



UHERO
THE ECONOMIC RESEARCH ORGANIZATION
AT THE UNIVERSITY OF HAWAII

**ESTIMATING DEMAND ELASTICITIES IN
NON-STATIONARY PANELS:
THE CASE OF HAWAII TOURISM**

BY

PETER FULEKY, CARL S. BONHAM AND QIANXUE ZHAO

Working Paper No. 2013-2R

February 2013

revised: August 15, 2013

This paper is funded (" in part " if appropriate) by a grant/cooperative agreement from the National Oceanic and Atmospheric Administration, Project R/TR-7PD, which is sponsored by the University of Hawaii Sea Grant College Program, SOEST, under Institutional Grant No. NA05OAR4171048 from NOAA Office of Sea Grant, Department of Commerce. The views expressed herein are those of the author(s) and do not necessarily reflect the views of NOAA or any of its subagencies. UNIHI-SEAGRANT-TT-13-02.

UNIVERSITY OF HAWAII AT MANOA
2424 MAILE WAY, ROOM 540 • HONOLULU, HAWAII 96822
WWW.UHERO.HAWAII.EDU

WORKING PAPERS ARE PRELIMINARY MATERIALS CIRCULATED TO STIMULATE DISCUSSION AND CRITICAL COMMENT. THE VIEWS EXPRESSED ARE THOSE OF THE INDIVIDUAL AUTHORS.

Estimating Demand Elasticities in Non-Stationary Panels: The Case of Hawaii Tourism

Peter Fuleky¹, Carl S. Bonham², and Qianxue Zhao

Abstract

It is natural to turn to the richness of panel data to improve the precision of estimated tourism demand elasticities. However, the likely presence of common shocks shared across the underlying macroeconomic variables and across regions in the panel has so far been neglected in the tourism literature. We deal with the effects of cross-sectional dependence by applying Pesaran's (2006) common correlated effects estimator, which is consistent under a wide range of conditions and is relatively simple to implement. We study the extent to which tourist arrivals from the US Mainland to Hawaii are driven by fundamentals such as real personal income and travel costs, and we demonstrate that ignoring cross-sectional dependence leads to spurious results.

Keywords: Panel Cointegration, Cross-Sectional Dependence, Tourism Demand, Hawaii

JEL: C23, C51, L83, R41

¹Corresponding address: University of Hawaii Economic Research Organization, and Department of Economics, University of Hawaii, 540 Saunders Hall, 2424 Maile Way, Honolulu, HI 96822, email: fuleky@hawaii.edu.

²This research was funded in part by a grant/cooperative agreement from the National Oceanic and Atmospheric Administration, Project R/TR-7PD, which is sponsored by the University of Hawaii Sea Grant College Program, SOEST, under Institutional Grant No. NA05OAR4171048 from NOAA Office of Sea Grant, Department of Commerce. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its subagencies. UNIHI-SEAGRANT-TT-13-02.

1. INTRODUCTION

The past several decades have seen tremendous growth in the literature which seeks to explain and forecast tourist flows (Song, Dwyer, Li & Cao, 2012; Li, Song & Witt, 2005). Following a wide variety of empirical methods applied across different countries and time periods, researchers have produced an even wider array of estimates for demand elasticities central to marketing, forecasting and policy work. While the income elasticity of tourism demand is generally expected to lie between one and two, Crouch (1995, 1996) found that nearly 5% of estimates from 80 international studies were negative. Analyzing 30 years of international tourism demand studies, Witt and Witt (1995) found income elasticity estimates ranging from 0.4 to 6.6 with a median value of 2.4.

Although an “inferior” tourist destination could explain a negative income elasticity, and an elasticity of less than 1 might be explained by some “necessary” short-haul international trips, such as those from the US to Canada, the large variation in estimates calls into question their validity and limits their usefulness to decision-makers. Estimates of price elasticities fare about the same. Witt and Witt (1995) found estimates ranging from -0.05 to -1.5, and Crouch (1995, 1996) found that about 29% of the estimates were positive. Finally, these studies found transportation price elasticity estimates ranging from 0.11 to -4.3. Crouch (1996) investigated a number of potential causes of such disparate results, and noted that model specification played an important role. We suspect that the wide range of elasticity estimates is due to the limited information and short samples used in time series models, and the use of panel estimation techniques that do not ade-

quately deal with important characteristics of panel the data.

There is a rich literature making use of a variety of approaches to explain and forecast tourist flows. Conventional approaches range from exponential smoothing to vector autoregressive and error correction models using time series data for a single origin-destination pair (Witt & Witt, 1995; Li, Song & Witt, 2005). Recently some alternative quantitative tools, such as artificial neural networks, fuzzy time series, and genetic algorithms, have been showing up in the literature. (For a comprehensive survey of recent developments in tourism demand modeling, see Song and Li (2008).) Unfortunately, the entire literature on tourism demand is at the mercy of short time series samples. And the limited data available for estimation has likely contributed to imprecise estimates of demand elasticities. For example, Bonham, Gangnes and Zhou (2009, p. 541) report an income elasticity for Hawaii tourism demand from the US that is “implausibly large and estimated quite imprecisely”. Fortunately, it may be possible to obtain better estimates of the parameters of interest by taking advantage of the variation in both the temporal and cross-sectional dimensions of panel data sets. This point has not been lost on the tourism literature, and as econometric tools have advanced, a trend to exploit the richness of panel data has emerged (Song & Li, 2008; Seetaram & Petit, 2012). In their review article, Song et al. (2012, p. 1657) suggest that “future studies should pay more attention to the dynamic version of panel data analysis and to more advanced estimation methods...”.

