

SPATIAL DIFFUSIONS OF THE TOURISM SYSTEM IN THE GREATER BAY AREA

ABSTRACT. abstract here.

CONTENTS

| | |
|---|---|
| 1. Aims and Objectives | 2 |
| 2. Background of Research | 2 |
| 2.1. Spatial Spillovers in Tourism | 2 |
| 3. Research Plan | 4 |
| 4. Methodology | 4 |
| 4.1. Agent-Based Modeling (ABM) | 4 |
| 4.2. Hierarchical Spatiotemporal Model | 4 |
| 4.3. Spatial Two-Regime Model | 4 |
| 4.4. Data | 6 |
| 5. Project Outcomes and Deliverables | 6 |
| 6. Budget and Budget Justification | 6 |
| 7. Contributions of the Project to SHTM | 6 |

1. AIMS AND OBJECTIVES

2. BACKGROUND OF RESEARCH

2.1. Spatial Spillovers in Tourism.

Tourism demand modeling and forecasting, in companies with the rapid development and increasing importance of international tourism in the global economy, has become an active field of research for a prolonged period since the 1970s. These models and estimations are crucial for tourism demand and economic development analysis, which in turn enables effective government policy planning. (jiaoTourismForecastingReview2019). High-quality analysis on tourism demand and development is also important for business practitioners and local authorities in the decision-making processes (pengMetaAnalysisInternationalTourism2015). Given the importance of tourism demand analysis, tourism researchers have put considerable efforts to address the problem. songReviewResearchTourism20 reviewed the four common types of models being used in tourism demand analysis, including time series models, econometrics models, artificial intelligence (AI) models and judgmental methods. As for the explainability and the potential for identifying causal relationships between economic indicators, econometric methods show their superiority. The estimated structural econometric equations can further be used to measure variables like the price and/or income elasticity of demand, which are essential inputs for economic analysis. In terms of forecasting performance, it is recognized that there does not exist a model that consistently outperforms others in every circumstance (liRecentDevelopmentsEconometric2005; wuNewDevelopmentsTourism2017).

Although sophisticated econometric methods have been implemented for the precision of prediction, most of them treat destinations as independent units. However, as any economic activity, the tourism demand of a destination is highly correlated and interdependent with that of the neighbors (longPoolingTourismDemand2019). Tourism development in one destination affects not only the economic outcomes within the territory, but also produces a spillover effect on other regions if they are close in a geographic or economic sense (or both) (liTourismRegionalIncome2016). The theoretical foundation of the spillover effects in tourism could be attributed to the cluster theory (kimVisitorFlowSpillover2022). porterCompetition2008 defines a *cluster* as a geographically proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and complementarities. The agglomeration of firms could be beneficial through the multiplier and positive externality effects for the local economy, as illustrated by berniniConventionIndustryDestination2009 in the study on the convention industry in Italy. Since the geographical proximity, potential consumers can reduce their searching costs significantly. The accumulation of reputation will also further attract customers to the location (kuahClusterTheoryPractice2002). On the supply side, the benefits include knowledge spillovers, infrastructure benefits and the ready availability of skilled labor in the area (kuahClusterTheoryPractice2002). The spatial spillover effect follows naturally if we think different tourism destinations in a region as firms within the same industry. For example, michaelTourismMicroClusters2003 highlights the

clustering effects driven by the development of tourism on economic opportunity in small communities. The increase in tourist arrivals (and hence, the tourism revenues) echoes the investment in infrastructure and human capital in a reciprocally beneficial way. Meanwhile, the nearby destinations may benefit from the demonstration (and learning) effect, the labor movement and the competition effect from that place (**longPoolingTourismDemand2019**). The synthesis of tourism destinations also happens through other mechanisms, for example, joint promotion activities, accidents (e.g., terrorist attacks) and multi-destination travel patterns (**yangSpatialEconometricApproach2012**). However, traditional econometrics assumes the tourism demand is endogenously determined within the region and therefore neglects spillover effects from others in the economic system. This problem could be substantial, especially for destinations with similar cultural backgrounds (e.g., cities in China). The confounding effects of underlying spatial structure lead to potentially biased estimators for the estimands of interest. Panel data could alleviate this problem as it takes into account both the time and space dimensions for the estimation. And these data can be used to estimate the dependence of the destinations in the system, as well as the spatial heterogeneity of them.

