Socio-Economic Planning Sciences 71 (2020) 100820

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

journal homepage: http://www.elsevier.com/locate/seps

Evaluating the value of an entrepreneurial city with a spatial hedonic approach: A case study of London

Hiroki Nakamura

Faculty of Commerce, Chuo University, 742-1 Higashi-nakano, Hachioji-shi, Tokyo, 192-0393, Japan

ARTICLE INFO

Keywords:
Entrepreneurship
Land price function
Spatial hedonic approach
London

ABSTRACT

This study assesses the value of a city by using the housing price function with a geographically weighted regression model, including various social, economic, and environmental factors. To assess these values, various specific data scores—such as those related to ethnic groups, green areas, crime rates, education, unemployment rates, number of entrepreneurs, and environmental quality—were considered in a case study of London. The results indicate that some variables such as time to city center and entrepreneurship have a positive impact on the local areas' value in London, as shown by housing prices, while those related to unemployment have a negative impact. Moreover, although the London city center has benefitted more in terms of value than its outskirts have, a few specific policies related to startups and entrepreneurship have succeeded in connecting these areas to existing companies and entrepreneurs. In general, large cities may be better equipped to promote such startup and entrepreneurship policies under local industry plans for future development.

1. Introduction

Land and housing prices are factors that can, to a certain extent, be used to evaluate economic values in urban areas and cities both objectively and quantitatively. Economic value evaluation in an area varies in response to environmental factors such as the convenience of the area, social and welfare factors, economic values, and amenities. These factors cannot be treated separately, as their interactions can affect the calculation. The price of land is an example of a metric that can be used to quantitatively evaluate cities and local areas. For example, while living in a high-class residential district with high housing prices may be desirable and confer a certain level of status, areas that lack proper infrastructure and face high unemployment and decay are not attractive, and so housing prices often go down.

Location is an important factor for entrepreneurial activities, which invigorate the economy and play a role in stimulating innovation and housing prices. For example, entrepreneurs may want to build credibility by having offices in prime locations, or aim to increase the efficiency of business and expand sales channels by setting up offices in business districts, both of which have high land values. This suggests that startup companies and entrepreneurs will be drawn toward areas where entrepreneurial activities are thriving and, therefore, toward high housing prices. Consequently, the land value may rise further.

At the same time, the gap between the rich and the poor widens, and the amenities and living conditions in some areas decline because of the housing shortage caused by rising housing prices in such areas. For example, Silicon Valley in the United States is globally renowned as a region that thrives with entrepreneurial activity. However, while housing rent is rising here owing to the influx of affluent people who are successful in business, there is an increase in the number of homeless people and a shortage of housing for people in lower income groups.

Beyond the United States, Tel Aviv, Israel; Beijing, China; and London, United Kingdom, are garnering attention as new hubs of entrepreneurial activity for startup companies. In its 2017 Global Startup Ecosystem Report, the San Francisco-based Startup Genome LLC [1], known for its evaluation of startup companies, ranked Silicon Valley, United States; New York, United States; and London, United Kingdom, as the top three cities in terms of the number of startups.

The complexity of recently developed urban areas and cities has motivated researchers to focus on the sustainable development of such areas and cities using a multidisciplinary approach, while embracing discourses on science, technology, and environmental policy [2,3]. As an area's reputation is reflected in land price, investigation and comparison of the economic value of an area (as expressed by land and housing prices) must be, at an urban level, related to entrepreneurial, environmental, social, and economic factors. From an economic perspective, many scholars have emphasized the importance of start-up firms stemming from entrepreneurship, which can stimulate economic

E-mail address: hnakamu@tamacc.chuo-u.ac.jp.

https://doi.org/10.1016/j.seps.2020.100820
Received 26 June 2018; Received in revised form 11 February 2020; Accepted 1 March 2020
Available online 5 March 2020
0038-0121/© 2020 Elsevier Ltd. All rights reserved.

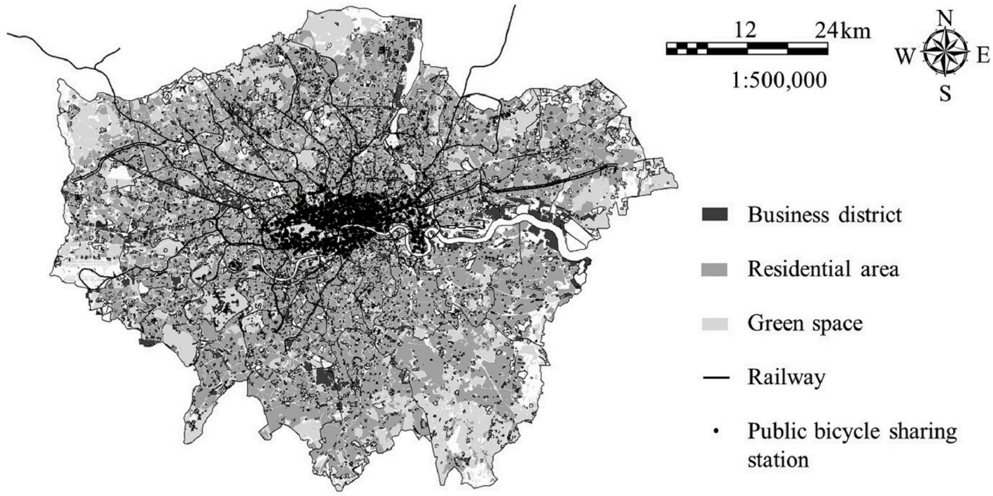

Fig. 1. Land utilization and public transportation in London.