While early panel studies ignored problems arising from nonstationarity and potential cointegration, the tourism literature has now begun to address such is-

sues. Among others, Seetanah, Durbarry and Ragodoo (2010) estimated a static model of demand for South African tourism using Fully Modified Ordinary Least Squares (FMOLS) developed by Pedroni (2001). Using the same technique, Lee and Chang (2008) investigated the long-run co-movements and causal relationships between tourism development and economic growth. Falk (2010) applied the dynamic heterogeneous panel technique of Pesaran, Shin and Smith (1999) to estimate the effects of snow fall on winter tourism in Austria.

One common thread running through this nascent literature is reliance on the assumption of cross-sectional independence, or that each unit contributes entirely new information to the dataset. Yet, cross-sectional units are almost certainly influenced by national or global shocks such as business cycles, technological innovations, terrorism events, oil crises or national fiscal and monetary policies. In fact, a large empirical macro and macro-finance literature (see Stock & Watson, 1989, 1998) and results presented here for Hawaii tourism show that cross-sectional dependence is very common. And, neglecting cross-sectional dependence can lead to substantial bias in conventional panel estimators (Kapetanios, Pesaran & Yamagata, 2011).

An increasingly common solution to the problem of cross-sectional dependence is to model such dependence using a factor structure. To the best of our knowledge, this approach has not been used in the tourism literature where cross-sectional dependence is usually ignored. But at least one study has included observed proxies for unobserved common factors. Nelson, Dickey and Smith (2011) used oil prices, indicator variables for the effects of the September 11, 2001 terror-

ist attacks, and a nonlinear time trend capturing the overall slow-down of tourism demand during recessions. Such proxy variables may be effective in mitigating the effects of cross-sectional dependence, but their choice involves judgement on the part of the researcher, and it is unclear whether they are adequate to capture all sources of common shocks. Alternatively, unobserved dynamic common factors can be approximated using the methods proposed by Bai, Kao and Ng (2009), Pesaran (2006), or Kapetanios et al. (2011). These approaches have the benefit that they do not require selection of a set of observed proxies.

We estimate tourism demand elasticities from a panel of tourist arrivals to Hawaii from 48 US states over 19 years using the common correlated effects (CCE) estimator of Pesaran (2006) and Kapetanios et al. (2011). This technique offers many advantages. First, the CCE estimator allows us to deal with the possibility of cross-sectional dependence caused by common factors. Second, it does not require ex ante information about the unobserved common factors, allows the factors to contain unit roots and to be correlated with the regressors. Finally, the CCE estimator offers good finite sample properties (Kapetanios et al., 2011; Westerlund & Urbain, 2011), and is relatively simple to implement.

The rest of this paper is organized as follows: in Section 2 we outline our tourism demand model and describe the CCE estimator we use to deal with cross-sectional dependence in panels; in Section 3 we present panel estimates of demand elasticities for Hawaii tourism; and Section 4 concludes.

2. ESTIMATING TOURISM DEMAND ELASTICITIES FROM PANELS WITH UNOBSERVED COMMON FACTORS

Empirical models of tourism demand borrow heavily from consumer theory which suggests that the most important factors affecting the tourist's budget constraint are income, the price of tourism services, and the prices of related goods (substitutes and complements). The demand for aggregate tourist flows from origin i to destination j can be written as

$$D_{ij} = f(Y_i, P_i, P_j, P_s), \quad (1)$$

where D_{ij} is a measure of tourism demand in destination j by consumers from origin i ; Y_i is the level of income at origin i ; P_i is the price of other goods and services at origin i ; P_j is the price of tourism goods and services at destination j ; P_s is the price of tourism products at competing destinations (Bonham et al., 2009). Assuming homogeneity, demand can be written as a function of real income, and relative prices

$$D_{ij} = f\left(\frac{Y_i}{P_i}, \frac{P_j}{P_i}, \frac{P_s}{P_i}\right). \quad (2)$$

If travel to destination j is assumed to compete with short distance trips near origin i , the tourism demand model can be simplified to

$$D_{ij} = f\left(\frac{Y_i}{P_i}, \frac{P_j}{P_i}\right). \quad (3)$$

It is common for researchers to augment income and price variables with deterministic variables such as time trends to capture evolving consumer tastes or secular growth or decline in an industry; a constant term to account for destination amenities such as natural assets or other factors that are time invariant; dummies to account for one-time events such as terrorism, natural disasters, major sporting events, and oil crises; seasonality; or changes in data definitions or collection methods. These types of events, if otherwise neglected, might lead to bias in the estimated parameters (Bonham et al., 2009). The method described in this section deals with such deterministic effects the same way it deals with non-deterministic common factors such as business cycles. As a result, we do not need to subjectively select deterministic proxies for such events.

Equation (3) can be written in the following log-linear form

$$y_{it} = \alpha_i + \beta_i' \mathbf{x}_{it} + u_{it}, \quad i = 1, 2, \dots, N, \quad t = 1, 2, \dots, T, \quad (4)$$

where $y_{it} = \log(D_{ij,t})$, $\mathbf{x}_{it} = \left(\log\left(\frac{Y_{i,t}}{P_{i,t}}\right), \log\left(\frac{P_{j,t}}{P_{i,t}}\right) \right)'$. The coefficients β_i represent the elasticities of demand with respect to the regressors \mathbf{x}_{it} . The dynamics and the common unobserved factors are modeled in the error terms u_{it} ,

$$u_{it} = \gamma_i' \mathbf{f}_t + \epsilon_{it}, \quad (5)$$

where \mathbf{f}_t is an $m \times 1$ vector of unobserved common effects, and ϵ_{it} are the individual-specific (idiosyncratic) errors assumed to be distributed independently of \mathbf{x}_{it} and \mathbf{f}_t . The ϵ_{it} are, however, allowed to be weakly dependent across i , and serially

correlated over time. A valuable feature of the model is that the error term, u_{it} , is allowed to be correlated with the regressors, \mathbf{x}_{it} , through the presence of the factors, \mathbf{f}_t , in both. Specifically, the $k \times 1$ vector \mathbf{x}_{it} is assumed to follow the factor structure

$$\mathbf{x}_{it} = \mathbf{a}_i + \mathbf{\Gamma}'_i \mathbf{f}_t + \mathbf{v}_{it}, \quad (6)$$

where \mathbf{a}_i is a $k \times 1$ vector of individual effects, and $\mathbf{\Gamma}_i$ is a $m \times k$ factor loading matrix. The idiosyncratic components \mathbf{v}_{it} are distributed independently of the common effects and across i , but assumed to follow general covariance stationary processes. Finally, the assumption that ϵ_{it} (in equation 5) is stationary implies that if \mathbf{f}_t contains unit root processes then y_{it} , \mathbf{x}_{it} , and \mathbf{f}_t must be cointegrated.

