



MONASH University

Department of Econometrics and Business Statistics

<http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/>

**Modelling Australian Domestic
and International Inbound Travel:
a Spatial-Temporal Approach**

Minfeng Deng, George Athanasopoulos

November 2009

Working Paper 10/09

Modelling Australian Domestic and International Inbound Travel: a Spatial-Temporal Approach

Minfeng Deng

Department of Econometrics and Business Statistics
Monash University, VIC 3800, Australia.
Email: Minfeng.Deng@buseco.monash.edu.au

George Athanasopoulos

Department of Econometrics and Business Statistics
and Tourism Research Unit,
Monash University, VIC 3800, Australia.
Email: George.Athanasopoulos@buseco.monash.edu.au

24 November 2009

JEL classification: C21,C51,L83

Modelling Australian Domestic and International Inbound Travel: a Spatial-Temporal Approach

Abstract:

In this paper Australian domestic and international inbound travel are modelled by an anisotropic dynamic spatial lag panel Origin-Destination (OD) travel flow model. Spatial OD travel flow models have traditionally been applied in a single cross-sectional context, where the spatial structure is assumed to have reached its long run equilibrium and temporal dynamics are not explicitly considered. On the other hand, spatial effects are rarely accounted for in traditional tourism demand modelling. We attempt to address this dichotomy between spatial modelling and time series modelling in tourism research by using a spatial-temporal model. In particular, tourism behaviour is modelled as travel flows between regions. Temporal dependencies are accounted for via the inclusion of autoregressive components, while spatial autocorrelations are explicitly accounted for at both the origin and the destination. We allow the strength of spatial autocorrelation to exhibit seasonal variations, and we allow for the possibility of asymmetry between capital-city neighbours and non-capital-city neighbours. Significant spatial dynamics have been uncovered, which lead to some interesting policy implications.

Keywords: Tourism demand, Dynamic panel models, Travel flow model.

1 Introduction

In two of the most recent and comprehensive reviews on tourism demand modelling and forecasting, [Li et al. \(2005\)](#) and [Song & Li \(2008\)](#) fail to identify any substantial studies using spatial methods. This finding is somewhat surprising, as tourism is a consumer product whose location of purchase and location of consumption are informative of consumer behaviour. For instance, one might reasonably expect that tourists from the same geographical region share similar values and travel interests, and that their travel patterns are similar in some way. One might also reasonably expect that tourists “package” their travels so that the number of destinations visited in one trip can be maximised. While studies of tourism demand have received much attention from a time series analytic perspective, spatial research into tourism demand has remained very limited.

In this paper, we model Australian domestic and international inbound tourism demand using a dynamic spatial panel Origin-Destination (OD) travel flow model. Time lags of the dependent variable are used to capture temporal dependencies, while contemporary spatial lags are used to capture spatial dependencies. We analyse tourism demand from an OD perspective, thus allowing spatial effects to differ between the origins of the tourists and the destinations of the tourists. Spatial OD models have traditionally been applied in a single cross-sectional setting [LeSage & Pace \(2008\)](#), where the spatial structure is assumed to have reached its long run equilibrium and temporal dynamics are not explicitly modelled. On the other hand, spatial effects are rarely accounted for in traditional tourism demand modelling and forecasting. Our current study is the first in formally applying spatial temporal methods in tourism research.

Before introducing our model, it is instructive to briefly review the specification and estimation of dynamic panel models, spatial panel models, and dynamic spatial panel models. An extensive literature exists on both dynamic panels and spatial panels. A small but growing literature exists on dynamic spatial panels, most notably [Elhorst \(2003a,b, 2005\)](#), [Beenstock & Felsenstein \(2007\)](#), and [Yu et al. \(2008\)](#). In dealing with dynamic panels, difficulties arise due to the correlation between lagged dependent variables and time-invariant individual effects. In dealing with spatial panels, the simultaneity of spatially lagged dependent variables is the main obstacle. In dealing with dynamic spatial panels, both sets of difficulties must be addressed. In our current study, we argue that the ML (maximum likelihood) estimator based on a mean-deviated equation is best suited for our data and will be used.

The paper is structured as follows. In Section 2 the specification and estimation of dynamic panels, spatial panels, and dynamic spatial panels is discussed. In Section 3 we present the dynamic spatial panel OD travel flow model used in our study and in Section 4 we present and discuss the ML estimation results and policy implications for Australian tourism industry. In Section 5 we conclude

the paper.

2 A Review of Some Panel Data Models

In this section, we provide a summary of the specification and estimation of dynamic panels, spatial panels, and dynamic spatial panels. As the literature on these panel models is sizable, this serves only as a brief review. We intend to highlight the most salient features of these panel models, compare the relative strengths and weaknesses of various estimators, and provide justifications for the estimation method used in our study.

2.1 Dynamic Panel Models

A dynamic panel model can be specified as

$$Y_t = \phi Y_{t-1} + X_t \beta + \mu + \varepsilon_t, \quad (1)$$

where $t = 1, 2, \dots, T$. Y_t is an $(N \times 1)$ vector of N cross-sectional observations at time t . X_t is an $(N \times K)$ matrix of exogenous explanatory variables observed at time t . ε_t is an $(N \times 1)$ vector of i.i.d. normal errors with $E(\varepsilon_t) = 0 \forall t$ and $E(\varepsilon_t \varepsilon_s^T) = \sigma^2 I_N \forall t$. Furthermore, the errors are assumed to be serially uncorrelated, i.e., $E(\varepsilon_t \varepsilon_s^T) = 0 \forall t \neq s$. ϕ is the first order autoregressive parameter of interest. μ is an $(N \times 1)$ vector of time-invariant individual effects, which can be specified either as fixed effects or as random effects. When they are specified as fixed effects, each cross-sectional unit is associated with a unique intercept. The standard estimator for a fixed effects panel is the LSDV (least squares dummy variable) estimator, which demeans the equation to eliminate the time-invariant fixed effects. When applied to a dynamic panel model, the demeaned equation

