Identifying Long-run Cointegrating Relations: An Application to the Hawaii Tourism Model

By
Allison Zhou, Carl Bonham and Byron Gangnes

October 11, 2004

University of Hawai'i
Economic Research Organization
2424 Maile Way, Room 542
Honolulu, Hawai'i 96822
www.uhero.hawaii.edu

University of Hawaii Economic Research Organization

Identifying Long-run Cointegrating Relations: An Application to the Hawaii Tourism Model

Allison Zhou, Carl Bonham and Byron Gangnes

University of Hawaii Economic Research Organization
University of Hawaii at Manoa

October 11, 2004

Abstract Cointegration analysis has gradually appeared in the empirical tourism literature. However, the focus has been exclusively on the demand side, neglecting supply influences and risking endogeneity bias. One reason for this may be the difficulty identifying structural relationships in a system setting. We estimate a demand-supply model of Hawaii tourism using a theory-directed sequential reduction methodology suggested by Hall, Henry, and Greenslade (2002). The resulting model illustrates the viability of such an approach as well its challenges.

Keywords: Cointegration, Identification, Weak exogeneity, Systems of equations, Tourism demand and supply analysis, Hawaii.

JEL: C320, L830

1 Introduction

The extraordinary growth in international tourism since World War II has spurred considerable empirical research on modelling and forecasting tourism flows. The bulk of these studies estimate tourism demand equations to explain either flows from various source markets into a particular tourism destination, or the allocation of outbound travel to alternative destinations. The overwhelming majority of extant studies use traditional econometric methods and ignore possible supply-side influences.[1]

The development of system-based cointegration methods permits more satisfactory analysis of tourism behavior where (as is usually the case) relevant time series are non-stationary, and where endogeneity among variables is expected. Integration and cointegration analysis can avoid the problem of spurious regressions among non-stationary series, while a system approach allows for important supply-demand interactions.

System methods of course introduce additional challenges, in particular the problem of identification of individual structural relationships. In a system with cointegrating rank r, Pesaran and Shin (2001) show that exact identification requires r restrictions in each of the r cointegrating vectors. The popular Johansen (1988, 1991, 1995) method uses a statistical approach to achieve the needed restrictions. Pesaran and Smith (1998) and Pesaran and Shin (2001) criticize this approach as a pure mathematical convenience, and instead have advocated a theory-based approach. Hall, Henry, and Greenslade (2002) argue that the different identification methods proposed in the literature are almost impossible to implement in practice due to the limited sample size available for most empirical research. They suggest imposing theory-based weak exogeneity assumptions at the earliest stage of the model reduction process.

In this paper, we estimate a demand-supply model for Hawaii tourism using a system-based cointegration approach. Hawaii is a particularly apt case for such analysis, because tourists from two markets—the mainland United States and Japan—represent a dominant 85% of the total market. Clearly in this case, demand parameters can not be estimated reliably without regard to supply constraints and potential price responses. And of course knowledge of supply behavior is of interest in its own right.

Following Hall, Henry, and Greenslade (2002), we use a pragmatic theory-based approach in identifying long-run cointegrating vectors, making use of exogeneity assumptions that are tested and imposed at an early stage of the model reduction process to increase the chance of identifying the "true" equilibrium relationships that govern tourism activity in Hawaii. This approach has not previously been applied to tourism modeling.[2]

The organization of the paper is as follows. Section 2 derives the tourism demand and supply equations, and identifies the variables to be used in the modeling exercise. Section 3 outlines our estimation methodology. Section 4 presents the empirical results of the

[1] For a review, refer to Lim (1997), Crouch (1994a, 1994b), and Witt and Witt (1992, 1995).

[2] A limited number of studies using cointegration methods exist, but some researchers still impose single-equation approach developed in the late 1980s with little or no mention of potential endogeneity problems. (See Kim and Song (1998), Vogt and Wittayakorn (1998), and Song, Romilly, and Liu (2000).) Other researchers have recently begun to adopt the system approach (e.g., Kulendran (1996), Lathiras and Siriopoulos (1998) Gangnes and Bonham (1998), and Song and Witt (2000), but the identification scheme is exclusively Johansen's reduced rank regression technique, despite the fact that alternative theory-based identification methods may be superior. Little consideration of supply side influences exists.

Hawaii tourism model. Section 5 evaluates the forecast performance of the model. Section 6 concludes.

2 A Demand-Supply Model of Tourism

There are a relatively limited number of theoretical studies of tourism economics and no well-established conceptual framework. Some early perspectives are reflected in Quandt (1970) and Gray (1970). More recently, Bull (1991) and Sinclair and Stabler (1997) attempt to give textbook overviews of tourism theory. Besides these, a few optimization-based models have been constructed (Copeland, 1989, 1990; Morely, 1992; and Taylor, 1995). While theoretical work is relatively sparse, empirical literature on tourism is substantial.

2.1 Tourism Demand

Empirical models of tourism demand borrow heavily from consumer theory (Varian, 1992) which predicts that the optimal consumption level depends on the consumer's income, the price of the good in question, the prices of related goods (substitutes and complements), and other demand shifters. Formally, the Marshallian demand for tourism product can be expressed as,

$$D_{ij} = F(Y_i, P_i, P_j, P_j^S, \mathbf{Z}); \tag{1}$$

where D_{ij} is the tourism product demanded in destination j by consumers from origin country i; Y_i is the income of origin country i; P_i is the price of other goods and services in the origin country i; P_j is the price of tourism product in destination country j; P_j^S is the price of tourism product in competing destinations; and $\mathbf{Z}$ is the vector of other factors affecting tourism demand. When homogeneity is assumed, demand can be expressed as a function of income in constant domestic prices and destination and substitute prices in relative terms,

$$D_{ij} = F(\frac{Y_i}{P_i}, \frac{P_j}{P_i}, \frac{P_j^S}{P_i}, \mathbf{Z}). \tag{2}$$

In the literature, there are at least two classes of tourism models, those explaining the distribution of outward flows from a single source market (outbound modeling) and those explaining aggregate tourism flows into a single destination (inbound modeling). For outbound modeling, market shares of visitors or expenditures are the typical dependent variables. For inbound modeling, the most appropriate measure is real expenditures on tourism-related goods and services. However, the unavailability and perceived poor quality of expenditure data confine the typical study to total visitor arrivals (Anastasopoulos, 1984; O'Hagan and Harrison, 1984). Of the 85 tourism studies reviewed in Crouch (1994a), 63% choose the number of visitor arrivals as the measure of demand while 48% use expenditure and receipts.

Proxies for the demand determinants vary considerably. Typical income measures include the gross domestic product (GDP), gross national product (GNP), national disposable income (NDI), personal income (PI) and consumption expenditure (CE), measured in either real, nominal, aggregate, or per capita form, depending on data availability and nature of tourism demand modelled. Generally speaking, PI and CE are used to model leisure and holiday travels, while GDP, GNP, and NDI are used to model business travel. As for the choice between nominal and real incomes, equations (1) and (2) make it clear that both are acceptable, provided that prices are specified accordingly. The per capita income specification is justified by Witt and Witt (1995) as a solution to the multicollinearity problem when both income and population are used to measure market size. Nevertheless, the inclusion of population as a separate variable distinct from aggregate income is itself questionable (Gangnes and Bonham, 1998).

Two types of prices appear in the demand specification. The first is the *own price* of tourism products, normally approximated by the consumer price index (CPI) in the destination country. The practice is sometimes criticized on the ground that, "the cost of living for local residents does not always reflect the cost of living for foreign visitors to that destination, especially in poor countries" (Song and Witt, 2000). So occasionally tourism-specific prices or indices are employed.[3] Martin and Witt (1987) report that tourism-specific indices do not perform any better than overall indices such as the CPI. Edwards (1995) prefers CPI because the coverage of the tourism-specific price is suspect (Witt and Witt, 1992).

The second is the *substitute price*. On the one hand, domestic travel is often found to be the strongest substitute for foreign travel, justifying the inclusion of a price index for other domestic goods and services, such as the CPI in the county of origin. On the other hand, competition among different overseas destinations may call for the inclusion of variables that represent the cost of substitute destinations. In the existing literature, most models employ an *exchange rate adjusted relative CPI* (or *real exchange rate*) to capture the substitution between domestic vacations and overseas holiday travel. For substitution among different overseas destinations, a number of studies include *real exchange rates* from a number of competing countries, while others use a *weighted real exchange rate* to capture the general effect (for examples of the former, see Kim and Song, 1998 and Song, Romilly and Liu, 2000; for the latter, see Vogt and Wittayakorn, 1998).[4]

Transportation cost is potentially another important factor in determining international travel. Song and Witt (2000) suggest using, "representative air fares between origin and destination for air travel," to approximate the true travel cost,[5] but Gangnes and Bonham (1998) reject such practice on the ground that, "frequent discounting and package trips" implies a significantly lower actual price than published fares. They recommend Edwards (1995) measure of International Air Transport Association (IATA) data on revenues per passenger ton/km. From a statistical perspective, the transportation cost variable may enter either separately or be merged with living costs to form a comprehensive cost measurement. The latter has the advantage of reducing the demands on limited data sets, but parameter

[3]For instance, Gangnes and Bonham (1998) use the hotel room price.

[4]The exchange rate is logically among the most influential factors in determining international travel. This effect is usually captured by converting destination prices into the currency of the tourism importing country. However, Lathiras and Siriopoulos (1998) and Vogt and Wittayakorn (1998) insist that exchange rates, either as a single bilateral rate or a composite index, must be considered separately because tourists respond very differently to changes in price levels and exchange rate fluctuations (which, some argue, is more obvious than destination cost-of-living fluctuations). Research finds that *exchange rate adjusted CPI*, either alone or together with a separate exchange rate variable, is a good proxy for tourism cost, while exchange rate itself is not (Martin and Witt, 1987).

[5]This is already put into practice by Fujii and Mak (1985) and Crouch (1991).

estimates may be more difficult to interpret.

Apart from the variables listed above, many studies also include a time trend (linear or nonlinear as in Witt and Witt, 1992) to capture evolving consumer tastes; a constant term to account for "utility image" that does not vary greatly with time; and dummies to account for various once-off events such as the Olympic Games, large-scale fairs, foreign currency/travel restrictions and oil crises. These types of events, if otherwise neglected, might lead to bias in the estimated parameters (Anastasopoulos, 1984; Crouch, Schultz and Valerio, 1992; Kliman, 1981; Mak, Moncur and Yonamine, 1977). In other cases, dummy variables are used to account for switches in data sources, inconsistency in recording methods, and seasonality.

For our Hawaii tourism model, we use the number of visitor arrivals as the dependent variable because high frequency expenditure data is unavailable. We seek to identify demand relationship for each of the two primary Hawaii tourism markets, U.S. mainland and Japanese visitors, as tourists from the two markets have consistently accounted for over 85% of all visitors to Hawaii during the last decade.

To keep the model size manageable, we are forced to choose only the principle determinants of tourism demand while leaving out influences that are deemed less central to behavior. In addition, some conceptually relevant factors are excluded because of difficult finding appropriate proxies. The model includes five demand determinants. They are the U.S. real personal income (nir_us), U.S. consumer price index (cpi_us), Japanese real personal income (nir_jp), Japanese *exchange rate adjusted* CPI (p_jp) and Hawaii average daily hotel room price (prm_hi). In a log linear form, the demand relations are:

$$vus\_hi = \alpha_0 + \alpha_1 * nir\_us + \alpha_2 * cpi\_us + \alpha_3 * prm\_hi + e_{us}; \tag{3}$$

$$vjp\_hi = \beta_0 + \beta_1 * nir\_jp + \beta_2 * p\_jp + \beta_3 * prm\_hi + e_{jp}; \tag{4}$$

The model variables are listed in Table 1. All series are at quarterly frequency and expressed as natural logarithms. The time series e_{us} and e_{jp} are regression residuals and are assumed to capture any unexplained movements in demand. For U.S. and Japanese visitor arrivals, continuous historical series were created from two periods (pre- and post-1989) over which the State changed reporting methods. Details are available from the authors upon request.

2.2 Tourism Supply

Both theoretical and empirical research on the supply of tourism services is scant (Crouch, 1994a). In much of the empirical tourism literature, supply is assumed to be perfectly elastic, and parameters of demand relations are estimated by Ordinary Least Squares (OLS). However, the infinite elasticity assumption is a convenient simplification rather than a tested hypothesis. Fujii, Khaled, and Mak (1985) estimate the supply elasticity of Hawaii lodging services to be close to two and it is not uncommon to observe sizable fluctuations in hotel room prices. The treatment of supply relations is therefore indispensable in deriving unbiased demand elasticities, and supply behavior is of interest in its own right.