Because the error term, u_{it} , contains common factors that are correlated with the regressors, failure to account for this correlation will generally produce biased estimates of the parameters of interest. Pesaran (2006) suggested using cross section averages of y_{it} and \mathbf{x}_{it} to deal with the effects of the unobserved factors. His CCE estimator is defined as,

$$\hat{\boldsymbol{\beta}}_i = (\mathbf{X}'_i \bar{\mathbf{M}} \mathbf{X}_i)^{-1} \mathbf{X}'_i \bar{\mathbf{M}} \mathbf{y}_i, \quad (7)$$

where $\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{iT})'$, $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{iT})'$, and $\bar{\mathbf{M}} = \mathbf{I}_T - \bar{\mathbf{H}}(\bar{\mathbf{H}}' \bar{\mathbf{H}})^{-1} \bar{\mathbf{H}}'$ with $\bar{\mathbf{H}} = (\boldsymbol{\iota}, \bar{\mathbf{X}}, \bar{\mathbf{y}})$, and $\boldsymbol{\iota}$ is a $T \times 1$ vector of ones. $\bar{\mathbf{X}}$ is a $T \times k$ matrix of cross-sectional means of the k regressors, and $\bar{\mathbf{y}}$ is a $T \times 1$ vector of cross-sectional means of the dependent variable.

The CCE estimator is equivalent to ordinary least squares applied to an auxil-

iary regression that is augmented with the cross-sectional means of the variables

$$y_{it} = \alpha_i + \beta_i' x_{it} + \theta_i' \bar{h}_t + \eta_{it}, \quad i = 1, 2, \dots, N, \quad t = 1, 2, \dots, T, \quad (8)$$

where $\bar{h}_t = (\bar{x}_t, \bar{y}_t)$, and the θ_i vector contains the individual specific loading coefficients of the cross-sectional averages. Thus, the CCE estimator of the β_i coefficients captures the effect of the demand determinants after controlling for co-movement across units. By allowing for heterogeneous loadings, θ_i , the CCE estimator is better able to control for common factors than estimators based on cross-sectional demeaning. This latter approach imposes unit loading coefficients on the cross-sectional means, which is inappropriate if the impact of global trends varies across regions.

While Pesaran (2006) derived the CCE estimator for stationary variables and factors, Kapetanios et al. (2011) proved that the CCE estimators are consistent regardless of whether the common factors, f_t , are stationary or non-stationary. In addition, they showed that the CCE estimator of the mean of the slope coefficients is consistent for any number of factors. These two results are important benefits of the CCE estimator: we do not need to know how many common factors exist, whether or not the common factors are stationary, or even provide estimates of the common factors and their loadings. In contrast, Bai et al. (2009) estimate homogeneous slope coefficients jointly with common factors using an iterative procedure. But the precision of that approach substantially depends on prior knowledge of the number of unobserved factors. Moreover, Kapetanios et al. (2011) and West-

erlund and Urbain (2011) have shown that the CCE estimators have lower bias than those of Bai et al. (2009) even if the true number of factors is known.

The objective of this paper is to obtain the best possible estimates for income, price, and travel cost elasticities of demand for Hawaii tourism. While the slope coefficients, β_i , are allowed to differ across origins, we use the CCE mean group estimator (CCE-MG) to estimate the overall effect of demand determinants. The CCE-MG estimator is a simple average of the individual CCE estimators, β_i ,

$$\hat{\beta}_{CCE-MG} = \frac{1}{N} \sum_{i=1}^N \hat{\beta}_i, \quad (9)$$

with variance

$$\widehat{Var}(\hat{\beta}_{CCE-MG}) = \frac{1}{N(N-1)} \sum_{i=1}^N (\hat{\beta}_i - \hat{\beta}_{CCE-MG})(\hat{\beta}_i - \hat{\beta}_{CCE-MG})'. \quad (10)$$

This approach is justified under the assumption of a random coefficient model where, $\beta_i = \beta + w_i$, and $w_i \sim IID(\mathbf{0}, \mathbf{V}_w)$, so that the overall demand elasticities are $\beta = E(\beta_i)$. When the slope coefficients, β_i , are homogeneous, efficiency gains can be achieved by pooling observations over the cross section units. The CCE pooled estimator (CCE-P) developed by Pesaran (2006) has the form

$$\hat{\beta}_{CCE-P} = \left(\sum_{i=1}^N X_i' \bar{M} X_i \right)^{-1} \sum_{i=1}^N X_i' \bar{M} y_i. \quad (11)$$

The CCE estimators are based on orthogonal projections of the variables onto proxies for common factors. That is, the $\hat{\beta}_{CCE-MG}$ and $\hat{\beta}_{CCE-P}$ coefficient estimates

are computed from idiosyncratic variation in the panels and are not affected by global fluctuations, such as business cycles. Consequently, the CCE estimator is able to isolate the marginal effects of the regressors, and avoid the bias that results when common factors are ignored.

In the next section we describe our empirical model and report estimation results.