$$\bar{Y}_t = \phi \bar{Y}_{t-1} + \bar{X}_t \beta + \bar{\varepsilon}_t, \quad (2)$$

is estimated by OLS where $\bar{Y}_t = Y_t - \frac{1}{T} \sum_{t=1}^T Y_t$, $\bar{Y}_{t-1} = Y_{t-1} - \frac{1}{T} \sum_{t=1}^T Y_{t-1}$, $\bar{X}_t = X_t - \frac{1}{T} \sum_{t=1}^T X_t$, and $\bar{\varepsilon}_t = \varepsilon_t - \frac{1}{T} \sum_{t=1}^T \varepsilon_t$. In the presence of lagged dependent variables, this procedure becomes problematic, as the demeaned lagged dependent variable and the demeaned error term are correlated of order $(1/T)$. Hsiao (1986) shows that the estimate of ϕ is biased downwards and the extent of the bias may not be negligible for small T . Only when $T \rightarrow \infty$ does this correlation disappear and the LSDV estimator will be consistent (Hsiao 1986, Baltagi 2001).

When the individual effects are specified as random effects, the variable intercepts are treated as random draws of an i.i.d. random variable. Correlation between the unobserved individual effect μ

and Y_{t-1} on the right hand side makes the OLS estimator biased and inconsistent. In cases like this, where the individual effects are treated as stochastic, and in cases where T is small and the LSDV estimator is biased and inconsistent, a number of estimators have been proposed. [Anderson & Hsiao \(1981, 1982\)](#) suggest first-differencing the equation, thus eliminating the individual effect μ ,

$$\Delta Y_t = \phi \Delta Y_{t-1} + \Delta X_t \beta + \Delta \varepsilon_t \quad (3)$$

and using either ΔY_{t-2} or Y_{t-2} as an instrument for ΔY_{t-1} . They show that their estimator for ϕ is consistent as $N \rightarrow \infty$ for any fixed T . Subsequently, [Arellano & Bond \(1991\)](#) and [Arellano & Bover \(1995\)](#) suggest using values of Y_{t-j} where $j \geq 2$ as instruments in the differenced equation. They argue that since both ΔY_{t-2} and Y_{t-2} are linear combinations of lagged values of Y_t , their GMM estimator is more efficient. In empirical studies using dynamic panels, GMM estimators of this type have been the most popular.

[Hsiao et al. \(2002\)](#) also suggest an unconditional ML estimator based on the first-differenced equation. They note that the first-differenced equation (3) is well-defined for $t \geq 2$, and they show that for $t = 1$ the differenced equation can be re-written as

$$\Delta Y_1 = \phi^m \Delta Y_{-m+1} + \sum_{j=0}^{m-1} \phi^j \Delta X_{1-j} \beta + \sum_{j=0}^{m-1} \phi^j \Delta \varepsilon_{1-j}, \quad (4)$$

where m is finite and needs to be chosen judiciously. When the distribution of ΔY_1 is completely specified, one can write down the unconditional log-likelihood function of the entire sample and estimate with ML. A strong assumption must be made about the initial values of the first period difference: either that they are the same for all cross-sectional units, or that the process started long ago and $E(\Delta Y_1) = 0$. Furthermore, the second term involving lagged differenced exogenous variables is also unobserved and it must be approximated following either [Bhargava & Sargan \(1983\)](#) or [Nerlove & Balestra \(1996\)](#). They show that their estimator is consistent as $N \rightarrow \infty$ for any size T .

2.2 Spatial Panel Models

Spatial models consist of: spatial lag models, where spatial effects are incorporated substantively via spatially lagged dependent variables; and spatial error models, where spatial autocorrelation is incorporated in the error term ([Anselin 1988](#)). A spatial lag panel model can be specified as

$$Y_t = \rho W Y_t + X_t \beta + \mu + \varepsilon_t, \quad (5)$$

where $t = 1, 2, \dots, T$. W is an $(N \times N)$ spatial weights matrix whose $ijth$ element specifies the spatial relationship between the ith and jth spatial unit. More specifically, W_{ij} satisfies that: $W_{ij} \geq 0$ for $i \neq j$, and $W_{ij} = 0$ for $i = j$. Therefore, nonzero W_{ij} 's are associated with cases where the ith and jth units are considered to be spatial neighbours (see [Anselin 1988](#), for a more detailed discussion on the specification of spatial weights matrices). ρ is known as the spatial autoregressive parameter and it specifies the extent of spatial autocorrelation. When the spatial weights matrix is row-standardised, i.e., $\sum_j W_{ij} = 1 \forall i$, which is almost always the case, WY_t gives the weighted average of spatial neighbours of Y at time t . Since the seminal work of [Ord \(1975\)](#) and [Anselin \(1988\)](#), ML is by far the most popular estimation method used in applied spatial econometric modelling. [Anselin \(1988\)](#) shows that the spatial lag panel model is a straightforward extension of the single cross sectional spatial lag model and it can be consistently estimated using ML. Since the individual effect μ is not correlated with any of the right hand side variables, its presence does not introduce additional complications.

On the other hand, a spatial error panel model can be specified as,

$$\begin{aligned} Y_t &= X_t\beta + \mu + \varepsilon_t \\ \varepsilon_t &= \rho W\varepsilon_t + u_t. \end{aligned} \tag{6}$$

[Baltagi & Koh \(2003\)](#) consider this model, and show that this model can also be consistently estimated with ML.