However, it is rather difficult to give a precise definition of tourism supply considering the variety of products tourists consume. The paper chooses hotel accommodations as a proxy because lodging service is the largest single product category in overall tourists expenditures

Table 1: Summary of Variables in the Hawaii Tourism Model

Mnemonic	Description	Units	Source
Hawaii Variables			
vus_hi	U.S. visitors to Hawaii	000s	DBEDT; authors' calc.
vjp_hi	Japanese visitors to Hawaii	000s	DBEDT; authors' calc.
prm_hi	Hawaii average daily hotel room rate	dollar	DBEDT
ocup_hi	Hawaii average daily hotel occupancy rate	%	DBEDT
trms_hi	Hawaii visitor plant inventory	000s	DBEDT
U.S. Variables			
nir_us	U.S. real personal income	bil 82-84$	BEA
cpi_us	U.S. CPI (1982-1984=100)	index	BLS
Japan Variables			
nir_jp	Japan real personal income	bil 95Yen	ESRI
cpi_jp	Japan CPI (1995=100)	index	SBSC
yxr_jp	yen/dollar exchange rate	yen/dollar	FED
Calculated Variables			
p_jp	*cpi_jp/yxr_jp*	–	Authors' calc.

DBEDT: Department of Business Economic Development and Tourism, State of Hawaii.
BEA: Bureau of Economic Analysis, U.S.
BLS: Bureau of Labor Statistics, U.S.
FED: Federal Reserve Bank at St. Louis.
ESRI: Economic and Social Research Institute, Japan.
SBSC: Statistics Bureau and Statistics Center, Japan.

in Hawaii and a measure for which their is reliable data.[6] Hotel products are nonstorable in nature. A hotel room not rented for a given day is lost forever as a potential source of revenue. This, together with heavy operating costs, results in a strong incentive for profit maximizing hotel owners to maintain high occupancy rates. In the short run, this leads hoteliers to price discriminate and offer off-peak discounts to fill rooms. Over longer horizons, capacity is adjusted through expansion and contraction of room inventory.

One approach to modeling hotel supply it to estimate an inverted tourism supply curve. Several examples appear in the hotel room tax literature (Fujii, Khaled and Mak, 1985; Bonham and Gangnes, 1996). The supply price of hotel rooms is assumed to be a mark-up over marginal cost,

$$P_H = markup \cdot MC = M \cdot H(Q_H, P_L, P_K, P_Z; K_H); \tag{5}$$

where Q_H is the total quantity supplied, i.e., the number of rooms rented; P_L, P_K and P_Z are the input prices of labor, capital and other inputs; M is the markup factor; and K_H is the short-run given hotel room supply. Assuming that marginal cost is homogeneous of degree one in input prices, equation (5) is equivalently written as,

$$\frac{P_H}{P_K} = M \cdot H(Q_H, \frac{P_L}{P_K}, \frac{P_Z}{P_K}; K_H). \tag{6}$$

Applying this to the Hawaii tourism model, we are looking for a supply relation of the form,

$$prm\_hi = L(Q^{Rented}, P^{Cost}, trms\_hi); \tag{7}$$

where Q^{Rented} is the total rooms rented and P^{Cost} measures the overall production cost. Since an exact number of hotel rooms occupied is not known, proxies are found in visitor arrivals to the islands (the sum of U.S. and Japanese tourists) and the hotel occupancy rate. As for production cost, it would be ideal to include Hawaii specific operation cost measurements, such as a Hawaii producer price index, but such measures do not exist. Considering the model size (9 variables) and limited data set (86 observations), we have opted to use the U.S. consumer price index (*cpi_us*) as a proxy for cost effects. Clearly, this is at best a partial measure of the true lodging input costs. In a linear form, the supply relation then is,

$$\begin{aligned} prm\_hi = {} & \gamma_0 + \gamma_1 * (vus\_hi + vjp\_hi) + \gamma_2 * ocup\_hi \\ & + \gamma_3 * cpi\_us + \gamma_4 * trms\_hi + e_{prm}. \end{aligned} \tag{8}$$

3 Empirical Methodology

We formulate the Hawaii tourism model as a cointegrated system. To identify structural relationshiops, we apply the sequential model reduction strategy advocated by Hall, Henry, and Greenslade (2002). This approach makes use of theory-guided weak exogeneity assumptions to increase the chance of discovering the true underlying behavior.

[6] Visitors to Hawaii spend an average of 33% of total expenditures on hotel lodging services for the past three decades.

3.1 Econometric Framework

Consider a VAR(k) model in an $m \times 1$ vector of I(1) variables, z_t,

$$z_t = \Phi_1 z_{t-1} + \cdots + \Phi_k z_{t-k} + c + \epsilon_t \ , \quad t = 1, 2, \ldots, T, \tag{9}$$

where c is an $m \times 1$ vector of unknown deterministic terms; Φ_i, $i = 1, 2, \ldots, k$, are $m \times m$ matrices of unknown parameters; ϵ_t is an $m \times 1$ vector of disturbances that is i.i.d.$(0, \Omega)$ and the initial values, $z_0, z_{-1}, \ldots, z_{-k+1}$ are given.[7]

The model specified in (9) can be reparameterized as a Vector Error Correction Model (VECM),

$$\Delta z_t = -\Pi z_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta z_{t-i} + c + \epsilon_t \ , \quad t = 1, 2, \ldots, T, \tag{10}$$

where $\Pi = I_n - \sum_{i=1}^{k} \Phi_i$, $\Gamma_i = -\sum_{j=i+1}^{k} \Phi_j$, $i = 1, \ldots, k-1$. The equilibrium properties of (10) are characterized by the rank of Π. If all elements of z_t are stationary, Π is a full rank $m \times m$ matrix. If the elements of z_t are I(1) but not cointegrated, Π is of rank zero and a VAR model in first differences is appropriate. If the elements of z_t are I(1) and cointegrated with $rank(\Pi) = r$, Π can be decomposed into two $m \times r$ full column rank matrices α and β where $\Pi = \alpha\beta'$. This implies that there are $r < m$ stationary linear combinations of z_t, $\xi_t = \beta' z_t$. The matrix of adjustment coefficients, α, measures how strongly deviations from the long-run equilibrium, ξ_t, feedback onto the system. The Johansen reduced rank regression technique involves maximizing the log likelihood function of equation (10), subject to the constraint that Π can be decomposed into two $m \times r$ full column rank matrices α and β such that $\Pi = \alpha\beta'$.

An identification problem arises because matrices α and β are not uniquely identified without additional information. To see this, note that for any $r \times r$ non-singular matrix Q we can define matrices $\alpha^* = \alpha Q$ and $\beta^{*'} = Q^{-1}\beta'$ such that $\Pi = \alpha^* \beta^{*'} = \alpha Q Q^{-1} \beta' = \alpha\beta'$. Pesaran and Shin (2001) show that r^2 restrictions are needed for exact identification. The restrictions must be evenly distributed across the cointegrating vectors, i.e., there must be r restrictions per vector.

The most common approach to imposing the r^2 indentifying restrictions is Johansen's statistical approach. Specifically, Johansen's just identified estimator of β is obtained by selecting the r largest eigenvectors of the system, subject to "ortho-normalization" and "orthogonalization" restrictions. This approach has been criticized, as a "pure mathematical convenience" (Pesaran and Shin 2001), rather than an economically justified approach.[8]

Recent developments in cointegration analysis emphasize the use of economic theory in guiding the search for long-run exact/over identification restrictions (Pesaran and Shin, 2001). The theory-guided approach takes Johansen's just identified vector β_J as given and replaces the "statistical" restrictions with ones that are economically meaningful. Typically the approach imposes exclusion and normalization restrictions to exactly identify the system and then uses χ^2 statistics to test over identifying restrictions. To illustrate, the Hawaii

[7] We assume that the roots of $| \ I_n - \Phi_1\lambda - \Phi_2\lambda^2 - \cdots - \Phi_k\lambda^k \ | = 0$ lie either on or outside of the unit circle, but rule out the possibility that one or more elements of z_t are I(2). A review of the econometric analysis of I(2) variables is provided in Haldrup (1998).

[8] Another non-theoretical method of indentification is Phillips' (1991, 1995) triangularization approach.

tourism model has nine variables $z_t = (vus\_hi,\ vjp\_hi,\ prm\_hi,\ ocup\_hi,\ trms\_hi,\ nir\_us,\ cpi\_us,\ nir\_jp,\ p\_jp)$ (see Table 1 for variable definitions). Tourism demand and supply theories suggest the existence of three long-run cointegrating vectors,

$$vus\_hi = \alpha_0 + \alpha_1 * nir\_us + \alpha_2 * cpi\_us + \alpha_3 * prm\_hi + e_{us}, \tag{11}$$

$$vjp\_hi = \beta_0 + \beta_1 * nir\_jp + \beta_2 * p\_jp + \beta_3 * prm\_hi + e_{jp}, \tag{12}$$

$$\begin{aligned} prm\_hi = \gamma_0 + \gamma_1 * (vus\_hi + vjp\_hi) + \gamma_2 * ocup\_hi \\ + \gamma_3 * cpi\_us + \gamma_4 * trms\_hi + e_{prm}. \end{aligned} \tag{13}$$

One set of exact identifying restrictions may: 1) exclude vjp_hi, nir_jp from the U.S. demand relation and normalize on vus_hi; 2) exclude vus_hi, nir_us from the Japan demand relation and normalize on vjp_hi; and 3) exclude nir_us, nir_jp from the Hawaii supply relation and normalize on prm_hi. Starting from the exactly identified system, over-identifying restrictions can be tested using χ^2 statistics either individually or in a group. Details on the tests performed are found in section 4.

An additional problem with the VAR system in (9) is that it has a prohibitively large number of parameters; each equation involves estimating mk lag coefficients plus one or more parameters for the deterministic components. Even moderate values of m and k quickly exhaust typical samples for macroeconometric research. For example, if all nine variables are treated as endogenous with a lag of four, each equation in the Hawaii tourism model in section 2 involves estimating 38 parameters and the system as a whole has 342 regression coefficients. With a sample size of 86 (1980Q1–2001Q2), the VAR approach quickly runs into the problem of severe lack of degrees of freedom. In-sample regression produces perfect fit, but out-of-sample forecasts are generally poor.

One way to address the over-parameterization problem is to test and impose weak exogeneity assumptions. For each series treated as weakly exogenous, the number of equations in the system is reduced by one and the number of parameters by $(mk + d)$, d being the number of deterministic components. For the Hawaii tourism model, if the external drivers (nir_us, nir_jp, cpi_us, p_jp) are treated as weakly exogenous, the number of equations is reduced from nine to five and parameters to be estimated from 342 to 190.

Partition the m-vector of I(1) random variables z_t into the n-vector y_t and the q-vector x_t such that $z_t = (y_t{}', x_t{}')'$ and $q = m - n$. The primary interest is the structural modeling of y_t conditional on its own past values, y_{t-1}, y_{t-2}, ..., and the current and past values of x_t. The parameters, matrices and the error terms in the VECM equation (10) can be partitioned conformably as $c = (c_y', c_x')', \alpha = (\alpha_y', \alpha_x')', \Gamma_i = (\Gamma_{yi}', \Gamma_{xi}')', i = 1, 2, \ldots, k-1, \epsilon_t = (\epsilon_{yt}', \epsilon_{xt}')'$ and the variance-covariance matrix as

$$\Omega = \begin{pmatrix} \Omega_{yy} & \Omega_{yx} \\ \Omega_{xy} & \Omega_{xx} \end{pmatrix}.$$

The model is transformed into a conditional model for y_t,

$$\Delta y_t = c_y - \omega c_x + \omega \Delta x_t + (\alpha_y - \omega\alpha_x)\beta' z_{t-1} + \sum_{i=1}^{k-1}(\Gamma_{yi} - \omega\Gamma_{xi})\Delta z_{t-i} + \epsilon_{yt} - \omega\epsilon_{xt}, \quad (14)$$

and a marginal model for x_t,

$$\Delta x_t = c_x + \alpha_x \beta' z_{t-1} + \sum_{i=1}^{k-1}\Gamma_{xi}\Delta z_{t-i} + \epsilon_{xt}, \quad (15)$$

where $\omega = \Omega_{yx}\Omega_{xx}^{-1}$.

For the system in equation (14) and (15), the cointegrating relations $\beta' z_{t-1}$ enter both the conditional and the marginal model, and the new adjustment coefficients $(\alpha_y - \omega\alpha_x)$ depend on the covariance matrix, Ω, and all the adjustment coefficients (α_y, α_x). In general parameters in the marginal and the conditional system are interrelated and a full system analysis is required.

