUS			
are zero. The maximum order of the lag(s) is 4.			

bi-directional causal relationships among the four variables using the same procedure.

6.6.3 Estimating the VAR models

Since the order selection criterion suggests that a VAR(1) model is appropriate for both the UK and USA, the two VAR(1) models are estimated. As discussed in Section 6.2, equations in the VAR system can be estimated individually by OLS. Tables 6.5–6.8 report the estimates of the VAR(1) model for UK tourism demand to South Korea. The results for the USA VAR(1) model are omitted to save space.

The estimated results of the VAR(1) model for the UK show that all the equations fit the data well as judged by the high R-squared values, although each equation fails at least one of the diagnostic statistics. Since the VAR model only relates the dependent variables to the lagged explanatory variables, the contemporaneous effects are not modelled. This may explain why in the first equation the only significant variable is lagged $\ln TA_{1t}$. However, this does not necessarily mean that the variables $\ln GDP_{it}$, $\ln TV_{it}$ and $\ln RCPI_{it}$ have no role to play in the determination of UK tourism demand to South Korea. In later sections, we will see that these variables are in fact cointegrated, and the changes in tourism demand can be

Table 6.5. OLS estimation of tourist arrivals equation in the unrestricted VAR model.

Dependent variable is LTAUK
32 observations used for estimation from 1963 to 1994

Regressor	Coefficient	Standard error	T ratio[prob]
LTAUK(−1)	0.919070	0.17987	5.10970[0.000]
LGDPUK(−1)	0.072108	0.48931	0.14736[0.884]
LTVUK(−1)	0.003055	0.09219	0.33140[0.974]
LRCPIUK(−1)	−0.084171	0.93956	−0.89585[0.378]
INPT	0.051942	1.27130	0.04086[0.968]

R-squared	0.9945	R-bar-squared	0.9937
S.E. of regression	0.1036	F-stat. F(4, 27)	1214.0
Mean of dependent variable	9.0256	S.D. of dependent variable	1.3004
Residual sum of squares	0.2898	Equation log-likelihood	29.8603
Akaike info. criterion	24.8603	Schwarz Bayesian criterion	21.1960
DW statistic	1.8893	System log-likelihood	134.1488

Diagnostic tests

Test statistics	LM version	F version
A: Serial correlation:	CHSQ(1) = 0.0596[0.807]	F(1, 26) = 0.0485[0.827]
B: Functional form:	CHSQ(1) = 6.7761[0.009]	F(1, 26) = 6.9846[0.014]
C: Normality:	CHSQ(2) = 1.3817[0.501]	Not applicable
D: Heteroscedasticity:	CHSQ(1) = 1.7315[0.188]	F(1, 30) = 1.7162[0.200]

A: Lagrange multiplier test of residual serial correlation.
B: Ramsey's RESET test using the square of the fitted values.
C: Based on a test of skewness and kurtosis of residuals.
D: Based on the regression of squared residuals on squared fitted values.

appropriately modelled by both the current (in differenced form) and lagged $\ln GDP_{it}$, $\ln TV_{it}$ and $\ln RCPI_{it}$ (as the error correction term).

6.6.4 Impulse response analysis and forecasting variance decomposition

The VAR(1) models are estimated over the period 1962 to 1994. In order to look at the dynamic effects of a unitary shock measured by one standard error to a particular regression equation in the VAR(1) model on the tourism demand variables $\ln TA_{it}$, $\ln GDP_{it}$, $\ln TV_{it}$ and $\ln RCPI_{it}$, the corresponding orthogonalised impulse response functions are plotted over 20 periods. To save space we plot only the orthogonalised impulse response functions of the two equations with $\ln TA_{it}$ and $\ln GDP_{it}$ as dependent variables for each country. The same method can also be used to examine

Table 6.6. OLS estimation of GDP equation in the unrestricted VAR model.

Dependent variable is LGDPUK
32 observations used for estimation from 1963 to 1994

Regressor	Coefficient	Standard error	T ratio[prob]
LTAUK(−1)	0.02269	0.05766	0.3936[0.697]
LGDPUK(−1)	0.70836	0.15685	4.5162[0.000]
LTVUK(−1)	−0.6675E − 3	0.02955	−0.0226[0.982]
LRCPIUK(−1)	−0.05427	0.03012	−1.8019[0.083]
INPT	0.77736	0.40752	1.9076[0.067]

R-squared	0.9777	R-bar-squared	0.9744
S.E. of regression	0.0332	F-stat. F(4, 27)	295.8493
Mean of dependent variable	4.3264	S.D. of dependent variable	0.2075
Residual sum of squares	0.0298	Equation log-likelihood	66.2674
Akaike info. criterion	61.2674	Schwarz Bayesian criterion	57.6031
DW statistic	2.4575	System log-likelihood	134.1488

Diagnostic tests

Test statistics	LM version	F version
A: Serial correlation:	CHSQ(1) = 3.6654[0.056]	F(1, 26) = 3.3634[0.078]
B: Functional form:	CHSQ(1) = 0.7931[0.373]	F(1, 26) = 0.6607[0.424]
C: Normality:	CHSQ(2) = 11.0401[0.004]	Not applicable
D: Heteroscedasticity:	CHSQ(1) = 2.4237[0.120]	F(1, 30) = 2.4584[0.127]

the impulse response functions in the third equation in which relative tourism prices are the dependent variable.

Figures 6.1 and 6.2 show the time path or impulse responses of ln TA_{it} to unitary 'shocks' in its own and other series for the UK and USA, respectively. Figure 6.1 suggests that the impact of a unit shock to the UK tourism demand equation tends to have a relatively large impact on trade volume and tourism demand itself for the first few periods, although the impact of the shock stabilises as the forecasting horizon increases. This makes sense as tourist arrivals in South Korea consist of a relatively large proportion of business travellers, and therefore there is likely to be a bi-directional causal relationship between tourism and trade – more tourists from the UK visiting Korea means more business and vice versa. This may also be observed by looking at the impulse response functions in the GDP equation (Figure 6.3). However, as expected, shocks to the demand for Korean tourism do not have much effect on the UK GDP level. This is logical because the income level in the UK (GDP) is likely to be exogenous

Table 6.7. OLS estimation of trade volume equation in the unrestricted VAR model.

Dependent variable is LTVUK
32 observations used for estimation from 1963 to 1994

Regressor	Coefficient	Standard error	T ratio[prob]
LTAUK(−1)	1.2875	0.4058	3.1724[0.004]
LGDPUK(−1)	−1.8913	1.1041	−1.7130[0.098]
LTVUK(−1)	0.2705	0.2080	1.3002[0.205]
LRCPIUK(−1)	0.3346	0.2120	1.5782[0.126]
INPT	3.2346	2.8685	1.1276[0.269]

R-squared	0.9829	R-bar-squared	0.9803
S.E. of regression	0.2338	F-stat. F(4, 27)	387.3689
Mean of dependent variable	6.3559	S.D. of dependent variable	1.6671
Residual sum of squares	1.4756	Equation log-likelihood	3.8205
Akaike info. criterion	−1.1795	Schwarz Bayesian criterion	−4.8438
DW statistic	1.5744	System log-likelihood	134.1488

Diagnostic tests

Test statistics	LM version	F version
A: Serial correlation:	CHSQ(1) = 1.9215[0.166]	F(1, 26) = 1.6609[0.209]
B: Functional form:	CHSQ(1) = 14.8842[0.000]	F(1, 26) = 22.6101[0.000]
C: Normality:	CHSQ(2) = 2.1527[0.341]	Not applicable
D: Heteroscedasticity:	CHSQ(1) = 1.7685[0.184]	F(1, 30) = 1.7550[0.195]

to tourism demand to Korea. The impulse response function in the USA model tells the same story. The policy implication from these impulse responses is that promoting tourism to South Korea will increase business opportunities and hence increase trade between Korea and the tourism-generating countries. Further examination of the income (GDP) equations for the UK and USA (Figures 6.3 and 6.4, respectively) shows that shocks to the income equations in the two tourism-generating countries tend to have large impacts on tourism demand ($\ln TA_{it}$), this being especially true for the USA.

We can also examine the forecasting error variance decomposition to determine the proportion of the movements in the time series which is due to shocks to its own series as opposed to shocks to other variables. If shocks to one series do not explain the forecasting error variance of the other, this suggests that the former series is exogenous to the latter. Figures 6.5 and 6.6 are the plots of the forecasting error variance decomposition for the two tourist arrivals equations.

Table 6.8. OLS estimation of tourist prices equation in the unrestricted VAR model.

Dependent variable is LRCPIUK
32 observations used for estimation from 1963 to 1994

Regressor	Coefficient	Standard error	T ratio[prob]
LTAUK(−1)	−0.2068	0.2160	−0.9574[0.347]
LGDPUK(−1)	−1.0610	0.5876	1.8057[0.082]
LTVUK(−1)	−0.0098	0.1107	−0.0883[0.930]
LRCPIUK(−1)	0.8223	0.1128	7.2885[0.000]
INPT	−3.6986	1.5265	−2.4228[0.022]

R-squared	0.9458	R-bar-squared	0.9378
S.E. of regression	0.1244	F-stat. F(4, 27)	117.8876
Mean of dependent variable	−5.6288	S.D. of dependent variable	0.4989
Residual sum of squares	0.4179	Equation log-likelihood	24.0058
Akaike info. criterion	19.0058	Schwarz Bayesian criterion	15.3414
DW statistic	1.2246	System log-likelihood	134.1488

Diagnostic tests

Test statistics	LM version	F version
A: Serial correlation:	CHSQ(1) = 5.3267[0.021]	F(1, 26) = 5.1923[0.031]
B: Functional form:	CHSQ(1) = 2.7912[0.095]	F(1, 26) = 2.4845[0.127]
C: Normality:	CHSQ(2) = 1.2297[0.541]	Not applicable
D: Heteroscedasticity:	CHSQ(1) = 0.0523[0.819]	F(1, 30) = 0.0491[0.826]

Figure 6.5 suggests that the largest proportion of the forecasting error variance in the UK tourist arrivals equation (more than 80%) is due to the shocks in the tourism arrivals series itself. The implication of this is that in order to produce better forecasts for the UK tourist arrivals series, it is crucially important to utilise correctly the information in the series itself. For the USA tourist arrivals equation, about 50% of the forecasting errors is due to the shocks to the relative price variable. Therefore, to improve the forecasting performance of the USA tourist arrivals equation, correctly forecasting the relative price variable is very important.

6.6.5 Ex post *forecasting performance*

In Chapter 5, we generated *ex post* forecasts over the period 1990–4 on the log changes of UK and USA tourist arrivals in South Korea using ECMs. To compare the forecasting performance of the VAR(1) models with that of the ECMs specified in Chapter 5, we also re-estimate our VAR(1) models

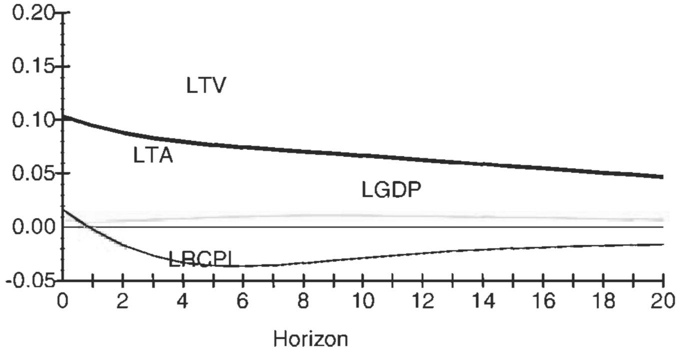

Figure 6.1 Impulse responses to one SE shock in the UK LTA equation.

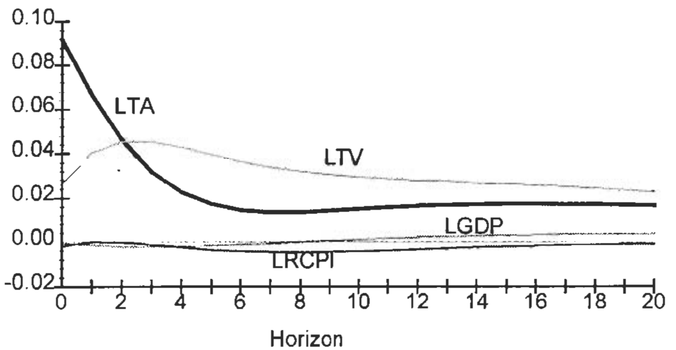

Figure 6.2 Impulse responses to one SE shock in the US LTA equation.

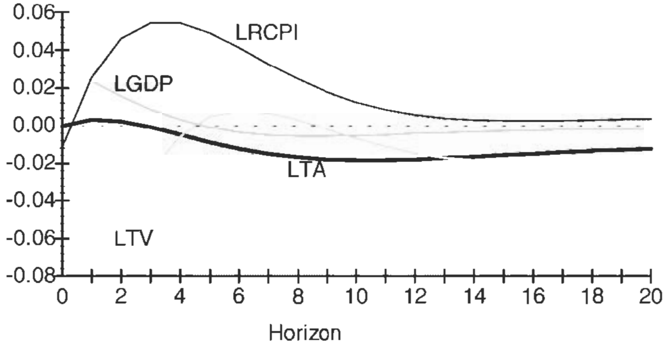

Figure 6.3 Impulse responses to one SE shock in the UK LGDP equation.

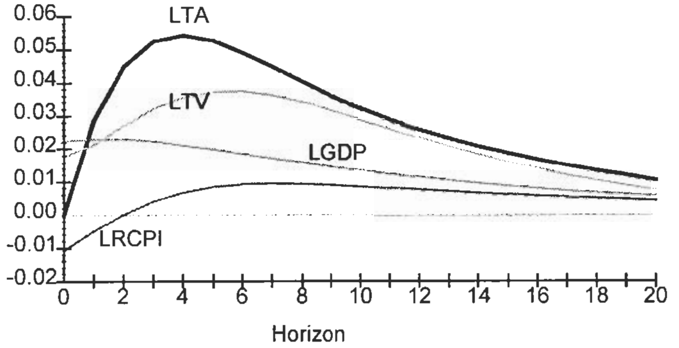

Figure 6.4 Impulse responses to one SE shock in the US LGDP equation.

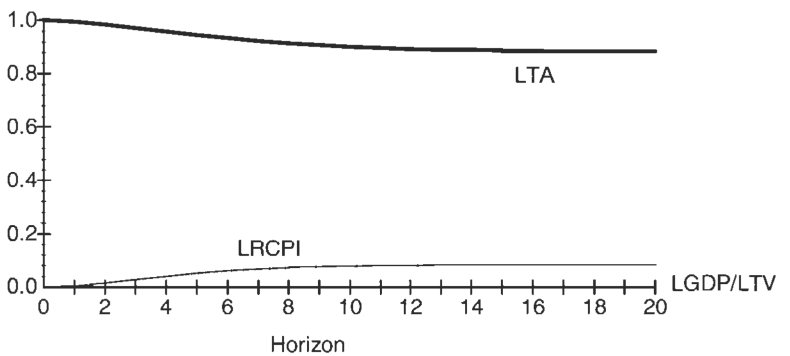

Figure 6.5 Forecast error variance decomposition for variable LTA (UK).

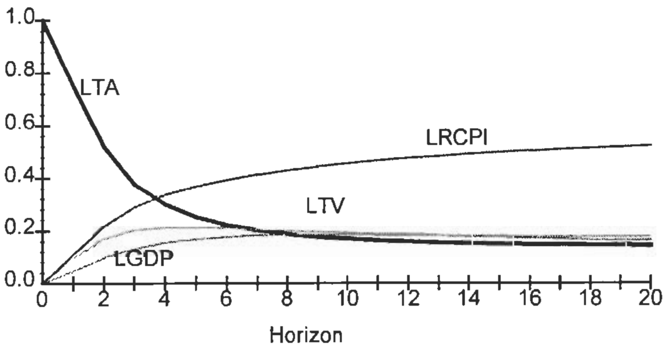

Figure 6.6 Forecast error variance decomposition for variable LTA (US).

Table 6.9. ***Ex post*** **forecast errors of selected models.**

Models	MAE	RMSE
UK:		
VAR(1)	0.0835	0.0853
Engle–Granger	0.0438	0.0469
Wickens–Breusch	0.0593	0.0617
ADLM	0.0322	0.0363
USA:		
VAR(1)	0.0934	0.0963
Engle–Granger	0.1041	0.1176
Wickens–Breusch	0.1002	0.1170
ADLM	0.0991	0.1020

over the period 1962–90 and reserve four data points in order to generate the forecasts. The MAE and RMSE are then calculated from the forecasting errors and these are presented together with the MAE and RMSE for the ECMs in Table 6.9. Since we are interested in forecasts of tourist arrivals, only the forecasting results from the tourist arrivals equations are presented.

The results show that in the case of USA tourism demand forecasting, the VAR(1) model outperforms the ECMs. By contrast, the UK VAR(1) model is the poorest among the competing specifications. These results are not surprising since empirical evidence shows that ECMs tend to outperform VAR models if the variables are cointegrated (Kim and Song, 1998), as is the case with the UK series. On the other hand, if the variables are not cointegrated according to the Engle and Granger two-stage approach (as is the case with the USA series) the single ECM cannot be used for forecasting.

6.7 VAR modelling and cointegration

6.7.1 Theoretical explanation

In Chapter 5 we introduced the Engle–Granger two-stage approach to cointegration analysis. The underlying assumption of this method is that there is only one cointegration relationship among a set of economic variables. However, this assumption is too restrictive and sometimes is unrealistic. In reality there may be more than one cointegration relationship if the long-run model involves more than two variables. What the Engle–Granger two-stage method detects is only an 'average' cointegrating vector over a number of cointegration vectors. Another serious problem with this approach is that it is based on a two-stage procedure, and any error introduced in the first stage will be carried over to the second stage. Fortunately, these problems can be overcome by the Johansen (1988) and Stock and Watson (1988) maximum likelihood estimators which have been developed within the VAR modelling framework. Not only are these estimators less prone to estimation bias, but also they can be used to test for and estimate multiple cointegration relationships. In this section we only introduce the Johansen procedure at a basic level, but for those readers who are interested in a more advanced exposition of the Johansen procedure and/or a discussion of the Stock and Watson procedure, as well as related topics on multivariate cointegration analysis, Hamilton (1994, pp. 630–53) is a good reference.

The Johansen cointegration procedure is an extension of the univariate Dickey–Fuller test to a multivariate VAR framework. In order to understand the Johansen cointegration technique, we start by extending the univariate first-order autoregressive process $y_t = \beta y_{t-1} + u_t$ to an m-variable VAR(1) process:

$$Y_t = BY_{t-1} + U_t \tag{6.25}$$

so that

$$\begin{aligned} \Delta Y_t &= (\mathrm{B} - \mathrm{I})\, Y_{t-1} + U_t \\ &= \Phi Y_{t-1} + U_t \end{aligned} \tag{6.26}$$

where Y_t is an $(m \times 1)$ vector of variables, U_t is an $(m \times 1)$ vector of errors, I is an $(m \times m)$ identity matrix and Φ is an $(m \times m)$ matrix of parameters.

We shall see later that the rank of Φ equals the number of cointegrating vectors. Therefore, the cointegration test within a multivariate VAR framework is to find out the rank of the matrix Φ. In order for the test statistics to be valid, the usual assumptions regarding the residual errors apply, that is U_t has to be normally distributed with zero mean and constant variance Σ. However, if the variables in Y_t are $I(1)$ processes, the error terms in U_t in Equation (6.26) are highly likely to be autocorrelated due to a common trend present in the data, and therefore any statistics calculated from the estimated Equation (6.26) will not be standard normal. To overcome this problem, as with the Dickey–Fuller test, Equations (6.25) and (6.26) can be generalised to allow for higher-order autoregressive processes:

$$Y_t = \mathrm{B}_1 Y_{t-1} + \mathrm{B}_2 Y_{t-2} + \ldots + \mathrm{B}_p Y_{t-p} + U_t \tag{6.27}$$

where Y_t is an $(m \times 1)$ vector of m potential endogenous variables and each of the B_i is an $(m \times m)$ matrix of parameters. By allowing a higher order of autoregression the residual term U_t is more likely to be well behaved.