2.3 Dynamic Spatial Panel Models

Finally, a dynamic spatial lag panel model is a combination of a dynamic panel and a spatial lag panel, and it can be specified as

$$Y_t = \phi Y_{t-1} + \rho WY_t + X_t\beta + \mu + \varepsilon_t, \tag{7}$$

and a dynamic spatial error model can be specified as

$$\begin{aligned} Y_t &= \phi Y_{t-1} + X_t\beta + \mu + \varepsilon_t \\ \varepsilon_t &= \rho W\varepsilon_t + u_t, \end{aligned} \tag{8}$$

where $t = 1, 2, \dots, T$. [Elhorst \(2003a,b, 2005\)](#) has provided the most complete discussion on the estimation of dynamic spatial panel models to date. He argues that in general asymptotics are easier to achieve in the cross-sectional dimension than in the time dimension, and estimators that rely on

$T \rightarrow \infty$ will generally suffer from sizable finite sample biases. He suggests a spatial extension of the unconditional ML approach of [Hsiao et al. \(2002\)](#), which is consistent in N for any size T . The first-differenced dynamic spatial lag panel model is

$$B\Delta Y_t = \phi \Delta Y_{t-1} + \Delta X_t \beta + \Delta \varepsilon_t, \quad (9)$$

where $B = I_N - \rho W$. The equation is well-defined for $t \geq 2$. For $t = 1$ the equation can be written as

$$B\Delta Y_1 = \phi^m B^{-m+1} \Delta Y_{-m+1} + \sum_{j=0}^{m-1} \phi^j B^{-j} \Delta X_{1-j} \beta + \sum_{j=0}^{m-1} \phi^j B^{-j} \Delta \varepsilon_{1-j}. \quad (10)$$

The spatial error version of the model can be similarly defined and will not be presented here. A set of assumptions and approximations similar to those of [Hsiao et al. \(2002\)](#) are required in order to derive the distribution of ΔY_1 . However, in light of the presence of spatial autocorrelation, the assumption of equal initial changes across all spatial units appears to be highly restrictive.

More recently, [Yu et al. \(2008\)](#) show that, if N is a nondecreasing function of T and if T goes to infinity, under a set of fairly general conditions, the demeaned dynamic spatial lag panel equation

$$B\bar{Y}_t = \phi \bar{Y}_{t-1} + \bar{X}_t \beta + \bar{\varepsilon}_t \quad (11)$$

where $B = I_N - \rho W$, can be consistently estimated with ML. They only presented their results for the spatial lag specification, but they expect those results to extend to the spatial error specification.

Finally, [Beenstock & Felsenstein \(2007\)](#) suggest a bias-corrected 2SLSDV approach. They first apply the LSDV estimator to equation (7) while omitting the spatial lag term and obtaining a set of fitted values \hat{Y}_t . They then use the spatially weighted $W\hat{Y}_t$ to instrument for WY_t in the original equation, and correct for the bias caused by the lagged dependent variable by using the downward asymptotic bias described in [Hsiao \(1986\)](#). Although computationally simple, when the exogenous explanatory variables are “weak”, it is likely that \hat{Y}_t gives a poor fit, which may result in the spatially weighted $W\hat{Y}_t$ being a “weak instrument”. Furthermore, as the asymptotic bias of [Hsiao \(1986\)](#) is derived in the absence of spatial effects, it is unclear how the presence of spatial autocorrelation affects the validity of this bias-correction.

3 An Anisotropic Dynamic Spatial Lag Panel Origin-Destination (OD) Travel Flow Model

3.1 Model Specification

We use the quarterly number of “visitor nights” as the indicator for tourism demand. In the time dimension, for domestic travellers we have observations from 1998Q1 to 2008Q2, giving us a total of $T = 42$ observations. For the international inbound travellers we have observations from 1999Q1 to 2008Q4, giving us a total of $T = 40$ observations. In the spatial dimension, Australia is sampled according to a total of 83 statistical local areas (SLA) defined by the Australian Bureau of Statistics (see [Australian Bureau of Statistics 2003](#), and [Athanasopoulos et al. 2009](#) for forecasts for a similar structure). Of all the SLA’s, 7 are capital cities. In terms of domestic travellers, since each region is potentially both an origin and a destination, in theory we have a total of $83^2 = 6724$ possible travel flows. In reality, the Australian populace concentrates heavily in and around the capital cities, and a large number of the travel flows between non-capital regions record zeros. After deleting the zero flows, we are left with 631 flows, i.e., a cross-sectional dimension of $N = 631$. In terms of the international inbound travels, all Australian regions are considered as destinations, while all foreign countries are considered as origins. After deleting the zero flows, we are left with a cross-sectional dimension of $N = 889$.

The OD flow model with spatially autoregressive error components has its origin in [Bolduc et al. \(1989\)](#) and [Bolduc et al. \(1992\)](#). [LeSage & Pace \(2008\)](#) generalise the model to incorporate both spatial lag components and spatial error components. In our study, we argue that spatial patterns in tourism demand are best interpreted as results of substantive human interaction, thus a spatial lag specification is favoured over a spatial error specification. A spatial lag OD travel flow model can be specified as:

$$Y_t = \rho_o W_o Y_t + \rho_d W_d Y_t + X_{o,t} \beta_o + X_{d,t} \beta_d + \alpha u_N + \varepsilon_t \quad (12)$$

where Y_t is an $(N \times 1)$ vector of observed travel flows between regions at time t . W_o and W_d are spatial weights matrices defining spatial relationships between flow origins and flow destinations respectively. Specifically, $W_{o,ij}$ is the ij th element of W_o , and it is nonzero if both the j th flow and the i th flow are heading to the same destination region and if the j th flow’s origin is a spatial neighbour of the i th flow’s origin. Similarly, $W_{d,ij}$ is the ij th element of W_d , and it is nonzero if both the j th flow and the i th flow are leaving from the same origin region and if the j th flow’s destination is a spatial neighbour of the i th flow’s destination. Graphically this is shown in [Figure 1](#).