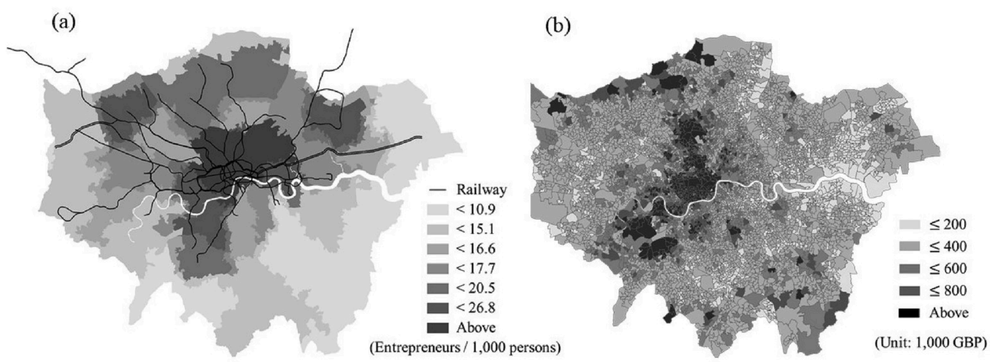

Fig. 2. Maps of the (a) number of entrepreneurs and (b) housing prices in London.

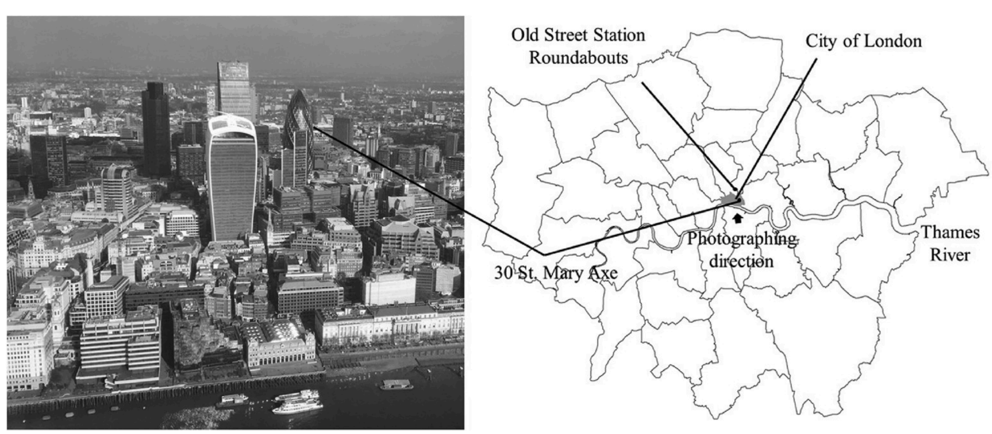

Fig. 3. Photo and locations in the vicinity of City of London.

growth through job creation and innovation, not only for countries but also for cities. Entrepreneurial ecosystems are expected to promote regional activities through the formation of new firms [4].

Assessing the value of local areas and cities, including examining the entrepreneurial, environmental, economic, and social dimensions of sustainability, has become key to fostering development, and such assessments can be conducted using a land-price function [5]. Nakamura [5] analyzed the land-price function using a geographically weighted regression model of land price, as well as explanatory variables related to entrepreneurial, environmental, economic, and social factors in the case of Japan's regional areas. The analysis results revealed that areas around large Japanese cities have benefited more from entrepreneurship and social vitality than have areas around small and mid-sized cities [5]. However, the study did not consider inner factors within only urban areas or cities but examined factors within larger areas in the country; moreover, the factors considered for the analysis were limited.

In the present study, London is the focus of analysis to determine the relationships among entrepreneurial activities, employment situations, and environmental/social/economic factors. London is one of the largest cities garnering attention as new hubs of entrepreneurial activity

Table 1
Summary of variables considered in this study.

Variable		Unit	Mean	Standard Deviation	Data Source
	Housing price (average)	£	396,155.4	356,421.8	HM Land Registry data 2012 from the GOV.UK website: https://www.gov.uk/topic/land-registration/data
Whi	Caucasian	%	60.02	20.48	2011 census data from the Office for National Statistics of the UK website: https://www.ons.gov.uk/census/2011census/2011censusdata
Asi	Asian	%	17.63	15.91	
Bla	Black	%	12.88	11.15	
Mix	Mixed race	%	4.84	1.92	
UK	British	%	64.11	14.47	
Chr	Christian	%	48.97	12.09	
Bud	Buddhist	%	0.99	0.73	
Jws	Jewish	%	1.84	5.68	
Isl	Muslim	%	11.95	10.59	
Emp	Employment rate	%	62.28	8.48	
Uep	Unemployment rate	%	2.00	1.34	
Crm	Crime	No. of Crimes	149.61	239.99	2012 data from the London Analysts Support Site (LASS): https://lass.london.gov.uk/lass/
Lan	Land area	1000 m^2	334.99	657.58	Generalised Land Use Database (GLUD) 2005 (Enhanced base map) statistics
Bld	Building occupancy rate	%	13.72	6.51	
Gre	Green coverage ratio	%	20.35	18.40	
Wat	Waterfront space	%	1.01	4.91	
Env	Environmental quality (environmentally hazardous substances)	mg/m^3	100	14.80	London Atmospheric Emissions Inventory (LAEI) data 2013 from the following website: https://data.london.gov.uk/dataset/london-atmospheric-emissions-inventory-2013
Off	Proportion of households with access to nearest employment centers within a reasonable travel time by public transport/walking	Min	7.74	2.75	Accessibility Indicators data 2011 by the Department for Transport, UK from the following website: https://data.gov.uk/dataset/6ce25e42-bdac-4a6c-9d75-736c7e7d1139/accessibility-indicators
Cut	Proportion of households with access to nearest city centers within a reasonable travel time by public transport/walking	Min	11.68	4.88	
Ent	Number of entrepreneurs per 1000 people	People	22.88	27.87	StartUp Britain 2014, UK Start-up Index data from StartUp website: http://startupbritain.org/startup-tracker/
Qcd	Academic score (The General Certificate of Secondary Education [GCSE] Qualifications and Curriculum Authority [QCDA] point score)	Average points per student	208.90	18.21	Neighbourhood statistics in England: academic year 2011–2012 by the UK Department of Education website: https://www.gov.uk/government/statistics/neighbourhood-statistics-small-area-pupil-attainment-and-absence-by-pupil-characteristics-in-england-academic-year-2010-to-2011

for startup companies, and it has various factors such as social, economic, and environmental factors to be considered for the analysis. The methodology for this study's analysis involves a spatial hedonic approach using a geographically weighted regression (GWR) model of housing price. This is because this study focuses on analyzing the inner condition of or relationship between various factors within urban areas, and GWR is an important local technique for exploring spatial heterogeneity in data relationships.