As with the transformation carried out in deriving the augmented Dickey–Fuller test, Equation (6.27) can also be further transformed into a more appropriate form for statistical testing. First, subtract Y_{t-1} from both sides of (6.27) to yield:

$$\Delta Y_t = (\mathrm{B}_1 - I)\, Y_{t-1} + \mathrm{B}_2 Y_{t-2} + \mathrm{B}_3 Y_{t-3} + \ldots + \mathrm{B}_p Y_{t-p} + U_t$$

Then add and subtract $(\mathrm{B}_1 - I)\, Y_{t-2}$ on the right-hand side of the above equation to obtain:

$$\Delta Y_t = (\mathrm{B}_1 - I)\, \Delta Y_{t-1} + (\mathrm{B}_1 + \mathrm{B}_2 - I)\, Y_{t-2} + \mathrm{B}_3 Y_{t-3} + \ldots + \mathrm{B}_p Y_{t-p} + U_t$$

Next add and subtract $(\mathrm{B}_1 + \mathrm{B}_2 - I)\, Y_{t-3}$ to arrive at

$$\begin{aligned}\Delta Y_t = {} & (\mathrm{B}_1 - I)\, \Delta Y_{t-1} + (\mathrm{B}_1 + \mathrm{B}_2 - I)\, \Delta Y_{t-2} + (\mathrm{B}_1 + \mathrm{B}_2 + \mathrm{B}_3 - I)\, Y_{t-3} \\ & + \ldots + \mathrm{B}_p Y_{t-p} + U_t\end{aligned}$$

A similar process is repeated until eventually Equation (6.28) is obtained:

$$\Delta Y_t = \sum_{i=1}^{p-1} \Phi_i \Delta Y_{t-i} + \Phi Y_{t-p} + U_t \tag{6.28}$$

where $\Phi_i = -(I - \mathrm{B}_1 - \mathrm{B}_2 - \ldots - \mathrm{B}_i)$, and $\Phi = -(I - \mathrm{B}_1 - \mathrm{B}_2 - \ldots - \mathrm{B}_p)$.

Equation (6.28) is known as a vector error correction model (VECM) and the error correction term is embodied in ΦY_{t-p}. The parameter matrices Φ_i and Φ in Equation (6.28) are short-run and long-run adjustments to the changes in Y_t, respectively. The $(m \times m)$ matrix Φ can be expressed as a product of two matrices:

$$\Phi = \alpha\beta'$$

where α represents the speed of adjustment to disequilibrium and β is a matrix of long-run coefficients such that the term $\beta' Y_{t-p}$ represents $r \leq (m-1)$ cointegration vectors which ensure that Y_t converges to its long-run equilibrium status. If the variables in Y_t are all $I(1)$, then ΔY_{t-i}

must be $I(0)$. For U_t to be a 'white noise' process in Equation (6.28), the term ΦY_{t-p} must satisfy one of the following three conditions.

First, all the variables in Y_t are stationary. This means that the spurious regression problem will not arise and a standard VAR model with levels variables would be a good strategy to model and forecast the economic relationships. Secondly, no cointegration relationships exist among the variables in Y_t. If this is the case, there is no need to incorporate the long-run information (level variables) into the system, and a standard VAR model with differenced series would be appropriate. The third situation is that there are $r \leq (m-1)$ cointegration relationships. In this case, the linear combination of the $I(1)$ variables in Y_t will be $I(0)$, and therefore Equation (6.28) is said to be a balanced equation, i.e., all the terms in Equation (6.28) are $I(0)$.

The number of cointegrating vectors can be obtained by looking at the significance of the characteristic roots of Φ. From linear algebra, we know that the rank of a matrix is the same as the number of characteristic roots that are different from zero. Suppose that we have found that the matrix Φ has r characteristic roots $\lambda_1, \lambda_2, \ldots, \lambda_r$ which are ranked in descending order. If the rank of Φ is zero, all the r characteristic roots are also zero, and this suggests that the variables in Y_t are not cointegrated.

The number of characteristic roots can be calculated from the following two statistics:

$$\lambda_{trace} = -T \sum_{i=r+1}^{m} \ln(1-\hat{\lambda}_i) \tag{6.29}$$

$$\lambda_{max} = -T \ln(1-\hat{\lambda}_{r+1}) \tag{6.30}$$

where $\hat{\lambda}_i$ are the estimated values of the characteristic roots or eigenvalues from the estimated Φ matrix in Equation (6.28), and T is the total number of observations. The first test is known as the trace test. The null hypothesis is that there are at most r cointegrating vectors, that is the rank of Φ is less than or equal to r, and the alternative hypothesis in this case is that there are more than r cointegrating vectors. The second statistic, known as the maximal eigenvalue test, assumes that the rank is r against the alternative that the rank is $(r+1)$. Johansen and Juselius (1990) report the critical values for these two statistics based on a Monte Carlo simulation approach. Although interested readers can consult the original paper for the critical values, most modern econometric packages report these values against the calculated statistics.

The VECM discussed above does not include any deterministic components (such as intercept, trend or seasonal dummies). Equation (6.28) can, however, be easily extended to include these deterministic terms:

$$\Delta Y_t = \sum_{i=1}^{p-1} \Phi_i \Delta Y_{t-i} + \Phi Y_{t-p} + \Psi D_t + U_t \tag{6.31}$$

where D_t is a vector of deterministic variables.

In Equation (6.31), we see that the cointegration relationship can now be represented by $\Phi Y_{t-p} + \Psi D_t$ or $\alpha\beta' Y_{t-p} + \Psi D_t$. Johansen (1995) considers five situations in which deterministic components could be incorporated in the cointegration equations. These are as follows:

Model I: The time series in Y_t does not have deterministic trends and the long-run cointegration equations do not have intercepts:

$$\Phi Y_{t-p} + \Psi D_t = \alpha\beta' Y_{t-p} \tag{6.32}$$

This type of cointegration model normally relates to differenced data. In practice this specification is uninteresting since most economic time series are integrated. Therefore the cointegrating equations tend to include either a trend or an intercept or sometimes both.

Model II: The time series in Y_t does not have deterministic trends but the cointegration equations have intercepts:

$$\Phi Y_{t-p} + \Psi D_t = \alpha(\beta' Y_{t-p} + \mu) \tag{6.33}$$

where μ is an $(m \times 1)$ vector of intercepts. This type of model normally relates to data that do not have linear trends, but may have stochastic trends (e.g., financial time series such as exchange rates).

Model III: The time series in Y_t has deterministic trends but the cointegration equations have intercepts only:

$$\Phi Y_{t-p} + \Psi D_t = \alpha(\beta' Y_{t-p} + \rho_0) + \alpha_\perp \gamma_0 \tag{6.34}$$

where $\alpha_\perp$ is an $m \times (m - r)$ matrix such that $\alpha' \alpha_\perp = 0$. Model III is the most commonly used specification in economics because in many long-run economic relationships, such as the relationship between consumption and income, trend components are normally not present. Tourism demand models tend to fall into this group of models as well (Kim and Song, 1998).

Model IV: Both the time series in Y_t and the cointegration models have linear trends:

$$\Phi Y_{t-p} + \Psi D_t = \alpha(\beta' Y_{t-p} + \rho_0 + \rho_1 t) + \alpha_\perp \gamma_0 \tag{6.35}$$

This type of model is normally associated with the demand for durable goods that involves habit persistence behaviour. For example, Romilly *et al.* (1998b) found a negative trend in their long-run car ownership model for Great Britain, and this negative trend reflects a change in consumers' attitudes towards not owning a car in the long run due to the increasing level of road congestion in the UK. However, there is not enough evidence to suggest that the long-run tourism demand model contains trend components.

Model V: The variables in levels form have quadratic trends, but the cointegration equations have linear trends:

$$\Phi Y_{t-p} + \Psi D_t = \alpha(\beta' Y_{t-p} + \rho_0 + \rho_1 t) + \alpha_\perp(\gamma_0 + \gamma_1 t) \tag{6.36}$$

This type of model is difficult to justify on economic grounds, especially when the variables enter the cointegration equations in log form, but it does imply that the long-run equilibrium relationship is evolving at an increasing or decreasing rate.

The critical values of the trace and maximal eigenvalue cointegration tests given by Equations (6.29) and (6.30) are not the same for each of these five models. Osterwald-Lenum (1992) provides simulated critical values for these models and they are normally reported in computing packages (e.g., Eviews 3.1 and Microfit 4.0).

A key question is which of the above five specifications should be selected. There is no clear-cut answer; however, as a rule of thumb, a researcher should be guided by both economic interpretations of the estimated long-run cointegrating vectors (such as signs and magnitudes) and statistical criteria.

6.7.2 *An empirical example*

The aim of this example is to provide a step-by-step illustration of the Johansen cointegration methodology. We use the same data set on the demand for Korean tourism by UK and USA residents as used in the previous section. We are interested in testing for cointegration relationships among the following variables: LTA_{it}, $LGDP_{it}$, LTV_{it} and $LRCPI_{it}$ where $i = 1, 2$ and $1 =$ UK and $2 =$ USA. Since the Engle–Granger two-stage approach identified a cointegration relationship between the demand variables for the UK but not for the USA, we will try to see whether the Johansen method confirms this or not. A comparison between the two procedures should be very useful.

Step 1. Pretest the data and select the lag length of the VAR model. In this step, it is advisable to plot the data to see how the variables evolve over time. Although the variables with higher orders of integration might also be included in the VAR model, it is better to include only the variables which are integrated of order one plus some of the deterministic variables which are normally $I(0)$. According to the Dickey–Fuller and Phillips–Perron tests we already discovered in Chapter 5 that all the tourism demand variables in question are integrated of order one.

The Johansen cointegration tests can be very sensitive to the choice of the lag length of the VAR model, and therefore the number of lags chosen will directly affect the validity of the test results. The criteria used to determine the lag structure of the VAR model are those discussed in the previous section, i.e., the *AIC*, *SBC* and the *LR* statistic. We already used these criteria in the determination of the order of the UK and USA VAR models in Section 6.6, and the results suggest that the lag length of both the UK and USA VAR models should be equal to one.

Step 2. Estimate the VECM and determine the rank of the parameter matrix Φ. The VECM in our case can be formulated as:

$$\Delta Y_t = \Phi_i \Delta Y_{t-1} + \Phi Y_{t-1} + U_t$$

or

$$\begin{bmatrix} \Delta LTA_{it} \\ \Delta LGDP_{it} \\ \Delta LTV_{it} \\ \Delta LRCPI_{it} \end{bmatrix} = \begin{bmatrix} \phi_{11}^1 & \phi_{12}^1 & \phi_{13}^1 & \phi_{14}^1 \\ \phi_{21}^1 & \phi_{22}^1 & \phi_{23}^1 & \phi_{24}^1 \\ \phi_{31}^1 & \phi_{32}^1 & \phi_{33}^1 & \phi_{34}^1 \\ \phi_{41}^1 & \phi_{42}^1 & \phi_{43}^1 & \phi_{44}^1 \end{bmatrix} \begin{bmatrix} \Delta LTA_{it-1} \\ \Delta LGDP_{it-1} \\ \Delta LTV_{it-1} \\ \Delta LRCPI_{it-1} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} & \phi_{14} \\ \phi_{21} & \phi_{22} & \phi_{23} & \phi_{24} \\ \phi_{31} & \phi_{32} & \phi_{33} & \phi_{34} \\ \phi_{41} & \phi_{42} & \phi_{43} & \phi_{44} \end{bmatrix}$$

$$\times \begin{bmatrix} LTA_{it-1} \\ LGDP_{it-1} \\ LTV_{it-1} \\ LRCPI_{it-1} \end{bmatrix} + \begin{bmatrix} u_{1it} \\ u_{2it} \\ u_{3it} \\ u_{4it} \end{bmatrix} \tag{6.37}$$

We see that OLS is not suitable for the estimation of the above VECM since we have to impose restrictions on the Φ matrix. Although we do not include an intercept vector in the formulation of the VECM for simplicity,

Table 6.10. Johansen cointegration test based on λ_{max} (UK).

32 observations from 1963 to 1994. Order of VAR = 1, chosen r = 2.
List of variables included in the cointegrating vector:

LTAUK	LGDPUK	LTVUK	LRCPIUK	Intercept

List of eigenvalues in descending order:

0.73681	0.49017	0.22565	0.16263	0.0000

Null	Alternative	Statistic	5% Critical value	10% Critical value
$r = 0$	$r = 1$	42.7166	28.27	25.80
$r \leq 1$	$r = 2$	21.5578	22.04	19.86
$r \leq 2$	$r = 3$	8.1832	15.87	13.81
$r \leq 3$	$r = 4$	5.6796	9.16	7.53

Table 6.11. Johansen cointegration test based on λ_{trace} (UK).

32 observations from 1963 to 1994. Order of VAR = 1, chosen r = 2.
List of variables included in the cointegrating vector:

LTAUK	LGDPUK	LTVUK	LRCPIUK	Intercept

List of eigenvalues in descending order:

0.73681	0.49017	0.22565	0.16263	0.0000

Null	Alternative	Statistic	5% Critical value	10% Critical value
$r = 0$	$r \geq 1$	78.1373	53.48	49.95
$r \leq 1$	$r \geq 2$	35.4206	34.87	31.93
$r \leq 2$	$r \geq 3$	13.8628	20.18	17.88
$r \leq 3$	$r = 4$	5.6796	9.16	7.53

when the model is estimated it is likely that intercept terms should be included since all of our series are trended.

Let us first look at the UK demand model. In the estimation of the UK VECM, an intercept vector is introduced into the long-run cointegrating equations. The estimation uses 32 observations from 1963 to 1994 since one data point is lost due to differencing. The results of the Johansen maximal eigenvalue and trace tests are reported in Tables 6.10 and 6.11 (printouts from Microfit 4.0).

The upper section of each table contains the information on the sample period, the names of the variables in the VECM and the calculated characteristic roots (eigenvalues). In the lower section of the tables, the calculated maximal eigenvalue and trace statistics and the corresponding critical values of the two statistics are shown. The inference is that if the calculated statistics are greater than the corresponding critical values at specific significance levels, the null hypotheses which are presented in the first

Table 6.12. Johansen cointegration test based on λ_{max} (USA).

32 observations from 1963 to 1994. Order of VAR = 1.
List of variables included in the cointegrating vector:

LTAUS	LGDPUS	LTVUS	LRCPIUS	Intercept

List of eigenvalues in descending order:

0.86810	0.51471	0.20891	0.089882	0.0000

Null	Alternative	Statistic	5% Critical value	10% Critical value
$r = 0$	$r = 1$	64.8231	28.27	25.80
$r \leq 1$	$r = 2$	23.1361	22.04	19.86
$r \leq 2$	$r = 3$	7.4991	15.87	13.81
$r \leq 3$	$r = 4$	3.0138	9.16	7.53

Table 6.13. Johansen cointegration test based on λ_{trace} (USA).

32 observations from 1963 to 1994. Order of VAR = 1.
List of variables included in the cointegrating vector:

LTAUS	LGDPUS	LTVUS	LRCPIUS	Intercept

List of eigenvalues in descending order:

0.86810	0.51471	0.20891	0.089882	0.0000

Null	Alternative	Statistic	5% Critical value	10% Critical value
$r = 0$	$r \geq 1$	98.4721	53.48	49.95
$r \leq 1$	$r \geq 2$	33.6491	34.87	31.93
$r \leq 2$	$r \geq 3$	10.5129	20.18	17.88
$r \leq 3$	$r = 4$	3.0138	9.16	7.53

columns of the two tables should be rejected. For example, if we look at the maximal eigenvalue test in Table 6.10, we can see that the hypothesis that there is no cointegration relationship ($r = 0$) is rejected since the calculated λ_{max} is 42.7166 which is greater than the critical values at both the 5% and 10% significance levels. The λ_{max} statistic suggests that there are two cointegration vectors at the 10% significance level, but not at the 5% level. The λ_{trace} test in Table 6.11 suggests two cointegrating vectors at the 5% significance level. Therefore, it is safe to conclude that two cointegrating relationships exist among the variables in the UK tourism demand model.

Following the same procedure we obtain the corresponding cointegration test statistics for the USA tourism demand model and they are presented in Tables 6.12 and 6.13. The results also suggest that there are likely to be two cointegration vectors, which contradicts the results from the Engle–Granger two-stage approach in which no cointegration relationship was found.

Table 6.14. Estimated cointegrated vectors in Johansen estimation (UK).

32 observations from 1963 to 1994. Order of VAR = 1, chosen $r = 2$.
List of variables included in the cointegrating vector:
LTAUK LGDPUK LTVUK LRCPIUK Intercept

	Vector 1	Vector 2
LTAUK	0.0856	−1.6713
	(−1.0000)	(−1.0000)
LGDPUK	0.4128	3.1380
	(−4.8248)	(1.8776)
LTVUK	−0.0516	0.7826
	(0.6031)	(0.4683)
LRCPIUK	0.1213	−0.0569
	(−1.4181)	(−0.3406)
Intercept	−1.7031	−6.6912
	(19.9052)	(−4.0035)

Table 6.15. Estimated cointegrated vectors in Johansen estimation (USA).

32 observations from 1963 to 1994. Order of VAR = 1, chosen $r = 2$.
List of variables included in the cointegrating vector:
LTAUS LGDPUS LTVUS LRCPIUS Intercept

	Vector 1	Vector 2
LTAUS	−0.1484	−0.8963
	(−1.0000)	(−1.0000)
LGDPUS	0.4743	1.4388
	(3.1973)	(1.6053)
LTVUS	−0.0078	0.4585
	(−0.0529)	(0.5115)
LRCPIUS	−0.6857	−0.5954
	(−4.6225)	(−0.6643)
Intercept	−4.6750	−3.9957
	(−31.5136)	(−4.4582)

According to the λ_{max} and λ_{trace} tests we know that the variables in both the UK and USA demand models have two cointegrating relationships, and these are presented in Tables 6.14 and 6.15. The values in brackets are the normalised cointegrating vectors, that is each value in the vector is divided by the first value in the same vector. As can be seen from the above equations, the income variable in the cointegration vector 1 in Table 6.14 and trade volume variable in the cointegration vector 1 in Table 6.15 do not have the expected signs. Therefore, although two cointegrating vectors for each of the two countries have been identified, only one is economically meaningful for each country.

Table 6.16. Long-run cointegration vectors based on different estimation methods.

	E-G		WB		ADLM		Johansen	
Variables	UK	US	UK	US	UK	US	UK	US
$\ln TA_{1t}$	−1.000	−1.000	−1.000	−1.000	−1.000	−1.000	−1.000	−1.000
$\ln GDP_{it}$	2.097	1.655	2.187	2.519	1.702	2.380	1.878	1.605
$\ln TV_{it}$	0.473	0.576	0.436	2.456	0.456	0.277	0.468	0.512
$\ln RCPI_{it}$	−0.213	−0.861	−0.187	−1.856	−0.322	−1.871	−0.341	−0.664
Intercept	−4.257	−6.158	−4.153	−13.640	−2.897	−13.430	−4.004	−4.458

The two economically acceptable cointegration relationships, together with those obtained from the Engle–Granger (EG), Wickens–Breusch (WB) and ADLM methods (see Chapter 5), are presented in Table 6.16. The long-run cointegrating vectors actually represent the long-run demand elasticities. Table 6.16 shows that the estimated income elasticities for the UK and USA models based on the Johansen method are lower than those estimated by the other three approaches (with one exception in the UK model). In terms of the elasticities of trade volume in the UK models the differences are relatively small no matter what methods are used (0.436–0.473), whilst for the USA the differences are much wider (from 0.277 based on the ADLM to 2.456 based on the WB). A similar pattern is also observed for the price elasticities between the two countries. These results suggest that the structure of the long-run UK model is relatively stable in comparison with the structure of the long-run USA model. In Chapter 7, we shall discuss how to model unstable relationships using a time varying parameter (TVP) approach.