3. ESTIMATING DEMAND ELASTICITIES FOR HAWAII

The goal of our study is to estimate demand elasticities for tourism from the US mainland to Hawaii using a panel of tourist arrivals from 48 states spanning 19 years. We approximate tourism demand by tourist arrivals ($TOUR_{it}$). Arrivals is frequently used in models explaining aggregate tourist flows to a single destination (inbound modeling) (Li, Song & Witt, 2005; Song & Li, 2008), and high frequency expenditure data by origin is not available for Hawaii visitors. While researchers modeling inbound tourist flows sometimes use tourist expenditures, tourist nights, or other proxies, of the 81 tourism demand studies reviewed by Song and Li (2008), nearly 70% choose the number of tourist arrivals as the measure of demand.

Empirical tourism demand studies commonly estimate the elasticities for a small set of demand determinants such as income at the origin country/region, the own price of a destination, and possibly substitute prices of alternative destinations (Song et al., 2012). We focus on two types of demand determinants: income and relative prices. Witt and Witt (1995) recommend using personal income to

predict holiday travel, and a more general income measure, such as national income, to predict business travel. Because the vast majority (over 70%) of tourists come to Hawaii for holiday, we use total personal income by state (Y_{it}) as the measure of income.

Several types of prices appear in the demand specification. The first is the own price of tourism products. Witt and Witt (1995) suggest that the price measures should include both, travel costs and the cost of living at the destination. Accordingly, we include two price variables in our model: the price of airfare ($PAIR_{it}$) and the price of renting a hotel room (PRM_{it}). The second type of prices used in tourism demand studies are measures of substitute prices. Because local travel in the origin region may substitute for travel to the destination, the price level at the origin is often included as a proxy for substitute prices. We deflate nominal variables using the consumer price index at the origin (CPI_{it}), so that prices enter the model in relative terms. The model to be estimated may be written as

$$\log TOUR_{it} = \alpha_i + \beta_{1i} \log Y_{it}^* + \beta_{2i} \log PAIR_{it}^* + \beta_{3i} \log PRM_{it}^* + u_{it} , \quad (12)$$

where $Y_{it}^* = \frac{Y_{it}}{CPI_{it}} \times 100$, $PAIR_{it}^* = \frac{PAIR_{it}}{CPI_{it}} \times 100$, and $PRM_{it}^* = \frac{PRM_{it}}{CPI_{it}} \times 100$.

3.1. Data

Our sample begins in the first quarter of 1993 and ends in the first quarter of 2012 for a total of 77 periods for each of the 49 states. The sample size is determined by the availability of data on tourist arrivals from the US mainland to Hawaii. These data are from various reports of the Hawaii Department of Busi-

ness, Economic Development and Tourism, Hawaii Tourism Authority, the Hawaii Visitors and Convention Bureau, and the Hawaii Visitors Bureau. Tourist arrivals data are available monthly for all years except 1995 and 1997, for which we use interpolated annual values. Data for total personal income is from the Bureau of Economic Analysis, and airfares to Hawaii are from the DB1B Market database of the Bureau of Transportation Statistics, which offers a 10% random sample of all domestic trips each quarter. From the available sample, we calculate the median airfare for each state and each quarter. We exclude Delaware and the District of Columbia from our analysis due to the lack of airfare data.

The Hawaii statewide average hotel room rate is from Hospitality Advisors LLC. The consumer price index is from the US Bureau of Labor Statistics. Because the CPI is only reported at the metropolitan level, we proxy state consumer prices using the CPI for metropolitan areas within the state. Where a metropolitan area CPI is not available, we use the CPI for the region as a proxy for state consumer prices. The CPI data is reported at a variety of frequencies, and we linearly interpolate the lower frequency series to approximate their values at the highest (monthly) frequency. We aggregate all monthly series to the quarterly frequency and seasonally adjust the data using the X-12 ARIMA method where necessary. (State level personal income is seasonally adjusted by the Bureau of Economic Analysis.)

The CCE estimator produces consistent estimates of the demand elasticities in equation (12) regardless of whether the series in the model are stationary or non-stationary. Nevertheless, before proceeding, we explore the stochastic prop-

erties of the arrivals, real income, and relative price panels. The variables used in the model are plotted in Figure 1, where they are standardized so that all cross-sectional units fit into a single plot. Estimation is carried out without this standardization. Table 1 reports descriptive statistics for each variable along with results from cross-sectional dependence, structural break, and unit root tests.

[Insert Figure 1 about here]

Franses and Haldrup (1994) show that additive outliers could produce spurious stationarity, and lead researchers to over-reject the null of a unit-root. To avoid such issues, we follow the procedure suggested by Perron and Rodríguez (2003) to test for and remove additive outliers in all series. After seasonal adjustment and outlier removal, we test the variables for cross-sectional dependence using Pesaran's (2004) CD test, which is based on the average pairwise correlation of individual cross-section units. The results in the bottom panel of Table 1 lead us to reject the null hypothesis of cross-sectional independence for each variable.

Next we test each variable for unit roots. A wide variety of panel unit root tests is available in the literature (Breitung & Pesaran, 2008). We chose to apply tests that have the best statistical properties for the hypothesis of interest. For example, because the CD test rejects the hypothesis of cross-sectional independence for each variable, the hypothesis of interest is whether the common factors contain unit roots. We first test the common components in each variable for structural breaks using the $Exp - W_{FS}$ test of Perron and Yabu (2009). The null hypothesis of no breaks is rejected for the common components of real income and relative

airfare. Therefore, we test these series for non-stationarity using the P_T^{GLS} unit root test of Carrion-i Silvestre, Kim and Perron (2009), which takes into account the presence of breaks.

For tourist arrivals and the relative room price, we cannot reject the null hypothesis of no breaks. Therefore we test these series for unit roots using Sul's (2009) *CRMA* test and Bai and Ng's (2004) principal components based *PANIC_f* test of the common factors. (To distinguish between the *PANIC* test applied to common factors versus the one applied to idiosyncratic components, we denote the former test with an *f* subscript and the later test with an *i* subscript.) Results are presented in the bottom panel of Table 1. We cannot reject the null hypothesis of a unit root for any of the variables.