However, in the case where x_t is weakly exogenous with respect to β, the conditional model (14) contains as much information about the cointegrating relations, $\beta' z_{t-1}$, as the full system so that analysis of the conditional model alone is efficient.[9] If the parameters of interest are the cointegrating vector β, x_t is weakly exogenous if and only if $\alpha_x = 0$ (Johansen 1991). This condition ensures that β does not appear in the marginal distribution for x_t (see equation (15)) , and that α_x does not appear in the conditional model (see equation (14)). While the condition $\alpha_x = 0$ is a necessary and sufficient condition for weak exogeneity of x_t with respect to β, this condition often proves to be too strong in practice because exogenous variables may form cointegrating relationships among themselves (Pesaran, Shin, and Smith, 2000) . Harbo, Johansen, Nielsen, and Rahbek (1998) propose an alternative weak exogeneity test. Instead of estimating the whole system and testing whether a subset of α is zero, they suggest estimating the conditional model alone and checking for weak exogeneity by adding the empirically derived cointegrating relations to the marginal model. Their approach avoids the common failure of weak exogeneity due to the cointegration among weakly exogenous variables.

3.2 Sequential System Reduction

The system in equation (14) is the general VECM with exogenous I(1) variables. It is rarely the final form, however, as significant simplification takes place to reduce the system to a most parsimonious representation. Four types of restrictions are relevant in the process,

1. Restrictions on the rank of long-run matrix (Π);
2. Restrictions on the short-run dynamic coefficients (Γ_i's);
3. Restrictions on the long-run cointegrating vectors, β;
4. Restrictions on the loading parameters, α.

[9]Two conditions must be satisfied for x_t to be weakly exogenous. 1) The parameters of interest are functions of the parameters in the conditional model alone. 2) The parameters in the conditional model and the parameters in the marginal model are variation-free; that is, they do not have any joint restrictions.

Researchers have proposed different ways to impose these restrictions (Johansen 1988, 1991, 1995; Phillips 1991, 1995; Saikkonen 1993a, 1993b; Pesaran and Shin 2001). Hall, Henry, and Greenslade (2002) argue that these approaches are almost impossible to implement successfully in a realistic situation when a fairly rich specification encounters limited sample size. The interaction of dynamic and long-run parameters has enormous effects on the size and power of the statistical tests conventionally adopted. Monte Carlo results of Pesaran and Shin (2001) reveal that imposing valid weak exogeneity restrictions before testing for the cointegrating rank generally improves the power of Johansen rank tests. Nevertheless, restricting the cointegrating rank has little impact on weak exogeneity tests, at least as long as the rank is not restricted to be less than the true rank.

We follow Hall, Henry, and Greenslade (2002) and apply the following pragmatic strategy in reducing our general VECM to a more parsimonious representation,

1. Make (and wherever possible test) weak exogeneity assumptions about the model;
2. Test the cointegrating rank;
3. Use Johansen's reduced rank procedure to estimate the cointegrating vectors. These vectors enter the VECM in an unrestricted fashion;
4. Follow Pesaran and Shin (2001) to impose long-run structural just (over) identifying restrictions on the β vector.
5. Estimate the complete dynamic model and simplify the dynamics. At this stage, the causality structure of the model can be established by eliminating unnecessary cointegrating vectors from an equation using likelihood ratio tests.

4 A Complete Hawaii Tourism Model.

The Hawaii tourism system under consideration has nine variables z_t=(*vus_hi*, *vjp_hi*, *prm_hi*, *ocup_hi*, *trms_hi*, *nir_us*, *cpi_us*, *nir_jp*, *p_jp*), among which y_t=(*vus_hi*, *vjp_hi*, *prm_hi*, *ocup_hi*, *trms_hi*) are endogenous and x_t=(*nir_us*, *cpi_us*, *nir_jp*, *p_jp*) are exogenous according to tourism demand and supply theories. The organization is as follows: Section 4.1 determines the order of integration of the series involved; Section 4.2 tests and imposes weak exogeneity restrictions; Section 4.3 establishes the rank of cointegration conditional on imposed weak exogeneity; Section 4.4 discusses the long-run cointegrating relations for the system; Section 4.5 reduces the model to the most parsimonious representation.

4.1 Unit Root Tests

In the literature, the most commonly used methodology to establish the order of integration of a series is to look for unit roots in the autoregressive processes using *Dickey-Fuller* (DF) and *augmented Dickey-Fuller* (ADF) tests (Dickey and Fuller, 1979, 1981). It has also been recognized that DF and ADF tests are sensitive to whether an intercept and/or a time trend is included. In addition, Schwert (1987) shows that unit root tests derived from pure autoregressive processes have different sampling distributions when the true process is a mixed autoregressive-integrated moving average (ARIMA) process. When the moving average parameter is close to 1, DF and ADF tests have true critical values that are far greater than the standard Dickey-Fuller distributions tabulated in Fuller (1976). That is, the DF and ADF tests tend to reject the null hypothesis of unit root too frequently.

The paper performs both the ADF and Schwert tests for unit root. Table 2 lists the standard DF and ADF(5) test statistics when a constant and a trend are included in the specification. From the table, the null of unit root cannot be rejected for all variables in levels. When variables are first differenced, ADF(5) test statistics reject the null of unit root for all except Hawaii hotel room price, U.S. and Japan real income series. Japanese price is borderline significant.

Schwert (1987) test specifies the lag length in ADF type tests using the rules of thumb $l_4 = [4*(N/100)^{1/4}]$ and $l_{12} = [12*(N/100)^{1/4}]$ with N being the total number of observations. For the Hawaii tourism model under discussion, $N = 86$, $l_4 = 4$ and $l_{12} = 12$. Table 3 reports the Schwert(l_4) and Schwert(l_{12}) test statistics both with and without a time trend. Readings from the table do not give consistent conclusions on the order of integration. When there is no time trend, the unit root hypothesis is rejected for all variables except Δnir_jp at 5% significance level with 4 lags in the specification. However, when the lag length is increased to 12, the unit root hypothesis is rejected only for Δp_jp. Similar conclusion yields when a time trend is included. From the *AR(1-5)* LM test, 4 lags seem sufficient for all variables under consideration. We therefore accept Schwert(l_4) test results and treat all variables except Δnir_jp as I(1) series.

A closer look at Δnir_jp yields that non-rejection of the unit root hypothesis is due to a slowdown in the growth of Japanese real income during the 1990s. Japanese real income growth averaged 1.2% during the 1980s, a number only to be followed by zero growth in the whole 1990s. Perron (1989, 1990) argue that a structural change in the mean of a stationary variable tend to bias the usual unit root tests towards non-rejection of the null of unit root. We therefore perform the Perron (1990) test for unit root on Δnir_jp, picking the break point at 1991Q1 because the Japanese economy peaked in February 1991 according to the Economic Planning Agency (EPA).

The Perron test is applied to the first difference of Japanese real income series Δnir_jp, cannot reject stationarity of real GDP once the structural break is taken into account.[10] Therefore, all variables used in the Hawaii tourism model are treated as I(1) in what follows. Figure 1 to figure 3 plot the series both in levels and first differences.

4.2 Weak Exogeneity

Tourism demand and supply theory and intuition indicate that four out of the nine variables are weakly exogenous: U.S. real income (*nir_us*), U.S. CPI (*cpi_us*), Japanese real income (*nir_jp*) and the *exchange rate adjusted* Japanese CPI (*p_jp*). To formally test these restrictions, we first adopt the routine of setting a whole row in the loading matrix to zero, i.e., $\alpha_x = 0$. Statistically this has the effect of excluding all cointegrating vectors from equations corresponding to the "theoretically" exogenous variables, i.e., cointegrating vectors do not explain variations in the first differences of exogenous variables. The χ^2 statistic and corresponding p-values are listed in Table 4.

For a system with nine variable, there can exist at the most eight cointegrating vectors.

[10]Details available upon request.

Table 2: Time Series Property of the Data – ADF Unit Root Test

$$\text{DF: } \Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \epsilon_t$$
$$\text{ADF: } \Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{5} \delta_i \Delta y_{t-i} + \epsilon_t$$
$$H_0 : \gamma = 0$$

Variables	DF	ADF(5)
vus_hi	-2.087	-2.023
vjp_hi	-1.682	0.431
prm_hi	-1.265	-1.627
trms_hi	-1.957	-1.707
ocup_hi	-3.898	-2.184
cpi_us	-5.576	-1.226
nir_us	-1.868	-2.143
p_jp	-0.711	-1.283
nir_jp	0.465	-0.584

Variables	DF	ADF(5)
Δ*vus_hi*	**-10.648**	**-3.525**
Δ*vjp_hi*	**-14.079**	**-4.709**
Δ*prm_hi*	**-11.601**	-2.931
Δ*trms_hi*	**-7.298**	**-3.731**
Δ*ocup_hi*	**-12.828**	**-4.138**
Δ*cpi_us*	**-6.157**	**-4.461**
Δ*nir_us*	**-8.482**	-2.724
Δ*p_jp*	**-6.849**	**-3.364**
Δ*nir_jp*	**-10.759**	-2.588

Note: Column 1 gives the target series (dependent variable in equations DF and ADF): *vus_hi* and *vjp_hi* are U.S. and Japanese visitor arrivals to Hawaii; *prm_hi*, *trms_hi* and *ocup_hi* are respectively average hotel room rate, total hotel room stock and average hotel occupancy rate in Hawaii; *cpi_us* is U.S. CPI; *nir_us* and *nir_jp* are U.S. and Japan real personal incomes; *p_jp* is *exchange rate adjusted* Japanese CPI. Column 2 is the Dick-Fuller test statistic for H_0. Column 3 is the Augmented Dick-Fuller statistic for H_0. All variables except *ocup_hi* are in logarithms. Boldness indicates significance at 5% level where critical value is -3.45.

Table 3: Time Series Property of the Data – Schwert Unit Root Test

$\Delta y_t = \alpha + \gamma y_{t-1} + \sum_{i=1}^{k} \delta_i \Delta y_{t-k} + \epsilon_t$

$k = 4$ for Schwert(l_4) and $k = 12$ for Schwert(l_{12})

$H_0 : \gamma = 0$

No Time trend in model

Variables	MA	Schwert(l_4)	CV(l_4)	AR(1-5)	Schwert(l_{12})	CV(l_{12})	AR(1-5)
Δvus_hi	-0.23	**-4.076**	-2.87	0.7652	-2.016	-2.82	0.4837
Δvjp_hi	-0.31	**-4.865**	-2.93	0.3567	-1.341	-2.85	0.7987
Δprm_hi	-0.20	**-3.434**	-2.87	0.9657	-1.813	-2.82	0.5255
$\Delta trms\_hi$	0.24	**-3.959**	-2.87	0.1901	-1.856	-2.82	0.4065
$\Delta ocup\_hi$	-0.37	**-5.348**	-2.93	0.8226	-1.867	-2.85	0.3811
Δnir_us	0.33	**-3.724**	-3.02	0.0874	-2.132	-2.82	0.1846
Δcpi_us	0.83	**-4.516**	-4.38	0.1845	-1.746	-2.92	0.2796
Δnir_jp	0	-1.758	-2.87	0.8874	-1.068	-2.82	0.4947
Δp_jp	0.35	**-3.389**	-3.02	0.9894	**-2.872**	-2.82	0.5615

$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{k} \delta_i \Delta y_{t-k} + \epsilon_t$

$k = 4$ for Schwert(l_4) and $k = 12$ for Schwert(l_{12})

$H_0 : \gamma = 0$

Time trend in model

Variables	MA	Schwert(l_4)	CV(l_4)	AR(1-5)	Schwert(l_{12})	CV(l_{12})	AR(1-5)
Δvus_hi	-0.23	**-4.267**	-3.41	0.6809	-1.956	-3.36	0.4416
Δvjp_hi	-0.31	**-5.389**	-3.49	0.7867	-2.305	-3.36	0.8646
Δprm_hi	-0.2	**-3.593**	-3.41	0.9856	-2.293	-3.36	0.4191
$\Delta trms\_hi$	0.24	**-4.556**	-3.41	0.1836	-2.863	-3.36	0.7987
$\Delta ocup\_hi$	-0.37	**-5.440**	-3.49	0.7893	-1.912	-3.36	0.2618
Δnir_us	0.33	**-3.641**	-3.61	0.0706	-2.151	-3.36	0.1915
Δcpi_us	0.83	-4.404	-5.09	0.2237	-1.869	-3.49	0.1697
Δnir_jp	0	-2.758	-3.41	0.8310	-3.023	-3.36	0.6757
Δp_jp	0.35	-3.525	-3.61	0.8925	-3.563	-3.36	0.3257

Note: Column 1 lists the series tested; Column 2 gives the MA parameter of the series; Column 3 and 6 are the Schwert(l_4) and Schwert(l_{12}) test statistics; Column 4 and 7 list the 5% critical values for Schwert(l_4) and Schwert(l_{12}) tabulated in *Table 7* of Schwert (1987) for the corresponding MA parameter. For instance, the estimated MA coefficient for Δvus_hi is -0.23, the closest case in Schwert (1987) has MA parameter of 0. The corresponding critical values are -2.87 for Schwert(l_4) and -2.82 for Schwert(l_{12}) without a time trend, and -3.41 for Schwert(l_4) and -3.36 for Schwert(l_{12}) with a time trend; Column 5 and 8 list the p-values of LM test for residual serial correlation with lags 1-5. All variables except *ocup_hi* are in logarithms. Boldness indicates significance at 5%.