Extensive literature exists on the specification of the spatial structure, a complete summary of which is beyond the scope of this paper. Interested readers can refer to [Anselin \(1988\)](#) for more detailed

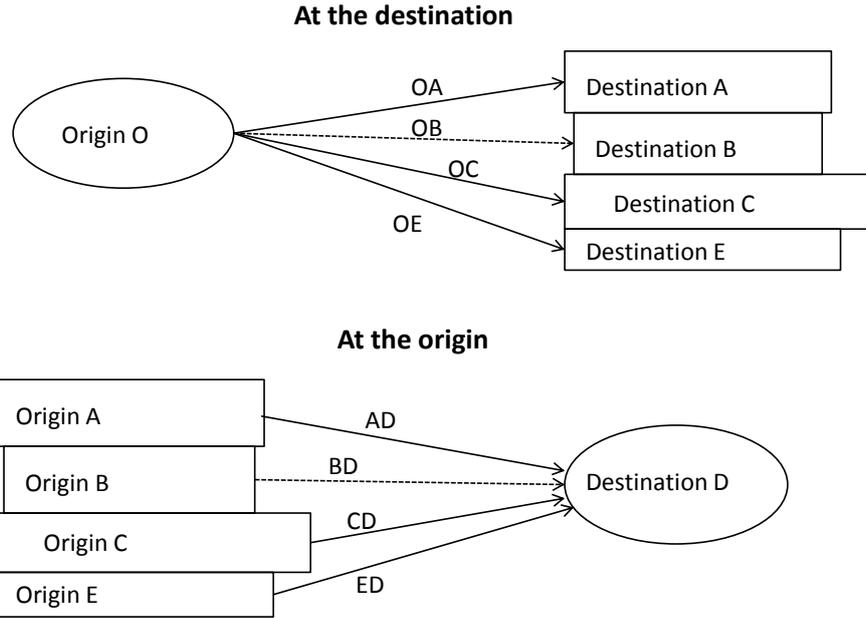


Figure 1: For travel flow *OB*, both *OA* and *OC* are considered to be spatial neighbours at the destination, while *OE* is not. For travel flow *BD*, both *AD* and *CD* are considered to be spatial neighbours at the origin, while *ED* is not.

treatments of the topic. Generally speaking, spatial structures can be classified into two types: binary contiguity structure, where spatial units sharing a border are considered neighbours, and distance decay spatial structure, where close-by units are given higher weights and distant units lower weights. In this paper, due to the high variability in sizes of the SLA's and sizes of the countries/regions outside Australia, we use the binary spatial contiguity structure, originally due to [Cliff & Ord \(1973\)](#).

$X_{o,t}$ and $X_{d,t}$ are $(N \times K_o)$ and $(N \times K_d)$ matrices of socioeconomic variables observed at the origin and at the destination respectively at time t . ι_N is an $(N \times 1)$ vector of ones and α is a constant. The advantage of a spatial OD model lies in its ability to separately identify spatial effects at the origin of the flows and at the destination of the flows. It has been applied in studies of transportation flows ([Bolduc et al. 1989](#)), trade flows ([Porojan 2001](#)), and migration flows ([LeSage & Pace 2008](#)). We consider it to be well-suited for the purpose of analysing tourism flows data also.

In addition to distinguishing between tourist origins and destinations, we also suspect the spatial behaviour of tourists in and out of capital cities to be different from those in and out of non-capital cities. To allow for the possibility of asymmetric spatial effects between capital-city neighbours and non-capital city neighbours, we use the anisotropic specification of [Deng \(2008\)](#). The OD model

becomes:

$$\begin{aligned}
 Y_t = & [(\rho_o D_{o1} + \rho_{oc} D_{oc}) \bullet W_o] Y_t + [(\rho_d D_{d1} + \rho_{dc} D_{dc}) \bullet W_d] Y_t \\
 & + X_{o,t} \beta_o + X_{d,t} \beta_d + \alpha_N + \varepsilon_t
 \end{aligned} \tag{13}$$

where D_{o1} is an $(N \times N)$ binary contiguity matrix whose $ijth$ element is equal to 1 when $W_{o,ij} \neq 0$, i.e., D_{o1} is W_o prior to row-standardisation. Similarly, D_{d1} is $(N \times N)$ and it is W_d prior to row-standardisation. D_{oc} is an $(N \times N)$ matrix, whose $ijth$ element is equal to 1 if the jth flow's origin is a spatial neighbour of the ith flow's origin and if the jth flow's origin is a capital city (c), otherwise it is 0. Similarly, D_{dc} is an $(N \times N)$ matrix, whose $ijth$ element indicates if the destination neighbour is a capital city neighbour. Therefore, while general spatial autoregressive effects at the origin (o) and at the destination (d) are captured by non-zero values of ρ_o and ρ_d respectively, if spatial effects are indeed distributed asymmetrically between capital-city neighbours and non-capital city neighbours, spatial parameters ρ_{oc} and ρ_{dc} (associated with the capital city indicators D_{oc} and D_{dc} respectively) will also be significantly different from zero. Note that $[(\rho_o D_{o1} + \rho_{oc} D_{oc}) \bullet W_o]$ is a Hadamard product of $(\rho_o D_{o1} + \rho_{oc} D_{oc})$ and W_o , and $[(\rho_d D_{d1} + \rho_{dc} D_{dc}) \bullet W_d]$ is a Hadamard product of $(\rho_d D_{d1} + \rho_{dc} D_{dc})$ and W_d .

Finally, we also suspect that spatial effects may exhibit seasonality. Tourists travelling during holiday seasons, such as during the new year and the Australian summer holiday period in quarter one, are likely to be predominantly holiday makers who aim to maximise their travel experiences and visit as many places as they can in one trip. As a result, one would expect significant spatial effects in quarter one. On the other hand, during low-seasons, travellers are likely to be travelling with a specific purpose and show fewer spatial patterns. To allow for possible seasonal variations, we allow the spatial coefficients to take on season-specific values.