[Insert Table 1 about here]

The simulations in Gengenbach et al. (2010) demonstrate that the finite sample performance of the *PANIC* test suffers from several issues which do not affect the *CRMA* test. The main drawback of Bai and Ng's (2004) *PANIC* test is its reliance on the assumption that the number of common factors and their paths are known. In practice both need to be estimated, and the estimation error can lead to misleading inference. The *CRMA* test avoids these complications by approximating all common factors in the data via a simple cross-sectional average. For these reasons we prefer the *CRMA* test, but we display both tests in Table 1.

3.2. Results

To illustrate the impact of cross-sectional dependence on estimates of demand elasticities, Table 2 compares Pedroni's (2001) FMOLS estimates, commonly used in the tourism literature (Seetaram & Petit, 2012), with Pesaran's (2006) mean group and pooled common correlated effects (CCE-MG and CCE-P) estimates. The FMOLS income elasticity estimate is small and the room price elasticity has the wrong sign.

[Insert Table 2 about here]

The bottom panel of Table 2 presents diagnostic tests for the FMOLS residuals. The t_{PP} , and t_{ADF} are the Pedroni (1999) tests for the null hypothesis of no-cointegration based on the Phillips and Perron (1988) t -statistic, and the augmented Dickey and Fuller (1979) t -statistic, respectively. Both tests assume cross-sectional independence, and both reject the null of no cointegration. However, Pesaran's (2004) CD test rejects the cross-sectional independence of the FMOLS residuals, suggesting that the FMOLS estimates are biased due to the presence of common factors in the residuals (Kao & Chiang, 2000). Moreover, failure to control for common factors shared across variables, such as national business cycles, confounds the relationship between the regressors and the dependent variable and is likely the source of the wrong sign associated with the FMOLS hotel room price elasticity.

The significant co-movement of the FMOLS residuals is illustrated in Figure 2. We test the common component in the FMOLS residuals for breaks using Perron

and Yabu's (2009) $Exp - W_{FS}$ test. Because we reject the presence of breaks, we apply Sul's (2009) $CRMA$ test and Bai and Ng's (2004) $PANIC_f$ test to the common factors in the residuals. We are unable to reject the null of a unit root in the FMOLS residuals using either test. The implication is that the FMOLS estimates are spurious, and ignoring cross-sectional dependence would lead to the acceptance of the invalid FMOLS results.

[Insert Figure 2 about here]

In addition to the biased FMOLS estimates based on a spurious panel regression, Table 2 also contains the CCE estimates (top panel) and CCE residual diagnostic tests (bottom panel). Kapetanios et al. (2011) showed that the CCE estimator of the mean slope coefficient is consistent for any number of factors, even if the common factors contain unit roots. However, consistent estimation of the model parameters requires that the regression residuals be stationary. Because the CCE regression filters out common factors, the hypothesis of interest is that the idiosyncratic components of the CCE residuals contain a unit root.

We test for unit roots in the CCE regression residuals using Pesaran's (2007) $CIPS$ test, a simple cross-sectional average of cross-sectionally augmented Dickey and Fuller (1979) t statistics, and Bai and Ng's (2004) $PANIC_i$ test for the idiosyncratic components. For both CCE estimators, we reject the null of a unit root in the residuals $\hat{\epsilon}_{i,t}$ of equation (5) at the 5% level or lower. The rejection of unit roots in the CCE regression residuals implies that the observed variables and the unobserved factors are cointegrated (Kapetanios et al., 2011). Note that as long

as the residuals are stationary, the CCE pooled (CCE-P) and CCE mean-group (CCE-MG) estimators are both consistent under the random coefficient model assumption (Pesaran & Smith, 1995).

The CCE estimates presented in the top panel of Table 2 are similar to the elasticities obtained by Nelson et al. (2011), who included in their model observed and deterministic proxies for common factors, such as oil prices and a non-linear time trend. The estimated income elasticity of demand for a trip from the US mainland to Hawaii is slightly greater than unity, implying that travel to Hawaii is regarded as a luxury good. Still, our result is close to the 0.996 income elasticity of Nelson et al. (2011), and much lower than the 3.5 estimated by Bonham et al. (2009) in a VECM that suffered from a short sample period and could not exploit variation in the cross-sectional dimension.

Our results indicate that demand for Hawaii travel is inelastic with respect to airfare. If airfare increases by 10%, arrivals to the state are expected to fall by a little more than 2%. Again, this value is fairly close to -0.211, the airfare elasticity estimate of Nelson et al. (2011). The estimated hotel room price elasticity suggests that tourists are more responsive to changes in room rates than to fluctuations in airfare. Facing a \$1000 airline ticket and a daily price of \$200 for a double occupancy room, a couple on a ten-day trip has to split its budget evenly between airfare and accommodation. Still, a 10% drop in the hotel room rate is expected to generate 12% higher tourist arrivals, over five times more than a corresponding drop in airfare.

This could be explained by a two stage decision making process on the part of travelers: in the first stage they choose a destination from a range of competing locations, and in the second stage they pick their flights. The choice in the first stage may be influenced by offerings of free nights, attraction packages, and the overall desirability of the destination. Because the choice of flights is relegated to the second stage, travelers largely focus on minimizing their airfare to the chosen destination, but do not necessarily switch to competing destinations. As a result, a shift in median airfare from a particular origin to Hawaii only has a modest impact on arrivals. The idea of a two-stage decision making process dates back to the monograph of Strotz (1957), who described the rationale for first deciding how to allocate a budget among several groups of goods and then making independent spending decisions within the groups. In a tourism context, multi-stage decision making has been discussed by Syriopoulos and Sinclair (1993), and Song et al. (2012). And, a number of studies have relied on this framework when studying tourism demand. (See Bonham and Gangnes (1996), Nicolau and Más (2005), and Eugenio-Martin and Campos-Soria (2011).)