2. Previous research on the analysis method

2.1. Regional/city evaluation method

In order to obtain a better picture of how society is doing, it is indispensable to consider beyond the ordinary income-based economic measures that are inadequate to capture the societal progress and shift the awareness to more comprehensive measures that incorporate multi-faceted human-centric criteria [6].

Therefore, many local governments use resident questionnaires as part of their regional/city evaluation and policy evaluation. This is called subjective region/city evaluation. In the questionnaire survey, satisfaction levels are ranked on a five-point scale for various evaluation items, and the results are tabulated. Thus, the items are often weighed and analyzed, not only for satisfaction levels but also for importance (Refer to Ref. [7] as an example of this analysis.). Residents' degree of satisfaction may be greatly influenced by their expectations for each item, thus the need for weighting using importance [8]. In other words, there is a possibility that satisfaction will increase (or vice versa) as a result of fewer expectations and requests for a certain item.

Meanwhile, regional and city evaluations based on objective data have also developed independently in several major countries, mainly for the purpose of local government policy evaluations [9–11]. Typical examples include the British Building Research Establishment Environmental Assessment Method for communities, the U.S. Leadership in Energy and Environmental Design for neighbourhood development, and Japan's Comprehensive Assessment System for Built Environment Efficiency for cities. The common feature of these evaluation ratings is that scores are calculated based on objective data for multiple items covering different fields such as the environment, society, and the economy, and the values are integrated and evaluated comprehensively from the viewpoint of urban efficiency and sustainability [9,12].

With the perspective that residents' subjective evaluation in cities is indispensable, it is important to conduct regional/city evaluation using both objective and subjective data. However, there are still many problems in terms of how to combine subjective and objective data, and research in this area is not extensive. Some studies, such as those by Hagerty [13,14], Kawakubo et al. [15], and Borseková et al. [16] and those dealing with subjective and objective data at the same time, have analyzed the relationship between these data types. Kawakubo et al. [15] enabled the evaluation of objective and subjective data in the same framework, and they found a high correlation between them. They showed that residents of cities with high objective evaluations related to the quality of the environment (i.e., cities with a good natural environment and no environmental pollution) have high satisfaction with the environment.

However, these subjective and objective evaluations do not always match. Some items are not significantly correlated or have a negative correlation [13,14,17,18]. A high objective evaluation related to a certain item does not necessarily mean that the residents' satisfaction level with the item will be high. For example, residents living in a city with a high literacy rate need not have a high level of satisfaction with the education system in the city [18].

Table 2
Correlation between variables.

	Whi	Asi	Bla	Mix	UK	Chr	Bud	Jws	Isl	Emp	Uep	Crm	Lan	Bld	Gre	Wat	Env	Off	Cit	Ent	Qcd
Whi	1.000																				
Asi	−0.747	1.000																			
Bla	−0.568	−0.057	1.000																		
Mix	−0.266	−0.234	0.621	1.000																	
UK	0.748	−0.584	−0.326	−0.282	1.000																
Chr	0.586	−0.715	0.058	−0.001	0.604	1.000															
Bud	−0.208	0.156	0.070	0.147	−0.428	−0.218	1.000														
Jws	0.092	−0.015	−0.171	−0.113	−0.041	−0.341	0.001	1.000													
Isl	−0.768	0.689	0.289	0.080	−0.613	−0.617	0.048	−0.081	1.000												
Emp	0.617	−0.383	−0.481	−0.139	0.352	0.287	−0.058	0.054	−0.639	1.000											
Uep	−0.466	0.181	0.527	0.224	−0.138	−0.136	−0.069	−0.143	0.490	−0.698	1.000										
Crm	−0.085	0.027	0.038	0.091	−0.210	−0.120	0.177	−0.031	0.063	−0.028	0.008	1.000									
Lan	0.149	−0.059	−0.133	−0.180	0.200	0.141	−0.087	0.028	−0.128	0.043	−0.067	0.051	1.000								
Bld	−0.029	−0.001	0.028	0.168	−0.221	−0.120	0.020	0.011	0.058	0.161	−0.127	−0.064	−0.392	1.000							
Gre	0.108	−0.092	−0.022	−0.105	0.226	0.179	−0.091	−0.007	−0.099	−0.081	0.075	−0.067	0.495	−0.693	1.000						
Wat	0.011	−0.019	−0.005	−0.028	−0.033	0.013	0.024	−0.017	0.003	0.016	−0.012	0.033	0.190	−0.220	0.042	1.000					
Env	−0.198	0.027	0.148	0.327	−0.541	−0.326	0.313	0.014	0.255	−0.087	−0.011	0.358	−0.224	0.236	−0.293	0.008	1.000				
Off	0.070	−0.036	−0.011	−0.104	0.274	0.168	−0.188	−0.026	−0.111	−0.012	0.072	−0.147	0.089	−0.098	0.160	−0.032	−0.373	1.000			
Cit	0.002	0.008	0.004	−0.085	0.075	0.106	−0.056	−0.081	−0.039	−0.082	0.092	−0.075	0.156	−0.165	0.192	0.020	−0.133	0.374	1.000		
Ent	−0.053	0.040	−0.032	0.072	−0.200	−0.263	0.111	0.075	0.145	−0.057	0.005	0.179	−0.073	0.035	−0.110	0.059	0.440	−0.151	−0.009	1.000	
Qcd	0.031	0.025	−0.086	−0.064	0.147	−0.039	−0.047	0.081	−0.049	0.093	−0.067	−0.087	0.034	−0.060	0.014	−0.053	−0.178	0.089	−0.052	−0.084	1.000

Table 3
Results of the principal component analysis for all variables.