We can now formally define the model used in our study. An anisotropic dynamic spatial lag panel OD travel flow model is:

$$\begin{aligned}
 Y_t = & \phi_1 Y_{t-1} + \phi_4 Y_{t-4} + \beta_{trend} t + \beta_{q1} Q_{1,t} + \beta_{q2} Q_{2,t} + \beta_{q3} Q_{3,t} \\
 & + \sum_{j=1}^4 q_{j,t} \{ [(\rho_{o,j} D_{o1} + \rho_{oc,j} D_{oc}) \bullet W_o] Y_t \} \\
 & + \sum_{j=1}^4 q_{j,t} \{ [(\rho_{d,j} D_{d1} + \rho_{dc,j} D_{dc}) \bullet W_d] Y_t \} \\
 & + X_{o,t} \beta_o + X_{d,t} \beta_d + \mu + \varepsilon_t
 \end{aligned} \tag{14}$$

where Y_t is an $(N \times 1)$ vector of observed travel flows between regions at time t . We allow Y_t to be

temporally correlated with its first lag Y_{t-1} and its seasonal fourth lag Y_{t-4} . A linear trend t is also included. $Q_{j,t}$, where $j = 1, 2, 3$, is an $(N \times 1)$ column of ones if time t corresponds to the j th quarter, and zeros otherwise. $X_{o,t}$ and $X_{d,t}$ are $(N \times K_o)$ and $(N \times K_d)$ matrices of socioeconomic characteristics observed at the origin and destination respectively at time t . μ is an $(N \times 1)$ vector of unobserved individual fixed effects. ε_t is an $(N \times 1)$ vector of i.i.d. normal errors with $E(\varepsilon_t) = 0 \forall t$, $E(\varepsilon_t \varepsilon_t^T) = \sigma^2 I_N \forall t$, and $E(\varepsilon_t \varepsilon_s^T) = 0 \forall t \neq s$. D_{o1} , D_{d1} , D_{oc} , D_{dc} , W_o , and W_d are $(N \times N)$ matrices that specify spatial structures and have already been defined earlier. Spatial contiguity is simply defined as binary contiguity, where two statistical regions/countries sharing the same boarder are considered as spatial neighbours. $q_{j,t}$ is a seasonal indicator that is equal to 1 if time t corresponds to the j th quarter, and zero otherwise. The spatial parameters $\rho_{o,j}$, $\rho_{d,j}$, $\rho_{oc,j}$, and $\rho_{dc,j}$ are indexed by the j th quarter. The combination of $q_{j,t}$ and seasonally indexed spatial parameters allows for potential seasonal variations in the spatial patterns. For instance, in Q1, spatial effects will be captured by the following two spatial terms:

$$\left\{ \left[(\rho_{o,1} D_{o1} + \rho_{oc,1} D_{oc}) \bullet W_o \right] Y_t \right\} \text{ and } \left\{ \left[(\rho_{d,1} D_{d1} + \rho_{dc,1} D_{dc}) \bullet W_d \right] Y_t \right\}$$

while all other spatial terms will be equal to zero.

3.2 Model Estimation

As summarised in Section 2, for dynamic spatial panels, ML estimators exist both in a differenced form and a mean-deviated form. A bias-corrected 2SLSDV estimator also exists. We argue that an estimator based on first-differencing the data, such as that of [Elhorst \(2003a,b\)](#), is not suitable in our current study. As our data is highly seasonal, differencing at the first lag while the seasonality is at the fourth lag could lead to large fluctuations in the differenced values. As one gets in and out of a high tourist quarter, the differenced dependent variable would fluctuate from a large positive to a large negative. Moreover, as the spatial parameters are season-specific, differencing at the first lag would result in differencing spatial terms that are associated with different spatial parameters, preventing collection of common terms and introducing further complications into the model. Differencing at the seasonal fourth lag is also not recommended, as it results in losing a large number of observations at the beginning. We also argue that the 2SLSDV estimator of [Beenstock & Felsenstein \(2007\)](#) is not suitable, as we question the validity of their suggested instrument in our context and the validity of their bias-correction measure in the presence of spatial effects.

On the other hand, we argue that, as T in our current study is sizable ($T = 42$ for the domestic data and $T = 40$ for the international inbound data), correlation between the mean-deviated lagged dependent variables and the mean-deviated error term is likely to be negligible, and it makes an

estimator based on the demeaned equation an attractive option. Based on the above considerations, we argue that the ML estimator of Yu et al. (2008) is most suitable for our current study. The log likelihood function of equation (14) can be written as:

$$\begin{aligned}
 \ln L = & -\left(\frac{NT}{2}\right) \ln(2\pi) - \left(\frac{NT}{2}\right) \ln(\sigma^2) \\
 & + \sum_{j=1}^4 \left\{ T_j \ln \left| I - [F_{o,j} \bullet W_o] - [F_{d,j} \bullet W_d] \right| \right\} \\
 & - \left(\frac{1}{2\sigma^2}\right) \left\{ \bar{Y} - [F_o \bullet W_o] \bar{Y} - [F_d \bullet W_d] \bar{Y} - \bar{X} \beta \right\}^T \\
 & \left\{ \bar{Y} - [F_o \bullet W_o] \bar{Y} - [F_d \bullet W_d] \bar{Y} - \bar{X} \beta \right\}
 \end{aligned} \tag{15}$$

where \bar{Y} and \bar{X} are \bar{Y}_t and \bar{X}_t stacked by t accordingly. T_j is the total number of the j th quarter in our sample. $F_{o,j} = (\rho_{o,j}D_{o1} + \rho_{oc,j}D_{oc})$ and $F_{d,j} = (\rho_{d,j}D_{d1} + \rho_{dc,j}D_{dc})$ are seasonally indexed and they define the anisotropic spatial structures. W_o and W_d are block diagonal matrices where each diagonal block is the same spatial weights matrix W_o and W_d respectively. F_o and F_d are block diagonal matrices where each diagonal block is the corresponding seasonal spatial structural matrix $F_{o,j}$ and $F_{d,j}$ respectively. The above log likelihood function is maximised with respect to σ^2 , β , and $(\rho_{o,j}, \rho_{oc,j}, \rho_{d,j}, \rho_{dc,j}) \forall j = 1, 2, 3, 4$.