The Hawaii hotel room rate is independent of trip origin, and it can be considered an observed common factor. Because the CCE estimation procedure is based on an orthogonal projection onto proxies for common factors, the exclusion of the real room rate only has a minor effect on the results: the CCE income elasticity and airfare elasticity estimates are similar for the model with and without lodging prices. Results with the room rate excluded are available in a working paper version or upon request.

4. CONCLUSION

The dramatic growth in tourism over the past several decades has led to an extensive literature which seeks to quantify the effect of income and prices on tourist flows. Although estimates of demand elasticities are central to marketing, forecasting and policy work, the literature on tourism demand has produced a range of elasticity estimates that are occasionally at odds with economic theory and reduce their usefulness in decision making. To improve the precision of estimates, it is natural to turn to the richness of panel data exhibiting variation in both the temporal and the cross-sectional dimension. This point has also been realized by scholars in the tourism literature, and as econometric tools have advanced, a trend to exploit the greater information content of panel data has emerged. However, panel estimation using non-stationary data requires careful attention to the likely presence of common shocks in the underlying macroeconomic variables.

Early panel studies of tourism demand have relied on the assumption of cross-sectional independence, or that each region contributes entirely new information to the dataset. Yet, cross-sectional units are generally influenced by national or global shocks such as business cycles, technological innovations, terrorism events, oil crises or national fiscal and monetary policies. We demonstrate that neglecting cross-sectional dependence leads to spurious estimation results. Our contribution to the literature lies in estimating tourism demand elasticities while accounting for unobserved non-stationary common factors in the data. We use the CCE estimators of Pesaran (2006) and Kapetanios et al. (2011) to deal with cross-sectional dependence in panel regressions. This technique offers several advantages over

competing methods: it does not require ex ante information about the unobserved common factors, allows them to contain unit roots and to be correlated with the regressors, allows for heterogeneous factor loadings, exhibits good finite sample properties, and is relatively simple to implement.

We apply the CCE estimators to US state level quarterly data spanning the period from the first quarter of 1993 to the first quarter of 2012. We obtain an income elasticity for Hawaii tourism demand that is slightly greater than one, elastic demand with respect to hotel room prices, and inelastic demand with respect to airfare. Our estimates are more plausible than those of Bonham et al. (2009) whose pure time series model did not benefit from cross-sectional variation, and are in line with the results of Nelson et al. (2011) who included in their model observed proxies for common factors, such as oil prices and a non-linear time trend.

However, the selection of observed variables as proxies requires some subjective judgement on the part of the researcher, and in general such proxies may not successfully capture all sources of unobserved common factors in the panels. In contrast, our CCE based approach does not require the selection of observable variables to capture the sources of common shocks. Instead, the method uses cross-sectional averages to directly approximate the unobserved factors in the analyzed series, and it consistently filters out those factors that are actually present in the panels.

It is important to note that our analysis is subject to some limitations. Driven primarily by data constraints we limit our analysis to a subset of relative price effects. For example, we do not consider price competition between Hawaii and

other long-haul destinations such as Mexico and the Caribbean. Furthermore, our pooled and mean group estimators produce elasticities of overall tourism demand from the US Mainland, and do not differentiate among market segments. It is likely that demand elasticities vary across regional and other types of market segments, and we intend to explore these issues in future research. Finally, while tourism demand may fluctuate with the business cycle, we are not analyzing that relationship. Because the influence of business cycles is filtered out in our estimation approach, the reported estimates represent elasticities, or isolated marginal effects of the regressors.

- Bai, J., Kao, C., & Ng, S. (2009). Panel cointegration with global stochastic trends. *Journal of Econometrics*, *149*, 82–99.
- Bai, J., & Ng, S. (2004). A panic attack on unit roots and cointegration. *Econometrica*, *72*, 1127–1177.
- Bonham, C., Gangnes, B., & Zhou, T. (2009). Modeling tourism: A fully identified vecm approach. *International Journal of Forecasting*, *25*, 531 – 549. Special Section: Time Series Monitoring.
- Bonham, C. S., & Gangnes, B. (1996). Intervention analysis with cointegrated time series: the case of the hawaii hotel room tax. *Applied Economics*, *28*, 1281–1293.
- Breitung, J., & Pesaran, M. (2008). Unit roots and cointegration in panels. *The Econometrics of Panel Data*, (pp. 279–322).
- Crouch, G. (1995). A meta-analysis of tourism demand. *Annals of Tourism Research*, *22*, 103–118.
- Crouch, G. (1996). Demand elasticities in international marketing: a meta-analytical application to tourism. *Journal of Business Research*, *36*, 117–136.
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, *74*, 427–431.

- Eugenio-Martin, J. L., & Campos-Soria, J. A. (2011). Income and the substitution pattern between domestic and international tourism demand. *Applied Economics*, *43*, 2519–2531.
- Falk, M. (2010). A dynamic panel data analysis of snow depth and winter tourism. *Tourism Management*, *31*, 912–924.
- Franses, P., & Haldrup, N. (1994). The effects of additive outliers on tests for unit roots and cointegration. *Journal of Business & Economic Statistics*, *12*, 471–478.
- Gengenbach, C., Palm, F., & Urbain, J. (2010). Panel unit root tests in the presence of cross-sectional dependencies: Comparison and implications for modelling. *Econometric Reviews*, *29*, 111–145.
- Kao, C., & Chiang, M. (2000). On the estimation and inference of a cointegrated regression in panel data. *Nonstationary Panels, Panel Cointegration, And Dynamic Panels*, *15*, 179–222.
- Kapetanios, G., Pesaran, M. H., & Yamagata, T. (2011). Panels with non-stationary multifactor error structures. *Journal of Econometrics*, *160*, 326–348.
- Lee, C., & Chang, C. (2008). Tourism development and economic growth: a closer look at panels. *Tourism Management*, *29*, 180–192.
- Li, G., Song, H., & Witt, S. (2005). Recent developments in econometric modeling and forecasting. *Journal of Travel Research*, *44*, 82–99.