Another point is that we did not find any cointegration relationship using the Engle–Granger two-stage approach in the case of the USA model, but the Johansen method suggests two long-run cointegration relationships and in one of them the coefficients are correctly signed and have the expected orders of magnitude. Since two cointegrating vectors have been identified for both the UK and USA long-run tourism demand models, the estimates from the Johansen approach are likely to be more reliable. As mentioned earlier, the Engle–Granger method tends to 'average' multiple cointegration relationships, and therefore there is a possibility that the cointegration relationships among the USA tourism demand variables have been 'averaged' out by the Engle–Granger approach.

Step 3. Generate forecasts with the VECM. Although we make our

Table 6.17. ***Ex post*** **forecast errors of the VECMs.**

	UK		US	
	CV I	CV II	CV I	CV II
MAE	0.0366	0.0446	0.0401	0.0564
RMSE	0.0395	0.0528	0.0489	0.0618

decisions about whether to accept a long-run cointegrating relationship based on the economic interpretations of the cointegration coefficients (demand elasticities in our case), for forecasting purposes this restriction does not apply. We should utilise all possible cointegrating vectors when generating forecasts.

We now re-estimate the VECMs using the sample period 1962–90 and then forecast the $\Delta \ln TA_{it}$s from 1991 to 1994 using both cointegrating vectors for each of the two generating countries. Table 6.17 presents the forecasting errors. These errors show that the cointegrating vectors (CV II for both countries) which make sense economically do not necessarily produce better forecasts. Furthermore, comparing the results in Table 6.17 with those in Table 6.9, we can see that the Johansen method generates the most accurate forecasts for the USA no matter which CV is used, whilst for the UK the Johansen approach outperforms the VAR(1), EG and WB, but not the ADLM. In general, we conclude that the Johansen approach gives more reliable estimates of the long-run coefficients and generates more accurate forecasts if more than one cointegrating vector exists among a set of economic variables. Moreover, the methods of testing for a single cointegration relationship, such as the Engle–Granger approach, are likely to fail to identify the long-run equilibrium relationships even when the variables are actually cointegrated. Therefore, our recommendation is to use the Johansen method whenever it is possible, especially in those cases where multiple cointegrating relationships are involved.

7

Time varying parameter modelling

7.1 Introduction

One of the recurrent features of causal tourism demand forecasting models compared with simple time series models, such as the no-change model, has been predictive failure (Witt and Witt, 1995). Predictive failure is normally associated with model structure instability, i.e., the parameters of the demand model vary over time. There are two basic reasons why a causal model may suffer from structural instability. The first, and traditional, view is that our knowledge about the structure of the model is limited. Thus, when a failure occurs we add this information into our knowledge base and produce a better model encompassing both the new and earlier experiences. In this case structural change is regarded merely as a manifestation of an inadequate or inappropriate specification. The second view is that predictive failure and the apparent coefficient changes exhibited in a tourism demand model may well be a reflection of underlying structural change in the data-generating process. This structural change is mainly related to important social, political and economic policy changes. Associated with these two views on the reasons for model instability there are two radically different approaches a modeller could adopt.

The first view requires re-specifying the unstable demand model to achieve a stable one encompassing the key characteristics of all existing models. This approach has many benefits, not least being the capacity to learn from the failures of the earlier models. In essence it is 'scientific' in the

Poperian sense of establishing hypotheses which are subsequently modified in the light of experimental evidence. However, this methodology fails if the DGP is not constant. According to the second view of structural instability, the appropriate way of dealing with the problem is to choose an alternative approach, e.g., the time varying parameter (TVP) regression, to model the structural change. Despite common recognition of these two causes of model instability, most people have adopted the former option of solving the problem of structural instability for two reasons. First, the traditional assumption of model parameter constancy, even if it is often unrealistic, has been accepted by most econometricians as doctrine. Second, the complexity of the TVP estimation approach and lack of readily available computer software have restricted its use. However, the number of studies that utilises the TVP approach to modelling and forecasting economic activities has been growing since the late 1980s. For example, the TVP method has been used successfully in modelling and forecasting rational expectations formation and inflation (Burmeister and Wall, 1982; Cuthbertson, 1988), demand for money (Swamy *et al.*, 1990), consumption (Song, 1995; Song *et al.*, 1998) and housing-prices (Brown *et al.*, 1997). However, very limited efforts have been made to use the TVP approach in tourism demand forecasting; notable exceptions are Song *et al.* (1999) and Riddington (1999).

The tourism industry has evolved from a supply-led industry to a demand-driven one since the 1970s. Political, economic and social shocks such as the two 'oil crises' in the 1970s, terrorism attacks in the 1980s and the Gulf War in the early 1990s have also had sustained impacts on the demand for international tourism. Although the dummy variable approach could be used to attempt to capture the structural change due to the above-mentioned shocks, the use of the TVP method is probably a more realistic alternative.

In Section 7.2 the parameter stability of the tourism demand model is examined using recursive OLS. Section 7.3 discusses the specification and estimation of the TVP model. Section 7.4 presents some empirical results relating to TVP tourism demand models, and Section 7.5 concludes the chapter.

7.2 Are tourism demand elasticities constant over time?

From Chapter 1, we know that the coefficients of the explanatory variables in the double log demand model are measures of demand elasticities. According to OLS these demand elasticities are time invariant. Is this is a reasonable assumption? Are the demand elasticities or the structure of the demand model really constant over time? This section sets out to answer these questions with the help of some statistical evidence.

A number of statistics may be used to test for parameter constancy. These include recursive OLS and the two Chow structural instability tests (Chow, 1960). Recursive OLS initially involves estimating the demand model using OLS from a small sub-sample of data $t = 1, 2, \ldots n$, where $n \geq k$ and k is the number of explanatory variables (including the constant term) in the demand model; the sample period is then extended by one observation to $t = 1, 2, \ldots n + 1$, and the model is re-estimated using OLS; this procedure continues until all the observations in the sample are used. Each time the regression is estimated, the estimated coefficients are plotted. Examination of the plotted coefficients should inform us whether the demand elasticities of the tourism demand model have changed over time or not. For example, from Equations (5.17) and (5.18) we can see that the long-run tourism demand models for the UK and USA are estimated as:

UK: $\ln TA_{1t} = -4.257 + 2.097 \ln GDP_{1t} + 0.473 \ln TV_{1t} - 0.213 \ln LRCPI_{1t}$

USA: $\ln TA_{2t} = -6.158 + 1.655 \ln GDP_{2t} + 0.536 \ln TV_{2t} - 0.861 \ln LRCPI_{2t}$

These two models show that the long-run income elasticity is 2.097 for the UK and 1.655 for the USA; the trade volume elasticity is 0.473 for the UK and 0.536 for the USA; and the own-price elasticity is −0.213 for the UK and −0.861 for the USA. These elasticities are estimated over the period 1964 to 1994 using OLS and they are assumed to be constant over the sample period.

However, if we re-estimate the above models using recursive OLS, we can see that the coefficient constancy assumption is too restrictive. Figures 7.1 and 7.2 present the recursive OLS estimates of the four elasticities in each of the two long-run models. The figures represent the estimated constant term, income elasticity, trade volume elasticity and own-price elasticity, respectively. The estimation starts from 1970 as the observations from 1964 to 1969 are used for the initial estimation. It is clear from the recursive estimates that some of the demand elasticities in the UK and USA models have changed significantly since the 1970s. For example, the income elasticity in the UK model changed from 5.0 in 1974 to 3.6 in 1976 and then dropped again to about 2.0 in 1985. The trade volume elasticity during the same period increased from 0.25 to 0.45. The change of income elasticity suggests that the demand for Korean tourism by UK residents was very sensitive to income changes in the 1970s, but the demand became relatively

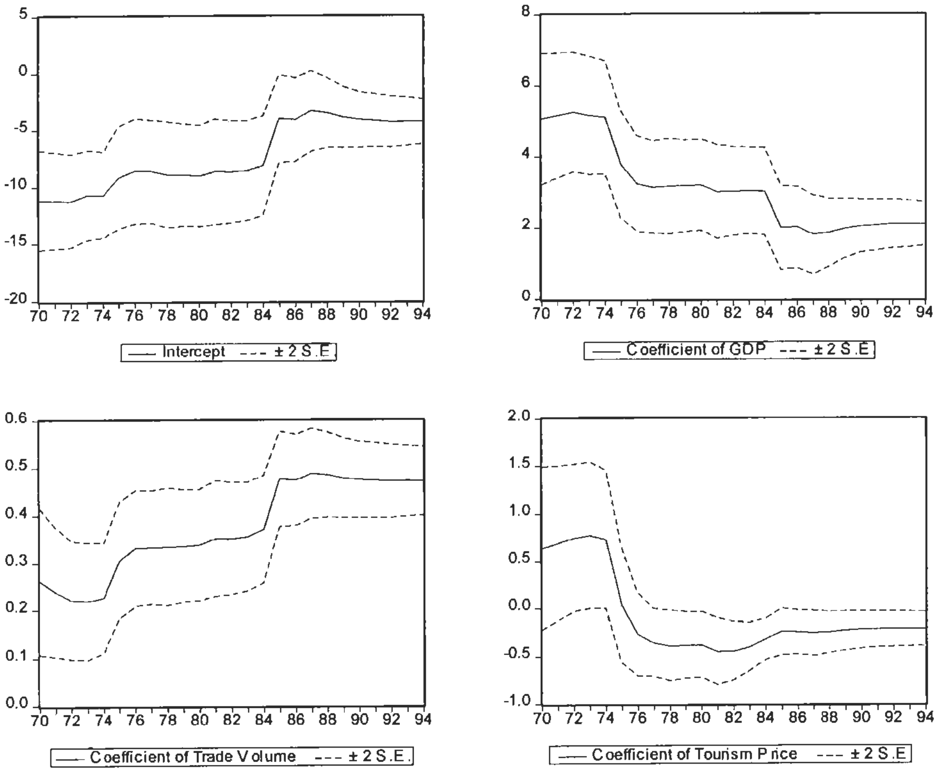

Figure 7.1 Recursive estimates of coefficients in UK model.

less sensitive to income changes in the 1980s and 1990s. The income and trade elasticities in the USA model experienced similar changes during the same period. Both models, however, show that the tourism price elasticity has been relatively stable since a sharp fall in 1974 (the first oil crisis).

The two Chow tests for structural instability were given in Chapter 3 (Equations 3.14 and 3.15). The first Chow test (the breakpoint test) is used here to see whether the elasticities are constant over time in the two demand models. In order to calculate the F statistic based on Equation 3.14, we need to specify the point at which the structure of the model breaks. By looking at the recursive OLS estimates of the elasticities in Figures 7.1 and 7.2, we can see that there were two significant changes in demand elasticities; one was in 1974 and the other in 1984. Therefore, two values of the Chow breakpoint statistic can be calculated for each of the two models in relation to the two break points, and they are given below.

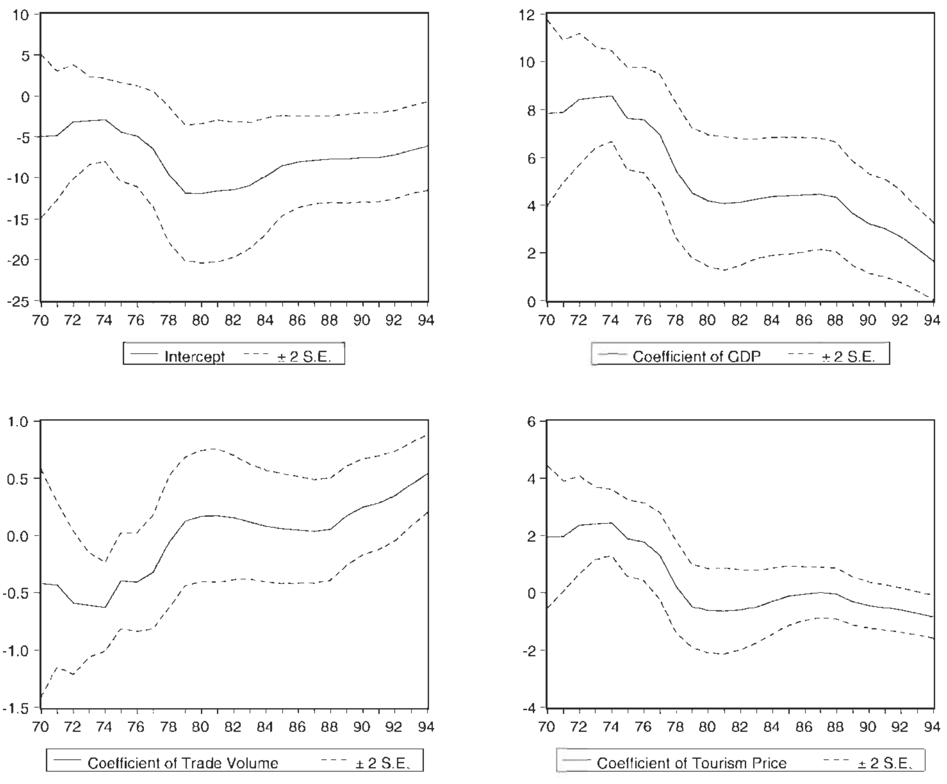

Figure 7.2 Recursive estimates of coefficients in USA model.

	1974	1984
UK model:	$F_{Chow1}(4, 24) = 7.56$	$F_{Chow1}(4, 24) = 2.64$
USA model:	$F_{Chow1}(4, 24) = 46.07$	$F_{Chow1}(4, 24) = 3.58$

The critical value of F_{Chow1} $(4, 24)$ at the 5% significance level is 2.54. From these figures we can see clearly that the demand elasticities have changed over the same period.

Recursive OLS is a useful tool to examine the structural stability of the regression model, especially to examine changes in the regression coefficients over the sample period due to persistent external shocks to the system. However, it suffers from the following problems in application. First, the validity of the recursive estimation depends on the length of the sample. If the sample is relatively small, the values of the parameters estimated by recursive OLS may exhibit structural change in the early part of the sample, but this does not necessarily mean that the structure of the model is unstable. The instability may simply be due to the small sample biases. Second, recursive OLS is not useful for forecasting if the structure of the model is indeed unstable due to external shocks, because when all observations are used in estimating the model the forecasts generated will be exactly the same as those obtained in the traditional constant parameter regression model, i.e., the structural instability cannot be taken into account in the forecasting process. The TVP approach, however, allows for richer specifications that go far beyond the recursive OLS method. The TVP method is capable of simulating different types of shocks that may influence the relationship between the dependent and explanatory variables, and the model is also unique in incorporating structural changes into the forecasting process.

7.3 The TVP model and Kalman filter

The TVP model may be specified in state space (SS) form:

$$\mathbf{y}_t = \mathbf{x}_t \boldsymbol{\beta}_t + \mathbf{u}_t \tag{7.1}$$

$$\boldsymbol{\beta}_t = \boldsymbol{\Phi} \boldsymbol{\beta}_{t-1} + \boldsymbol{R}_t \boldsymbol{e}_t \tag{7.2}$$

where $\mathbf{y}_t$ is a $T \times 1$ vector of tourism demand, $\mathbf{x}_t$ is a $T \times k$ matrix of the explanatory variables, $\boldsymbol{\beta}_t$ is a $k \times 1$ vector of parameters known as the *state vector*, $\boldsymbol{\Phi}$ is a $k \times k$ matrix, $\mathbf{R}_t$ is a $k \times g$ matrix, $\mathbf{u}_t$ is a $T \times 1$ vector of residuals with zero mean and constant covariance matrix $\boldsymbol{H}_t$, and $\boldsymbol{e}_t$ is a $g \times 1$ vector of serially uncorrelated residuals with zero mean and constant covariance matrix $\boldsymbol{Q}_t$.

Equation (7.1) is called the measurement or system equation while Equation (7.2) is known as the transition equation. There are two additional assumptions related to Equations (7.1) and (7.2). The first is that the initial vector $\boldsymbol{\beta}_0$ has a mean of $\boldsymbol{b}_0$ and a covariance matrix $\boldsymbol{P}_0$. The second assumption is that the residual terms $\mathbf{u}_t$ and $\boldsymbol{e}_t$ in the measurement and transition equations are not correlated. The matrices $\mathbf{x}_t$, $\boldsymbol{H}_t$, $\boldsymbol{R}_t$ and $\boldsymbol{Q}_t$ are also assumed to be non-stochastic.

If the components of the matrix $\boldsymbol{\Phi}$ in Equation (7.2) equal unity, the transition equation becomes a *random walk*:

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{R}_t \boldsymbol{e}_t \tag{7.3}$$

If the transition equation is a random walk, the parameter vector $\boldsymbol{\beta}_t$ is said to be non-stationary.

Another possible form of the transition equation is:

$$\boldsymbol{\beta}_t = \boldsymbol{\mu} - \boldsymbol{\Phi}(\boldsymbol{\beta}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{R}_t \boldsymbol{e}_t \tag{7.4}$$

where $\boldsymbol{\mu}$ is the mean of $\boldsymbol{\beta}_t$. Equation (7.4) suggests a stationary process.

The specification of the transition equation is normally determined by experimentation. The criteria used to determine the structure of the transition equation are the goodness of fit and the predictive power of the model. Once the SS model is formulated a convenience algorithm, known as the *Kalman filter* (KF), can be used to estimate the SS model (Kalman, 1960). The KF is a recursive procedure for calculating the optimal estimator of the state vector given all the information available at time t.

The derivation of the KF is now briefly demonstrated, but for a detailed exposition see Harvey (1987). Let $\boldsymbol{b}_t$ and $\boldsymbol{P}_t$ denote the optimal estimators of the state vector and the covariance of $\boldsymbol{\beta}_t$ in Equation (7.2), respectively. Suppose the estimation starts at time $t-1$; then $\boldsymbol{b}_{t/t-1}$ and $\boldsymbol{P}_{t/t-1}$ can be calculated from:

$$\boldsymbol{b}_{t/t-1} = \boldsymbol{\Phi}\boldsymbol{b}_{t-1} \tag{7.5}$$

and

$$\boldsymbol{P}_{t/t-1} = \boldsymbol{\Phi}\boldsymbol{P}_{t-1}\boldsymbol{\Phi}' + \boldsymbol{R}_t\boldsymbol{Q}_t\boldsymbol{R}_t' \tag{7.6}$$

Given Equations (7.5) and (7.6), we can estimate $\mathbf{y}_t$ based on the information at $t-1$:

$$\hat{\mathbf{y}}_{t/t-1} = \mathbf{x}_t\boldsymbol{b}_{t/t-1} \tag{7.7}$$

The prediction error of $\mathbf{y}_t$ is

$$\boldsymbol{r}_t = \mathbf{y}_t - \hat{\mathbf{y}}_{t/t-1} = \boldsymbol{\Phi}(\boldsymbol{\beta}_t - \boldsymbol{b}_{t/t-1}) + \mathbf{u}_t \tag{7.8}$$

and the mean squared error of $\mathbf{y}_t$ is:

$$\boldsymbol{F}_t = \boldsymbol{\Phi}\boldsymbol{P}_{t/t-1}\boldsymbol{\Phi}' + \boldsymbol{H}_t \tag{7.9}$$

Once a new observation becomes available, the estimator of the state vector can be updated, and the updating process is written as:

$$\boldsymbol{b}_t = \boldsymbol{b}_{t/t-1} + \boldsymbol{P}_{t/t-1}\mathbf{x}_t'\boldsymbol{F}_t^{-1}(\mathbf{y}_t - \boldsymbol{\Phi}\boldsymbol{b}_{t/t-1}) \tag{7.10}$$

and

$$\boldsymbol{P}_t = \boldsymbol{P}_{t/t-1} - \boldsymbol{P}_{t/t-1}\boldsymbol{\Phi}'\boldsymbol{F}_t^{-1}\boldsymbol{\Phi}\boldsymbol{P}_{t/t-1} \tag{7.11}$$

Equations (7.5)–(7.11) are together called the KF. Given the initial values of $\boldsymbol{b}_0$ and $\boldsymbol{P}_0$, the Kalman filter produces the optimal estimator of the state vector as each observation becomes available. Therefore, one of the important steps in the KF estimation is to determine the initial values of $\boldsymbol{b}_0$ and $\boldsymbol{P}_0$, and we will return to this later. When all the T observations are exhausted, $\boldsymbol{b}_T$ should contain all the information needed for forecasting the future values of $\boldsymbol{b}$, $\boldsymbol{P}$ and $\mathbf{y}$.