4 Australian Domestic and International Inbound Tourism

Demand

We measure tourism demand by the number of “visitor nights”, which is the total number of nights spent away from home. Our domestic data comes from the National Visitor Survey (NVS), which is a quarterly survey of approximately 15,000 Australian households each year. Our international inbound data comes from the International Visitor Survey (IVS), which is a quarterly survey of approximately 20,000 international tourists at points of departure from Australia each year. While the NVS covers a recording period from 1998:Q1 to 2008:Q2, the IVS covers a recording period from 1999:Q1 to 2008:Q4. The number of households sampled within each geographical region is weighted according to population size of the region. The only exogenous socioeconomic variables available are the estimated regional incomes in Australia.

4.1 Estimation Results

For domestic tourism ML estimates are presented in Table (1). In terms of the temporal dependencies, the coefficient of the first lag of the dependent variable (ϕ_1) is found to be negative (-0.0437) and statistically significant, while the seasonal fourth lag (ϕ_4) is found to be positive and also statistically significant. The significant negative coefficient of the first lag is consistent with our expectations. If someone from origin A has travelled to destination B in the last quarter, it is less likely that they will travel to the same destination again this quarter. The significant and positive coefficient of the fourth lag is also consistent with our expectations and captures what is described in the tourism literature as “habit persistence” (see for example [Song & Witt 2000](#), page 7). Tourists often become “comfortable” with destinations they have visited and are more likely to return to the same destination again for holidays. In terms of the quarterly seasonality, only the coefficient of the first quarter (β_{q1}) is found to be positive (9.0630) and significant. The positive estimate is consistent with the fact that the first quarter of the calendar year in Australia coincides with the New Year and school holiday period and is also the peak summer holiday season. In terms of the overall trend in domestic travels, the coefficient of the trend (β_{trend}) is found to be negative (-0.3603) and significant. This is in line with an earlier study of [Athanasopoulos & Hyndman \(2008\)](#), where a general decline in Australian domestic tourism has been identified. Finally, in terms of the income effects at the origin (β_{o_income}) and at the destination (β_{d_income}), only the origin income is found to have a positive (19.0332) and significant impact. The positive estimate at the origin suggests that high income regions are more likely to generate tourist outflows, a result that is within our expectations.

All estimated spatial coefficients are highly significant in both quarter 1 and quarter 3, while statistically insignificant in the other two quarters. Note that, in Australia, quarter 1 corresponds with the peak summer tourist season, while quarter 3 corresponds with the peak winter season, during which “cold-weather-adverse” individuals would often travel north for the sunshine; while “cold-weather-loving” individuals would often travel south for some skiing in the alpine region. Recall our initial hypothesis is such that, during peak tourist seasons, travellers are more likely to exhibit strong spatial patterns. Significant spatial effects only in peak summer season and peak winter season are precisely what one would expect to find if our hypothesis was true. More specifically, significant overall spatial autocorrelation at the origin ($\rho_{o,q1}, \rho_{o,q3}$) suggests that the number of visitor nights from tourists leaving adjacent regions are likely to be similar. A number of explanations can be suggested here. Adjacent regions often share similar social values and experience similar economic conditions, and hence are more likely to be similar to each other. Regional tourism operators might also design their business operation in such a way that their target customers involve people from neighbouring regions, which is both logistically sensible and allows for economies of scale. Finally, Australia has a large migrant population, and migrant groups tend to

Table 1: Estimates for Australian domestic tourism demand

Temporal		Spatial	
Coefficient	Estimate	Coefficient	Estimate
ϕ_1	-0.0437*** (0.0057) ^a	$\rho_{o,q1}$	0.1608*** (0.0216)
ϕ_4	0.4344*** (0.0056)	$\rho_{oc,q1}$	-0.1464*** (0.0277)
β_{trend}	-0.3603*** (0.0911)	$\rho_{d,q1}$	0.1450*** (0.0160)
β_{q1}	9.0630*** (1.1755)	$\rho_{dc,q1}$	-0.1939*** (0.0379)
β_{q2}	-1.8054 (1.1662)	$\rho_{o,q2}$	0.0158 (0.0253)
β_{q3}	-1.4777 (1.1795)	$\rho_{oc,q2}$	-0.0005 (0.0396)
β_{o_income}	19.0332*** (6.7589)	$\rho_{d,q2}$	0.0318 (0.0222)
β_{d_income}	3.4979 (5.2815)	$\rho_{dc,q2}$	0.0011 (0.0546)
		$\rho_{o,q3}$	0.1553*** (0.0239)
		$\rho_{oc,q3}$	-0.1397*** (0.0322)
		$\rho_{d,q3}$	0.1156*** (0.0194)
		$\rho_{dc,q3}$	-0.1015*** (0.0517)
		$\rho_{o,q4}$	0.0128 (0.0249)
		$\rho_{oc,q4}$	-0.0051 (0.0395)
		$\rho_{d,q4}$	0.0259 (0.0221)
		$\rho_{dc,q4}$	-0.0225 (0.0461)

*** significant at the 1% level

** significant at the 5% level

* significant at the 10% level

^a standard errors are shown in parentheses

live in close-by regions. Strong spatial autocorrelation at the origin of the travel flows might simply reflect the significant clustering of ethnic groups.