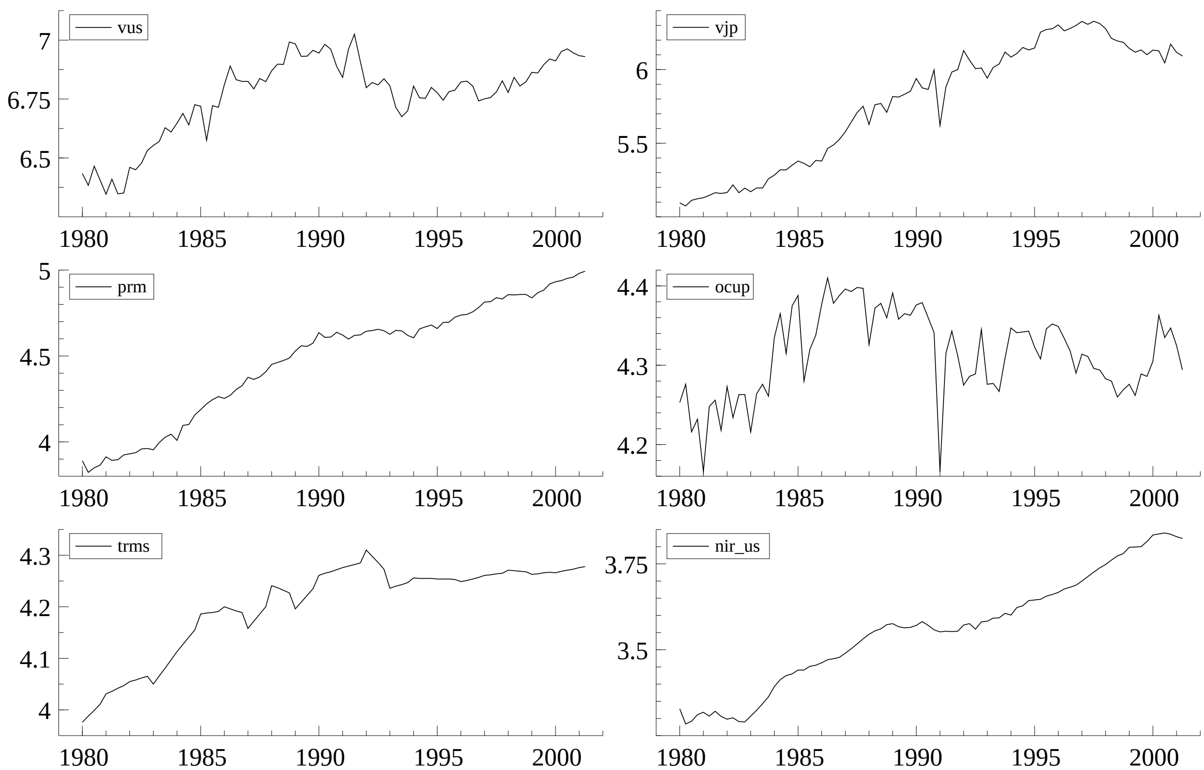

Figure 1: Variables in the Hawaii tourism model

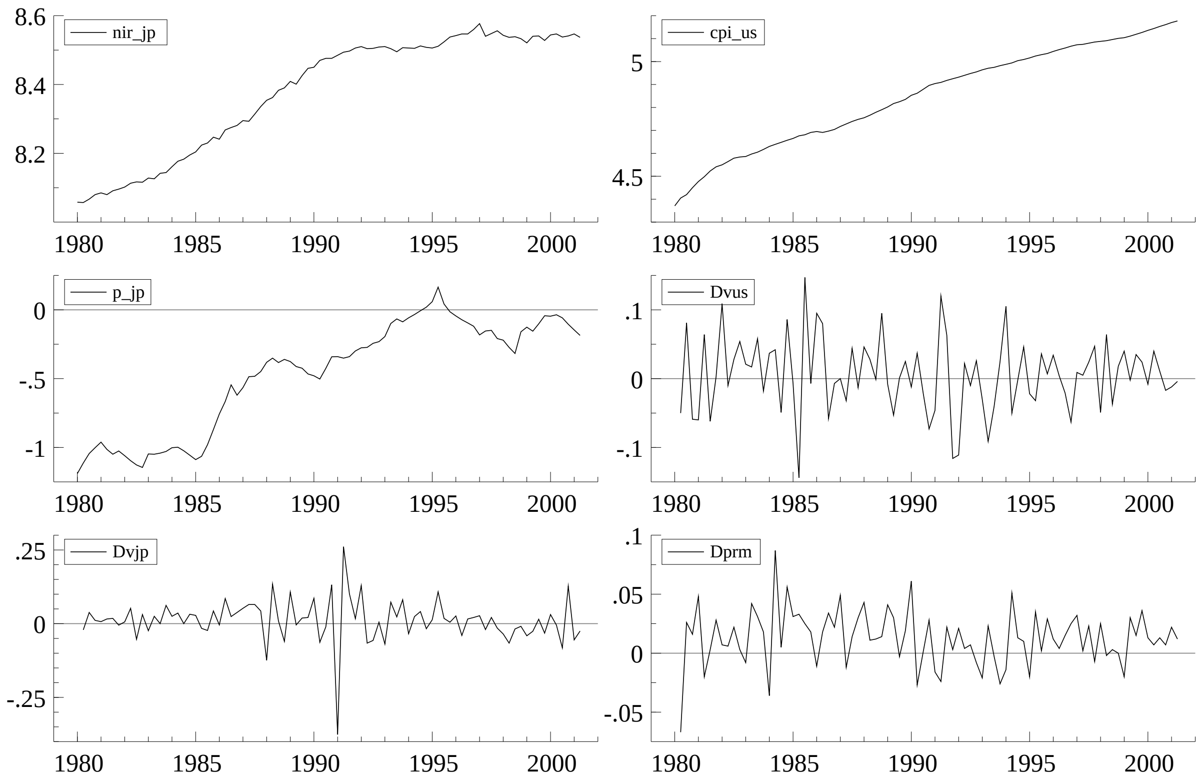

Figure 2: Variables in the Hawaii tourism model

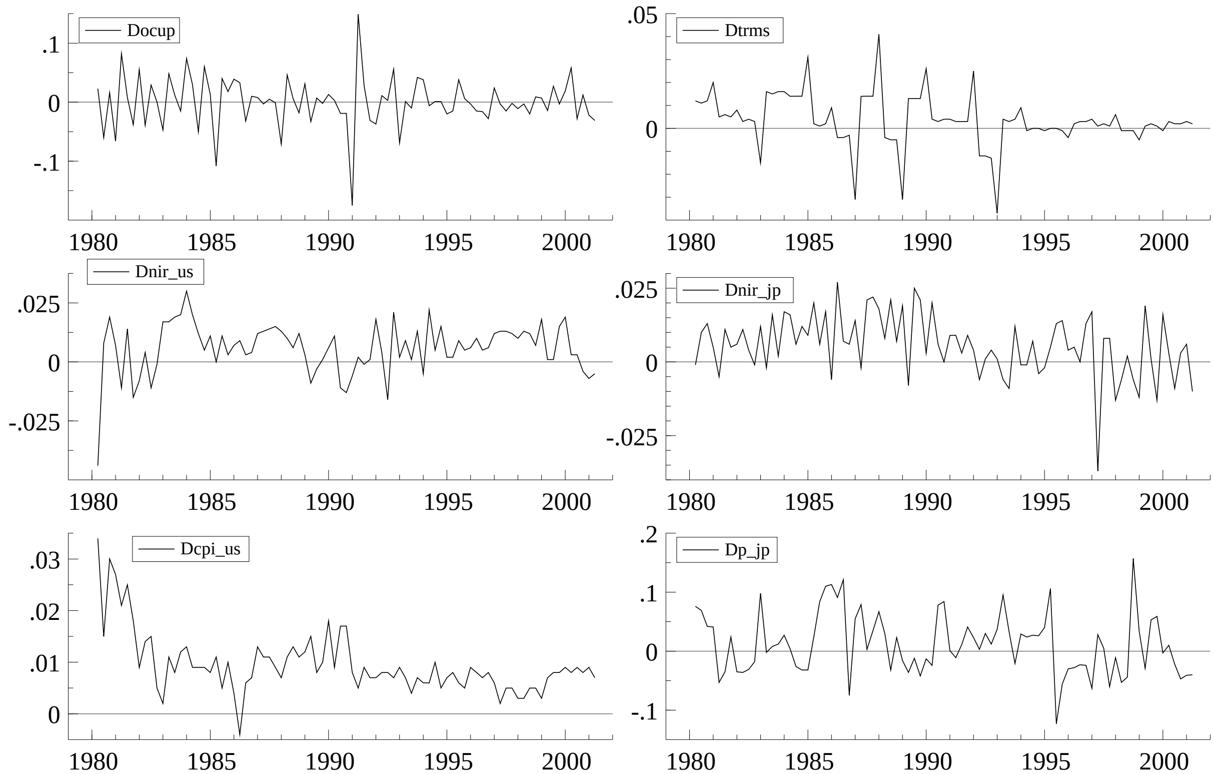

Figure 3: Variables in the Hawaii tourism model

Table 4: Weak Exogeneity Test ($\alpha_x = 0$)

A) eight cointegrating vectors		
	$\chi^2(8)$	p-value
nir_us	14.14	0.0781
cpi_us	**42.70**	0.0000
nir_jp	**41.78**	0.0000
p_jp	12.60	0.1264
B) six cointegrating vectors, *nir_us* and *p_jp* exogenous		
	$\chi^2(6)$	p-value
cpi_us	**40.53**	0.0000
p_jp	**38.37**	0.0000
C) three cointegrating vectors		
	$\chi^2(3)$	p-value
nir_us	1.78	0.6198
cpi_us	**15.19**	0.0017
nir_jp	**22.17**	0.0001
p_jp	4.81	0.1860
D) three cointegrating vectors, *nir_us* and *p_jp* exogenous		
	$\chi^2(3)$	p-value
cpi_us	**20.52**	0.0001
p_jp	**26.74**	0.0000

Note: Column 1 lists the variables tested; Column 2 gives the χ^2 statistics and column 3 is the corresponding probability values for the test. Bold entries indicate significance at 5% level.

Table 5: Harbo et. al. Weak Exogeneity Test

Variables	F-test	p-value
nir_us	0.20	0.90
cpi_us	1.24	0.31
nir_jp	2.09	0.12
p_jp	0.37	0.78

Note: Column 1 lists the variables studied; Column 2 presents the F-test statistics and column 3 are the corresponding critical values.

Weak exogeneity under the assumption of eight cointegrating vectors is rejected for *cpi_us* and *nir_jp* at 5% level, but not for *nir_us* or *p_jp*. The literature has found that weak exogeneity depends on model specification and Hall, Henry, and Greenslade (2002) suggest exogenizing the non-rejecting variables and re-test weak exogeneity for the remaining variables. We therefore treat *nir_us* and *p_jp* as weakly exogenous and re-estimate the system with seven endogenous variables, two exogenous variable and six cointegrating vectors. The test statistic still strongly rejects the weak exogeneity of *cpi_us* and *nir_jp* at 5% significance level. When three cointegrating relations are assumed (as depicted by equations (11)–(13)), the results are unchanged.

By setting $\alpha_x = 0$, it is implicitly assumed that variables in the x_t vector (*nir_us*, *cpi_us*, *nir_jp*, *p_jp*) are not cointegrated among themselves. For the Hawaii tourism model, however, it is highly like that these exogenous variables are cointegrated. Using a restricted trend, unrestricted intercept VAR(5) specification, we can not reject the hypothesis that there is at least one cointegrating relation among the four variables. Therefore, the rejection of $\alpha_x = 0$ is likely to be due to cointegration among exogenous variables, not to their endogeneity. According to Harbo, Johansen, Nielsen, and Rahbek (1998), weak exogeneity under such circumstances can be tested by estimating the conditional model with assumed exogeneity and inserting the estimated cointegrating vectors back into the marginal model. Weak exogeneity is established by statistical insignificance of the cointegrating vectors in the marginal model.