The m-period-ahead forecasts of $\boldsymbol{b}$, $\boldsymbol{P}$ and $\mathbf{y}$ are based on Equations (7.12)–(7.14):

$$\boldsymbol{b}_{T+m/T} = \boldsymbol{\Phi}\boldsymbol{b}_{T+m-1} \tag{7.12}$$

$$\boldsymbol{P}_{T+m/T} = \boldsymbol{\Phi}\boldsymbol{P}_{T+m-1/T}\boldsymbol{\Phi}' + \boldsymbol{R}_{T+m}\boldsymbol{Q}_{T+m}\boldsymbol{R}'_{T+m} \tag{7.13}$$

$$\hat{\mathbf{y}}_{T+m/T} = \mathbf{x}_{T+m}\boldsymbol{b}_{T+m/T} \tag{7.14}$$

The mean square forecasting error is

$$\boldsymbol{F}_{T/m} = \mathbf{x}_{T+m}\boldsymbol{P}_{T+m/T}\mathbf{x}'_{T+m} + \boldsymbol{H}_{T+m} \tag{7.15}$$

The KF process estimates the state vector $\boldsymbol{\beta}_t$, given information available up to time t. After the $\boldsymbol{\beta}_t$ is estimated using the KF, the Kalman smoother (KS) is then applied to re-estimate $\boldsymbol{\beta}_t$ using the information after time t; the KS estimator of $\boldsymbol{\beta}_t$ is denoted $\boldsymbol{b}_{t/T}$. Since the KS utilises more information than the KF, the estimated $\boldsymbol{b}_{t/T}$ will always have a smaller covariance than the KF estimator.

To initialise the KF, the starting-values of $\boldsymbol{b}_0$ and $\boldsymbol{P}_0$, known as the initial conditions of the KF, need to be determined first. There are a number of ways in which the initial values $\boldsymbol{b}_0$ and $\boldsymbol{P}_0$ may be determined (Harvey, 1993, pp. 88–9). The first way of determining $\boldsymbol{b}_0$ and $\boldsymbol{P}_0$ is to use a method called *diffuse priors*. This procedure assigns very large initial values to the covariance matrix $\boldsymbol{P}$, while the initial values for the coefficients of the explanatory variables are chosen arbitrarily. If we assume that the coefficients follow a random walk process, an initial value of 1 should be assigned to those coefficients. The starting-value of $\boldsymbol{P}_0$ is then set to equal to $\kappa\boldsymbol{I}$, where κ is a large but finite number and $\boldsymbol{I}$ is a $k \times k$ identity matrix. After the values of $\boldsymbol{b}_0$ and $\boldsymbol{P}_0$ are determined, the KF can then be used to estimate the TVPs. An alternative approach to choosing $\boldsymbol{b}_0$ and $\boldsymbol{P}_0$ is to estimate Equation (7.1) using OLS and then equate $\boldsymbol{b}_0$ and $\boldsymbol{P}_0$ to the corresponding OLS constant parameters (Eviews 3.1 uses this approach to initialise the Kalman filter estimation).

Over the last decade, the cointegration, error correction and vector autoregressive models have been widely used, and sometimes even over-used, in the areas of modelling and forecasting virtually all macroeconomic and financial activities. Although these modelling methods are different in the ways in which the models are constructed, they are all derivatives of traditional OLS. According to Bomhoff (1994), if time series are stationary their first and second moments are well defined and there is no conceptual problem in computing the unconditional means, variances and covariances based on the observations over the same period. However, when time series are non-stationary, as with most economic time series, OLS is invalid since the properties of the time series depend on the length of the sample period. Therefore, unconditional means, variances and covariances cannot be calculated using OLS. To overcome this problem, the data need to be differenced. This often results in the loss of the long-run characteristics of the model. The TVP approach, on the other hand, estimates the parameters of the model sequentially using the forwards KF and backwards KS, and produces conditional distributions for means and variances. It is therefore more useful in analysing non-stationary series. Moreover, the TVP approach does not require the data to be stationary before model estimation, and so the procedure of model specification and estimation is drastically simplified since one does not have to worry about unit root testing and data differencing.

In their comparative study of the forecasting performance of a number of econometric models of international tourism demand, Song *et al.* (1999) found that the TVP model generated the most accurate short-run forecasts. This finding is consistent with previous studies on forecasting other macroeconomic activities (Song, 1995; Song *et al.*, 1998; and Brown *et al.*, 1997).

7.4 A worked example

In this section we illustrate the process of specifying and estimating the tourism demand Equations (5.17) and (5.18) using the TVP approach. The specification of the demand model in SS form is given as:

$$\ln TA_{it} = \alpha_{0t} + \alpha_{1t}\ln GDP_{it} + \alpha_{2t}\ln TV_{it} + \alpha_{3t}\ln LRCPI_{it} + u_{it} \tag{7.16}$$

$$\alpha_{jt} = \alpha_{jt-1} + \varepsilon_{jt} \tag{7.17}$$

where $i = 1$ and 2, 1 refers to the UK and 2 to the USA, $j = 1, 2, 3$ and 4 representing the number of TVPs in the system equation, and the variables are the same as those described in Chapter 5.

Equation (7.16) defines the relationship between the dependent and explanatory variables, while Equation (7.17) expresses the process according to which the parameters vary. In this example, we assume that the parameters follow a random walk process. Although the transition equation (7.17) can also be specified in other ways, the random walk process should

Table 7.1 Kalman filter estimates of the TVP models.
(Dependent variable: $\ln TA_{it}$)

Variable	UK model	USA model
Intercept	0.568 (0.159)	2.047 (1.807)
$\ln GDP_{it}$	1.022* (0.418)	0.146 (0.490)
$\ln TV_{it}$	0.549* (0.045)	0.586* (0.094)
$\ln LRCPI_{it}$	−0.140 (0.17)	−0.573* (0.172)
R^2	0.999	0.998
S.E.	$0.57E-5$	$0.91E-4$

Notes: Values in parentheses are the standard errors; * denotes that the estimates are significant at the 5% level.

always be tried first as this is the simplest process of parameter evolution. In fact researchers have found that the random walk process can adequately capture the effects of parameter changes in many economic models (see, for example, Bohara and Sauer, 1992; Greenslade and Hall, 1996; Kim, 1993).

Equations (7.16) and (7.17) are estimated using the Kalman filter which is available in the Eviews 3.1 program. The results are reported in Table 7.1. Since the transition equations all follow a random walk process, that is $\alpha_{jt} = \alpha_{jt-1}$, their estimates are not reported.

The results show that the two models fit the data very well according to the very high R^2 values and the very low regression standard errors compared with the OLS equivalents. However, the relative price variable in the UK model and the GDP variable in the USA model are insignificant at the 5% level.[1] It is also worth noting that the reported estimates of the demand elasticities are the final values from the whole sample. Researchers, however, may also wish to examine how the elasticities evolve over time. By looking at the patterns of the changes in model parameters over time, a researcher can gain useful insights into the behavioural changes of potential tourists in relation to policy and regime changes during the sample period. From basic economic theory we know that the magnitudes of the demand elasticities are unlikely to remain constant over time, since the nature of the product demanded and the tastes of consumers are time varying. One example may be that holiday travel to an overseas destination which was a

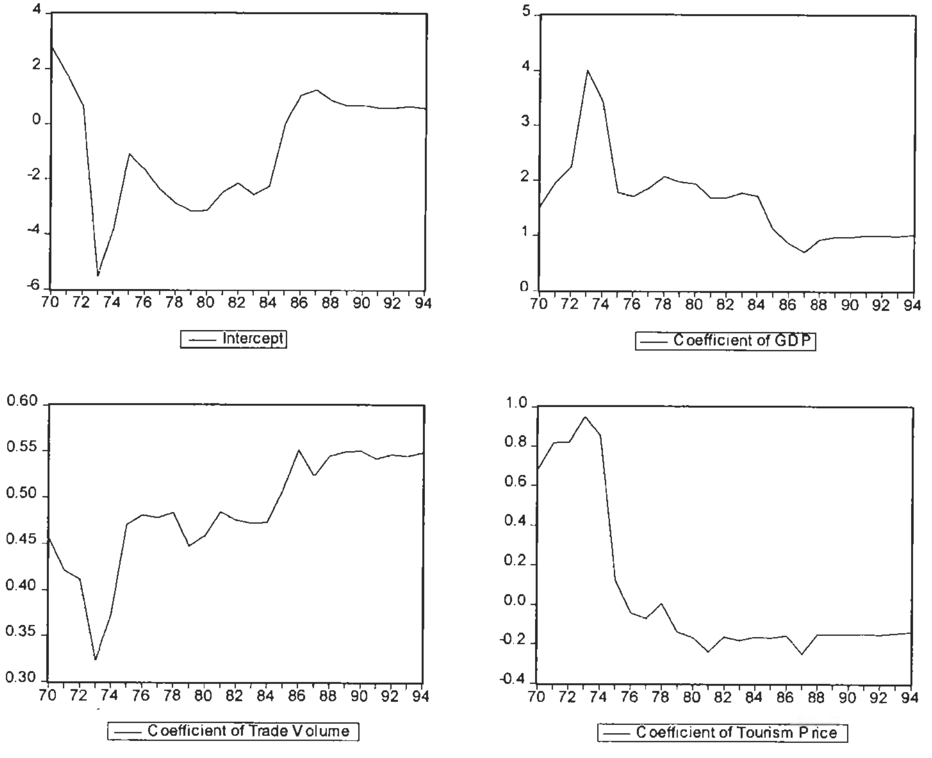

Figure 7.3 Estimated coefficients in UK model based on Kalman filter approach.

luxury product for many UK residents in the 1970s (high income elasticity) had become a necessity for a considerable proportion of the UK population in the late 1980s and the 1990s (low income elasticity). The demand elasticities are also likely to vary with changes in market structure and the increasing competition within the tourism industry. To examine how the demand elasticities have evolved over time, it is therefore useful to plot the Kalman filter estimates of the elasticities during the sample period. Figures 7.3 and 7.4 give such plots.

If we ignore the first few years in the sample, we can see that the income elasticity in the UK model dropped from around 2 in the 1970s to about 1 in the 1990s, and this supports our discussion regarding the decline of income elasticity above. However, the change in income elasticity in the USA was much smaller. The tourism demand elasticities with respect to trade volume in both models increased by about 0.2 during the sample

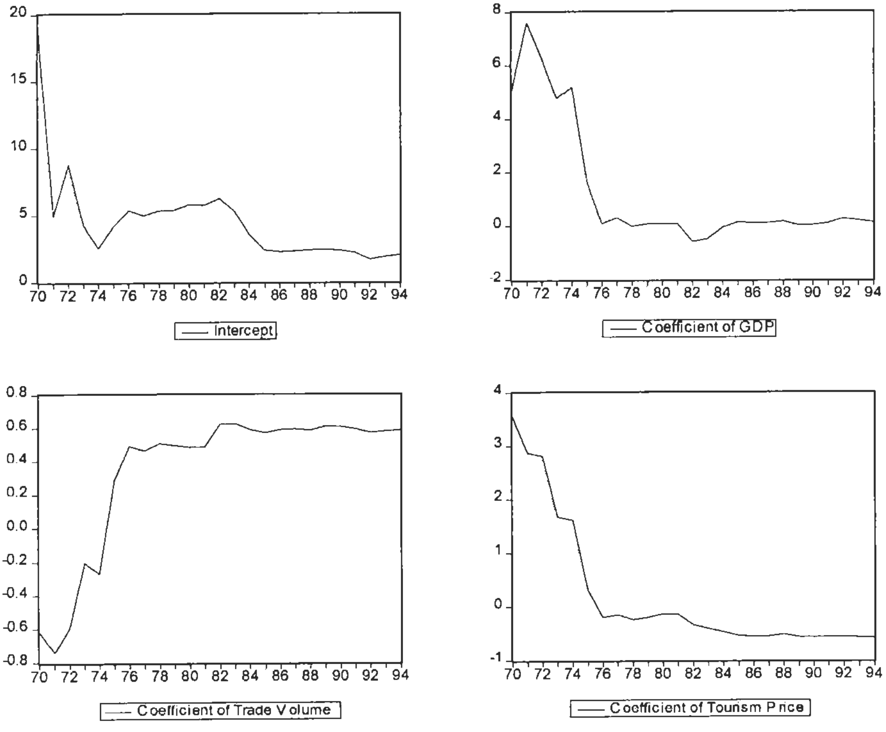

Figure 7.4 Estimated coefficients in USA model based on Kalman filter approach.

period. This suggests that business links have become more important in the determination of demand for Korean tourism by UK and USA residents. The Kalman filter estimates of the price elasticities in the two models appear to be stable over the last 20 years, and this is consistent with the estimates of the recursive OLS in section 7.2.

7.5 Summary and conclusions

Demand elasticities in tourism demand models are likely to be time varying due to policy changes, evolution of consumers' tastes and economic regime shifts. The traditional econometric modelling approach that assumes parameter constancy is unable to take into account parameter instability, and therefore often generates poor forecasts. The TVP method is an alternative method that simulates structural changes in tourism demand models and may well therefore produce more accurate forecasts. The TVP model can be written in the SS form that includes a measurement/system equation and a transition equation. The specification of the transition equation is normally determined by experimentation and the criteria used for the selection of the transition equation are the goodness of fit and the ex post forecasting accuracy. In most cases, the random walk process is adequate in capturing the instability in the parameters.

Once the TVP model is written in SS form, an algorithm known as the Kalman filter may be applied to estimate the parameters of the model. To start the KF estimation initial values of the parameters and the corresponding covariances need to be determined. Once these initial values have been decided based on either the *diffusion priors* or the OLS estimates, the estimation can proceed sequentially with a forwards filter and backwards smoother.

The TVP model is likely to outperform the OLS regression model when the structure of the model is time dependent, and the forecasting performance of the TVP model will be evaluated in Chapter 9.

Note

1. It is possible to improve the significance of these two variables by using other specifications for the transition equations, such as allowing the transition equations to follow an AR or ARMA process. However, for the purpose of simplicity, we report only the estimates based on the random walk specification for the transition equations.

8

Panel data analysis*

8.1 Introduction

The increasing importance of international tourism in the balance of payments and in the creation of employment for both developed and developing economies has led to a large number of econometric studies on the determinants of international tourism demand over the past three decades. Witt and Witt (1995) provide a survey of the major studies. As noted by Lim (1997), however, although time series data are widely utilised in estimating tourism demand, few studies use cross-section or panel data. Exceptions are Carey (1991), Tremblay (1989) and Yavas and Bilgin (1996), who pool their cross-section and time series data to estimate elasticities with respect to income, exchange rates, relative prices, transport costs and other explanatory variables for selected Caribbean countries, 18 European countries and Turkey, respectively.

The purpose of this chapter is to quantify tourism demand relationships using the panel data approach and to ascertain whether variables of a social as well as an economic nature are important in the determination of international tourism demand. One of the advantages of panel data analysis over pure time series or pure cross-sectional modelling is the relatively large number of observations and the consequent increase in degrees of

* This chapter was written jointly with Peter Romilly of the University of Abertay Dundee and Xiaming Liu of Aston University Business School.

freedom. This reduces collinearity and improves the efficiency of the estimates. This is particularly important in tourism demand models where social variables, such as age structure, gender and education, are included in the models, as collinearity problems are likely to arise as a result of the lack of variation in the social variables over time. The use of panel data is not unproblematic, however, since the choice of an appropriate model depends *inter alia* on the degree of homogeneity of the intercept and slope coefficients and the extent to which any individual cross-section effects are correlated with the explanatory variables. These issues are discussed later in this chapter.

The remainder of this chapter is structured as follows. Section 8.2 provides an introduction to the panel data approach, Section 8.3 presents an example of panel data analysis based on a data set of 138 countries over the period 1989–95, and Section 4 concludes the chapter.

8.2 Model specification and estimation

8.2.1 Test for poolability

Given observations on the tourism demand variable y across $i = 1, \ldots, N$ regions over $t = 1, \ldots, T$ time periods and $k = 1, \ldots, K$ explanatory variables denoted by the $K \times 1$ vector $\mathbf{x}$, the classical linear regression model is of the form:

$$y_{it} = \alpha_i + \mathbf{b}_i' \mathbf{x}_{it} + \varepsilon_{it} \tag{8.1}$$

where the intercept and slope coefficients are allowed to vary across regions. The error term is identical and independently distributed or iid $(0, \sigma_\varepsilon^2)$ for all i and t. The first stage in the estimation procedure is to test for the poolability of the data. Although there are many examples of estimation using pooled data, poolability is by no means a foregone conclusion (see, for example, the results for tourism demand equations given in Tremblay (1989)). A starting-point is to estimate a restricted model in which both the intercept and slope coefficients are assumed homogeneous across regions:

$$y_{it} = \alpha + \mathbf{b}' \mathbf{x}_{it} + \varepsilon_{it} \tag{8.2}$$

and compare this with the unrestricted model above by means of an F test. The first F test is:

$$F_1 = \frac{(RSS_{R1} - RSS_U)/[(N-1)(K+1)]}{RSS_U/[N(T-K-1)]} \tag{8.3}$$

where RSS_{R1} and RSS_U are the residual sums of squares of the restricted and unrestricted models, respectively, and $(N-1)(K+1)$ and $N(T-K-1)$ are the degrees of freedom. If the calculated value of F_1 is smaller than the critical value, the null hypothesis of homogeneous slopes and intercepts across the N regions should be accepted. If the null hypothesis is accepted, this means that the data can be pooled and the panel data modelling approach is appropriate. However, if the null hypothesis is rejected, i.e., either the intercept or the slope or both are heterogeneous, we then should proceed to test for homogeneous slopes and heterogeneous intercepts with the use of a second F test:

$$F_2 = \frac{(RSS_{R2} - RSS_U)/[(N-1)(K)]}{RSS_U/[N(T-K-1)]} \tag{8.4}$$

where RSS_{R2} is the residual sum of squares for the model in which the intercept terms are allowed to vary across regions but the slope coefficients are restricted to be equal. If the calculated F_2 value is smaller than the critical value, we say that the intercepts of the models are heterogeneous but the slope coefficients are homogeneous. If this happens, we can still pool the data using the panel data approach.

If we accept this null hypothesis, it is also possible to test the null hypothesis of homogeneous intercepts conditional on homogeneous slopes by a third F test:

$$F_3 = \frac{(RSS_{R1} - RSS_{R2})/(N-1)}{RSS_{R2}/[NT-N-K)]} \tag{8.5}$$

To summarise, the decision rules based on the above statistics are given as follows.