Variables	Principal Components					
	1	2	3	4	5	6
Whi	−0.40	−0.17	0.02	0.08	0.05	0.12
Asi	0.30	0.12	−0.40	−0.16	0.05	−0.23
Bla	0.22	0.16	0.48	0.02	−0.14	0.06
Mix	0.15	−0.07	0.48	0.11	−0.15	0.13
UK	−0.38	0.11	0.07	−0.09	−0.04	0.14
Chr	−0.32	0.07	0.37	0.03	0.04	−0.19
Bud	0.15	−0.16	−0.03	0.25	0.00	−0.28
Jws	−0.01	−0.10	−0.27	0.01	−0.24	0.63
Isl	0.38	0.15	−0.12	−0.09	0.01	−0.01
Emp	−0.29	−0.32	−0.09	0.02	0.02	−0.14
Uep	0.22	0.33	0.23	−0.07	−0.03	0.19
Crm	0.09	−0.12	0.01	0.38	0.10	−0.03
Lan	−0.13	0.26	−0.17	0.34	−0.08	−0.05
Bld	0.09	−0.38	0.13	−0.39	0.15	−0.02
Gre	−0.12	0.39	−0.08	0.33	−0.15	0.06
Wat	−0.01	0.06	−0.06	0.30	0.01	−0.19
Env	0.23	−0.32	0.09	0.30	0.14	0.13
Off	−0.11	0.26	0.01	−0.23	0.39	0.17
Cit	−0.05	0.24	−0.01	0.01	0.64	0.11
Ent	0.12	−0.17	−0.06	0.28	0.28	0.45
Qcd	−0.06	0.04	−0.13	−0.18	−0.41	0.15
Eigenvalue	4.88	2.72	2.13	1.76	1.17	1.08
Proportion	0.23	0.13	0.10	0.08	0.06	0.05
Cumulative	0.23	0.36	0.46	0.55	0.60	0.65

Based on the above points, a method has been developed to estimate the subjective evaluation using objective and economic data. For example, public property characteristics in the region, such as regional networks, safety, the environment, and forms of livelihood, are non-market goods, or goods that are not traded in the market. A method for estimating the value of these characteristics has been developed in environmental economics.

Virtual evaluation methods include virtual market valuation methods and conjoint laws, in which virtual scenarios are created and the value of non-market goods is estimated. In order to quantitatively evaluate the value of public goods in the region using a virtual market evaluation method, it is necessary to receive a response about the value in the questionnaire.

Another method that uses a proxy market, for example, is known as the hedonic approach by Rosen [19]. In this approach, the value of non-market goods is measured from proxy markets such as land, housing, and labor markets. The hedonic approach hypothesizes that various characteristics such as the surrounding environment (which is a component of goods) are capitalized by land or housing prices. Thus, the relational expression called the land or housing price function is estimated, and the characteristic value is evaluated by regressing the land or housing prices on various characteristics. For example, variables representing the characteristics of the area include unemployment rate, number of criminal cases, land area, building occupation rate, labor force population, green area ratio, workplace access, and academic score. In addition, some analyses also consider the number of immigrants, proportion of each race, proportion of various religious beliefs, and so on.

The features of this approach will be examined in detail in the next section in order to conduct a regional evaluation based on the spatial hedonic approach as the basis of this hedonic approach.

2.2. Spatial hedonic approach

In the 1970s, the hedonic approach was introduced as a theory consistent with the microeconomic theory put forth by Rosen [19] and others, and it has developed greatly since then. The hedonic housing price function is expressed by the following equation.

Table 4
Analysis results and models based on explanatory regression.

Variables	Model 1	Model 2	Model 3	Model 4	Model 5
Whi	3021.4*** (253)	–	–	2961.8*** (252.8)	–
Jws	10,130*** (816.5)	10,734.2*** (828)	10,474.6*** (829.7)	10,383.8*** (816.9)	10,765.2*** (829.6)
Emp	–	–	347.5 (767.8)	–	465.8 (766.3)
Uep	−66,474.1*** (3875.1)	−89,169.5*** (3495)	−86,204.6*** (4882.3)	−68,141.1*** (3885.6)	−87,105.2*** (4783.2)
Cit	–	4779.7*** (954.7)	–	4194.3*** (942.7)	4797.6*** (955.2)
Ent	2561.3*** (164.9)	2447.1*** (166.6)	2449*** (167.5)	2562.1*** (164.6)	2454.3*** (167)
_cons	268,824*** (21,086.2)	441,309*** (14,248.1)	470,011*** (55,820.7)	226,253*** (23,117.6)	407,733*** (57,043.3)
N	4765	4765	4765	4765	4765
VIF	1.29	1.03	1.10	1.31	1.10
F	329.4	293	285.3	268.6	234.4
AIC	134,202	134,318	134,343	134,185	134,320
Adj R^2	0.216	0.197	0.193	0.219	0.197

Table 5
Analysis results of the principal component (PC) scores and PC model.

PC	Mean PC Score (10^{-10})	Standard Deviation	Min	Max	PC Model Result
PC1: Immigrant component	5.49	2.21	−7.09	7.66	−17,031*** (1944)
PC2: De-urbanization component	−0.81	1.65	−7.32	8.67	−90,899*** (2606)
PC3: Racial polarization component	−3.51	1.46	−5.88	7.07	−32,332*** (2946)
PC4: Mixed-land use component	−3.01	1.33	−3.61	14.96	60,764*** (3242)
PC5: City access component	4.63	1.08	−3.99	9.44	67,760*** (3980)
PC6: Entrepreneurial component	−1.42	1.04	−4.75	8.87	36,502*** (4129)
_cons	–	–	–	–	394,160*** (4296)
N	–	–	–	–	4765
F	–	–	–	–	355.5
AIC					133,607
Adj R^2	–	–	–	–	0.309

$$p_j = \beta_0 + \sum \beta_p {}^* x_{j,p} + \sum \beta_q {}^* y_{j,q} + \varepsilon_{i,j}. \tag{1}$$

The composition of each quality of goods and services affects the price of real estate and land. However, real estate also has social and regional cohesion. Therefore, price formation shows regional cohesion, and the prices of land and real estate belonging to neighboring areas are expected to have a strong correlation. This is called spatial autocorrelation.