The Harbo *et. al.* exogeneity test requires fully specified cointegrating vectors and can only be performed after the cointegrating vectors are identified. We present the test statistics in this section, however, for completeness and note that the cointegrating vectors included are identified in the following sections. To perform the test, the first differences of exogenous variables (Δ*nir_us*, Δ*cpi_us*, Δ*nir_jp* and Δ*p_jp*) are each regressed onto the current and lagged first differences of all variables, the three identified cointegrating vectors and a constant. Weak exogeneity test amounts to a joint F-test that coefficients on all three cointegrating vectors are zero. The test results are listed in Table 5. As expected, weak exogeneity is not rejected for any of the variables.

4.3 Cointegrating Rank

With weak exogeneity for *nir_us*, *cpi_us*, *nir_jp* and *p_jp*, we proceed to test for the cointegrating rank using Johansen's reduced rank methodology. Table 6 reports the test

Table 6: Cointegration Rank Statistics

	$LR(H_r\|H_n)$			$LR(H_r\|H_{r+1})$		
Hr	Statistic	0.05CV	0.10CV	Statistic	0.05CV	0.10CV
$r = 0$	**197.6**	130.6	125.1	**56.51**	49.76	46.74
$r \leq 1$	**141.1**	99.11	93.98	**54.76**	43.75	41.01
$r \leq 2$	**86.29**	69.84	65.90	**40.88**	37.44	34.66
$r \leq 3$	**45.41**	45.10	41.57	28.15	30.55	27.86
$r \leq 4$	17.25	23.17	20.73	17.25	23.17	20.73

Note: Column 1 lists the null hypothesis of zero, at least one, two, three, four cointegrating vectors; Column 2 lists the *trace statistic*; Column 3 and 4 are the critical values for *trace statistic* at 5% and 10% significance levels; Column 5 lists the *maximum eigenvalue statistic*; Column 6 and 7 are the critical values for *maximum eigenvalue statistic* at 5% and 10% significance levels; Bolded numbers indicate significance at 5% level.

statistics and the corresponding asymptotic critical values at the 5% and 10% significance levels, as tabulated in *Table T.4* of Pesaran, Shin, and Smith (2000) with four exogenous variables.

From the table, the null of zero, one, and two cointegrating relations are rejected at both the 5% and 10% levels. When it comes to the null of three cointegrating vectors, the trace statistic rejects at both 5% and 10% significance levels (though the rejection at 5% is borderline). Nevertheless, the maximum eigenvalue statistic rejects the null at 10% level, but not at 5% level. We therefore conclude that the system has three long-run cointegrating vectors.

4.4 Long-run Cointegrating Vectors

We denote the three cointegrating vectors associated with z_t=(vus_hi, vjp_hi, prm_hi, $ocup\_hi$, $trms\_hi$, nir_us, cpi_us, nir_jp, p_jp, $t)'$ as β_1^*, β_2^*, and β_3^*, and treat them as explaining the U.S. tourism demand, Japanese tourism demand and Hawaii tourism supply respectively.

	vus_hi	vjp_hi	prm_hi	$ocup\_hi$	$trms\_hi$	nir_us	cpi_us	nir_jp	p_jp	t
β_1^*	β_{11}	β_{21}	β_{31}	β_{41}	β_{51}	β_{61}	β_{71}	β_{81}	β_{91}	β_{01}
β_2^*	β_{12}	β_{22}	β_{32}	β_{42}	β_{52}	β_{62}	β_{72}	β_{82}	β_{92}	β_{02}
β_3^*	β_{13}	β_{23}	β_{33}	β_{43}	β_{53}	β_{63}	β_{73}	β_{83}	β_{93}	β_{03}

Exact identification requires imposing three restrictions per vector, accomplished by two exclusion restrictions and one normalization restriction. Specifically, we exclude Japanese real income and Japanese visitor arrivals from the U.S. demand relation and normalize on U.S. visitor arrivals; exclude U.S. real income and U.S. visitor arrivals from the Japanese demand relation and normalize on Japanese visitor arrivals; exclude both real income variables from the Hawaii tourism supply relation and normalize on Hawaii room price.

	vus_hi	vjp_hi	prm_hi	$ocup\_hi$	$trms\_hi$	nir_us	cpi_us	nir_jp	p_jp	t
β_1^*	1	0	β_{31}	β_{41}	β_{51}	β_{61}	β_{71}	0	β_{91}	β_{01}
β_2^*	0	1	β_{32}	β_{42}	β_{52}	0	β_{72}	β_{82}	β_{92}	β_{02}
β_3^*	β_{13}	β_{23}	1	β_{43}	β_{53}	0	β_{73}	0	β_{93}	β_{03}

This is the exactly identified system and serves as the basis for over-identifying restrictions. When estimated, we obtain the following cointegrating vectors with the value of the log-likelihood function $LL_E = 1682.65$. Asymptotic standard errors are given in parentheses.[11]

	vus_hi	vjp_hi	prm_hi	$ocup\_hi$	$trms\_hi$	nir_us	cpi_us	nir_jp	p_jp	t
β_1^*	1	0	0.14	1.60	35.54	-36.13	-47.50	0	-0.12	0.54
			(3.90)	(4.09)	(8.10)	(9.12)	(12.47)		(0.95)	(0.13)
β_2^*	0	1	0.88	-2.13	3.79	0	-0.12	-6.07	0.55	-0.01
			(0.76)	(0.70)	(1.37)		(1.84)	(1.45)	(0.21)	(0.02)
β_3^*	-0.25	-0.23	1	-0.73	-1.12	0	1.04	0	0.09	-0.02
	(0.09)	(0.08)		(0.20)	(0.35)		(0.38)		(0.06)	(0.002)

The three final cointegrating vectors to be identified are,

$$vus\_hi = \beta_{01} * t + \beta_{61} * nir\_us + \beta_{71} * cpi\_us + \beta_{31} * prm\_hi + e_{us}, \tag{16}$$

$$vjp\_hi = \beta_{02} * t + \beta_{82} * nir\_jp + \beta_{92} * p\_jp + \beta_{32} * prm\_hi + e_{jp}, \tag{17}$$

$$\begin{aligned} prm\_hi = {} & \beta_{03} * t + \beta_{13} * (vus\_hi + vjp\_hi) + \beta_{43} * ocup\_hi \\ & + \beta_{73} * cpi\_us + \beta_{53} * trms\_hi + e_{prm}. \end{aligned} \tag{18}$$

All equations are over-identified. Based on the exact-identifying restrictions, we test the following over-identifying restrictions using the long-run structural modelling techniques advanced in Pesaran and Shin (2001),

$$\begin{aligned} &Equation(16): \quad \beta_{41} = \beta_{51} = \beta_{91} = 0, \quad \beta_{31} = -\beta_{71}; \\ &Equation(17): \quad \beta_{42} = \beta_{52} = \beta_{72} = 0, \quad \beta_{32} = -\beta_{92}; \\ &Equation(18): \quad \beta_{93} = 0, \quad \beta_{13} = \beta_{23}. \end{aligned}$$

4.4.1 U.S. Tourism Demand

To identify a U.S. tourism demand relation, we test the four over-identifying restrictions specified above, i.e., three exclusion restrictions $\beta_{41} = \beta_{51} = \beta_{91} = 0$ (corresponding to *ocup_hi*, *trms_hi* and *p_jp*), and the homogeneity restriction $\beta_{31} = -\beta_{71}$ (parameters on *prm_hi* and *cpi_us* are equal in size but opposite in sign) . Witt and Witt (1995) find that income elasticities tend to exceed unity, consistent with the notion that international travel is a luxury good. For a sample of fourteen models from four studies, they report a median income elasticity of 2.4. In a separate study, Sheldon (1993) surveys ten econometric studies of tourism expenditures from 1966 to 1987 for a wide range of source-destination

[11] Computation results reported in this section are carried out using *Pc-Fiml 9.10.*

pairs including U.S. travel to Canada, Europe, and Mexico, Canadian tourism to the U.S. and other countries and U.S. destination tourism by major foreign countries. He finds a large range for income elasticities (from -0.15 to 6.6) with a median of 2.2. We therefore restrict the income elasticity (β_{61}) to be 2.5 initially. With these five over-identifying restrictions, the following β estimate yields:

	vus_hi	*vjp_hi*	*prm_hi*	*ocup_hi*	*trms_hi*	*nir_us*	*cpi_us*	*nir_jp*	*p_jp*	*t*
β_1^*	1	0	0.11	0	0	-2.50	-0.11	0	0	0.01
							(0.33)			(0.002)
β_2^*	0	1	-2.91	1.75	6.55	0	-5.82	-1.75	-0.13	0.07
			(0.65)	(0.71)	(1.24)		(1.72)	(1.15)	(0.20)	(0.01)
β_3^*	-0.35	-0.24	1	-0.77	-0.84	0	0.71	0	0.10	-0.01
	(0.07)	(0.07)		(0.19)	(0.31)		(0.38)		(0.05)	(0.002)

The (log-) likelihood ratio test has value 13.59 and a p-value of 1.84% for $\chi^2(5)$ distribution. These over-identifying restrictions are therefore rejected at 5% significance level, but not at 1% level. Edwards (1995) obtains an income elasticity of 5 for U.S. travellers to Asia-Pacific region. To study the impacts of different income elasticity numbers, we set the value to 3, 4 and 5 respectively. The resulting price elasticities are -0.32, -0.73, and -1.16, with p-values 3.18%, 6.64% and 10.13%. Price elasticity increases when income elasticity rises, but the change is small. We therefore set income elasticity to 5 to be consistent with Edwards (1995). The following U.S. demand estimate (β_1^*) yields where all parameter estimates have theoretically correct signs:

	vus_hi	*vjp_hi*	*prm_hi*	*ocup_hi*	*trms_hi*	*nir_us*	*cpi_us*	*nir_jp*	*p_jp*	*t*
β_1^*	1	0	1.16	0	0	-5	-1.16	0	0	0.025
							(0.45)			(0.003)
β_2^*	0	1	-2.76	0.74	7.96	0	-6.31	-2.86	0.056	0.069
			(0.73)	(0.75)	(1.38)		(1.9)	(1.36)	(0.22)	(0.02)
β_3^*	-0.23	-0.23	1	-0.75	-1.16	0	1.10	0	0.09	-0.016
	(0.08)	(0.07)		(0.19)	(0.33)		(0.38)		(0.06)	(0.002)

4.4.2 Japanese Tourism demand

A Japanese demand relation is harder to identify. We initially employ restrictions similar to those used in the U.S. tourism demand. The income elasticity of Japanese visitors are left unrestricted as there is no indication in the literature of a good estimate. Altogether four over-identifying restrictions are applied: three exclusion restrictions on *ocup_hi*, *trms_hi*, and *cpi_us* ($\beta_{42} = \beta_{52} = \beta_{72} = 0$) and one homogeneity restriction on *prm_hi* and *p_jp* ($\beta_{32} = -\beta_{92}$). With these restrictions, we obtain a $\chi^2(4)$ statistic of 15.68 and a p-value of 0.35%. The restrictions are strongly rejected.

The literature has found very different responses of overseas tourism demand to exchange rate movement than to consumer price index simply because exchange rate changes are much easier to observe. The first modification to the Japanese demand relation, therefore, is to break the price homogeneity assumption. This leads to a $\chi^2(3)$ statistic of 8.50 and a p-value of 3.68%. The remaining restrictions are rejected at 5% significance level, but not at 1%.

To relax one step further, we re-include *trsm_hi* in the cointegrating space. Empirically this is justified by the "You build, I come" mentality of travellers, especially travellers to foreign countries in pursuit of exotic experience. When this is done, the $\chi^2(2)$ reduces to 2.32 with a p-value of 31.33%. We do not reject the two restrictions at 5% significance level. A closer look at the estimated vector reveals that trend (t) can be safely excluded. The final three restrictions have $\chi^2(3)$ statistic at 2.35 and a p-value of 50%. The estimated vector (β_2^*) and corresponding standard errors listed in the table below. The coefficient estimate on room stock variable is negative and contrary to expected. One explanation might come from the special sample period under study. The sample period chosen is 1980Q1 to 2001Q2. During this time, Hawaii experienced an economic cycle brought about by the booming, over-heating and eventual collapsing of the Japanese economy. Between 1985 and 1995, Japanese invested no less than $12 billion in Hawaii compared with a total of $850 million in the preceding 10 years. Total visitor plant inventory correspondingly increased from roughly 65,000 transient rental accommodations in the years 1985-87 to nearly 74,000 by the end of 1993.[12] With the bust of Japanese economic bubble, visitor counts from the country dropped. But the hotel accommodation infrastructure, once installed, takes time to diminish. Statistically this may show as a negative correlation. Apart from the room stock, *p_jp* has the wrong sign. A rise in Japanese *exchange-rate adjusted* price level increases the competitiveness of Hawaii travel and the variable should have a positive sign. However, the estimated parameter is negative.[13]

	vus_hi	*vjp_hi*	*prm_hi*	*ocup_hi*	*trms_hi*	*nir_us*	*cpi_us*	*nir_jp*	*p_jp*	*t*
β_1^*	1	0	0.33	2.27	37.7	-38.8	-50.5	0	-0.20	0.58
			(4.13)	(4.32)	(8.56)	(9.73)	(13.24)		(1.0)	(0.14)
β_2^*	0	1	0.29	0	3.33	0	0	-5.2	0.29	0
			(0.21)		(0.94)			(0.78)	(0.14)	
β_3^*	-0.23	-0.23	1	-0.89	-1.11	0	0.99	0	0.10	-0.015
	(0.09)	(0.08)		(0.20)	(0.37)		(0.38)		(0.06)	(0.003)

4.4.3 Hawaii Tourism Supply

The Hawaii tourism supply relation has two over-identifying restrictions: exclusion restriction on *p_jp* ($\beta_{93} = 0$) and an equal parameter restriction on U.S. and Japanese arrival counts ($\beta_{13} = \beta_{23}$). When imposed, the long-run cointegrating relations becomes: The $\chi^2(2)$ statistic associated with the test is 1.54 and well below the 5% critical value. The supply restrictions are not rejected. Nevertheless, the estimated parameter on *cpi_us* has the wrong sign. U.S. CPI in the supply relation supposedly approximates hotel operating cost. Theory prescribes a positive relationship between cost and price of final product—a markup

[12] Numbers come from "Japanese Investment in Hawaii: Past and Future", University of Hawaii Economic Research Organization (UHERO) research report, 1998.