On the other hand, significant overall spatial autocorrelation at the destination ($\rho_{d,q1}, \rho_{d,q3}$) suggests that the number of visitor nights are likely to be similar for adjacent destination regions. This may be explained by the fact that tourist destinations adjacent to each other often offer similar tourist attractions and activities, which result in similar tourist behaviour in those regions. Moreover, multiple destination travels during one holiday trip is common, and tourists may also exhibit “diffusion behaviour” within the neighbouring destinations once they have reached their main destination. Both these scenarios would result in similarities in visitor nights amongst a group of close-by regions.

The estimates of the anisotropic spatial effects for both capital-city origin neighbours and capital-city

Table 2: Estimates for Australian international tourism demand

Temporal		Spatial	
Coefficient	Estimate	Coefficient	Estimate
ϕ_1	0.0729*** (0.0049)	$\rho_{o,q1}$	0.0731*** (0.0059)
ϕ_4	0.4801*** (0.0051)	$\rho_{d,q1}$	0.0914*** (0.0225)
β_{trend}	0.1518*** (0.0364)	$\rho_{dc,q1}$	-0.0783*** (0.0255)
β_{q1}	2.3091*** (0.7318)	$\rho_{o,q2}$	0.1129*** (0.0094)
β_{q2}	-5.1757*** (0.7384)	$\rho_{d,q2}$	0.0326 (0.0243)
β_{q3}	-1.2263* (0.7254)	$\rho_{dc,q2}$	-0.0355 (0.0322)
β_{d_income}	-1.0459 (2..3746)	$\rho_{o,q3}$	0.0865*** (0.0107)
		$\rho_{d,q3}$	0.0433** (0.0239)
		$\rho_{dc,q3}$	-0.0219 (0.0342)
		$\rho_{o,q4}$	0.0628*** (0.0093)
		$\rho_{d,q4}$	0.0423 (0.0241)
		$\rho_{dc,q4}$	-0.0435 (0.0318)

*** significant at the 1% level

** significant at the 5% level

* significant at the 10% level

^a standard errors are shown in parentheses

destination neighbours ($\rho_{oc,q1}, \rho_{dc,q1}, \rho_{oc,q3}, \rho_{dc,q3}$) are all negative and highly significant in quarter 1 and quarter 3. Recall that the anisotropic spatial effects are associated with capital-city neighbours. A negative and statistically significant estimate suggests that the spatial effect of a capital-city neighbour is significantly reduced. Moreover, we note that the absolute values of these estimates are similar to the absolute values of the overall spatial effects ($\rho_{o,q1}, \rho_{d,q1}, \rho_{o,q3}, \rho_{d,q3}$). This suggests that the spatial effect of travel flows both in and out of a capital-city neighbour is almost reduced to zero. This is consistent with our prior expectation that tourists travelling in and out of capital cities are likely to be different from the rest. The negative and statistically significant estimates associated with these capital-city indicators suggest that the key spatial patterns identified thus far are driven predominantly by travellers in and out of non-capital city regions.

For international inbound tourism demand, we do not have a reliable measure of income levels at the origin (i.e., overseas countries), as some of the countries are grouped together due to a general lack of observations from those countries. Also, we do not specify anisotropic structures at the origin of the flows, as we cannot clearly identify foreign countries into “capital countries” and “non-capital countries”. The estimated coefficients are presented in Table 2.

In terms of temporal dependencies, both the first lag (ϕ_1) and the fourth lag (ϕ_4) coefficients are found to be positive and significant, at (0.0729) and (0.4801) respectively. In terms of the seasonalities, the first quarter (β_{q1}) is found to be positive and significant, the second quarter (β_{q2}) to be negative and significant, and the third quarter (β_{q3}) to be negative and slightly significant. In comparison with domestic travels, where only the first quarter is found to be positive and significant, in international inbound travels, the second and third quarter negative effects are also significant. The overall trend (β_{trend}) is found to be positive (0.1518) and significant. Finally, the income effect at the destination (β_{d_income}) is found to be insignificant.

The overall spatial effects at the origin are found to be positive and significant in all four quarters ($\rho_{o,q1}, \rho_{o,q2}, \rho_{o,q3}, \rho_{o,q4}$), while overall spatial effects at the destination are significant in only the first quarter ($\rho_{d,q1}$) and the third quarter ($\rho_{d,q3}$). Thus, just like domestic tourism, visitor nights of overseas tourists in close-by regions are also likely to be similar in both peak summer and peak winter seasons. However, unlike domestic tourists, overall spatial similarity at the origin is now significant in all quarters. This is likely to be due to similarities in socioeconomic conditions in nearby countries. Finally, the estimate of the anisotropic spatial effect in quarter one ($\rho_{dc,q1}$) is negative and highly significant. The absolute value of this estimate is also comparable to that of the overall spatial effect ($\rho_{d,q1}$). This suggests that, similar to domestic tourists, international tourists travel patterns to capital cities might also be very different from those to non-capital cities.

4.2 Policy Implications

The temporal and spatial dynamics discovered in this study have several important policy implications. Many of the results from our study are well-understood by tourism operators through experience, even though they have never been numerically substantiated. In terms of the temporal dynamics, seasonal dependence is found to be highly positive and significant for both domestic travellers and international inbound travellers, suggesting the need for tourism operators to maintain a strong focus on revisiting seasonal travellers. Moreover, while the first quarter is undeniably the high quarter for both domestic and international inbound travels, the second and third quarters appear to be significantly low quarters for international inbound activities. This suggests that greater efforts are needed in promoting those two quarters to the international market. Finally, while a significant positive trend has been identified for international inbound travels, a significant negative trend has been identified for domestic travels. More research is needed in understanding and maintaining the positive trend in inbound travels, as well as identifying the causes of the decreasing domestic travels.