(a) If the null hypothesis of homogeneous intercepts and slopes is accepted based on F_1, this means that the cross-sectional and time series data may be pooled and further tests are not necessary.
(b) However, if the null hypothesis in (a) is rejected, we need to perform the F_2 test. This tests whether the slope coefficients are homogeneous across regions given heterogeneous intercepts. Acceptance of this null hypothesis indicates that we can still use the panel data approach to estimate the demand model. Rejection of this null hypothesis suggests that the data cannot be pooled, and therefore the panel data approach is not appropriate.
(c) If the null hypothesis in (b) is accepted, the F_3 statistic can be used to

test for the null hypothesis of homogeneous intercepts conditional on homogeneous slopes. Acceptance of the null hypothesis indicates that the intercepts are homogeneous if the slopes are restricted to be equal.

More detail concerning these procedures is given in Hsiao (1986).

8.2.2 Fixed and random effects

There are three ways in which the estimation of a pooled tourism demand model can take place, depending on the assumptions made about the intercept term. In the pooled ordinary least squares (POLS) model the intercept is treated as a constant across all cross-section units. Using appropriate dummy variables, the fixed effects (FE) model allows the intercept to vary between cross-section units, so that each unit has a fixed intercept specific to that unit. Differences in the intercepts reflect the unobserved differences between cross-section units. The random effects (RE) model also allows the intercept to vary between units, but treats the variation as randomly determined. The determination and estimation of the FE and RE models are described below.

Each model has its advantages and disadvantages. The POLS model is straightforward to estimate, but the assumption that the unit-specific effects do not differ is very restrictive and usually unrealistic. The FE model captures the unobserved differences between units, but the inclusion of dummy variables in the OLS estimation reduces the degrees of freedom and makes the estimates less efficient than those of the RE model. On the other hand, the RE model assumes that the unobserved unit-specific effects are uncorrelated with the regressors, since these effects are randomly determined. This assumption may not be appropriate, and its violation may cause the RE model to produce biased and inconsistent estimates.

In order to explore further the nature of the FE and RE models, we assume that the error term in Equation (8.1) can be re-specified as

$$w_{it} = u_i + \varepsilon_{it} \tag{8.6}$$

where u_i denotes the unobservable cross-section-specific effects and ε_{it} is the usual disturbance term which varies across regions and time. There are two assumptions which can be made about u_i. First, u_i is correlated with $\mathbf{x}_{it}$, in which case we have the FE model. Second, u_i is uncorrelated with $\mathbf{x}_{it}$, so that we have the RE model. Some care must be exercised in choosing between these models, since they can give widely differing coefficient estimates in panel data sets where, typically, the number of time series observations is small and the number of cross-section units is large (Hausman, 1978).

Because the FE model assumes that u_i and $\mathbf{x}_{it}$ are correlated, the model must be estimated *conditionally* on the presence of u_i, which is treated as a fixed parameter to be estimated:

$$y_{it} = u_i + \mathbf{b}'\mathbf{x}_{it} + \varepsilon_{it} \tag{8.7}$$

OLS estimation of this model will yield biased estimators because of the correlation between u_i and $\mathbf{x}_{it}$. Unbiased OLS estimates can be obtained by either first-differencing the variables or, alternatively, differencing them by cross-section-specific means. In both cases the fixed effect u_i is eliminated and so is the correlation problem. A further implication is that the effect of any *observed* variable which is time invariant is also eliminated. Thus the FE model is robust to the omission of any relevant time-invariant regressors.

A priori considerations suggest that the FE model is one from which inferences can be made conditional on the observed sample. If the focus of attention is on the sample itself, then the FE model is appropriate. If one wishes to make inferences about the population from which the sample is drawn, however, then it is more appropriate to view u_i as randomly distributed across regions:

$$y_{it} = \alpha + \mathbf{b}'\mathbf{x}_{it} + u_i + \varepsilon_{it} \tag{8.8}$$

where u_i and ε_{it} are independently and identically distributed (iid) as iid $(0, \sigma_u^2)$ and iid $(0, \sigma_\varepsilon^2)$, respectively. Given that u_i is assumed to be uncorrelated with $\mathbf{x}_{it}$, OLS is asymptotically unbiased but inefficient compared with feasible generalised least squares (FGLS). A straightforward Lagrange multiplier test for the existence of random effects is based on the OLS residuals from the model in which both slope and intercept terms are assumed constant (Breusch and Pagan, 1980). The test is based on

$$\begin{aligned} &H_0: \sigma_u^2 = 0 \qquad (\text{or } \mathrm{Corr}[w_{it}, w_{is}] = 0) \\ &H_1: \sigma_u^2 \neq 0 \end{aligned}$$

where the test statistic is

$$LM = \frac{nT}{2(T-1)}\left[\frac{\sum_{i=1}^{n}\left[\sum_{t=1}^{T} e_{it}\right]^2}{\sum_{i=1}^{n}\sum_{t=1}^{T} e_{it}^2} - 1\right]^2 \tag{8.9}$$

and e_{it} is the residual from the pooled OLS regression. Under the null hypothesis LM is distributed as chi-squared with one degree of freedom.

There are a number of alternatives to the Breusch–Pagan LM test above, although it is not always clear which one should be used. Baltagi (1995) provides a detailed discussion of these tests. Alternatively, the RE model can be tested directly against the FE model. One of these tests is that devised by Hausman (1978). It is based on the idea that if u_i is uncorrelated with $\mathbf{x}_{it}$, there should be no systematic difference between the estimates from the FE and RE models, since both OLS in the FE model and FGLS in the RE model are consistent. The Hausman test is

$$H = (\hat{\mathbf{b}}_{\mathbf{fe}} - \hat{\mathbf{b}}_{\mathbf{re}})'(\hat{\mathbf{\Sigma}}_{\mathbf{fe}} - \hat{\mathbf{\Sigma}}_{\mathbf{re}})(\hat{\mathbf{b}}_{\mathbf{fe}} - \hat{\mathbf{b}}_{\mathbf{re}}) \tag{8.10}$$

where $\hat{\mathbf{\Sigma}}_{\mathbf{fe}}$ and $\hat{\mathbf{\Sigma}}_{\mathbf{re}}$ are the estimated slope covariance matrices for the FE and RE models, respectively. Under the null hypothesis that u_i is uncorrelated with $\mathbf{x}_{it}$, the Hausman test statistic is distributed asymptotically as chi-squared with K degrees of freedom. If the calculated value of $\chi^2(K)$ is greater than the critical value, this suggests that the RE model is relevant.

8.3 An empirical study of a tourism demand model based on the panel data approach

Instead of investigating the role of traditional determinants in influencing international tourism demand from one or a few origin countries to one or a few destination countries, the current section focuses on the determinants of international tourism expenditure by 138 origin countries for the period 1989–95.

8.3.1 The data and the model

The tourism expenditure variable is related to three economic explanatory variables and five 'social' variables, where the latter are included to control for changes in the demographic or socio-cultural structure of the country concerned. Specifically, we are interested in examining the following relationship:

$$TSP = f(\underset{+}{GDP}, \underset{?}{EXC}, EXV, \underset{+}{AGE}, \underset{?}{GEN}, HHS, \underset{+}{LIT}, \underset{+}{URB}) \tag{8.11}$$

where the variable definitions for country i (the tourism-generating country) are:

TSP = real per capita international tourism spending (US$) by country i, where the deflator is the price level index for country i

GDP = real per capita gross domestic product (US$) in country i, where the deflator is the price level index for country i

EXC = real exchange rate = nominal exchange rate (units of country i's currency per US$) multiplied by the relative price level (world price level/country i price level)

EXV = real exchange rate volatility = $\left|\dfrac{EXC_{it} - \mu_i}{\mu_i}\right|$ where μ_i = the arithmetic mean value of the real exchange rate for country i

AGE = the proportion of 15–64-year-olds to the total population in country i (%)

GEN = the proportion of males in the total population of country i (%)

HHS = average household size of country i (number of persons)

LIT = adult literacy rate in country i (%)

URB = the proportion of urban population to total population in country i (%)

Details of the variable sources are provided in Appendix 8.1. As in the traditional economic literature on the determination of international tourism, we include real income and real exchange rates as explanatory variables. Given the extreme fluctuations in some real exchange rates, however, we also include a measure of real exchange rate volatility as an explanatory variable. The relationship between TSP and GDP is expected to be positive, whilst that between TSP and EXV is expected to be negative because exchange rate volatility increases uncertainty and the transactions costs associated with holidays abroad. The real exchange rate variable EXC captures two effects on the price of international tourism: first, the influence of the nominal exchange rate which is defined as the number of units of country i's currency per US dollar; and second, the influence of the relative price level between country *i* and the world, where the USA price level is used as a proxy for the world price level.[1] The relationship between *TSP* (which is measured in US dollars) and the nominal exchange rate as defined above is expected to be negative. An increase in the nominal exchange rate implies a fall in the value of the currency of country *i*, making overseas tourism more expensive for country *i*'s residents, and reducing the volume of outbound tourism trips from country *i*. Assuming the dollar price of an outbound tourist trip remains unchanged, *TSP* will fall. On the other hand, the relationship between *TSP* and the relative price level can be positive or negative, depending on the price elasticity of demand for outbound tourist trips. For a given nominal exchange rate and domestic price level, an increase in the relative price of outbound tourist trips will reduce the volume of outbound tourist trips. If demand is price-inelastic, *TSP* will increase and *vice versa*. Thus the relationship between *TSP* and *EXC* depends on whether the demand for outbound tourist trips is price-inelastic and, if so, whether this positive effect on *TSP* outweighs the negative effect associated with the nominal exchange rate. The estimation results in Section 4 illustrate that this is indeed the case.

The age structure variable consists of those of working age who are likely to be income earners and more likely to consume international tourism services. The proportion of 15–64-year-olds to total population ranges from around 65% in developed countries to below 50% in some developing countries, so that it is necessary to control for the differences in demographic structure across countries in the sample. It is expected that there is a positive relationship between *TSP* and *AGE*.

The gender variable attempts to discover whether men or women have more influence on international tourism spending. In general, it does not seem possible to determine *a priori* whether the sign of the coefficient of the gender variable should be negative or positive, although there is a tendency for men to travel more than women (particularly for business travel) and one might expect a weak positive association between *GEN* and *TSP*.

The household size variable enters the model because it may affect international tourism in the following ways. First, it may be inversely related to per capita income and thus inversely related to international tourist spending. Second, it may be positively associated with tourism demand if family holidays abroad are popular or negatively associated if singles or couples holidays abroad are popular. On balance, one would expect that the first effect is greater than the second, so that overall a negative relationship between *HHS* and *TSP* is likely.

The rate of literacy in a nation reflects its level of education. As education levels increase, so too does awareness of different countries and cultures. This is likely to promote greater interest in visiting other countries, so a positive relationship between *LIT* and *TSP* is expected.

Finally, the process of urbanisation of a nation is closely related to the process of industrialisation and civilisation. One basic feature of this process, apart from the associated rise in living-standards, is the nation's closer contact with the outside world. Industrialisation both enables and provides an incentive for political, economic and cultural exchanges at an international level. Urbanisation not only confers advantages, but also creates social pressures. Living or working in cities can be an unpleasant experience, and international tourism is one way of providing an escape from the alienation and stress induced by urban living (Krippendorf, 1987, p. 17). For these reasons one would expect a positive relationship between *URB* and *TSP*.

The data set is based on editions one to four of the *World Economic Factbook* and covers 138 countries, from Afghanistan to Zimbabwe, over the period 1989 to 1995. Not all countries are included in the sample because of missing data. This problem is not confined to low-income countries, but also affects other income groups. Hong Kong and Brunei, for example, have no data on their international tourism spending. If there are significant gaps in the data for any of the variables used in this study, then the country concerned is dropped from the sample. The countries remaining in the sample constitute two-thirds of all possible countries. The issue of whether these remaining countries can be regarded as a random sample from the population is an important question which is tested statistically later.

The explanatory variables are chosen both for their expected importance in determining *TSP* and their availability. This latter criterion is particularly relevant for low-income countries, where the availability of data for other potential explanatory variables such as infant mortality can vary considerably. Including these other variables would reduce the sample size for low-income countries in particular because of the missing-data problem. For variables such as sales promotion there is almost no data available for low-income countries. There is also the problem that inclusion of more variables, even if this were feasible for all countries, would increase the possibility of multicollinearity in the estimation procedures and reduce the precision of the coefficient estimates. Finally, the inclusion of dummy variables to account for events such as currency restrictions, political disturbances, terrorism and so on, would be a daunting task for a sample of 138 countries. In any case, one of the advantages of panel data estimation

Table 8.1 Low-, medium- and high-income country groups: descriptive statistics.

Variables	Low income (Group 1) GDP < US$650 $N = 63$	Medium income (Group 2) US$650 < GDP < US$4258 $N = 36$	High income (Group 3) GDP > US$4258 $N = 39$	Overall $N = 138$
TSP	5.1	34.0	393.2	122.3
	(9.4)	(30.5)	(351.5)	(253.3)
GDP	283.2	1566.3	13163.8	4258.1
	(285.0)	(910.8)	(7275.2)	(6835.3)
EXC	172.8	106.3	73.8	127.5
	(337.1)	(281.6)	(241.6)	(301.3)
EXV	0.29	0.15	0.10	0.20
AGE	54.4	57.7	64.2	58.0
	(5.6)	(6.3)	(5.8)	(7.1)
GEN	49.8	47.1	49.7	49.1
	(1.3)	(10.8)	(1.9)	(5.8)
HHS	4.8	4.5	3.4	4.3
	(0.8)	(0.9)	(1.2)	(1.1)
LIT	61.9	77.6	91.6	74.4
	(25.1)	(18.2)	(11.5)	(23.8)
URB	37.4	47.2	73.9	50.2
	(21.1)	(17.1)	(18.4)	(24.7)

Note: The upper figure in each cell is the arithmetic mean value of the variable. The lower figure in parentheses is the standard deviation of the variable.

is that such country-specific effects can be taken into account in the estimation procedures.

The countries in the sample range from the very rich to the very poor, so that there is considerable variation in per capita gross domestic product and international tourism spending across countries. Table 8.1 divides the sample into three groups of countries with low, medium and high incomes, where the division is based on the average real per capita GDP of each country. This average is derived as the average real per capita GDP for each country over the seven years of the sample. The division is based on the OECD classification where 'low-income' (or Group 1) countries are those with a 1993 per capita income less than $650. In general, the results in Table 8.1 conform to expectations, although certain features of the economic variables are noteworthy. The variables *TSP* and *GDP* exhibit much greater sample variability, as measured by the standard deviation relative to the mean value, for Group 1 countries compared to Groups 2 and 3. For the real exchange rate variable *EXC* this situation is reversed, with Group 3 countries having the largest sample variability. Since these variables are measured in real terms, their variability arises not only from nominal changes but also from changes in the price level deflators. These changes are sometimes very large for countries with high inflation rates. In terms of exchange rate fluctuations for a given group over time, however, real exchange rate volatility (*EXV*) is greatest in Group 1 countries.

The social variables have much less sample variability. They also exhibit a lack of variability within a given country over time. This is hardly surprising, since the social structure of a given country is unlikely to change significantly over a seven-year period. It does mean, however, that some estimation procedures are not possible because of matrix near-singularity caused by almost perfect collinearity between some of the variables. Nonetheless, there is still sufficient variability across countries within a given time period to estimate the fixed and random effects models, and it is these models which form the basis of our methodological approach discussed in the next sub-section.

Our aim is to construct a general model which explains international tourism spending in terms of both economic and social variables. The model is estimated in double log form. Estimation results from the ordinary linear regression were unsatisfactory, and the double log form enables direct comparisons to be made with many other studies in the area of tourism. The estimated slope coefficients also have the convenient property of representing constant elasticities. In the context of the panel data approach, given the stochastic multiplicative demand function:

$$TSP_{it} = \alpha_i GDP_{it}^{\beta_1} EXC_{it}^{\beta_2} EXV_{it}^{\beta_3} AGE_{it}^{\beta_4} GEN_{it}^{\beta_5} HHS_{it}^{\beta_6} LIT_{it}^{\beta_7} URB_{it}^{\beta_8} e_{it} \quad (8.12)$$

where i and t represent the cross-section and time series observations, respectively, α varies across countries but is constant over time, the βs are constant across countries and over time, and e is a multiplicative error term which is iid $(1, \delta)$, then the double log form is

$$\begin{aligned}\ln TSP_{it} = {} & \ln \alpha_i + \beta_1 \ln GDP_{it} + \beta_2 \ln EXC_{it} + \beta_3 \ln EXV_{it} + \beta_4 \ln AGE_{it} \\ & + \beta_5 \ln GEN_{it} + \beta_6 \ln HHS_{it} + \beta_7 \ln LIT_{it} + \beta_8 \ln URB_{it} + \varepsilon_{it} \quad (8.13)\end{aligned}$$

where $\varepsilon_{it} = \ln e_{it}$ which is assumed iid $(0, \sigma^2)$. This assumption implies that the observations are serially uncorrelated for a given country, and that the errors have constant variance across countries and over time.

Imposing the restriction of intercept and slope coefficient homogeneity (either across countries, or over time, or both) is in effect 'pooling' or combining the time series and cross-section data. An increase in ln *GDP*, *ceteris paribus*, will have the same impact on ln *TSP* in every country and in every time period, where this impact is measured by β_1.

Table 8.2 Simple double log regression of TSP on GDP.

	Low income (Group 1)	Medium income (Group 2)	High income (Group 3)	Overall
Pooled OLS:				
β_1	1.0100	1.3232	0.9178	1.1121
t	(40.5343)	(18.5775)	(10.4267)	(80.1614)
$\overline{R}^2$	0.7887	0.5782	0.2837	0.8694
FE (no weights):				
β_1	0.9229	0.8127	0.2486	0.9115
t	(42.6308)	(20.3703)	(2.5019)	(57.1259)
$\overline{R}^2$	0.9582	0.9559	0.9102	0.9778
FE (weights):				
β_1	0.9356	0.7715	0.4786	0.9155
t	(58.9573)	(20.3719)	(9.4065)	(68.0974)
$\overline{R}^2$	0.9958	0.9964	0.9082	0.9997
RE (FGLS)				
β_1	0.9331	0.8387	0.3620	0.9495
t	(45.0085)	(20.9136)	(3.9053)	(64.0295)
$\overline{R}^2$	0.9581	0.9539	0.9077	0.9770

Notes: β_1 is the (constant) elasticity of TSP with respect to GDP.
t is the t value.
$\overline{R}^2$ is the adjusted R^2.

8.3.2 Estimating income elasticities in a simple model

Before proceeding to the full model, it may be of some value to consider the results of a simple double log regression of *TSP* on *GDP* in order to determine the (constant) elasticity of *TSP* with respect to *GDP* for each of the three groups. The FE model is estimated with and without weights, where the weighted least squares procedure[2] allows for heteroscedasticity. The results are shown in Table 8.2.

In all cases β_1 is significant and positive. Apart from the pooled OLS case, it decreases as GDP increases across the three groups. This is consistent with economic theory and the empirical results on consumption spending functions generally, where the marginal propensity to spend tends to decrease at higher levels of income. The $\overline{R}^2$ from the POLS for Groups 2 and 3 is well below that for the other estimation procedures because POLS does not take into account country-specific effects. For the overall sample, all four methods yield an elasticity of around unity. But this latter result raises a very important question: can the sample observations be pooled in this way? The next section discusses this issue, as well as the estimation results.[3]

8.3.3 *Testing for poolability*

In this section we present the test results from the F_1 and F_2 statistics only, since F_3 does not change the overall conclusion.