[13] It is acknowledged that the structural break identified in the Japanese income series (*nir_jp*) may have contributed to the weak demand cointegration relation for Japanese visitors.

	vus_hi	*vjp_hi*	*prm_hi*	*ocup_hi*	*trms_hi*	*nir_us*	*cpi_us*	*nir_jp*	*p_jp*	*t*
β_1^*	1	0	-2.88	1.38	52.3	-44.3	-67.9	0	0.63	0.75
			(4.86)	(5.30)	(10.2)	(11.08)	(15.50)		(1.14)	(0.16)
β_2^*	0	1	0.14	-2.13	5.49	0	-2.04	-5.9	0.68	0.01
			(0.82)	(0.84)	(1.61)		(2.12)	(1.59)	(0.23)	(0.02)
β_3^*	-0.17	-0.17	1	-0.49	-1.56	0	1.58	0	0	-0.018
		(0.06)		(0.19)	(0.32)		(0.34)			(0.002)

over cost. Obviously the approximation is far less than satisfactory. U.S. CPI may very well be excluded from the supply relation. Apart from this, the coefficient on *trms_hi* does not have the correct sign either. A rise in hotel room stock should increase accommodation availability and puts downward pressure on the average room rate. However, the estimated coefficient is positive. This might reflect that expansion rate of tourism accommodations is not catching up with the growth rate of room price.

4.4.4 Joint Restrictions

Finally, we combine all three hypotheses and perform a joint test. The log-likelihood ratio statistic for the joint test has value 21.15, which is below the 1% critical level of $\chi^2(10)$ distribution. We do not reject the joint hypotheses of two demand and one supply relations. The parameter estimates of the joint test change only slightly from those of individual tests.

	vus_hi	*vjp_hi*	*prm_hi*	*ocup_hi*	*trms_hi*	*nir_us*	*cpi_us*	*nir_jp*	*p_jp*	*t*
β_1^*	1	0	1.2	0	0	-5	-1.2	0	0	0.024
							(0.41)			(0.003)
β_2^*	0	1	0.23	0	3.42	0	0	-5.2	0.31	0
			(0.18)		(0.82)			(0.64)	(0.1)	
β_3^*	-0.15	-0.15	1	-0.86	-1.45	0	1.4	0	0	-0.018
		(0.07)		(0.17)	(0.36)		(0.33)			(0.002)

4.5 The dynamic Model

The model contained in (16)–(18) is static and represents the long-run equilibria of the system. Dynamic adjustment to the equilibria is captured by lagged differences in each equation. In tourism demand equations, these are due to the delayed response of travel demand to income and price changes. In the supply equation, it reflects the gradual adjustments in hotel room price to changes in demand and supply factors.

Based on the long-run cointegrating vectors, we have:

$$\begin{aligned} vus\_hi_t &= 5 * nir\_us_t - 1.2 * (prm\_hi_t - cpi\_us_t) - 0.024 * t + ECM_{vus}; \\ vjp\_hi_t &= 5.2 * nir\_jp_t - 0.23 * prm\_hi_t - 3.42 * trms\_hi_t \\ &\qquad - 0.31 * p\_jp_t + ECM_{vjp}; \\ prm\_hi_t &= 0.15 * (vus\_hi_t + vjp\_hi_t) + 0.86 * ocup\_hi_t \\ &\qquad + 1.45 * trms\_hi_t - 1.4 * cpi\_us_t + 0.018 * t + ECM_{prm}; \end{aligned}$$

Table 7: Estimates of the Error Correction Coefficients and Diagnostic Statistics

Equation	α_{1y}	α_{2y}	α_{3y}	R^2	AR(1-5)	χ^2_N	$ARCH_4$	X^2_i	Reset
Δvus_hi	-0.11	-	-	0.54	0.89	2.13	0.26	0.89	0.21
	(-3.20)				[0.50]	[0.34]	[0.90]	[0.59]	[0.65]
Δvjp_hi	-	-0.17	-0.36	0.67	1.56	2.82	1.26	0.33	3.46
		(-2.48)	(-2.90)		[0.19]	[0.24]	[0.30]	[0.99]	[0.07]
Δprm_hi	-0.07	-0.12	-0.39	0.54	1.07	0.22	0.66	0.74	0.37
	(-4.42)	(-3.73)	(-7.02)		[0.39]	[0.89]	[0.62]	[0.75]	[0.54]
$\Delta ocup\_hi$	-0.01	-	0.12	0.61	1.69	0.16	1.03	0.87	4.76
	(-2.23)		(2.18)		[0.15]	[0.92]	[0.40]	[0.61]	[0.03]
$\Delta trms\_hi$	-0.02	-0.07	-	0.70	1.26	1.01	0.55	0.78	1.06
	(-2.11)	(-4.66)			[0.30]	[0.60]	[0.70]	[0.73]	[0.31]

Note: Column 1 lists the dependent variable of individual equations in the system; Column 2 to 4 gives the regression coefficient and the corresponding *Student* t-statistic for the three identified cointegrating vectors; Column 5 presents the coefficient of determination R^2; Column 6 lists the test statistics for autocorrelated residuals, performed through the auxiliary regression of the residuals on the original variables and lagged residuals. Column 7 is the χ^2 normality test on regression residuals. Column 8 is the autoregressive conditional heteroscedasticity (ARCH) test following Engle (1982). It is done by regressing the squared residuals on a constant and lagged squared residuals and testing the significance of the lagged squared residuals. Column 9 is another heteroscedasticity test based on White (1980), which involves an auxiliary regression of the squared residuals on a constant, the original regressors and the original regressors squared. Column 10 is the functional form mis-specification test. It amounts to adding powers (2,3,4) of the fitted values to the original regression. The null is no functional mis-specification, which would be rejected if the test statistic is too high. figures in parenthesis (.) are the *Student* t-statistics corresponding to the loading parameters whereas those in brackets [.] are p-values for individual tests.

where ECM_{vus}, ECM_{vjp}, and ECM_{prm} are the three equilibrium errors, with which the following VECM is estimated:

$$\begin{aligned}\Delta y_t = c_0 + \omega \Delta x_t + \sum_{i=1}^{4} \Gamma_i \Delta z_{t-1} + \alpha_{1y} ECM_{vus} + \alpha_{2y} ECM_{vjp} \\ + \alpha_{3y} ECM_{prm} + u_t\end{aligned} \tag{19}$$

and α_{1y}, α_{2y}, and α_{3y} are 3-dimensional vectors of loading parameters. At this stage, dynamics are simplified by dropping statistically insignificant terms. This involves excluding first differenced terms that has t-statistic less than 2, starting from the smallest. The error correction terms are eliminated by the same criterion. A total of 134 zero restrictions are applied. The χ^2 statistic is 166.7 with a p-value equal to 2.91%. We do not reject these exclusion restrictions at 1% level. The estimated loading parameters and corresponding diagnostic test statistics are shown in Table 7.

The estimated system appears to be an adequate model for Hawaii tourism activity. All equations perform fairly well, explaining 54%, 67%, 54%, 61% and 70% of the variation in Δvus_hi, Δvjp_hi, Δprm_hi, $\Delta ocup\_hi$, and $\Delta trms\_hi$ respectively. All equations pass all diagnostic tests at 1% significance level. The existence of long-run equilibrium error terms (ECMs) in model equations allows for temporary disequilibrium between causal variables and the demand and supply variables. The adjustment factor (α's) captures the speed of adjustment toward the equilibrium relationship. For example, if U.S. arrivals are less than predicted by U.S. real income growth and the relative cost of Hawaii vacation, arrivals would increase over time to eliminate the disequilibrium error. The three long-run ECMs enter the five equations differently. Δvus_hi equation contains only ECM_{vus} with a loading parameter of -0.11, so 11% of the equilibrium error is corrected per period. Δvjp_hi equation contains both ECM_{vjp} and ECM_{prm} with the coefficient on the latter twice of that on the former. This implies that *dis*-equilibrium in hotel room price has a larger dampening effects on visitors from Japan. All three ECMs enter Δprm_hi equation significantly. The estimated parameter on ECM_{prm} is relatively large, reflecting the high incentive to adjust markup in maintaining occupancy. The hotel occupancy equation contains both ECM_{vus} and ECM_{prm}. The presence of ECM_{vus} signifies the importance of U.S. visitor arrivals in determining hotel occupancy due to its size.[14] The last equation, hotel room stock, responds to both ECM_{vus} and ECM_{vjp} with a much higher coefficient on Japanese tourists. The dynamic model is picked over sample period 1980Q1-1997Q4, leaving 14 observations for *ex post* forecasts. Figure 4 plots Δvus_hi, Δvjp_hi, Δprm_hi, $\Delta ocup\_hi$, $\Delta trms\_hi$, the corresponding dynamic *ex post* forecasts, together with the 95% confidence error bands.

5 Forecast Encompassing

This section evaluates the forecasting capability of the newly identified Hawaii tourism model, referred to as *Coint-tourmod* below. In order to provide some benchmark, the evaluation is done in comparison to two rival models: One, a simple no-change model (hereafter *RW-tourmod*) in which the best forecast for next quarter and all future quarters is the observed value this quarter; Two, the tourism forecasting model maintained by the University of Hawaii Economic Research Organization (hereafter, *UHERO-tourmod*). The UHERO tourism model is a structural multi-equation forecasting model with single equation estimation.[15] The goal here is to test whether any forecast accuracy is gained by fully accounting for the structural relationships using system estimation.

The criterion used is forecast encompassing. It concerns whether h-step ahead forecasts of one model can explain the forecast errors made by a rival model. The test was originally proposed by Chong and Hendry (1986) as a feasible way to evaluate large-scale econometric models. In the literature, forecasting encompassing is closely related to forecast combination and forecast "conditional efficiency" (see Bates and Granger (1969) for the former and Nelso (1972) and Granger and Newbold (1973) for the latter).

[14] To understand the ECMs in the occupancy equation, recall that they enter in first lags. So if in the previous quarter U.S. arrivals are higher than prescribed by U.S. real income and relative cost of a Hawaii vacation, U.S. arrivals this period tend to go down, resulting in a downward pressure on occupancy rate. Similarly, if hotel room rate in the last period is higher than warranted by the equilibrium relation, it has a tendency to decrease this quarter, leading to higher occupancy rate.

[15] Details of the model are available from the authors upon request.