Researchers have begun to apply spatial econometric methods to model and estimate tourism demand and development (**dengModellingAustralianDomestic2011**; **marrocuDifferentTouristsDifferent2013**), as well as the spatial spillover effect of tourism growth (**yangSpatialEffectsRegional2014**; **maTourismSpatialS**). Spatial econometrics has also been used in industrial-level research. **adamPerceivedSpatialAgglomeration201** evaluated the perceived spatial agglomeration effects on hotel location choice. **gunterModelingAirbnbDemand2** used a spatial Durbin model (SDM) to estimate the spillover effect of Airbnb demand in New York City. As for the tourism resorts, **kimVisitorFlowSpillover2022** explored the visitor flow spillover effects in London on attraction using a spatial econometric model. The interdependence of tourism activities between cities in China has also been confirmed. **yangSpatialEffectsRegional2014** found a significant spatial correlation between the regional tourism developments in China. **liuEffectsTourismDevelopment2022** constructed both static and dynamic spatial Durbin models to explore the effects of tourism development on economic growth. These studies show the great power of spatial econometrics models on tourism economics research.

Based on the existing literature, this project identifies the limitations of current research and the unique challenges for spatial tourism econometrics research in the case of the Greater Bay Area.

- (1) The spillover effects are assumed to be homogeneous in traditional spatial models. This is an unrealistic assumption, and incapable of capturing the dynamics in the time dimension for complex tourism systems like the GBA.
- (2) The homogeneity assumption is also insufficient in terms of modeling the asymmetry of spillover effects across the cities in the GBA, which is expected to be significant due to the different development stages of these regions.
- (3) The geographic and political relationships are complicated in the GBA. However, a single spatial weighting matrix is limited in terms of explanatory power. Meanwhile, a single-level spatial

model cannot capture the effects from higher levels (e.g., provincial and special administrative region) variables.

This project develops cutting-edge spatial econometrics models to address these limitations. The specific model settings will be further discussed in the methodology section.

3. RESEARCH PLAN

4. METHODOLOGY

4.1. Agent-Based Modeling (ABM).

4.2. Hierarchical Spatiotemporal Model.

Given the particularity of the administrative and geographic division of the GBA, the single spatial weighting matrix in traditional spatial econometrics models are inadequate in terms of the relationships between cities in the area and those near the border. Therefore, this project decomposes the spatial spillover effects inside and outside the region and across them by using multiple weighting matrices. As an example, in this project, the higher-order dynamic spatial Durbin model (SDM) could be written as

$$y_{it} = \rho y_{it-1} + \lambda_1 \sum_{j \neq i}^N w_{ij}^{\text{non-GBA}} y_{jt} + \lambda_2 \sum_{j \neq i}^N w_{ij}^{\text{GBA}} y_{jt} + \lambda_3 \sum_{j \neq i}^N w_{ij}^{\text{Cross}} y_{jt} + \beta x_{it} + \theta \sum_{j \neq i}^N w_{ij} x_{jt} + \mu_i + \epsilon_{it}.$$

In this equation, y_{it} is the tourism income of city i in period t . λ_1 is the coefficient on the weighted average of the tourism income of cities *outside* the GBA, λ_2 is the coefficient on the weighted average of the tourism income of cities *inside* the GBA, and λ_3 is the coefficient on the weighted average of the tourism income of cities *across* the sub-systems (i.e., the GBA and non-GBA areas). β and θ give information about the effects of x on the dependent variable y . These will be the estimands of our main interest. w_{ij}^{GBA} measures the strength of dependence (defined by distance or commute time) between city i and j *inside the GBA*. If city i or j is not in the GBA, then w_{ij}^{GBA} is defined to be 0. $w_{ij}^{\text{non-GBA}}$ and w_{ij}^{Cross} are defined similarly, with $w_{ij}^{\text{non-GBA}} = 0$ if city i or j (or both) is in the GBA, and $w_{ij}^{\text{Cross}} \neq 0$ only if i and j are in different sub-systems. x_{it} is a vector of control variables, μ_i is a city fixed effect and ϵ_{it} is the error term.

We use Two-stage least-squares regression (2SLS) for estimation. The estimated $\hat{\lambda}_1$, $\hat{\lambda}_2$ and $\hat{\lambda}_3$ measures the interdependence and synthesis effects of the cities inside, outside and across the GBA, respectively. Different from the traditional econometric models, we cannot take $\hat{\beta}$ simply as the effect of x on the panel system and the sub-systems. Instead, we average over the diagonal terms in the resulting Jacobian matrix to get the direct effect of x on tourism revenue. The indirect effects are measured by the average of the off-diagonal elements. Moreover, these indirect effects can further be decomposed into sub-system and cross-system effects. The estimation of these parameters helps us understand the mechanism of tourism development in the GBA and the country.