Based on the above, there is a criticism that the premises for regression analysis have not satisfied the hedonic approach's requirement for sample independence, and econometric models in previous studies take spatial correlation into account (e.g., refer to Refs. [20–27]). Representative models in these spatial econometrics include the spatial autocorrelation regression model and spatial error model.

With spatial interdependence, a feature of house price data, estimation is complicated by the presence of a spatially lagged dependent variable, which is typically correlated with the disturbance terms [28]. Spatial econometrics sometimes includes all-encompassing specifications involving various autoregressive spatial lags [29]. When there is an obvious spatial correlation and a spatial spillover effect, the traditional econometric methods may lead to a biased estimation [30]. Regular spatial econometric models include Spatial Lag Model (SLM), Spatial Error Model (SEM), and Spatial Durbin Model (SDM) [31].

In any of the above cases, parameter estimation for the whole area is performed using sample data pertaining to the area to be researched. Emphasis is placed on the spatial homogeneity of the target areas; information based on sample data is aggregated to find the general rules of the area. Therefore, when the price of the land is regressed on each explanatory variable, for example, only one estimation of the coefficient of the time distance to the central business district can be obtained for the whole area.

Therefore, the analysis in the present study involved the use of a regression model assuming spatial non-stationarity. The hedonic approach using these spatial statistical analysis models is called the spatial hedonic approach.

2.3. Geographically weighted regression model

Research on the spatial hedonic approach performs local parameter estimation using sample data for the studied areas and aims to discover differences between the research areas and exceptional areas to determine the general rules of the target area. Therefore, such a model will estimate coefficient parameters, among other factors, locally. For example, the GWR model used in this study is a representative model used in Fotheringham et al. [32] and Kestens et al. [33]. Details of GWR are provided below.

In the usual linear regression model, the regression coefficient is assumed to be constant regardless of the point as in Eq. (1), but in the GWR model, regression coefficients and constant terms are given for each point, as in Eq. (2), and are characterized by points.

$$p_j = \beta_{0,j} + \sum \beta_{p,j} {}^* x_{j,p} + \sum \beta_{q,j} {}^* y_{j,q} + \varepsilon_{i,j} \tag{2}$$

For the sake of simplicity, if the above equation is expressed by the following equation, β_j is a regression coefficient at point *j*.

$$p_j = \sum \beta_j {}^* V_j + \varepsilon_j \tag{3}$$

Now, assuming β_j, all observed values are weighted by the distance to *j* as follows.

$$W_j^{1/2} p = W_j^{1/2} V \beta_j + W_j^{1/2} \varepsilon \tag{4}$$

Table 6
Spatial lag models based on previous models and results.

Variables	Lag model 1	Lag model 2	Lag model 3	Lag model 4	Lag model 5	PC lag model
Spatial lag of housing price	3.26×10^{-4}*** (1.22)	2.74×10^{-4}*** (1.22×10^{-5})	2.55×10^{-4}*** (1.21×10^{-5})	3.44×10^{-4}*** (1.22×10^{-5})	2.74×10^{-4}*** (1.22×10^{-5})	3.03×10^{-4}*** (1.66×10^{-5})
Whi	4944.4*** (246.5)	–	–	4936*** (244.3)	–	–
Jws	8648.1*** (763.2)	9869.9*** (788.4)	9478.4*** (795.1)	9062.4*** (757.8)	9909.2*** (789.9)	–
Emp	–	–	384.8 (734.5)	–	592.8 (728.8)	–
Uep	−46,628.2*** (3687.8)	−84,922.5*** (3329.4)	−81,107.0*** (4676.8)	−48,778*** (3662.4)	−82,294.1*** (4639.7)	–
Cit	–	8296.7*** (921.5)	–	8231.1*** (884.4)	8320.5*** (921.9)	–
Ent	1207.8*** (164.9)	1250.7*** (167.2)	1332.4*** (168.8)	1132.8*** (160.6)	1259.5*** (167.6)	–
PC1: Immigrant component	–	–	–	–	–	−41,561.1*** (2310.3)
PC2: De-urbanization component	–	–	–	–	–	−53,545.2*** (3245.5)
PC3: Racial polarization component	–	–	–	–	–	−47,657.9*** (2969.6)
PC4: Mixed-land use component	–	–	–	–	–	40,083.7*** (3332.8)
PC5: City access component	–	–	–	–	–	67,586.9*** (3847.5)
PC6: Entrepreneurial component	–	–	–	–	–	23,966.6*** (4050.5)
_cons	−199,742*** (26,312.7)	129,350*** (19,418.4)	213,412*** (54,776.5)	−309,767*** (28,633.6)	86,529.1 (56,109.3)	70,933.2*** (18,180.6)
N	4765	4765	4765	4765	4765	4765
AIC	133,535	133,842	133,922	133,452	133,843	133,286
Adj R^2	0.319	0.274	0.261	0.331	0.274	0.354

$$W_j = \begin{pmatrix} a_{j1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{jn} \end{pmatrix} \tag{5}$$

W_j is an $n \times n$ matrix expressed by Equation (5), and the diagonal component a_{jk} ($k = 1, 2, \ldots, N$) of W_j is a weight given to point k. The estimated regression coefficient at point j is given by the following equation.

$$\widehat{\beta}_j = (V' W_j V)^{-1} V' W_j p \tag{6}$$

Here, the most widely used method to assign $\boldsymbol{a_{jn}}$ is to assign a continuous function that depends on distance, as follows.

$$a_{jn} = \exp\left(\frac{-d_{jn}^2}{\varphi^2}\right) \tag{7}$$

$$a_{jn} = \begin{cases} \left[1 - \left(\dfrac{d_{jn}}{\varphi}\right)^2\right]^2, & if\ d_{jn} < \theta \\ 0, & otherwise \end{cases} \tag{8}$$