The F*₁ test.* Due to computational constraints, a maximum of six explanatory variables can be used in this test so it is necessary to drop a minimum of two of them from the model specification. A variety of model specifications are tested, and the variables which are consistently significant and of the expected sign are *GDP*, *EXC*, *AGE* and *URB*. Accordingly, these variables are used to derive the restricted residual sum of squares. Only *GDP* could be used to derive the unrestricted residual sum of squares, since inclusion of the other explanatory variables results in the problem of a near-singular matrix in the estimation process. The result is $F_{1_{(685,276)}} = 3.53$ which is greater than the critical F value of $F_{1_c} = 1.4$ at the 1% significance level with $K = 4$. The null hypothesis of homogeneous slopes and intercepts is therefore rejected.

The F*₂ test.* When the intercepts are allowed to vary across countries and the slope coefficients are restricted to be equal (unrestricted model), the F_2 statistic is calculated based on Equation (7.4), and the result is $F_{2_{(548,276)}} = 0.19$ which is less than the corresponding critical F value of $F_{2_c} = 1.33$ at the 1% significance level. This implies that we cannot reject the null hypothesis of homogeneous slopes and heterogeneous intercepts.

From the above statistics, we can see that the overall conclusion is that the data can be appropriately pooled in the form of a homogeneous slope and heterogeneous intercept model.

8.3.4 *Testing for fixed effects and random effects models*

After establishing the fact that the cross-sectional and time series data can be pooled, we need to decide whether the fixed effect or random effect model should be used in modelling tourism expenditure. Based on Equation (8.9), a double log regression of *TSP* on all eight explanatory variables gives LM = 1574.83, which is greater than the critical value of 6.64 at the 1% level of significance. The null hypothesis that the FE model is appropriate is clearly rejected, implying that REs are present. Based on Equation

Table 8.3 Fixed and random effects models: estimation results.

	Fixed effects (weights)		Fixed effects (no weights)		Random effects	
	(i)	(ii)	(i)	(ii)	(i)	(ii)
Coefficient:						
Constant					−11.500	−10.979
					(−5.894)*	(−7.279)*
β_1	0.826	0.830	0.852	0.857	0.896	0.890
	(43.193)*	(42.949)*	(57.235)*	(59.206)*	(50.616)*	(51.028)*
β_2	0.115	0.113	0.075	0.068	0.073	0.077
	(7.268)*	(7.089)*	(5.093)*	(4.661)*	(4.860)*	(5.223)*
β_3	0.012		0.003		−0.014	
	(0.820)		(1.056)		(−0.904)	
β_4	1.250	1.268	0.792	0.842	1.401	1.401
	(3.136)*	(3.152)*	(4.808)*	(5.161)*	(3.671)*	(3.797)*
β_5	−0.099		−1.058		−0.150	
	(0.000)		(0.000)		(−0.736)	
β_6	−0.008		−0.088		0.246	
	(−0.024)		(−0.956)		(0.948)	
β_7	−1.866		−0.756		0.138	
	(−4.843)*		(−3.143)*		(0.648)	
β_8	0.224	0.212	0.078	0.088	0.444	0.419
	(1.747)**	(1.671)**	(1.374)***	(1.771)**	(4.101)*	(4.192)*
$\overline{R}^2$	0.983	0.982	0.999	0.999	0.978	0.978
SSR	183.1	188.4	172.4	182.5	236.6	229.1
F	6698.4	15259.2	297591.4	652708.3		
DW	1.103	1.076	1.361	1.357	0.897	0.911

Notes: Figures in parentheses are t values.

*, ** and *** indicate significance at the 10%, 5% and 1% levels respectively. The t tests for β_2 and β_5 are two-tailed; the remainder are one-tailed.

SSR is the sum of the squared residuals.

(8.10), the calculated Hausman statistic is $\chi^2_{(8)} = 64.60$ which is greater than the critical value of 15.51 at the 5% significance level. This result implies that the null hypothesis of no correlation should be rejected, that is the demand model has FEs.

The result of the Breusch–Pagan test, which strongly indicates the existence of REs, is not supported by that from the Hausman test, which finds in favour of the FE model. The choice of model, however, must be guided by economic theory as well as statistical tests. Table 8.3 gives the estimation results for the FE (both with and without cross-sectional weights) and the RE model for two cases: case (i) in which all eight explanatory variables from equation (3) are included; and case (ii) where only those explanatory variables from equation (3) which tend to be significant and of the expected sign are included. These four variables are *GDP*, *EXC*, *AGE* and *URB*.

For the case (i) results, the *GDP*, *EXC*, *AGE* and *URB* variables are all significant and have the expected sign (assuming, in the case of *EXC*, that the demand for international tourism is price inelastic). The first three variables are all significant at the 1% level. In the FE models *LIT* is also highly significant but has the wrong sign. *EXV* and *HHS* are not significant at all; *EXV* is incorrectly signed in the FE models and correctly signed in the RE model, whilst the sign of *HHS* is negative and positive in the FE and RE models, respectively. *GEN* is also insignificant, although its sign is always negative, so there is a weak suggestion that the amount of international tourism spending is inversely related to the proportion of males in the country's population. Although the diagnostic statistics are somewhat better for the FE models, the RE model is more consistent with economic theory since it manages to give the expected sign for the *EXV* and *LIT* variables. All the models have extremely high $\overline{R}^2$ values.

The insignificant variables are dropped for case (ii) estimation and, as expected, the coefficient values and significance of the remaining variables do not change very much from the case (i) results. Concentrating on the case (ii) variables of *GDP*, *EXC*, *AGE* and *URB*, Table 8.3 shows that the income elasticities are all very similar in value. There is less agreement between the models for the other coefficient values. Taking the averages of the coefficient values for *GDP*, *EXC*, *AGE* and *URB* yields average elasticity values of $\beta_1 = 0.86$, $\beta_2 = 0.09$, $\beta_4 = 1.17$ and $\beta_8 = 0.24$, respectively.

The income elasticity value of 0.86 is lower than the value of around 1.8 reported in the study by Crouch and Shaw (1992), which was based on an average of 777 different estimates. It is also lower than the short- and long-run values of 2.22 and 2.37, respectively, found by Song *et al.* (2000). But the findings from these studies, in contrast to the present study, are based primarily on time series estimates. Naturally, there is no reason why one should necessarily expect to find similar values for time series and cross-section elasticities, since the two types of elasticity are conceptually different. This conceptual difference is underlined by the results from other panel data studies, where income elasticities of 1.09 (Carey, 1991), 1.18 (Tremblay, 1989), and 0.58 (Yavas and Bilgin, 1996) were obtained.[4]

Irrespective of the model specification used, the coefficient of *EXC* is always positive and significant. This implies that the positive effect on *TSP* arising from price inelasticity outweighs the negative effect from the nominal exchange rate. This hypothesis is easily verified by splitting *EXC* into its constituent parts (*NEXC* = the nominal exchange rate, and *RELP* = the relative price level) and including them as separate explanatory variables in the regression. Using the FE and RE models, the coefficient on *RELP* is always positive and highly significant, and the coefficient on *NEXC* is always negative (but not significant). These results imply that price levels in the origin and destination countries are more important than nominal exchange rates in determining international tourism spending. In terms of estimation, however, it is preferable to use the composite variable *EXC*; it is a more parsimonious specification, and use of the separate variables *NEXC* and *RELP* results in an unrealistically low income elasticity.

Own-price demand inelasticity for international tourist visits is consistent with a two-stage decision-making process, where consumers choose a holiday on the basis of their income in the first instance and then subsequently on secondary factors such as price. Given this two-stage process, consumer demand is likely to be less sensitive to changes in the secondary factors. It is also consistent with the empirical results of Song *et al.* (2000) where short- and long-run average own-price elasticities of −0.62 and −0.69, respectively, were obtained.

8.4 Summary and conclusions

This chapter introduces the application of panel data techniques to tourism demand analysis. The advantages and limitations of this approach are also highlighted. Using the panel data modelling approach, an empirical tourism demand model is developed based on both economic and social variables for a total of 138 countries over seven years. A preliminary analysis is conducted by splitting the data set by income level into three groups and estimating the income elasticity of tourism for each group on the basis of a simple regression of real per capita tourism spending on real per capita GDP. Using the fixed and random effects models, the income elasticity diminishes as per capita GDP increases; the estimated income elasticities range from around 0.9 for the low-income countries to 0.25–0.48 for the high-income countries. Pooled OLS does not follow the same pattern, since the income elasticity for medium-income countries is higher than that for low-income countries.

A key question is whether the data for the three country groups can be pooled into a single group, thereby allowing the estimation of a general set of results. Analysis shows that the data are poolable in the form of a heterogeneous intercept, homogeneous slope coefficient model. Further analysis attempts to determine whether the fixed or random effects model is appropriate, although the results of these tests are inconclusive. The variables which are consistently significant (and of the expected sign) in determining international tourism spending are real per capita GDP, the real exchange rate, age structure and urbanisation. The variable which is by far the most significant is real per capita GDP, and the pooled income elasticity of international tourism spending is 0.86. This is lower than the values which are usually reported from pure time series analyses.

The use of panel data allows the influence of social variables to be investigated. In time series data these variables do not normally display sufficient variability for estimation purposes and are omitted from the model specification. This study includes five social variables in the model specification, and concludes that age structure and urbanisation are significant determinants of international tourism spending across countries. Although it is not necessarily the case that these variables will exert the same influence in a given country over time, it seems reasonable to assume that their temporal influence cannot be neglected.

The policy and methodological implications from this study follow accordingly. Demographic as well as economic changes affect international tourism spending, and national governments and tourism agencies, whether they are in an origin or destination country, should take these changes into account. This is particularly relevant in the area of forecasting international tourism demand, where a properly specified econometric forecasting model should control for the influence of social, as well as economic, variables. Indeed, the quantitative influence of social variables on international tourism spending needs to be analysed in greater detail. The present study is based on a relatively small number of social variables, and it is possible that other measures of social change could yield additional insights. The use of the adult literacy rate in this study, for example, gives unsatisfactory results, and it could be that the use of another measure, such as the proportion of the population enrolled in further and higher education institutions, would be more appropriate.

Appendix 8.1: the data set and variable definitions

The data are compiled from the *World Economic Factbook* (WEF), editions one to four. This publication is updated regularly and provides a reasonably consistent set of observations over the sample period. Edition one (1993) is used for 1989 and 1990 data, edition two (1994/5) for 1991, edition three (1996) for 1992, 1993 and 1994, and edition four (1996/7) for 1995. The WEF collects data on international tourism spending directly from government agencies, national statistical bodies and customs and excise bodies. Where necessary, this is supplemented with data from the World Tourism Organization and the Organization for Economic Co-operation and Development. Data on gross domestic product and demographic characteristics are based mainly on national statistics, whilst data on inflation and exchange rates are derived mainly from International Monetary Fund statistics.

Not all countries are included in this study because of missing observations. In a number of cases, missing values are estimated where reliable estimates can be made and where the number of missing values is small. Over the sample period there are a number of changes to the status of certain countries, which prevented their inclusion in the sample. It was decided to retain the data for Germany and the Czech Republic in the sample, however. The West and East German data pre 1990 are combined, using per capita GDP weightings where necessary. The values for the *URB* variable, for example, cannot simply be added together and then averaged. This would imply that East German urbanisation has equal importance to that in West Germany in the determination of *TSP* in Germany, even though West German *TSP* and *GDP* are much higher than East German. Weighting by per capita GDP is one way, albeit somewhat crude, of allowing for the relative importance of East and West Germany in the combined data. Since East German *TSP* is small relative to West German *TSP*, any measurement error resulting from combining the data should also be small. The data for the Czech Republic pre 1991 are also adjusted by eliminating the data for Slovakia, where necessary by appropriate per capita GDP weightings.

The measurement of the real exchange rate is problematic. The tourism spending data are for country i to all overseas destinations, rather than for country i tourism in country j, so a measure of the *general* cost of tourism is needed. The variable definition above measures the cost of international tourism from country i as a function of the nominal exchange rate (defined in relation to US$) and the relative price level (defined in relation to the USA price level). An alternative measure of the nominal exchange rate which in theory might more accurately capture the general cost of international tourism is the effective exchange rate of country i, which would be based on a weighted average of all other countries' exchange rates. But this would only be an accurate measure of the general cost of international tourism if the weights used in the calculation of the effective exchange rate were similar to the weights of the tourist destinations in international tourism spending. For example, if the USA accounted for 20% of the effective exchange rate of Mexico, but 40% of Mexican tourist spending is in the USA, then this measure of general tourism cost would not be appropriate. Another problem is that the data for international tourism spending are for country i on all other countries in the world, but the sample includes only two-thirds of total countries. Thus, even if weightings are computed, there is an inevitable disparity between the coverage of the tourism spending data and the weights used for the sample countries.

An alternative measure of the relative price level variable is a relative price index analogous to the effective exchange rate, i.e., the weighted price level for the rest of the world relative to the price level in country i. This would be given by P_r/P_i where P_r and P_i are the price levels in the rest of the world and country i, respectively. But this measure is open to the same objection as that for the effective exchange rate; the price level weightings for the rest of the world might not reflect the tourism spending weightings. For most countries in the sample, however, P_r would be similar in value since each country would be small in relation to the rest of the world. Given that P_r can be treated as a constant across countries (but obviously not over time), a reasonable proxy for P_r is the USA price level. Differences in the levels of these two variables will be absorbed in the constant term, and will not affect the estimated slope coefficients.

Notes

1. The appropriate measurement of *EXC* is discussed further in Appendix 8.1.
2. This procedure weights the regression variables by w, where w is a series which is proportional to the reciprocals of the standard deviations of the error terms obtained from the unweighted regression.
3. A number of other tests were conducted as part of a pre-modelling data analysis, including a test of whether the model specification should be in nominal rather than real terms (or, alternatively, whether money illusion is a significant factor in the determination of tourism spending). Specifying the model in nominal terms gave poor estimation results.

4. Each of the cited studies reports more than one estimated income elasticity value. The values given above are the unweighted averages of the values reported.

9

Evaluation of forecasting performance

9.1 Introduction

So far we have explained the construction and estimation of various econometric models and their applications to tourism demand analysis. The advantages and disadvantages associated with each of these models have also been discussed. Table 9.1 summarises the applicability of the various models under different conditions.

In this chapter we compare the forecasting performance of the models based on the data set given in Chapter 5. The models compared include the three ECMs based on the Engle–Granger (1987), Wickens and Breusch (1988) and Johansen (1988) approaches, an unrestricted VAR model, an autoregressive distributed lag model (ADLM) and a TVP model. The panel data technique is not considered because of data constraints. It is useful to have a yardstick or benchmark against which to compare forecasting accuracy, and so we include the no-change model as a comparator. It will be interesting to see to what extent the recommendations in Table 9.1 are supported by the empirical results.

The organisation of this chapter is as follows. Section 9.2 discusses measures of forecasting accuracy. Section 9.3 compares the forecasting performance of different tourism demand models based on the accuracy measures discussed in 9.2, and Section 9.4 concludes the chapter.

Table 9.1 Summary of model selection criteria.

Model	When to use
Engle–Granger ECM	1. When the sample size is large (minimum 40 annual or 80 quarterly observations for each of the variables in the model). 2. When there are a maximum of two explanatory variables in the cointegration regression or if more than two only one cointegration relationship is identified. 3. When it is necessary to examine both the long-run equilibrium relationship and short-run dynamics of the tourism demand model. 4. When both policy evaluation and forecasting are of interest.
Wickens–Breusch ECM	1. When the sample size is small (less than 40 annual or 80 quarterly observations for each of the variables in the model). 2–4. Same as above.
Johansen method	1. Ideally, the sample size should be a minimum of 40 annual or 80 quarterly observations for each of the variables in the model. 2. When there are more than two explanatory variables in the model. 3–4. Same as above.
VAR model	1. When the distinction between endogenous and exogenous variables is not clear. 2. When the effects of policy 'shocks' on forecasting are of interest (impulse response analysis). 3. When the direction of causality of tourism demand variables is of interest (Grange block causality test).
TVP model	1. When the structure of the model is unstable (based on the two Chow structural instability tests and recursive OLS). 2. When the effects of policy changes on demand elasticities are of interest. 3. When only short-term forecasts are of interest.
Panel data method	1. When both time series and cross-sectional data are available. 2. When only the relationship between the tourism demand variable and its determinants is of interest, but not forecasts.
Non-causal models	1. When it is difficult to obtain data on the explanatory variables in the demand model. 2. When only forecasting, but not the causal relationship, is of interest. 3. When the forecasting performance of various models is compared (non-causal time series models are normally used as benchmarks).

9.2 Measures of forecasting accuracy

9.2.1 Ex post *and* ex ante *forecasting*

Before explaining the different measures of forecasting accuracy, two concepts related to forecasting need to be explained; *ex post* forecasting and *ex ante* forecasting. These two concepts are illustrated by Figure 9.1. t and Y in Figure 9.1 represent the time span and the time series data, respectively. The figure shows that the data are available from period 1 to N (the historical data). If a tourism demand model is estimated based on the data from t_1 to t_n (the sample data) and the forecasts from t_n to t_N are generated, these forecasts are called *ex post* forecasts. In this case, we know the values of the explanatory and dependent variables over the forecasting period, and the comparison of *ex post* forecasts between different models allows us to decide which model produces the best forecasts. In practice, practitioners and managers are more interested in the forecasts beyond time t_N, and the forecasts beyond t_N are termed *ex ante* forecasts. In this case, the values of the explanatory variables are not known and have to be determined (forecast) before the forecasts of the tourism demand variable can be obtained. Since there is no information available on the time series variables after t_N, it is impossible to evaluate the forecasting accuracy of the models. The purpose of this chapter is to examine the forecasting performance of the different models under consideration. It is possible to generate *ex ante* forecasts over the period t_n to t_N by not utilising any

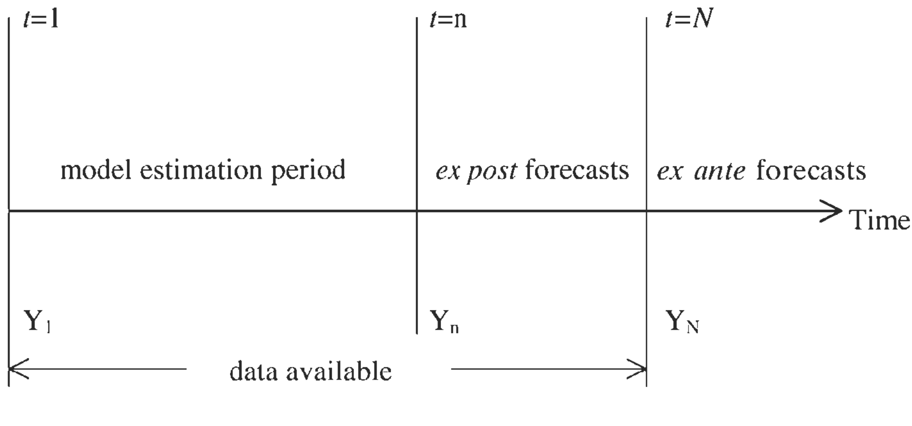

Figure 9.1 Time horizon of forecasts.

information which becomes available after period n, and then to evaluate these forecasts using the known values of the dependent variable. However, this is much more complicated than using known values of the explanatory variables, and so we generate *ex post* forecasts in this chapter.

9.2.2 The forecasting errors

The accuracy of a forecasting model depends on how close the forecast values of Y ($\hat{Y}$) are to the actual values of Y. The differences between the actual and forecast values are known as the forecasting errors, which are defined by:

$$e_t = Y_t - \hat{Y}_t \tag{9.1}$$

In theory, if the model is correctly specified, the forecasting errors are expected to be a random series with a mean of zero. Therefore, when the forecasting accuracy of a model over a certain forecasting horizon is evaluated, the sum of the forecasting errors is likely to be zero. Models which generate very poor forecasts can give very small forecasting errors simply because the positive and negative forecasting errors cancel out. To overcome this problem, the forecasting errors of Equation (9.1) are transformed either to $|e_t|$ or e_t^2. As a general rule, the smaller the $\Sigma_{t=1}^{m}|e_t|$ or $\Sigma_{t=1}^{m} e_t^2$ (where m is the length of the forecasting horizon), the better the forecasts.