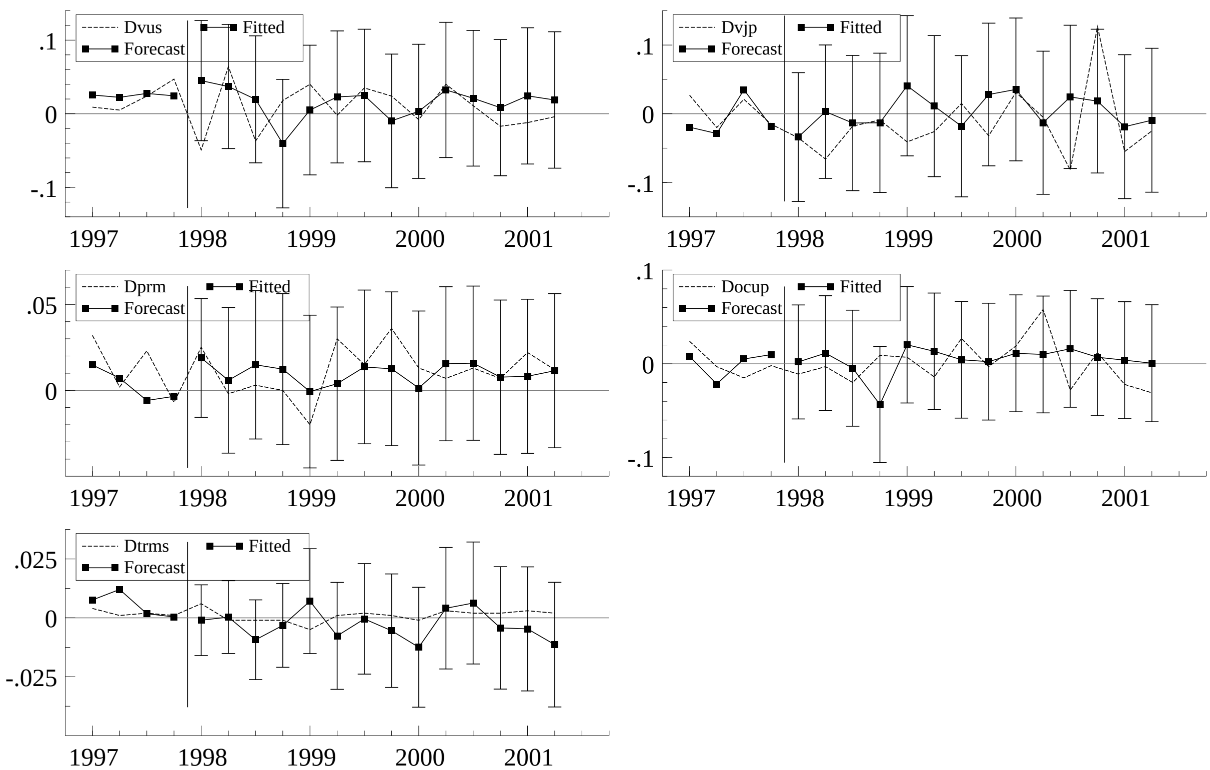

Figure 4: Dynamic forecasts with error bands

Rolling samples are used in the exercise except for the *UHERO-tourmod* which is estimated over full sample. In this sense, *UHERO-tourmod* forecasts are in sample dynamic forecasts and are expected to perform better.[16] *Coint-tourmod* is estimated over sample period 1980Q1–1997Q4, leaving data over period 1998Q1–2001Q2 for *ex post* forecasts. To execute, the estimation ending period is set to 1997Q4 initially and *one*- through *eight*-step ahead dynamic forecasts are generated for 1998Q1–1999Q4. The sample end is then moved one-step ahead to 1998Q1 and another set of *one*- through *eight*-step ahead dynamic forecasts generated for 1998Q2-2000Q1. We continue the process till the sample end is reached. As such, we obtain 14 *one*-step, 13 *two*-step, 11 *four*-step and 7 *eight*-step ahead dynamic forecasts. Forecast encompassing is then performed on forecast errors of Δvus_hi, Δvjp_hi, Δprm_hi, $\Delta ocup\_hi$, $\Delta trms\_hi$ in pairs: *Coint-tourmod* against *UHERO-tourmod*, *Coint-tourmod* against *RW-tourmod*, and *UHERO-tourmod* against *RW-tourmod*. For instance, to test if *Coint-tourmod* forecast encompasses *RW-tourmod*, we regress forecast errors from the *Coint-tourmod* onto the difference between the forecast errors from *Coint-tourmod* and those from *RW-tourmod* and test whether coefficients on the error differences are zeros. The test statistic follows t-distribution under the null.

All encompassing test results are listed in Table 8 through Table 12. To read the table, each model listed by row is the control model and every model listed by column is a rival model. For instance, row one in Table 8 presents encompassing test results of *one*-step ahead forecasts for Δvus_hi when *Coint-tourmod* is the control model and the rival models are *UHERO-tourmod* and *RW-tourmod* respectively. The diagonal line shows not applicable (n.a.) as a model can not forecast encompass itself.

From Table 8 to Table 12, we cannot reject the null hypothesis that *Coint-tourmod* forecast encompasses *RW-toumod* for all endogenous variables at all forecast horizons except $\Delta ocup\_hi$ at *four*-step ahead and $\Delta trms\_hi$. For $\Delta trms\_hi$, the null hypothesis of *Coint-tourmod* forecast encompassing *RW-tourmod* is rejected at 1% significance at all forecast horizons. The converse, *RW-tourmod* forecast encompassing *Coint-tourmod* for $\Delta trms\_hi$ is not rejected at any forecast horizon. The evidence indicates that a simple no-change model in first differences offers a better specification for hotel room stock. This might be explained by the fact that Hawaii hotel room inventory stays relatively flat for the forecast period. When comparing *Coint-tourmod* to *UHERO-tourmod*, we reject the null hypothesis that *Coint-tourmod* forecast encompasses *UHERO-tourmod* except for Δvjp_hi and Δprm_hi at *eight*-step ahead, and $\Delta ocup\_hi$ at *two*-, *four*-, and *eight*-step ahead. The converse, *UHERO-tourmod* forecast encompasses *Coint-tourmod*, is not rejected for any of the variables at any horizon. This may be explained by the fact that *UHERO-tourmod* is estimated using full sample while *Coint-tourmod* is estimated using only sample 1980Q1–1997Q4. Overall, the newly identified *Coint-tourmod* performs relatively well against *UHERO-tourmod* and *RW-tourmod*, especially at longer forecast horizons.

[16] For fair comparison, it is ideal to apply rolling sample to all models involved. However, the *UHERO-tourmod* is a multi-equation system containing hundreds of equations. It is difficult to re-estimate the system.

Table 8: Forecast Encompassing Test – Δvus_hi

Model	Coint-tourmod	UHERO-tourmod	RW-tourmod
	One-step ahead		
Coint-tourmod	n.a.	**41.7**	0.03
	–	[0.00]	[0.86]
UHERO-tourmod	0.43	n.a.	1.5
	[0.52]	–	[0.24]
RW-tourmod	**10.8**	**94.9**	n.a.
	[0.00]	[0.00]	–
	Two-step ahead		
Coint-tourmod	n.a.	**11.6**	7.12
	–	[0.005]	[0.02]
UHERO-tourmod	0.18	n.a.	0.89
	[0.68]	–	[0.37]
RW-tourmod	**9.39**	**16.0**	n.a.
	[0.01]	[0.00]	–
	Four-step ahead		
Coint-tourmod	n.a.	**25.3**	1.06
	–	[0.00]	[0.33]
UHERO-tourmod	0.03	n.a.	3.64
	[0.86]	–	[0.09]
RW-tourmod	**17.25**	**108.1**	n.a.
	[0.00]	[0.00]	–
	Eight-step ahead		
Coint-tourmod	n.a.	**31.3**	7.96
	–	[0.001]	[0.03]
UHERO-tourmod	0.00009	n.a.	2.46
	[0.99]	–	[0.17]
RW-tourmod	**13.44**	**22.04**	n.a.
	[0.00]	[0.003]	–

Note: Column 1 lists the control model; Column 2 to 4 give the rival models and the corresponding t-statistic on forecast encompassing at different horizons. Numbers in brackets [.] are corresponding p-values. All bold inputs indicate significance at 1% significance level.

Table 9: Forecast Encompassing Test – Δvjp_hi

Model	Coint-tourmod	UHERO-tourmod	RW-tourmod
	One-step ahead		
Coint-tourmod	n.a.	**28.1**	0.01
	–	[0.00]	[0.91]
UHERO-tourmod	0.16	n.a.	1.27
	[0.70]	–	[0.28]
RW-tourmod	**14.1**	**79.8**	n.a.
	[0.00]	[0.00]	–
	Two-step ahead		
Coint-tourmod	n.a.	**15.7**	1.19
	–	[0.00]	[0.30]
UHERO-tourmod	0.0005	n.a.	0.007
	[0.98]	–	[0.94]
RW-tourmod	6.34	**26.6**	n.a.
	[0.03]	[0.00]	–
	Four-step ahead		
Coint-tourmod	n.a.	**21.6**	2.8
	–	[0.00]	[0.13]
UHERO-tourmod	0.41	n.a.	0.18
	[0.54]	–	[0.68]
RW-tourmod	2.4	**20.01**	n.a.
	[0.15]	[0.00]	–
	Eight-step ahead		
Coint-tourmod	n.a.	5.81	1.17
	–	[0.05]	[0.32]
UHERO-tourmod	0.2	n.a.	0.00002
	[0.67]	–	[0.99]
RW-tourmod	**15.85**	**14.92**	n.a.
	[0.00]	[0.00]	–

Note: Column 1 lists the control model; Column 2 to 4 give the rival models and the corresponding t-statistic on forecast encompassing at different horizons. Numbers in brackets [.] are corresponding p-values. All bold inputs indicate significance at 1% significance level.

Table 10: Forecast Encompassing Test – Δprm_hi

Model	Coint-tourmod	UHERO-tourmod	RW-tourmod
	One-step ahead		
Coint-tourmod	n.a.	**14.9**	5.54
	–	[0.00]	[0.03]
UHERO-tourmod	0.42	n.a.	0.02
	[0.53]	–	[0.88]
RW-tourmod	**14.9**	**27.6**	n.a.
	[0.00]	[0.00]	–
	Two-step ahead		
Coint-tourmod	n.a.	**11.47**	5.78
	–	[0.005]	[0.03]
UHERO-tourmod	0.5	n.a.	0.47
	[0.49]	–	[0.51]
RW-tourmod	7.74	**14.0**	n.a.
	[0.02]	[0.00]	–
	Four-step ahead		
Coint-tourmod	n.a.	**10.4**	0.0046
	–	[0.009]	[0.95]
UHERO-tourmod	0.01	n.a.	1.74
	[0.92]	–	[0.22]
RW-tourmod	**10.44**	**38.9**	n.a.
	[0.009]	[0.00]	–
	Eight-step ahead		
Coint-tourmod	n.a.	7.15	1.91
	–	[0.04]	[0.22]
UHERO-tourmod	0.29	n.a.	0.63
	[0.61]	–	[0.45]
RW-tourmod	**15.6**	**31.9**	n.a.
	[0.007]	[0.00]	–

Note: Column 1 lists the control model; Column 2 to 4 give the rival models and the corresponding t-statistic on forecast encompassing at different horizons. Numbers in brackets [.] are corresponding p-values. All bold inputs indicate significance at 1% significance level.

Table 11: Forecast Encompassing Test – $\Delta ocup\_hi$

Model	Coint-tourmod	UHERO-tourmod	RW-tourmod
	One-step ahead		
Coint-tourmod	n.a.	**16.9**	0.01
	–	[0.00]	[0.92]
UHERO-tourmod	2.25	n.a.	1.07
	[0.16]	–	[0.32]
RW-tourmod	7.25	**29.9**	n.a.
	[0.02]	[0.00]	–
	Two-step ahead		
Coint-tourmod	n.a.	5.76	1.09
	–	[0.03]	[0.32]
UHERO-tourmod	0.0019	n.a.	0.12
	[0.97]	–	[0.73]
RW-tourmod	4.43	**10.5**	n.a.
	[0.06]	[0.007]	–
	Four-step ahead		
Coint-tourmod	n.a.	3.71	1.10
	–	[0.08]	[0.32]
UHERO-tourmod	0.002	n.a.	**18.5**
	[0.97]	–	[0.00]
RW-tourmod	**16.45**	**83.0**	n.a.
	[0.00]	[0.00]	–
	Eight-step ahead		
Coint-tourmod	n.a.	0.007	0.33
	–	[0.94]	[0.59]
UHERO-tourmod	0.89	n.a.	0.80
	[0.38]	–	[0.41]
RW-tourmod	**17.73**	**14.3**	n.a.
	[0.00]	[0.00]	–

Note: Column 1 lists the control model; Column 2 to 4 give the rival models and the corresponding t-statistic on forecast encompassing at different horizons. Numbers in brackets [.] are corresponding p-values. All bold inputs indicate significance at 1% significance level.

Table 12: Forecast Encompassing Test – $\Delta trms\_hi$

Model	Coint-tourmod	UHERO-tourmod	RW-tourmod
	One-step ahead		
Coint-tourmod	n.a.	**58.5**	**51.8**
	–	[0.00]	[0.00]
UHERO-tourmod	3.38	n.a.	3.82
	[0.09]	–	[0.07]
RW-tourmod	0.10	1.86	n.a.
	[0.75]	[0.20]	–
	Two-step ahead		
Coint-tourmod	n.a.	**81.4**	**81.04**
	–	[0.00]	[0.00]
UHERO-tourmod	0.92	n.a.	1.67
	[0.36]	–	[0.22]
RW-tourmod	7.52	8.73	n.a.
	[0.018]	[0.012]	–
	Four-step ahead		
Coint-tourmod	n.a.	**375.2**	**155.2**
	–	[0.00]	[0.00]
UHERO-tourmod	1.28	n.a.	0.21
	[0.28]	–	[0.65]
RW-tourmod	1.9	**15.1**	n.a.
	[0.20]	[0.00]	–
	Eight-step ahead		
Coint-tourmod	n.a.	**909.5**	**62.4**
	–	[0.00]	[0.00]
UHERO-tourmod	0.9	n.a.	8.93
	[0.38]	–	[0.02]
RW-tourmod	2.32	**235.2**	n.a.
	[0.18]	[0.00]	–

Note: Column 1 lists the control model; Column 2 to 4 give the rival models and the corresponding t-statistic on forecast encompassing at different horizons. Numbers in brackets [.] are corresponding p-values. All bold inputs indicate significance at 1% significance level.