4.3. Spatial Two-Regime Model.

After basic modeling setting and calibration, this project introduces two-regime spatial econometric models to further formalize the spillover effects in time and space dimensions. In the time dimension, we investigate the effect of the establishment of the GBA on tourism development in terms of the spillover and agglomeration effects. We choose March 2017 (the first time when the GBA plan appears in the government report) as a cutoff point for our analysis. In order to overcome the potential omitted variable bias and the shortcomings of the two-regime spatial lag model (which does not control for the WX variables), **elhorstEvidencePoliticalYardstick2009** suggests including the spatial lags of both the dependent and independent variables for the analysis. We also include the space and time fixed effects to correct for the time-persistent and spatial-persistent variables in the error terms. The structural equation of our two-regime spatial Durbin model could be written as (**wangStrategicInteractionIndustrial2020**)

$$y_{it} = \delta_1 d_{it} \sum_{j=1}^N w_{ij} y_{jt} + \delta_2 (1 - d_{it}) \sum_{j=1}^N w_{ij} y_{jt} + \beta x_{it} + \theta \sum_{j=1}^N w_{ij} x_{jt} + \mu_i + \lambda_t + \alpha + \epsilon_{it}.$$

In this equation, y_{it} is the tourism development indicator of city i at time t . In this project, we use tourist arrivals and tourism revenues as the indicators. w_{ij} measures the relationship of city i and j , and these give us a N -by- N weighting matrix W . There are various ways to define w_{ij} , and we will provide more details in the later sections. x_{it} is a vector of independent variables affecting tourism development, such as income and price levels. Although we include all cities in a single panel, the heterogeneity of each city cannot be neglected. μ_i represents the city fixed effect that controls for all space-specific, time-invariant variables. And λ_t represents the time-period fixed effects which control for all time-specific, spatial-invariant variables. $\sum_{j=1}^N w_{ij} y_{jt}$ is the spatial lag of the dependent variable, and its coefficients δ_1 and δ_2 capture the spatial spillover effect in different contexts (more on this later). Similarly, $\sum_{j=1}^N w_{ij} x_{jt}$ is the spatial lag of the independent variables, and θ represents the spillover effect of the socio-economic conditions of other cities on the tourism development of i .

We define the dummy variable $d_{it} \in \{0, 1\}$ in two different approaches. First, we let $d_{it} = 1$ if city i is in the GBA and t is after 2017, and $d_{it} = 0$ otherwise. With this definition, we can evaluate the effect of the establishment of the GBA on tourism development for cities in the area. $(1 - d_{it}) \sum_{j=1}^N w_{ij} y_{jt}$ and $d_{it} \sum_{j=1}^N w_{ij} y_{jt}$ represent the spatial interdependence of the tourism development in city i with other cities before (i.e., the first regime) and after (i.e., the second regime) the establishment of the GBA, respectively. The coefficients δ_1 and δ_2 suggest the level of spatial interdependence. If δ_1 is greater than δ_2 both statistically and substantively, we may conclude that the establishment of the GBA has a positive effect on the tourism destinations agglomeration effect.

We can also define $d_{it} \in \{0, 1\}$ as follows. Let

$$d_{it} = \begin{cases} 1, & \text{if } y_{it} > \sum_{j=1}^N y_{jt}; \\ 0, & \text{otherwise.} \end{cases}$$

In words, if the tourism development level of city i is higher than other cities at period t , then we set $d_{it} = 1$. If city i is a “less-developed/weak” city, then $d_{it} = 0$. Now, we can investigate the spatial spillover effects that account for the heterogeneity of development stages of different cities. According to the significance level and signs of δ_1 and δ_2 , we can further identify different forms of the spillover of the cities in the GBA. We list these nine forms in table (1).

TABLE 1. Nine forms of the spillover effects.

| | $\delta_2 < 0$ | δ_2 insignificant | $\delta_2 > 0$ |
|--------------------------|--|---|--|
| $\delta_1 < 0$ | completely suppressed by the development of neighbors | suppressed by the development of weak neighbors | suppressed by the development of weak neighbors, reinforced by the development of strong neighbors |
| δ_1 insignificant | suppressed by the development of strong neighbors | independent development | reinforced by the development of strong neighbors |
| $\delta_1 > 0$ | reinforced by the development of weak neighbors, suppressed by the development of strong neighbors | reinforced by the development of weak neighbors | completely reinforced by the development of neighbors |

Sources: **diaoSpatialSpilloverEffect2021** and the authors’ own compilation.

4.4. Data.

5. PROJECT OUTCOMES AND DELIVERABLES

6. BUDGET AND BUDGET JUSTIFICATION

7. CONTRIBUTIONS OF THE PROJECT TO SHTM