Although there are a number of error measures that can be used in model evaluation (Martin and Witt, 1989; Witt and Witt, 1992), in this chapter we use only the following two measures to evaluate the forecasting performance of the tourism demand models.

The mean absolute percentage error (MAPE):

$$MAPE = \frac{\sum_{t=0}^{m} \frac{|e_t|}{Y_t}}{m} \tag{9.2}$$

The root mean square percentage error (RMSPE):

$$RMSPE = \sqrt{\frac{1}{m}\sum_{t=1}^{m}\left(\frac{e_t}{Y_t}\right)^2} \tag{9.3}$$

where m is the length of forecasting horizon, and e_t and Y_t are the forecasting error and actual value of the dependent variable, respectively.

The reasons why the MAPE and RMSPE are good measures to use in evaluating the forecasting performance of tourism demand models are explained in Witt and Witt (1992). One of them is that these measures do not depend on the magnitudes of the demand variables being predicted. In practice, we are normally interested in forecasting either outbound demand for tourism from a given origin to a number of destinations or inbound tourism demand to a given destination from several tourist-generating countries, and the magnitudes of the demand variables are likely to vary from country to country (unit to unit). The use of unit-independent measures allows us to compare forecasting accuracy not only between different models but also across countries (units).

9.3 Forecasting evaluation of different models

The forecasting performance of tourism demand models is examined based on the data set relating to the demand for Korean tourism by UK and USA residents. Seven models are specified:

(1) and (2) *The Engle–Granger and the Wickens–Breusch ECMs*
The specifications of these two ECMs are given in Chapter 5, Section 5.3.2.

(3) *The Johansen VECM*
Since two cointegrating vectors were identified for the UK model as well as for the USA model based on the Johansen approach (see Table 6.14 and Table 6.15 in Chapter 6), we can specify two VECMs for each of the two models according to the two cointegrating vectors.

(4) *An autoregressive distributed lag model*
The general ADLM is first specified based on the lagged dependent variable and the current and lagged explanatory variables (with lag length set to equal 1). The test down procedure explained in Chapter 3 is then applied to the initial ADLM. After completing the testing-down procedure, a smaller and parsimonious model is obtained, and this model is used to generate forecasts.

(5) *An unrestricted VAR model*
The unrestricted VAR models are specified in Sections 6.6.2 and 6.6.3 in Chapter 6.

(6) *A TVP model*
The TVP models are estimated in Chapter 7.

(7) *No-change model*
Forecasts from the no-change model are generated and used as benchmarks.

All these models are estimated based on data from 1962 to 1990[1] and the *ex post* forecasts are generated for the period 1991 to 1994. In order to assess the forecasting performance of the models, we first generate four one-year-ahead forecasts and three two-years-ahead forecasts using recursive forecasting techniques. We then generate the overall static forecasts based on the estimated models using the data just from 1962 to 1990 (a combination of one-year-ahead to four-years-ahead forecasts). The MAPE and RMSPE are then calculated for each of these forecasting horizons. Tables 9.2–9.4 present the empirical results.

The following conclusions may be drawn from the forecasting error measures in Tables 9.2–9.4. First, in terms of one-year-ahead forecasting, the UK RMSPE and the USA MAPE and RMSPE show that the TVP model outperforms all other competing models. The no-change model is ranked first for the UK in terms of MAPE and second in terms of RMSPE. For the USA the no-change model is ranked third (MAPE and RMSPE). The worst forecasting model for the UK is the VAR model while for the USA it is the Engle–Granger ECM. The superior performance of the TVP model for one-year-ahead forecasting is consistent with the results obtained by Song *et al.* (1999).

Second, the two forecasting error measures show that the no-change model is ranked best for two-years-ahead forecasts in the USA case while the best model in the UK case is the Johansen VECM1. The TVP model is in second place for the USA (MAPE and RMSPE) and in third place for

Table 9.2 One-year-ahead forecasting error measures.

	UK		USA	
The model	MAPE	RMSPE	MAPE	RMSPE
Engle–Granger ECM	4.775	6.711	4.068	4.978
Wickens–Breusch ECM	5.052	7.766	3.353	3.861
Johansen VECM1	3.599	5.917	2.054	2.103
Johansen VECM2	5.384	7.501	3.449	4.157
ADLM	4.726	7.067	3.560	4.174
VAR	11.590	14.443	3.725	4.191
TVP	3.661	4.688	1.422	1.530
No change	3.459	4.761	2.188	2.271

Table 9.3 Two-years-ahead forecasting error measures.

	UK		USA	
The model	MAPE	RMSPE	MAPE	RMSPE
Engle–Granger ECM	1.983	2.934	4.310	5.452
Wickens–Breusch ECM	2.591	3.172	3.535	4.541
Johansen VECM1	0.964	1.318	2.180	2.268
Johansen VECM2	2.584	3.100	3.382	4.396
ADLM	2.591	3.240	3.750	4.908
VAR	8.126	9.989	3.918	4.509
TVP	2.095	2.441	1.467	1.598
No change	1.412	1.740	0.759	0.968

Table 9.4 One-to-four-years-ahead forecasting error measures.

	UK		USA	
The model	MAPE	RMSPE	MAPE	RMSPE
Engle–Granger ECM	3.253	4.906	5.018	5.929
Wickens–Breusch ECM	3.566	5.668	4.358	5.075
Johansen VECM1	2.206	3.963	5.724	11.841
Johansen VECM2	3.379	5.234	4.052	4.979
ADLM	2.761	4.939	4.569	5.542
VAR	8.554	11.624	4.203	4.647
TVP	2.437	3.377	1.494	1.668
No change	2.533	3.645	1.570	1.836

the UK according to the RMSPE, but fourth place in terms of MAPE. The no-change model is ranked second for the UK (MAPE and RMSPE).

Third, the error measures with respect to the overall (one-year-ahead to four-years-ahead) forecasting performance show clearly that the TVP model is the best, followed by the no-change model, for the USA (MAPE and RMSPE). For the UK the RMSPE suggests that the TVP model is the best and no-change the second best. Although the UK MAPE indicates that the ranking order is the Johansen VECM1, followed by the TVP model and the no-change model, we can still conclude that the TVP model is the most accurate forecasting model overall among the alternative forecasting techniques.

Fourth, the generally superior forecasting performance of the no-change model compared with the ECMs is consistent with the findings of Song *et al.* (1999) and Kulendran and Witt (2001), although other studies have shown that ECMs generate more accurate forecasts than 'no change' (Kim and Song, 1998; Song *et al.*, 2000).

9.4 Concluding remarks

The overall superiority of the TVP model in forecasting suggests that it is important to take structural instability into consideration when generating tourism demand forecasts. The ability of the TVP model to generate the most accurate short-term (one-year-ahead) forecasts is consistent with prior expectations (see Table 9.1).

With regard to the choice of ECM, the small sample size (29–31 annual observations used in estimation) suggests the use of the WB model in preference to the EG model, but the empirical results show that the WB ECM outperforms the EG ECM in only 50% of cases for one-, two- and one-to-four-years-ahead forecasts. However, since two cointegration relationships were identified for both the USA and UK models, we would expect the Johansen method to generate the most accurate ECM forecasts (see Table 9.1). This is supported by the empirical results in Tables 9.2–9.4; the Johansen VECM1 outperforms the EG and WB models in all four cases (UK/USA, MAPE/RMSPE) for one-year-ahead and two-years-ahead forecasts, and in both UK cases for one-to-four-years-ahead forecasts.

The no-change model performs well, and this is in agreement with several previous empirical studies. Of the 12 cases considered in Tables 9.2–9.4 (two origin countries, two error measures, three forecasting horizons), the no-change model is ranked first or second (out of eight models) in 75% of cases, and is ranked first, second or third in 100% of cases.

The relatively poor forecasting performance of the VAR model (ranked seventh or eighth in 75% of cases) suggests that the distinction between endogenous and exogenous variables in tourism demand forecasting models is fairly clear, or at least that allowing for the distinction not to be clear does not result in more accurate forecasts (see Table 9.1).

Although econometric models (such as ECMs) may not necessarily generate more accurate forecasts than simple non-causal models (e.g. the no-change model), the popularity of econometric models is still evident. In particular, these models allow us to examine the short-run and long-run causal relationships between tourism demand and its determinants, and to evaluate government policies related to tourism.

Note

1. The estimation results are omitted.

References

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, AC-19, 716–23.

Artus, J.R. (1972). An Econometric Analysis of International Travel. *IMF Staff Papers*, 19, 579–614.

Baltagi, B.H. (1995). *Econometric Analysis of Panel Data*. Chichester: Wiley.

BarOn, R. (1989). *Travel and Tourism Data*. Cambridge: Euromonitor Publications Ltd.

Bewley, R. (1979). The Direct Estimation of the Equilibrium Response in a Linear Dynamic Model. *Economics Letters*, 55, 251–76.

Bohara, A.K. and Sauer, C. (1992). Competing Macro-Hypotheses in the United States: A Kalman Filter Approach. *Applied Economics*, 24, 389–99.

Bomhoff, E.J. (1994). *Financial Forecasting for Business and Economics*, London: Dryden Press.

Box, G. and Jenkins, G. (1976). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Breusch, T. (1978). Testing for Autocorrelation in Dynamic Linear Models. *Journal of Australian Economic Papers*, 17, 334–55.

Breusch, T. and Pagan, A. (1980). The LM Test and its Applications to Model Specification in Econometrics. *Review of Economic Studies*, 47, 239–54.

Brown, J., Song, H. and McGillivray, A. (1997). Forecasting UK House Prices: A Time Varying Coefficient Approach. *Economic Modelling*, 14, 529–48.

Burmeister, E. and Wall, K.D. (1982). Kalman Filtering Estimation of Unobserved Rational Expectations with an Application to the German Hyperinflation. *Journal of Econometrics*, 20, 255–84.

Carey, K. (1991). Estimation of Caribbean Tourism Demand: Issues in Measurement and Methodology. *Atlantic Economic Journal*, 19, 32–40.

Chow, G. (1960). Test of Equality between Sets of Coefficients in Two Linear Regressions. *Econometrica*, 28, 591–605.

Cochrane, E. and Orcutt, G.H. (1949). Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms. *Journal of the American Statistical Association*, 44, 32–61.

Crouch, G. and Shaw, R.N. (1992). International Tourism Demand: A Meta-Analytical Integration of Research Findings, in P. Johnson and B. Thomas (eds), *Choice and Demand in Tourism*. London: Mansell.

Crouch, G.I., Schultz, L. and Valerio, P. (1992). Marketing International Tourism to Australia: A Regression Analysis. *Tourism Management*, 13, 196–208.

Cuthbertson, K. (1988). Expectations, Learning and the Kalman Filter. *The Manchester School*, Vol. LVI, 223–46.

Davidson, J., Hendry, D.F., Saba, F. and Yeo, S. (1978). Econometric Modelling of the Aggregate Time Series Relationships between Consumers Expenditure and Income in the United Kingdom. *Economic Journal*, 88, 661–92.

Davidson, R. and MacKinnon, J.G. (1981). Several Tests for Model Specification in the Presence of Alternative Hypotheses. *Econometrica*, 49, 781–93.

Deaton, A.S. and Muellbauer, J. (1980). Econometric and Consumer Behaviour. Cambridge: Cambridge University Press.

Dickey, D.A. and Fuller, W.A. (1979). Distribution of the Estimates for Autoregressive Time Series with Unit Roots. *Journal of the American Statistical Association*, 74, 427–31.

Dickey, D.A. and Fuller, W.A. (1981). Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica*, 49, 1057–72.

Dickey, D.A., Hasza, D.P. and Fuller, W.A. (1984). Testing for Unit Roots in Seasonal Time Series. *Journal of the American Statistical Association*, 79, 355–67.

Doan, T., Littleman, R.B. and Sims, C.A. (1984). Forecasting and Conditional Projection Using Realistic Prior Distributions. *Econometric Review*, 3, 1–100.

Edwards, A. (1985). *International Tourism Forecasts to 1995: EIU Special Report No. 188*. London: Economist Publications.

Engle, R.F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation. *Econometrica*, 50, 987–1007.

Engle, R.F. and Granger, C.W.J. (1987). Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica*, 55, 251–76.

Fuller, W.A. (1976). *Introduction to Statistical Time Series*. New York: Wiley.

Geweke, J. (1984). Inference and Causality in Economic Time Series Models, in Z. Griliches and M.D. Intriligator (eds), *Handbook of Econometrics*, Vol. 2, Amsterdam: North Holland.

Godfrey, L.G. (1978). Testing for Higher Order Serial Correlation in Regression Equations when the Regressors Contain Lagged Dependent Variables. *Econometrica*, 46, 1303–10.

Granger, C.W.J. (1969). Investigating Causal Relationships by Econometric Models and Cross Spectral Methods. *Econometrica*, 37, 424–38.

Gray, H.P. (1966). The Demand for International Travel by the United States and Canada. *International Economic Review*, 7, 83 92.

Greenslade, J.V. and Hall, S.G. (1996). Modelling Economies Subject to Structural Change: The Case of Germany. *Economic Modelling*, 13, 545–59.

Hall, R.E. (1978). Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Evidence. *Journal of Political Economy*, 86, 971–87.

Hamilton, J.D. (1994). *Time Series Analysis*, Princeton, New Jersey: Princeton University Press.

Harvey, A.C. (1987). Applications of the Kalman Filter in Econometrics, in T.F. Bewley (eds), *Advances in Econometrics — Fifth World Congress*, Vol. 1, pp. 285–313, Cambridge: Cambridge University Press.

Harvey, A.C. (1993). *Time Series Models*, 2nd Edition. London: Harvester Wheatsheaf.

Hausman, J.A. (1978). Specification Tests in Econometrics. *Econometrica*, 46, 251–72.

Hendry, D.F. (1995). *Dynamic Econometrics: Advanced Text in Econometrics*. Oxford: Oxford University Press.

Hendry, D.F. and von Ungern-Sternberg, T. (1981). Liquidity and Inflation Effects on Consumers Expenditure, in A.S. Deaton (ed.), *Essays in the Theory and Measurement of Consumer Behaviour*, Cambridge: Cambridge University Press.

Hsiao, C. (1986). *Analysis of Panel Data*, Cambridge: Cambridge University Press.

Hylleberg, S., Engle, R.F., Granger, C.W.J. and Yoo, B.S. (1990). Seasonal Integration and Cointegration. *Journal of Econometrics*, 44, 215–18.

Jarque, C.M and Bera, A.K. (1980). Efficient Tests for Normality, Homoskedasticity and Serial Independence of Regression Residuals. *Economic Letters*, 6, 255–9.

Johansen, S. (1988). A Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, 12, 231–54.

Johansen, S. (1995). *Likelihood Based Inference on Cointegration in the Vector Autoregressive Model.* Oxford: Oxford University Press.

Johansen, S. and Juselius, K. (1990). Maximum Likelihood Estimation and Inference on Cointegration with Application to Demand for Money. *Oxford Bulletin of Economics and Statistics*, 52, 169–210.

Kalman, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions ASME Journal of Basic Engineering*, D82, 35–45.

Kim, C.J. (1993). Sources of Monetary Growth, Uncertainty and Economic Activity: The Time Varying Parameter Model with Heteroskedastic Disturbances. *The Review of Economics and Statistics*, 74, 483–92.

Kim, S. and Song, H. (1998). An Empirical Analysis of Demand for Korean Tourism: A Cointegration and Error Correction Approach. *Tourism Analysis*, 3, 25–41.

Kliman, M.L. (1981). A Quantitative Analysis of Canadian Overseas Tourism. *Transportation Research*, 15, 487–97.

Krippendorf, J. (1987). *The Holiday Makers: Understanding the Impact of Leisure and Travel.* Oxford: Butterworth-Heinemann.

Kulendran, N. (1996). Modelling Quarterly Tourism Flows to Australia Using Cointegration Analysis. *Tourism Economics*, 2 , 203–22.

Kulendran, N. and King, M.L. (1997). Forecasting International Quarterly Tourist Flows Using Error Correction and Time Series Models. *International Journal of Forecasting*, 13, 319–27.

Kulendran, N. and Witt, S.F. (2001). Cointegration Versus Least Squares Regression. *Annals of Tourism Research*, 28.

Kwack, S.Y. (1972). Effects of Income and Prices on Travel Spending Abroad 1960III–1967IV. *International Economic Review*, 13, 245–56.

Lee, C.K., Var, T. and Blaine, T.W. (1996), Determinants of Inbound Tourism Expenditure. *Annals of Tourism Research*, 23, 527–42.

Lim, C. (1997). An Econometric Classification and Review of International Tourism Demand Models. *Tourism Economics*, 3, 69–81.

Little, J.S. (1980). International Travel in the US Balance of Payments. *New England Economic Review* (May/June) 42–55.

Loeb, P. (1982). International Travel to the United States: An Economic Evaluation. *Annals of Tourism Research*, 9, 5–20.

MacKinnon, J.G. (1991). Critical Values for Cointegration Tests. In R.F. Engle and C.W. Granger (eds), *Long Run Equilibrium Relationships*. Oxford: Oxford University Press.

Mak, J., Moncur, J. and Yonamine, D. (1977). Determinants of Visitor Expenditures and Visitor Lengths of Stay: A Cross-section Analysis of U.S. Visitors to Hawaii. *Journal of Travel Research*, 15, 5–8.

Martin, C.A. and Witt, S.F. (1987). Tourism Demand Forecasting Models: Choice of Appropriate Variable to Represent Tourists' Cost of Living. *Tourism Management*, 8, 233–46.

Martin, C.A. and Witt, S.F. (1988). Substitute Prices in Models of Tourism Demand. *Annals of Tourism Research*, 15, 255–68.

Martin, C.A. and Witt, S.F. (1989). Forecasting Tourism Demand: A Comparison of the Accuracy of Several Quantitative Methods. *International Journal of Forecasting*, 5, 7–19.

Mizon, G.E. and Richard, J.F. (1986). The Encompassing Principle and its Application to Testing Non-nested Hypotheses. *Econometrica*, 54, 657–78.

O'Hagan, J.W. and Harrison, M.J. (1984). Market Shares of US Tourist Expenditure in Ireland: Some Econometric Findings. *Applied Economics*, 16, 919–31.

Osborn, D.R., Chui, A.P.L., Smith, J.P. and Birchenhall, C.R. (1988). Seasonality and the Order of Integration for Consumption. *Oxford Bulletin of Economics and Statistics*, 50, 361–77.

Osterwald-Lenum, M. (1992). A Note with Quantiles of the Asymptotic Distribution of the ML Cointegration Rank Test Statistics. *Oxford Bulletin of Economics and Statistics*, 54, 461–72.

Pesaran, M.H. and Shin, Y. (1995). An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis. In S. Strom, A. Holly and P. Diamond (eds), *Centennial Volume of Rangar Frisch*, Econometric Society Monograph. Cambridge: Cambridge University Press.

Phillips, C.P.B. and Perron, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrica*, 75, 335–46.

Quandt, R.E. and Young, K.H. (1969). Cross-sectional Travel Demand Models: Estimation and Tests. *Journal of Regional Science*, 9, 201–14.

Ramsey, J.B. (1969). Test for Specification Errors in Classical Linear Least Squares Regression Analysis. *Journal of the Royal Statistical Society*, Series B, 31, 350–71.

Riddington, G.L. (1999). Forecasting Ski Demand: Comparing Learning Curve and Varying Parameter Coefficient Approaches. *Journal of Forecasting*, 18, 205–14.

Romilly, P., Liu, X. and Song, H. (1998a). Economic and Social Determinants of International Tourism Spending: A Panel Data Analysis. *Tourism Analysis*, 3, 3–15.