Here, d_{jn} is the Euclidean distance between j and n, and φ is a parameter called kernel bandwidth. Here, when using functions such as Equations (7) and (8), φ gives stronger weight to observation values close to the distance to j. When φ is relatively small, that is, when considering only the point in the vicinity of j, the standard error of the estimate of the coefficient increases, and, conversely, when φ is large, the bias becomes large when a large number of points is considered. Although φ is affected by standard error and bias, the most frequently used method is to adopt φ that minimizes the cross validation score as follows.

$$CV(\varphi) = \sum \{p_j - \widehat{p}_{\neq j}(\varphi)\}^2 \tag{9}$$

GWR has recently been applied several times in various fields, such as the environment (e.g., Ref. [34]), real estate (e.g., Ref. [35,36]), urban infrastructure (e.g., Ref. [37]), and the spatial hedonic model. GWR is often used in Geographic Information Systems (GIS) software, which represents a valid tool in addressing complex, multi-dimensional problems at various scales [38], and is useful tool for exploring spatial heterogeneity in data relationships. However, the GWR model in many studies was established based on screened predictors through preliminary correlation analysis, and this process might result in the omission of necessary factors [39].

Although many prior studies using the spatial hedonic approach attempted to control for spatial effects by increasing the sample size [37], some studies in the field of environmental research, such as Qu et al. [40], Sabokbar et al. [41], and Zhai et al. [39], integrated the principal component analysis (PCA) with GWR and developed a best subsets regression modeling approach for the estimation by fully considering all the potential variables' contributions simultaneously. PCA is a statistical analysis method to manage large data sets through understanding the relationship between data by extracting the main component of those data [42]. The main component is the representative indicator that explains the characteristics of each data item properly.

3. Description of focus area and analysis data: London

3.1. London and Tech City concept

In this study, Greater London, referred to as London in this paper, is adopted as the subject of analysis. London includes the City of London and its 32 boroughs; the data used in this study include the detailed administrative unit for every district.

Fig. 1 shows a map of the land utilization in London. In London, the transportation networks are spread mainly in the city center; the figure shows that the transport infrastructure in the areas north of the Thames River, which runs through the middle of London, is more advanced than that in the south. Furthermore, in recent years, availability of the public

Table 7
Result statistics of the geographically weighted regression of Model 4.

		Average	Min	Max	Std. dev.
(a)	Local R^2	0.39	0.22	0.66	0.09
(b)	Estimated value of the dependent variable	396,907	−438,147	3,651,829	238,573
(c)	Coefficient value of explanatory variable related to Caucasian	4975	410	13,520	3675
(d)	Coefficient value of explanatory variable related to Jewish	96,394	−1240	454,729	122,570
(e)	Coefficient value of explanatory variable related to unemployment rate	−32,369	−111,724	667	2,0624
(f)	Coefficient value of explanatory variable related to time to city center	4448	−630	20,398	4356
(g)	Coefficient value of explanatory variable related to entrepreneur	8434	−4614	47,048	13,496
(h)	Contributing value of explanatory variable related to Caucasian	296,937	5186	1,154,338	252,885
(i)	Contributing value of explanatory variable related to Jewish	50,981	−1688	3,259,453	102,834
(j)	Contributing value of explanatory variable related to unemployment rate	−60,146	−652,830	2124	55,492
(k)	Contributing value of explanatory variable related to time to city center	71,956	−22,004	340,224	63,617
(l)	Contributing value of explanatory variable related to entrepreneur	94,324	−92,031	1,089,174	169,275
N		4765			
Bandwidth		17,481			
AIC		131,959			
R^2 adj.		0.52			

bicycle-sharing system, which is rapidly becoming a new form of public transportation, has increased intensively in central London.

Fig. 2 shows the (a) number of entrepreneurs and (b) housing prices in London. The figure shows that housing prices are high in the city and central areas as well as in the north and western Thames River areas. The central areas tend to have higher values for all these variables. Particularly, areas with high values are scattered in the eastern part of the city along the Thames River. It is assumed that this is because several industrial and commercial sites are established in those areas. There are many entrepreneurs in the central area, and the number increases along the public transportation network.

In London, the then Prime Minister David Cameron announced in November 2010 the East London Tech City concept (Tech City concept) to provide aggressive support for infrastructure and so on, so that startups and tech companies can come together in this area. According to the Tech City concept, the areas surrounding the financial districts on the eastern side of the central part of London known as the City of London—specifically, the areas from Old Street Station to the eastern Olympic stadium—will further accelerate entrepreneurship as a Tech City, with the aim to develop as the British version of Silicon Valley (Fig. 3).

Through the Tech City concept, the government provides supportive measures in the form of tax incentives for businesses and tax reduction measures for investors. Further, a non-profit organization called Tech City UK provides technical support to new entrepreneurs trying to establish businesses, and it links regional environments and related stakeholders such as the government, education officials, and investors in areas where activities will be centered. After these measures were adopted, the first global companies, including Google and Facebook, opened offices in Tech City. The number of companies based in Tech City increased rapidly, and Tech City is now a core strand in the vision for the further redevelopment of East London [43].

Ten IT and related companies were already located on Old Street in what would become London's Tech City by 2008, before the Tech City concept even emerged. As these companies gradually gained attention during this period when businesses were growing rapidly, the area became known as Silicon Roundabout, a play on Silicon Valley that referred to the traffic roundabouts on Old Street. There are now thousands of companies in that area. Google purchased a seven-story building near Old Street Station in 2011, and Google Campus London opened in 2012, making it the base for co-working space and other workshop events.

However, although the Silicon Roundabout region is close to central London, it is not in the center of the financial district, and it was originally not a highly reputable area. Therefore, the rent was cheap, and empty houses were prominent. Thus, artists who did not have much money gathered in the area, followed by people from various fields. When people gathered, shops, cafes, and a creative environment focusing on youth came to the area. The UK government continues to focus early on such changes in the specific regional environment and carries out unified work and provides entrepreneurship support to promote cooperation between the financial district and central urban areas of London for revitalizing entrepreneurial activities.