6 Concluding Remarks

This paper estimates a complete model of Hawaii tourism using a system cointegration approach. Unlike existing models that are exclusively demand oriented, we model both demand and supply variables. For tourism activities in Hawaii, the paper identifies one demand relation each for the U.S. and Japanese visitors and an inverse supply curve depicting average hotel room prices. By formally incorporating the supply side, the Hawaii tourism model is less vulnerable to regression biases caused by demand and supply interactions.

To our knowledge, ours is the first paper in this literature to tackle the important problem of identification in a cointegrated system. We follow a theory-guided approach advocated by Pesaran and Smith (1998) and Pesaran and Shin (2001), rather than relying on the purely statistical identification method of Johansen. Compared to the Johansen procedure, the theory-guided approach has intuitive appeal and identifies economically meaningful cointegrating vectors. For the Hawaii tourism model, over-identifying restrictions are imposed on each of the three long-run cointegrating vectors.

The paper also follows Hall, Henry, and Greenslade (2002) method in system reduction. Rather than estimating a full model, the paper tests and imposes weak exogeneity assumptions at the earliest stage in the model reduction process. By doing so, the number of parameters to be estimated is greatly reduced, saving degrees of freedom and improving the efficiency of estimated coefficients.

Not every aspect of the identified model conforms with expectations. The estimated income elasticities for both the U.S. and Japanese visitors come out on on the high side. Coefficients on some other variables have signs contrary to expectations. These deficiencies to some extent demonstrate the difficulties faced by empirical modelers when estimating a large system with limited macro data set. This reinforces the need for careful theoretical and empirical design from the start.

Despite the "wrong signs" and large income elasticites, the system peforms reasonably well in out of sample forecast comparisons. In future work we will test the model against a wider range of competing forecasting methods, make use of statistical tests of forecast accuracy, and experiment with imposing more reasonable but statistically rejected values for elasticities.

References

Abel, A. (1983). Optimal investment under uncertainty. *American Economic Review 73*, 228–233.

Abel, A. (1984). The effects of uncertainty on investment and the expected long-run capital stock. *Journal of Economic Dynamics and Control 7*, 39–53.

Abel, A. (1985). A stochastic model of investment, marginal q and the market value of the firm. *International Economic Review 26*, 305–322.

Anastasopoulos, P. (1984). *Interdependencies in International Travel: The Role of Relative Prices - A Case Study of the Mediterranean Region*. Ph. D. thesis, New School for Social Research.

Askari, H. (1975). Investment by the tourist industry in four European countries. In S. P. Ladany (Ed.), *Management Science Applications to Leisure-time Operations*. New York: American Elsevier Publishing Company.

Bates, J. and C. Granger (1969). The combination of forecasts. *Operations Research Quarterly 20*, 451–468.

Bonham, C. and B. Gangnes (1996). Intervention analysis with cointegrated time series: the case of the Hawaii hotel room tax. *Applied Economics 28*, 1281–1293.

Brewbaker, P. (1994). A dynamic model of hotel investment in Hawaii. University of Hawaii Department of Economics mimeo.

Bull, A. (1991). *The Economics of Travel and Tourism*. Pitman.

Caballero, R. (1999). Aggregate investment. In J. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1B, pp. 813–862. Amsterdam; New York and Oxford: Elsevier Science, North-Holland.

Carey, K. (1992). Optimal hotel capacity: the case of Barbados. *Social and Economic Studies 41*, 103–126.

Chirinko, R. (1993). Business fixed investment spending: modeling strategies, empirical results and policy implications. *Journal of Economic Literature 31*, 1875–1911.

Chong, Y. and D. Hendry (1986). Econometric evaluation of linear macro-economic models. *Review of Economic Studies 53*, 671–690. Reprinted in Granger, C.W. J. (ed.) (1990) *Modelling Economic Series*. Oxford:Clarendon Press.

Copeland, B. (1989). Taxing tourists: optimal commodity taxation and public goods provision in the presence of international tourism. University of British Columbia mimeo.

Copeland, B. (1990). Tourism, welfare and de-industrialization in a small open economy. University of Alberta research paper No. 90-11.

Crouch, G. (1991). Determinants of international tourist flows: a multi-variable regression analysis. Australian Tourist Commission Report.

Crouch, G. (1994a). The study of international tourism demand: a survey of practice. *Journal of Travel Research 32*, 41–55.

Crouch, G. (1994b). The study of international tourism demand: a review of findings. *Journal of Travel Research 33*, 12–23.

Crouch, G., L. Schultz, and P. Valerio (1992). Marketing international tourism to Australia: a regression analysis. *Tourism Management 13*, 196–208.

Dickey, D. and W. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association 74*, 427–431.

Dickey, D. and W. Fuller (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica 49*, 1057–1072.

Dixit, A. and R. Pindyck (1995). *Investment Under Uncertainty*. Princeton, New Jersey: Princeton University Press.

Edwards, A. (1995). Asia-pacific travel forecasts to 2005. Economist Intelligence Unit Research Report.

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica 50*, 987–1007.

Engle, R. and C. Granger (1987). Cointegration and error correction representation: estimation and testing. *Econometrica 55*, 251–276.

Eymann, A. and G. Ronning (1994). Microeconomic models of tourists' destination choice. University of California, Berkeley, mimeo.

Fujii, E., M. Khaled, and J. Mak (1985). The exportability of hotel occupancy and other tourist taxes. *National Tax Journal 38*, 169–177.

Fujii, E. and J. Mak (1981). Forecasting travel demand: some methodological issues. *Annals of Regional Science 15*, 73–83.

Fuller, W. (1976). *Introduction to Statistical Time Series*. Wiley, New York.

Gangnes, B. and C. Bonham (1998). A tourism forecasting model for the state of hawai'i: Final report. Technical report, University of Hawaii Economic Research Organzation. Consultant's report to the Research and Economic Analysis Division, State of Hawai'i Department of Business, Economic Development and Tourism.

Granger, C. and P. Newbold (1973). Some comments on the evaluation of economic forecasts. *Applied Economics 5*, 35–47.

Granger, C. and P. Newbold (1974). Spurious regressions in econometrics. *Journal of Econometrics 2*, 111–120.

Gray, H. (1970). *International Travel-International Trade*. Lexington, Massachusetts: Heath Lexington Books.

Haldrup, N. (1998). An econometric analysis of I(2) variables. *Journal of Economic Surveys 12*, 595–650.

Hall, S. G., S. Henry, and J. Greenslade (2002). On the identification of cointegrated systems in small samples: a modelling strategy with an application to UK wages and prices. *Journal of Economic Dynamics and Control 26*, 1517–1537.

Harbo, I., S. Johansen, B. Nielsen, and A. Rahbek (1998). Asymptotic inference on cointegrating rank in partial systems. *Journal of Business & Economic Statistics 16*, 388–399.

Hartman, R. (1972). The effects of price and cost uncertainty on investment. *Journal of Economic Theory 5*, 258–266.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control 12*, 231–254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models. *Econometrica 59*, 1551–1580.

Johansen, S. (1992). Cointegration in partial systems and the efficiency of single-equation analysis. *Journal of Econometrics 52*, 389–402.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. London: Oxford University Press.

Kim, S. and H. Song (1998). An empirical analysis of demand for Korean tourism: a cointegration and error correction approach. *Tourism Economics 3*, 25–41.

Kliman, M. (1981). A quantitative analysis of Canadian overseas tourism. *Transportation Research 15*, 487–497.

Kulendran, N. (1996). Modeling quarterly tourism flows to Australia using cointegration analysis. *Tourism Economics 2*, 203–222.

Lathiras, P. and C. Siriopoulos (1998). The demand for tourism to Greece: a cointegration approach. *Tourism Economics 4*, 171–185.

Lim, C. (1997). An econometric classification and review of international tourism demand models. *Tourism Economics 3*, 69–81.

Maddala, G. (1992). *Introduction to Econometrics* (2nd ed.). Macmillan Publishing Company.

Mak, J. (1997). Japanese investment in Hawaii: past and future. University of Hawaii Economic Research Organization (UHERO) research report.

Mak, J., J. Moncur, and D. Yonamine (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. and S. Witt (1987). Tourism demand forecasting models: choice of appropriate variable to represent tourists' cost of living. *Tourism Management 8*, 223–246.

McDonald, R. and D. Siegel (1986). The value of waiting to invest. *Quarterly Journal of Economics 101*, 707–727.

Morely, C. (1992). A microeconomic theory of international tourism demand. *Annals of Tourism Research 19*, 250–267.

Nelso, C. (1972). The prediction performance of the FRB-MIT-PENN model of the US economy. *American Economic Review 62*, 902–917.

O'Hagan, J. and M. Harrison (1984). Market shares of U.S. tourist expenditures in Europe: an econometric analysis. *Applied Economics 16*, 919–931.

Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis. *Econometrica 57*, 1361–1401.

Perron, P. (1990). Testing for a unit root in a time series with a changing mean. *Journal of Business and Economic Statistics 8*, 153–162.

Pesaran, M. and Y. Shin (1995). An autoregressive distributed lag modeling approach to cointegration analysis. In S. Strom, A. Holly, and P. Diamond (Eds.), *Centennial Volume of Rangar Frisch*. Cambridge: Cambridge University Press. Econometric Society Monograph.

Pesaran, M. and Y. Shin (2001). Long-run structural modeling. University of Cambridge DAE Working Paper No.9419.

Pesaran, M., Y. Shin, and R. Smith (2000). Structural analysis of vector error correction models with exogenous I(1) variables. *Journal of Econometrics 97*, 293–343.

Pesaran, M. and R. Smith (1998). Structural analysis of cointegrating VARs. *Journal of Economic Surveys 12*, 471–505.

Phillips, P. (1991). Optimal inference in cointegrated systems. *Econometrica 59*, 283–306.

Phillips, P. (1995). Fully modified least squares and vector autoregressions. *Econometrica 63*, 1023–1078.

Quandt, R. (1970). *The Demand for Travel: Theory and Measurement*. Lexington, Mass.: Heath Lexington Books.

Saikkonen, P. (1993a). Estimation of cointegration vectors with linear restrictions. *Econometric Theory 9*, 19–35.

Saikkonen, P. (1993b). Continuous weak convergence and stochastic equicontinuity results for integrated processes with an application to the estimation of a regression model. *Econometric Theory 9*, 155–188.

Sheldon, P. (1993). Forecasting tourism: expenditure versus arrivals. *Journal of travel research 32*, 13–20.

Sims, C. (1980). Macroeconomics and reality. *Econometrica 48*, 1–48.

Sinclair, M. and M. Stabler (1997). *The Economics of Tourism*. London: Routledge.

Song, H., P. Romilly, and X. Liu (2000). An empirical study of outbound tourism demand in the UK. *Applied Economics 32*, 611–24.

Song, H. and S. Witt (2000). *Tourism Demand Modelling and Forecasting*. Kidlington, Oxford, UK: Elsevier Science Ltd.

Taylor, L. (1995). Taxing sales to tourists over time. Tulane University mimeo.

Varian, H. (1992). *Microeconomic Analysis* (3rd ed.). W. W. Norton & Company.

Vogt, M. and C. Wittayakorn (1998). Determinants of the demand for Thailand's exports of tourism. *Applied Economics 30*, 711–715.

White, H. (1980). A heteroskedastic-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica 48*, 817–838.

Wickens, M. and T. Breusch (1988). Dynamic specification, the long-run and the estimation of transformed regression models. *Economic Journal 98*, 189–205.

Witt, S. and C. Martin (1995). Forecasting tourism demand: a review of empirical research. *International Journal of Forecasting 11*, 447–475.

Witt, S. and C. Witt (1992). *Modelling and Forecasting Demand in Tourism*. London: Academic Press.

Witt, S. and C. Witt (1995). Forecasting tourism demand: a review of empirical research. *International Journal of Forecasting 11*, 447–475.

Yule, G. (1926). Why do we sometimes get nonsense correlations between time-series? a study in sampling and the nature of time series. *Journal of the Royal Statistical Society 60*, 812–854.