Romilly, P., Song, H. and Liu, X. (1998b). Modelling and Forecasting Car Ownership in Britain: Cointegration and General to Specific Approach. *Journal of Transport Economics and Policy*, 32, No. 2, 165–85.

Sargan, J.D. (1964). Wages and Prices in the United Kingdom: A Study of Econometric Methodology. In P.E. Hart and J.K. Whittaker (eds), *Econometric Analysis for National Economic Planning*, London: Butterworth.

Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, 6, 461–4.

Sims, C. (1972). Money, Income and Causality. *American Economic Review*, 62, 540–52.

Sims, C. (1980). Macroeconomics and Reality. *Econometrica*, 48, 1–48.

Smeral, E., Witt, S.F. and Witt, C.A. (1992). Econometric Forecasts: Tourism Trends to 2000. *Annals of Tourism Research*, 19, 450–66.

Song, H. (1995). A Time Varying Parameter Consumption Model for the UK. *Applied Economics Letters*, 2, 339–42.

Song, H., Liu, X. and Romilly, P. (1997). A Time Varying Parameter Approach to the Chinese Aggregate Consumption Function. *Economics of Planning*, 29, 185–203.

Song, H., Romilly, P. and Liu, X. (1998). The UK Consumption Function and Structural Instability: Improving Forecasting Performance Using a Time Varying Parameter Approach. *Applied Economics*, 30, 975–83.

Song, H., Romilly, P. and Liu, X. (2000). An Empirical Study of UK Outbound Tourism Demand. *Applied Economics*, 32, 611–24.

Song, H., Witt, S.F. and Jensen, T. (1999). Forecasting Performance of Tourism Demand Models: The Case of Denmark. Paper presented at the 19th International Symposium on Forecasting, Washington, DC.

Stewart, J. and Gill, L. (1998). *Econometrics*, pp. 129–33, London: Prentice Hall.

Stock, J. and Watson, M. (1988). Testing for Common Trends. *Journal of the American Statistical Association*, 83, 1097–107.

Stone, J.R.N. (1954). Linear Expenditure Systems and Demand Analysis – An Application to the Pattern of British Demand. *Economic Journal*, 64, 511–27.

Swamy, P.A.V.B., Kennickell, A.B. and von zur Muechien, P. (1990). Comparing Forecasts from Fixed and Variable Coefficient Models: The Case of Money Demand. *International Journal of Forecasting*, 6, 469–77.

Syriopoulos, T.C. and Sinclair, M.T. (1993). An Econometric Study of Tourism Demand: the AIDS Model of US and European Tourism in Mediterranean Countries. *Applied Economics*, 25, 1541–52.

Theil, H. (1965). The Information Approach to Demand Analysis. *Econometrica*, 33, 67–87.

Thomas, R.L. (1993). *Introductory Econometrics: Theory and Applications*. London: Longman.

Thomas, R.L. (1997). *Modern Econometrics: An Introduction*. Harlow: Addison-Wesley.

Tremblay, P. (1989). Polling International Tourism in Western Europe. *Annals of Tourism Research*, 16, 677–91.

Turner, L., Kulendran, N. and Fernando, H. (1997). The Use of Composite National Indicators for Tourism Forecasting. *Tourism Economics*, 3, 309–17.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test of Heteroskedasticity. *Econometrica*, 48, 817–38.

Wickens, M.R. and Breusch, T.S. (1988). Dynamic Specification, the Long-Run and the Estimation of Transformed Regression Models. *Economic Journal*, 98 (Conference): 189–205.

Witt, S.F. (1980a). An Abstract Mode-Abstract (Destination) Node Model of Foreign Holiday Demand. *Applied Economics*, 12, 163–80.

Witt, S.F. (1980b). An Econometric Comparison of UK and German Foreign Holiday Behaviour. *Managerial and Decision Economics*, 1, 123–31.

Witt, S.F. and Martin, C.A. (1987a). International Tourism Demand Models: Inclusion of Marketing Variables. *Tourism Management*, 8, 33–40.

Witt, S.F. and Martin, C.A. (1987b). Measuring the Impacts of Mega-events on Tourism Flows. *The Role and Impact of Mega-Events and Attractions on Regional and National Tourism Development: Proceedings of International Association of Scientific Experts in Tourism, 37th Congress.* St Gallen, Switzerland: AIEST, 213–21.

Witt, S.F. and Witt, C.A. (1992). *Modeling and Forecasting Demand in Tourism*. London: Academic Press.

Witt, S.F. and Witt, C.A. (1995). Forecasting Tourism Demand: A Review of Empirical Research. *International Journal of Forecasting*, 11, 447–75.

World Economic Factbook (edns 1–4), Cambridge: Euromonitor Publishing Ltd.

Wu, D.M. (1973). Alternative Tests of Independence Between Stochastic Regressors and Disturbances. *Econometrica*, 41, 733–50.

Yavas, B.F. and Bilgin, Z. (1996). Estimation of Tourism Demand in Turkey: A Pooled Cross Section and Time Series Analysis. *Tourism Analysis*, 1, 19–27.

Indices

Author Index

Akaike, H., 23
Artus, J.R., 29

Baltagi, B.H., 142
BarOn, R., 19
Bera, A.K., 37
Bewley, R., 79, 87
Bilgin, Z., 19, 137, 151
Birchenhall, C.R., 64
Blaine, T.W., 10, 31, 81
Bohara, A.K., 132
Bomhoff, E.J., 130
Box, G., 30
Breusch, T., 35, 77, 141, 157
Brown, J., 124, 131
Burmeister, E., 124

Carey, K., 19, 137, 151
Chow, G., 38, 125
Chui, A.P.L., 64
Cochrane, E., 15
Crouch, G.I., 7, 151
Cuthbertson, K., 124

Davidson, J., 28
Davidson, R., 41
Deaton, A.S., 16
Dickey, D.A., 57, 59, 60, 62, 63, 64, 68
Doan, T., 101

Edwards, A., 9
Engle, R.F., 30, 37, 55, 56, 64, 68, 70, 73, 75, 76, 157

Fernando, H., 31
Fuller, W.A., 57, 59, 60, 62, 63, 64, 68

Geweke, J., 101
Gill, L., 51
Godfrey, L.G., 35
Granger, C.W.J., 30, 55, 56, 64, 68, 70, 73, 75, 76, 99, 157
Gray, H.P., 29
Greenslade, J.V., 132

Hall, R.E., 32, 132
Hamilton, J.D., 112
Harrison, M.J., 16
Harvey, A.C., 129, 130
Hasza, D.P., 64
Hausman, J.A., 40, 141, 142
Hendry, D.F., 28, 29, 31
Hsiao, C., 140
Hylleberg, S., 64

Jarque, C.M., 37
Jenkins, G., 30
Jensen, T., 124, 131, 162, 164
Johansen, S., 88, 92, 112, 114, 115, 157
Juselius, K., 114

Kalman, R.E., 129
Kennickell, A.B., 124
Kim, C.J., 132
Kim, S., 30, 32, 111, 115, 164
King, M.L., 30, 32, 92
Kliman, M.L., 5, 31
Krippendorf, J., 145
Kulendran, N., 30, 31, 32, 92, 164
Kwack, S.Y., 29

Lee, C.K., 10, 31, 81
Lim, C., 20, 137
Little, J.S., 30
Littleman, R.B., 101
Liu, X., 6, 19, 30, 32, 116, 124, 131, 151, 152, 164
Loeb, P., 29

McGillivray, A., 124, 131
MacKinnon, J.G., 41, 69, 70
Mak, J., 18
Martin, C.A., 5, 7, 8, 26, 31, 160
Mizon, G.E., 28, 41
Moncur, J., 18
Muellbauer, J., 16

O'Hagan, J.W., 16
Orcutt, G.H., 15
Osborn, D.R., 64
Osterwald-Lenum, M., 116

Pagan, A., 141
Perron, P., 61
Pesaran, M.H., 77, 79, 86
Phillips, C.P.B., 61

Quandt, R.E., 18

Ramsey, J.B., 37
Richard, J.F., 28, 41
Riddington, G.L., 124
Romilly, P., 6, 19, 30, 32, 116, 124, 131, 151, 152, 164

Saba, F., 28
Sargan, J.D., 28
Sauer, C., 132
Schultz, L., 7
Schwarz, G., 23

Shaw, R.N., 151
Shin, Y., 77, 79, 86
Sims, C., 91, 100, 101
Sinclair, M.T., 16
Smeral, E., 9, 16
Smith, J.P., 64
Song, H., 6, 19, 30, 32, 111, 115, 116, 124, 131, 151, 152, 162, 164
Stewart, J., 51
Stock, J., 76, 112
Stone, J.R.N., 16
Swamy, P.A.V.B., 124
Syriopoulos, T.C., 16

Theil, H., 16
Thomas, R.L., 34, 41, 42
Tremblay, P., 19, 137, 138, 151
Turner, L., 31

Valerio, P., 7
Var, T., 10, 31, 81
von Ungern-Sternberg, T., 28
von zur Muechien, P., 124

Wall, K.D., 124
Watson, M., 112
White, H., 36
Wickens, M.R., 77, 157
Witt, C.A., 2, 9, 10, 16, 26, 28, 30, 123, 137, 160, 161
Witt, S.F., 2, 5, 7, 8, 9, 10, 16, 26, 28, 30, 31, 32, 92, 123, 124, 131, 137, 160, 161, 162, 164
Wu, D.M., 40

Yavas, B.F., 19, 137, 151
Yeo, S., 28
Yonamine, D., 18
Yoo, B.S., 64
Young, K.H., 18

Subject Index

ADF regression, 65–7, 70
ADF statistic, 66, 71
ADF test, 59–61, 71
  critical values, 69
ADLM method, 28–9, 33, 34, 48, 77–9, 85–8, 121, 161
Akaike information criterion (AIC), 61, 66, 86
Alternative hypothesis, 25
ARIMA, 30
Augmented Dickey–Fuller test, *see* ADF test
Autoregressive conditional heteroscedasticity (ARCH) test, 37
Autoregressive distributed lag model, *see* ADLM
Autoregressive model, 30
  testing, 46

Best linear unbiased estimates (BLUE), 22
Block Granger causality test, 100, 103–4
Breakpoint test, 38–9
Breusch–Godfrey LM test, 45
Breusch–Pagan test, 150

Chow parameter constancy test (structural instability test), 38–9, 125, 126
Chow predictive failure test, 39, 45
Civilisation, 145
Coefficient of determination, 22
Cointegrated vectors in Johansen estimation, 120
Cointegration, 53–71
  concept, 55
  defined, 56
  and error-correction mechanism (ECM), 73–5
  relationship, 55, 57, 68, 71, 87–8, 116–17
  test for, 68–71, 79–81
  and VAR modelling, 111–22
  see also Johansen cointegration; Long-run cointegration
Cointegration Durbin–Watson (CIDW), 70–1, 80
Cointegration regression, 57
Common factor (*COMFAC*) model, testing for, 48
Consumer price index (CPI), 4, 53, 54, 79
Consumer tastes, 6
Cost of living, 4, 5
Cost of travel, 4
Country-specific effects, 146
Cross-sectional data, 18

Data collection, 18–20
Data-generating process (DGP), 28, 93, 124
Data sources, 19–20
Dead start model, 32, 49, 50
  testing for, 48
Decision-making process, 152
Demand elasticities, 8–13, 134
Demand function for tourism product, 2–3
Dependent variables, 3, 22
Destination preference index, 6
Destinations, 12–13
DF statistic, 63, 67
DF test, 59–61, 80, 112, 113, 117
  critical values, 62, 69
  see also Augmented Dickey–Fuller (ADF) test
Diagnostic checking, 34–42
Diagnostic statistics, 82–3
Dickey–Fuller test, *see* DF test
Diffusion priors, 130, 135
Disequilibrium error, 56
Dummy variables, 8, 145
Durbin–Watson (DW) statistic, 34–5, 70

ECMs, 32, 34, 73–89
  and cointegration, 73–5
  estimation, 76–9, 81–8
  forecasting errors, 88
  forecasting performance, 88
  worked examples, 79–88
Effective exchange rate, 155
Encompassing test, 41–2, 50
Engel curve, 16
Engle–Granger two-stage approach, 76–7, 81–5, 111, 121, 158, 161
Engle test for autoregressive conditional heteroscedasticity, 45
Error correction models, *see* ECMs
Error variance decomposition, 97–8
Estimated general model, 45
Ex ante forecasting, 159
Exchange rate, 5
Exchange rate index, 44
Exchange rate volatility (EXV), 147
Exogeneity, 98–101
  testing for, 40–1
Expectations, 7–8
Explanatory variables, 3–8, 25, 36, 40
Ex post forecasting, 159, 162
Ex post forecasting errors, VECMs, 122
Ex post forecasting performance, 108–11

F statistic, 25, 33, 46, 47, 64, 66, 67
F test, 33, 34, 41, 45, 66, 138–9, 149
Feasible generalised least squares (FGLS), 141
Finite distributed lag model, 32
  testing for, 48
Fixed effects models, 140–2, 149–52
Force of equilibrium, 55
Forecasting accuracy measures, 159–61
Forecasting errors, 97–8, 160–1
  ex post, VECMs, 122
Forecasting performance, *ex post*, 108–11
Forecasting performance evaluation, 157–65
  comparison of models, 161–4
Forecasting variance decomposition, 105–8
Forecasts, time horizon of, 159
Functional form, 8–13

General-to-specific modelling, 27–51
  methodology, 84
  model selection, 42–3
  worked example, 43–51
Goldfeld–Quandt test, 35–6
Granger causality
  concept, 98–9
  tests for, 99–101
  see also Block Granger causality test
Gross domestic product (GDP), 15, 79, 80, 106, 132, 143, 146–8, 152–4
Growth rate model, 30
  testing for, 46

Habit persistence, 7–8
Hannan–Quinn criterion (HQC), 86
Hausman statistic, 150
HEGY seasonal unit root test, 64
Heteroscedasticity
  testing for, 35–7
  White test for, 45
Hypothesis formulation, 16–17
Hypothesis testing, 23–5

Identical and independently distributed (IID) regression, 61
Impact parameter, 74
Impulse response analysis, 94–7, 105–8
Impulse response functions, 96–7
Impulse responses to SE shock, 109–10
Income, 4
Income elasticities, 12, 148, 151
Independent variables, 9
Inflation, 5
Information criteria, 23
Instrumental variables, 40
Integrated process, 56
Integration order, test for, 57–68
International tourism
  demand, 3
  expenditure, 142

J-test, 41
Jarque–Bera (J–B) test, 37, 45
Johansen cointegration methodology, 112, 158, 161
  empirical example, 116–22

Johansen cointegration tests, 117–19
Johansen estimation, cointegrated vectors in, 120
Joint hypothesis, 62
Joint null hypothesis, 63
Joint significance, 24–5

Kalman filter, 128–34

Lagged dependent variables, 7, 30
Lagrange Multiplier (LM) test, 35, 141
Leading indicator model, 31
  testing for, 47
Likelihood ratio statistic, 93
Linear relationship, 9
Long-run cointegration coefficients, 75
Long-run cointegration regression, 75
Long-run cointegration vectors, 121
Long-run equilibrium regression, 56

Marginal revenue (*MR*), 12
Marketing, 6–7
Maximum log-likelihood value, 103
Mean absolute forecasting error (MAE), 88
Mean absolute percentage error (MAPE), 160–1
Mean square forecasting error (MSE), 130
Medium-income country groups, 146
Mis-specification, testing for, 37–8
Model estimation, 20–3
Model selection criteria, 42, 158
Model specification, 17–18
Model structure instability, 123
Monte Carlo simulations, 64, 70
Multiple regression model, 36, 40

No-change model, 162
Nominal exchange rate (NEXC), 152
Non-causal models, 158
Non-cointegration, 69–70
Non-stationary time series, 56, 57
  and unit roots, 59
Normal equations, 22
Normality, testing for, 37
Null hypothesis, 35, 39, 40, 42, 45, 61, 63, 69–70, 80, 139, 140, 142, 149

OLS, 27, 33, 38, 40, 46, 49, 53, 56, 57, 65, 70, 78, 93, 108, 117, 124–5, 130–1, 141
Omitted variables/functional form, Ramsey test for, 45
One-tailed statistic, 24
Ordinary least squares (OLS), 15, 20–2
Outbound tourism demand model, 27
Own price, 4–5
Own-price elasticity, 10, 11

Panel data analysis, 19, 137–71, 158
  data set and variable definitions, 154–5
  empirical study, 142–52
  model specification and estimation, 138–42
Partial adjustment model, 31, 49, 50
Partial adjustment process, testing for, 47
Personal disposable income (PDI), 44
Phillips–Perron tests, 61–2, 80, 117
Point elasticity, 9
Policy evaluations, 25–6
Poolability test, 138–40, 149
Pooled data, 19
Pooled ordinary least squares (POLS) model, 140, 148
Population, 3–4
Power function, 10–11
Power relationship, 9
Predictive failure, 123
Price elasticity of demand, 11
Price measures, 4

Qualitative effects, 8

Ramsey RESET statistic, 45, 47
Ramsey RESET test, 37–8
  for omitted variables/functional form, 45
Random effects, 140–2
Random effects models, 149–52
Random walk, 128–9, 131
Recursive least squares procedure, 39–40
Recursive OLS estimates, 126–8

Regression parameters, restriction tests on, 33–4
Relative price index, 155
RESET test, *see* Ramsey RESET test
Residual sum of squares (*RSS*), 139
Restriction tests, 45
  on regression parameters, 33–4
Root mean square forecasting error (RMSE), 88
Root mean square percentage error (RMSPE), 160–1

Schwarz Bayesian criterion (SBC), 61, 86
Seasonal time series, 63
Seasonal unit roots, 63–5
Seemingly unrelated regression estimator (SURE), 93
Short-run error correction model, 83
Specific-to-general approach, 28
Spurious correlation, 53
Spurious regression, 53–5
Stability condition, 95
Standard error, 24
Static model, 29, 45
Stationary time series, 57
Structural instability, testing for, 38–40
Substitute destinations, 13
Substitute prices, 5–6
Sum of squared residuals (SSR), 21, 22, 33, 67

t statistic, 24, 61, 64, 70
Time horizon of forecasts, 159
Time series, 57
Time series data, 18
Time varying parameter modelling, *see* TVP model
Total revenue (*TR*), 12
Tourism demand
  analysis, 1–13
  determinants, 2
  elasticities, constant over time, 124–8
  forecasting, 25–6
  international, 3
  modelling
    failure of traditional approach, 26
    traditional methodology, 15–27
Tourism expenditure, 142
Tourism product, demand function for, 2–3
TVP model, 121, 123–35, 158, 161
  state space (SS) form, 128–9
  worked example, 131–4
Two-tailed statistic, 24

Unit root tests
  examples, 65–8
  procedure, 62–3
Unit roots, 59
  DF and ADF tests for, 59–61
Unrestricted VAR model, 161
Urbanisation, 145

VECM, 114, 117, 118, 121–2, 161
  ex post forecast errors, 122
Vector autoregressive (VAR) analysis, 92
Vector autoregressive (VAR) forecasting, 97–8
Vector autoregressive (VAR) modelling, 40, 91–122, 158
  choice criteria, 102
  and cointegration, 111–22
  estimation, 104–5
  example, 101–11
  framework, 56
  order selection, 101–2
Vector moving average (VMA), 95

Wald statistic, 45, 48
Wald test, 41
Weighted least squares (WLS), 18
Weighted price level, 155
White test, 36–7
  for heteroscedasticity, 45
Wickens–Breusch one-stage approach, 77, 83–5, 121, 158, 161
World Economic Factbook (WEF), 154
Wu–Hausman statistic, 50–1
Wu–Hausman test, 50