Therefore, there is a strong connection between the new entrepreneurial activities and existing large enterprises in the areas of Tech City and Silicon Roundabout. This is a major difference from the advanced region of Silicon Valley. While Silicon Valley is more than 50 km away from the San Francisco city center, Tech City and Silicon Roundabout are right next to central London. As these financial districts are close, many startups in the field integrate financial and informational technologies in Tech City (Fin Tech).

3.2. Analysis data for London

Various data related to the social, economic, and environmental factors are collected to analyze the relation between housing price and these factors. The data include housing prices, nationality, race, religion, and employment and unemployment rates from the UK census; and environmental quality and number of entrepreneurs in each area within London. A detailed summary of all the data considered in this study is presented in Table 1. Because of data constraints, the data are not all from the same year; where required, data have been taken from the closest available year. These data points seem to have a considerable relationship with housing price and include the proportion of each race, proportion of religious beliefs, employment rate, unemployment rate, number of crimes, land area, building occupancy rate, green area ratio, waterfront space, amount of environmentally harmful substances, workplace access, number of entrepreneurs, and academic score.

However, it is possible that, for example, areas inhabited by many Caucasians are also inhabited by many Christians, or areas inhabited by many non-British citizens have a high unemployment rate; thus, the features may overlap. As mentioned above, in the spatial hedonic approach, strong correlation between the variables is a serious analytical problem because of the spatial nature of the area. Therefore, the correlation between these variables was analyzed (Table 2). A high correlation was found between multiple variables. Here, it is also assumed that the housing price in areas with high crime rate is high. However, for several reasons, including a strong correlation between the variables, the crime rate variable can be excluded from the analytical model that needs to be considered in the preliminary regression analysis. For example, according to Table 2, crime rate has a high correlation with employment rate.

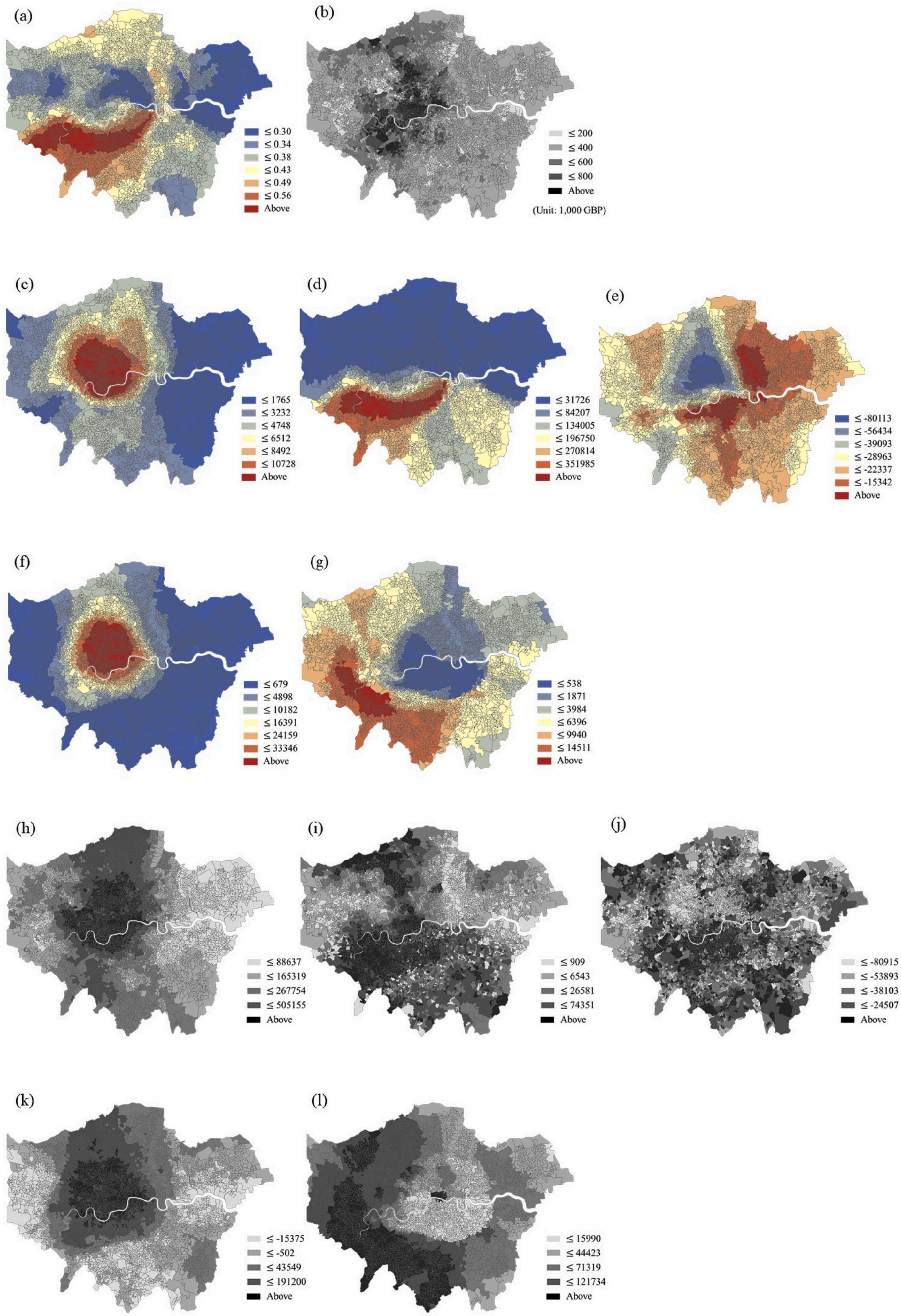

Fig. 4. Maps with results of the geographically weighted regression of Model 4: (a) local R^2 values; (b) estimated values of the dependent variable; (c)-(g) coefficient values of explanatory variables related to Caucasian, Jewish, unemployment rate, time to city center, and entrepreneur, respectively; (h)-(l) contributing values of each explanatory variable (Caucasian, Jewish, unemployment rate, time to city center, and entrepreneur, respectively) to estimated values of the dependent variable.