In terms of the spatial dynamics, significant spatial autocorrelation is found both at the origin and at

the destination, and for both domestic travellers and international inbound travellers. At tourist destinations, offering greater “regional connections” may facilitate the “dispersion” of tourists. For instance, easier access to rental vehicles and/or more regional bus/train connecting routes could potentially encourage travellers to visit areas neighbouring their main destination. At tourist origins, we suspect that promotions targeting overseas travellers from a “wider region” might be more effective. For instance, one might treat Europe as a region distinct from Asia, which is in turn distinct from the Americas. Given the significant origin effects identified in our study, it is likely that an increase in tourists from parts of a wider region could induce greater tourist numbers from neighbouring countries too. The significant seasonal variations in the spatial effects also suggest that effective utilisation of these spatial patterns may require tourism operators to alter their operational strategies depending on the season. Finally, travellers to capital cities appear to be different from those travelling to non-capital cities. This may be explained by the fact that capital cities often offer different tourist activities from non-capital cities. Moreover, it is likely that travellers that choose to visit non-capital cities tend to be more adventurous and exhibit a higher degree of spatial “dispersion”. The distinction between capital city travellers and non-capital city travellers needs to be recognised and better understood.

5 Conclusion

Our study is the first in formally incorporating both temporal and spatial dynamics into tourism demand modelling. Our study is also the first in formulating tourism demand from the view point of origin-destination travel flows. Since tourism is essentially a consumption of locations, we argue that the spatial dimension of tourism demand deserves more attention. Via the estimation of a dynamic spatial panel model, we have uncovered a rich set of temporal and spatial patterns. We believe that these results have important policy implications for the Australian tourism industry. They also form the basis for building a large scale forecasting system for Australian tourism demand in which both temporal and spacial dynamics will be incorporated.

References

- Anderson, T. W. & Hsiao, C. (1981), ‘Estimation of dynamic models with error components’, *Journal of the American Statistical Association* **76**, 598–606.
- Anderson, T. W. & Hsiao, C. (1982), ‘Formulation and estimation of dynamic models using panale data’, *Journal of Econometrics* **18**, 47–82.

- Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Kluwer, Dordrecht.
- Arellano, M. & Bond, S. (1991), 'Some tests of specification for panel data: Monte carlo evidence and an application to employment equations', *Review of Economic Studies* **58**, 277–297.
- Arellano, M. & Bover, O. (1995), 'Another look at the instrumental variable estimation of error-component models', *Journal of Econometrics* **68**, 29–51.
- Athanasopoulos, G., Ahmed, R. A. & Hyndman, R. J. (2009), 'Hierarchical forecasts for Australian domestic tourism', *International Journal of Forecasting* **25**(1), 146–166.
- Athanasopoulos, G. & Hyndman, R. J. (2008), 'Modelling and forecasting Australian domestic tourism', *Tourism Management* **29**, 19–31.
- Australian Bureau of Statistics (2003), *Tourism Region Maps and Concordance Files*, Technical report, ABS Cat No. 9503.0.55.001.
- Baltagi, B. H. (2001), *Econometrics analysis of panel data*, 2nd edn, Wiley, Chichester, England.
- Baltagi, B.H., S. S. & Koh, W. (2003), 'Testing panel data models with spatial error correlation', *Journal of Econometrics* **117**, 123–150.
- Beenstock, M. & Felsenstein, D. (2007), 'Spatial vector autoregressions', *Spatial Economic Analysis* **2**, 167–196.
- Bhargava, A. & Sargan, J. D. (1983), 'Estimating dynamic random-effects models from panel data covering short time periods', *Econometrica* **51**, 1635–1659.
- Bolduc, D., Dagenais, M. G. & Gaudry, M. J. I. (1989), 'Spatially autocorrelated errors in origin-destination models: A new specification applied to aggregate mode choice', *Transportation Research B* **5**, 361–372.
- Bolduc, D., Laferriere, R. & Santarossa, G. (1992), 'Spatial autoregressive error components in travel flow models', *Regional Science and Urban Economics* **22**, 371–385.
- Cliff, A. D. & Ord, J. K. (1973), *Spatial Autocorrelation*, Pion, London.
- Deng, M. (2008), 'An anisotropic model for spatial processes', *Geographical Analysis* **40**, 26–51.
- Elhorst, J. P. (2003a), 'Specification and estimation of spatial panel data models', *International Regional Science Review* **26**, 244–268.
- Elhorst, J. P. (2003b), *Unconditional maximum likelihood estimation of dynamic models for spatial panels*, Technical report, Research School SOM, Groningen.

- Elhorst, J. P. (2005), 'Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial panels', *Geographical Analysis* **37**, 62–83.
- Hsiao, C. (1986), *Analysis of panel data*, Cambridge University Press, Cambridge.
- Hsiao, C., Pesaran, M. H. & Tahmiscioglu, A. K. (2002), 'Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods', *Journal of Econometrics* **109**, 107–150.
- LeSage, J. P. & Pace, R. K. (2008), 'Spatial econometric modelling of origin-destination flows', *Journal of Regional Science* **48**, 941–967.
- Li, G., Song, H. & Witt, S. F. (2005), 'Recent developments in econometric modelling and forecasting', *Journal of Travel Research* **44**, 82–99.
- Nerlove, M. & Balestra, P. (1996), Formulation and estimation of econometric models for panel data, in 'The Econometrics of Panel Data: A Handbook of the Theory and Applications, 2nd Revised Edition.', Kluwer Academic Publishers, Dordrecht.
- Ord, J. K. (1975), 'Estimation methods for models of spatial interaction', *Journal of the American Statistical Association* **70**, 120–126.
- Porojan, A. (2001), 'Trade flows and spatial effects: The gravity model revisited', *Open Economics Review* **12**, 265–280.
- Song, H. & Li, G. (2008), 'Tourism demand modelling and forecasting – A review of recent literature', *Tourism Management* **29**, 203–220.
- Song, H. & Witt, S. F. (2000), *Tourism demand modelling and forecasting: modern econometric approaches*, Oxford, Pergamon.
- Yu, J., de Jong, R. & Lee, L. (2008), 'Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and t are large', *Journal of Econometrics* **146**, 118–134.