- Nelson, L., Dickey, D., & Smith, J. (2011). Estimating time series and cross section tourism demand models: Mainland united states to hawaii data. *Tourism Management*, 32, 28–38.
- Nicolau, J. L., & Más, F. J. (2005). Stochastic modeling: a three-stage tourist choice process. *Annals of Tourism Research*, 32, 49–69.
- Pedroni, P. (1999). Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics*, 61, 653–670.
- Pedroni, P. (2001). Purchasing power parity tests in cointegrated panels. *Review of Economics and Statistics*, 83, 727–731.
- Perron, P., & Rodríguez, G. (2003). Searching for additive outliers in nonstationary time series. *Journal of Time Series Analysis*, 24, 193–220.
- Perron, P., & Yabu, T. (2009). Testing for shifts in trend with an integrated or stationary noise component. *Journal of Business & Economic Statistics*, 27, 369–396.
- Pesaran, M., Shin, Y., & Smith, R. (1999). Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association*, 94, 621–634.
- Pesaran, M., & Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, 68, 79–113.

- Pesaran, M. H. (2004). General diagnostic tests for cross section dependence in panels. *CESifo Working Paper Series No. 1229; IZA Discussion Paper No. 1240*, .
- Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, *74*, 967–1012.
- Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, *22*, 265–312.
- Phillips, P. C., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*, 335–346.
- Seetanah, B., Durbarry, R., & Ragodoo, J. (2010). Using the panel cointegration approach to analyse the determinants of tourism demand in south africa. *Tourism Economics*, *16*, 715–729.
- Seetaram, N., & Petit, S. (2012). Panel data analysis. *Handbook of Research Methods in Tourism: Quantitative and Qualitative Approaches*, (pp. 127–144).
- Carrion-i Silvestre, J. L., Kim, D., & Perron, P. (2009). Gls-based unit root tests with multiple structural breaks under both the null and the alternative hypotheses. *Econometric Theory*, *25*, 1754–1792.
- Song, H., Dwyer, L., Li, G., & Cao, Z. (2012). Tourism economics research: A review and assessment. *Annals of Tourism Research*, *39*, 1653–1682.

- Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. *Tourism Management*, 29, 203–220.
- Stock, J., & Watson, M. (1989). New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*, (pp. 351–394).
- Stock, J. H., & Watson, M. W. (1998). Diffusion indexes. *NBER Working Paper*, No. 6702, .
- Strotz, R. H. (1957). The empirical implications of a utility tree. *Econometrica: Journal of the Econometric Society*, 25, 269–280.
- Sul, D. (2009). Panel unit root tests under cross section dependence with recursive mean adjustment. *Economics Letters*, 105, 123–126.
- Syriopoulos, T. C., & Thea Sinclair, M. (1993). An econometric study of tourism demand: the aids model of us and european tourism in mediterranean countries. *Applied Economics*, 25, 1541–1552.
- Westerlund, J., & Urbain, J. (2011). *Cross sectional averages or principal components?*. Maastricht: METEOR, Maastricht Research School of Economics of Technology and Organization.
- Witt, S., & Witt, C. (1995). Forecasting tourism demand: A review of empirical research. *International Journal of Forecasting*, 11, 447–475.

Table 1: Descriptive Statistics and Diagnostic Tests

Descriptive Statistics	$TOUR$	Y^*	$PAIR^*$	PRM^*
Mean	21177	102266.0	371.2	81.1
Max	510013	725758.6	809.6	111.7
Min	641	6430.5	210.5	64.0
S.D.	55666	116523.1	58.0	8.0
S.D.(WG)	9854	20719.9	37.1	7.2
S.D.(BG)	54791	114674.4	44.7	3.6
Variable Diagnostics	$\log TOUR$	$\log Y^*$	$\log PAIR^*$	$\log PRM^*$
CD	155.35*	285.51*	193.23*	130.12*
$Exp - W_{FS}$	1.49	7.24*	7.09*	1.72
P_T^{GLS}	-	25.53	5.95	-
$CRMA$	-1.13	-	-	-0.03
$PANIC_f$	-1.35	-	-	-10.99
Unit Root?	YES	YES	YES	YES

Note: $TOUR$, Y^* , $PAIR^*$, and PRM^* are tourist arrivals, real income, relative airfare, and relative room price, respectively. SD, S.D.(WG), and S.D.(BG) are the overall, within group, and between group standard deviation, respectively. Within group standard deviation is an indicator of average variability over time within states, and between group standard deviation is a measure of variability across states. CD is Pesaran's (2004) cross-sectional independence test. $Exp - W_{FS}$ is Perron and Yabu's (2009) structural break test (H_0 : no break). P_T^{GLS} is Carrion-i Silvestre et al.'s (2009) unit root test allowing for a structural break. $CRMA$ is Sul's (2009) unit root test of the cross-sectional means. $PANIC_f$ is Bai and Ng's (2004) unit root test of the common factors. The lag lengths for the unit root tests are determined by the Bayesian Information Criterion. Statistical significance at the 5% level or lower is denoted by *.

Table 2: Regression Results

$$\log TOUR_{it} = \alpha_i + \beta_{1i} \log Y_{it}^* + \beta_{2i} \log PAIR_{it}^* + \beta_{3i} \log PRM_{it}^* + u_{it} \quad (12)$$

Coefficient Estimates			
	β_1	β_2	β_3
FMOLS	0.34*	-0.37*	0.52*
CCE-MG	1.20*	-0.23*	-1.23*
CCE-P	1.27*	-0.26*	-1.20*
Residual Diagnostics			
	t_{PP}	t_{ADF}	CD
FMOLS	-6.90*	-4.65*	68.93*
	$CRMA$	$PANIC_f$	$Exp - W_{FS}$
	-1.56	-1.47	1.43
	$CIPS$	$PANIC_i$	
CCE-MG	-19.25*	24.97*	
CCE-P	-14.59*	24.97*	

Note: See also definitions in Table 1. $CIPS$ is Pesaran's (2007) unit root test of the idiosyncratic components. The lag length for the $CIPS$ test was set to $T^{1/3} = 4$ in each cross-sectional unit. $PANIC_i$ is Bai and Ng's (2004) unit root test of the idiosyncratic components. The lag length for the $PANIC_i$ test is determined by the Bayesian Information Criterion. Statistical significance at the 5% level or lower is denoted by *.

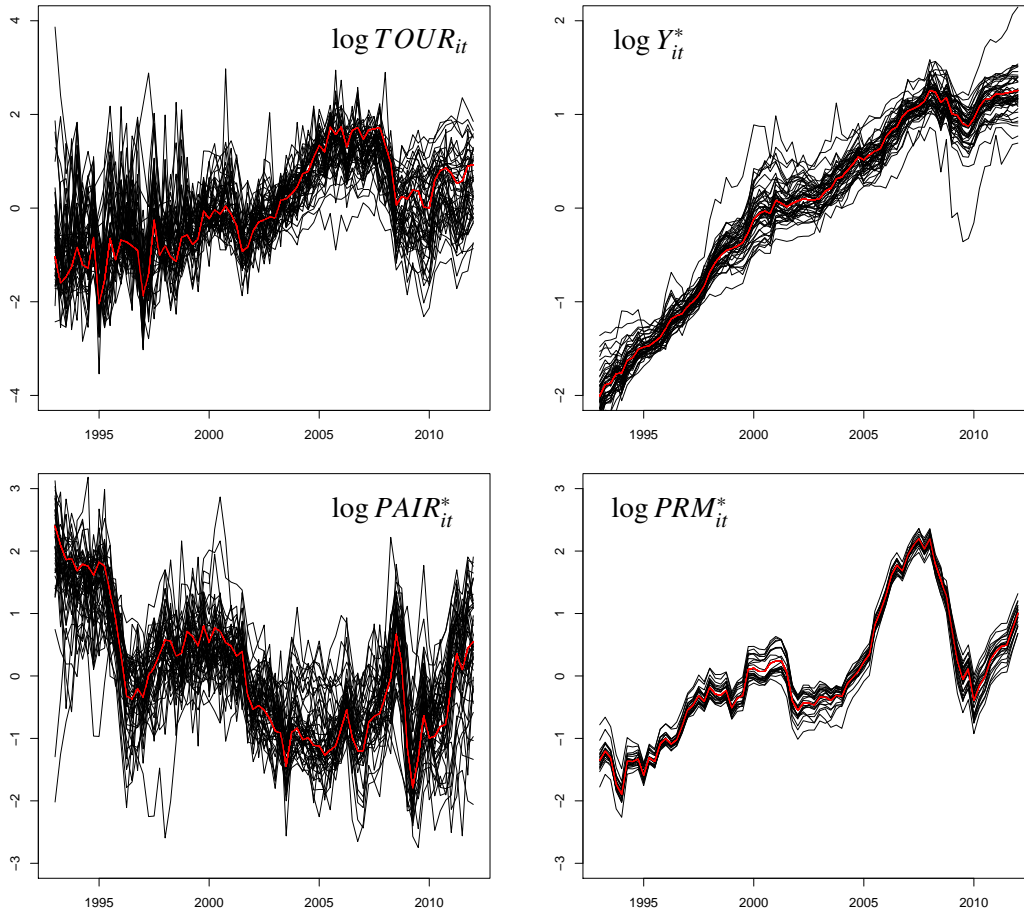


Figure 1: Time plots (1993Q1 - 2012Q1) of standardized logarithms of quarterly tourist arrivals ($TOUR_{it}$), real income (Y_{it}^*), relative airfare ($PAIR_{it}^*$) and relative room rate (PRM_{it}^*) by state of origin. The red line in each graph represents the cross-sectional average of the series.

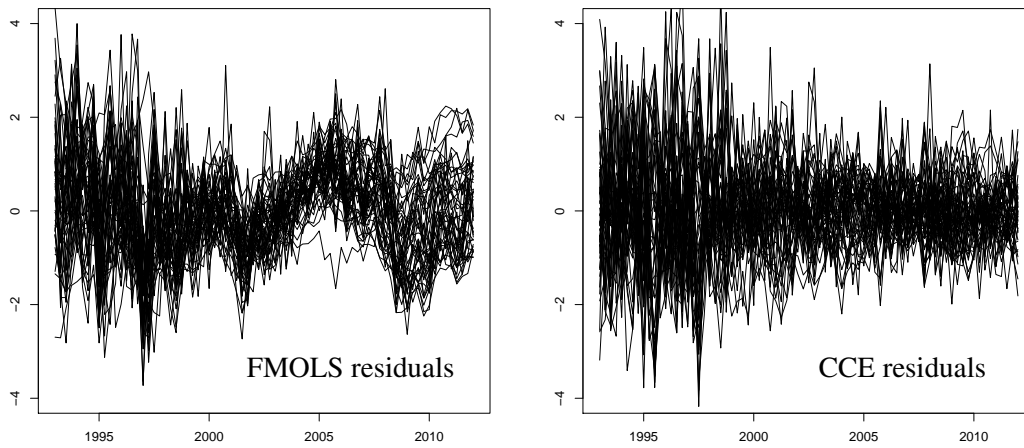


Figure 2: Time plots of standardized FMOLS and CCE residuals.