Table 8
Result statistics of the geographically weighted regression of the principal component (PC) model.

		Average	Min	Max	Std. dev.
(a)	Local R^2	0.36	0.15	0.61	0.08
(b)	Estimated value of the dependent variable	400,869	−348,827	2,001,935	233,249
(c)	Coefficient value of PC1	−34,089	−143,661	274	28,387
(d)	Coefficient value of PC2	−55,620	−163,341	56,094	39,202
(e)	Coefficient value of PC3	−58,748	−144,459	−22,401	26,544
(f)	Coefficient value of PC4	7592	−117,364	156,561	42,229
(g)	Coefficient value of PC5	35,197	−48,084	210,281	62,956
(h)	Coefficient value of PC6	27,363	−117,980	152,505	49,525
(i)	Contributing value of PC1	−9033	−749,378	286,384	83,088
(j)	Contributing value of PC2	29,571	−555,015	798,549	113,668
(k)	Contributing value of PC3	5526	−370,510	496,492	82,314
(l)	Contributing value of PC4	−5877	−925,394	471,312	58,213
(m)	Contributing value of PC5	3168	−465,838	932,088	91,011
(n)	Contributing value of PC6	−9449	−508,127	609,977	52,669
N		4765			
Bandwidth		8526			
AIC		132,019			
R^2 adj.		0.51			

Then, a preliminary regression analysis considering all the variables was conducted prior to the actual analysis in order to omit endogenous variables. The preliminary regression analysis was performed on all possible combinations of the chosen explanatory variables. The specified adjusted R^2, coefficient p-value, variance inflation factor, and threshold condition of spatial autocorrelation p-value were evaluated, along with the minimum and maximum number of explanatory variables included in each model. Then, a spatial autocorrelation analysis tool was run for the residual error of that model, especially if the spatial autocorrelation p-value was also greater than the value specified by the minimum allowable p-value of spatial autocorrelation. In the next section, several models are presented for a comparative study based on the results.

3.3. Principal component analysis for data

In this analysis, the number of variables to be considered for analysis is large, and it is not appropriate to examine how the selection process is conducted based on only statistical rules and not theory. To check for robustness and compare the models, PCA is additionally adopted for the variables, and another model is constructed based on the analysis. PCA is a mathematical procedure that uses an orthogonal transformation. This transformation is defined in such a way that the first principal component (PC) has the highest variance possible, and each succeeding component, in turn, has the highest variance possible under the constraint that it is orthogonal or uncorrelated to the preceding components. PCA also calculates component scores and loadings. The component scores are the transformed variable values, and the component loadings are correlation coefficients of variable values and component scores.

PCA was applied to all the variables in this study to determine the PCs. The contribution ratios of the six components reached around 65%. Table 3 shows the corresponding eigenvectors, which are the PCs and have unit length; the column-wise sum of the squares of the loadings is 1. Mathematically, the PCs are the eigenvectors of the covariance matrix of the original dataset. Because the covariance matrix is symmetric, the eigenvectors are orthogonal. The PCs (eigenvectors) correspond to the direction in the original dimensional space with the greatest variance in the data.

An important issue in PCA is the interpretation of the components after the reduction of the observation data. For instance, the first component has positive loadings on Asian, black, and mixed-race people as well as Muslim variables and negative loadings for the UK nationality and Caucasian race. Thus, the first PC distinguishes sensitivity for immigrant-related characteristics versus UK nationality and is named as the immigrant component. Details and names of all PCs are as follows:

PC1: The first component showed negative effects of variables related to UK nationality and Caucasian race, and positive effects of the Asian, black, and mixed-race people as well as Muslim variables (immigrant component).

PC2: The second component showed not only positive effects of unemployment rate and green area variables but also negative effects of the urbanization field, such as environmentally hazardous substances and building occupation rate variables (de-urbanization component).

PC3: The third component showed a strong positive effect of the black and mixed-race people and Christian variables but negative effects of the Asian and Jewish people and total land area variables (racial polarization component).

PC4: The fourth component showed not only a positive effect of the variables of total land area, green area, water front area, and the environment variable but also negative effects of the building occupation rate variable (mixed-land use component).

PC5: The fifth component showed strong positive effects of the variables of time to city center and employment center (city access component).

PC6: The sixth component showed strong positive effects of the entrepreneurial and Jewish variables (entrepreneurial component).

4. Analysis models and results

4.1. Exploratory regression models and results

When there are many potential explanatory variables, the exploratory regression analysis can help to find a properly specified model. Exploratory regression is a data-mining tool that attempts all possible combinations of explanatory variables to determine which models are appropriate. Evaluating all possible combinations of the candidate explanatory variables helps to find the best model for the analysis, because it involves identifying models that meet all the requirements and assumptions. The candidate models must exceed the specified threshold of adjusted R^2, coefficient p-values, and coefficient variance inflation factor (VIF) values (the VIF measures redundancy among explanatory variables) for all explanatory variables when all the coefficients are statistically significant.

In this study, five models were selected from the results of the explanatory regression by running the statistical tool in ArcGIS. The results obtained through regression analysis, in which the dependent variable is housing price, from the five analysis models using these variables are shown in Table 4. The results indicate that only the unemployment rate has a negative coefficient; all other variables have a positive coefficient. In London, people with similar economic, occupational, and educational backgrounds have a strong tendency to live in the same area. Similarly, there is a strong tendency for immigrants and people belonging to the same race/ethnicity to live together. Here, the fact that the coefficients of variables related to the ratio of Caucasians and Jewish people are positive means that the housing prices in areas in which these people live tend to be higher.

Likewise, the fact that the coefficients of variables related to time to city center and number of entrepreneurs are positive means that areas with many business districts with commuters, dense urban population, and entrepreneurs tend to have high housing prices. On the other hand, the fact that the coefficient of the unemployment variable is negative indicates that the housing price is lower as the unemployment rate is higher.

Among the five models, the results in Table 4 indicate that Model 4

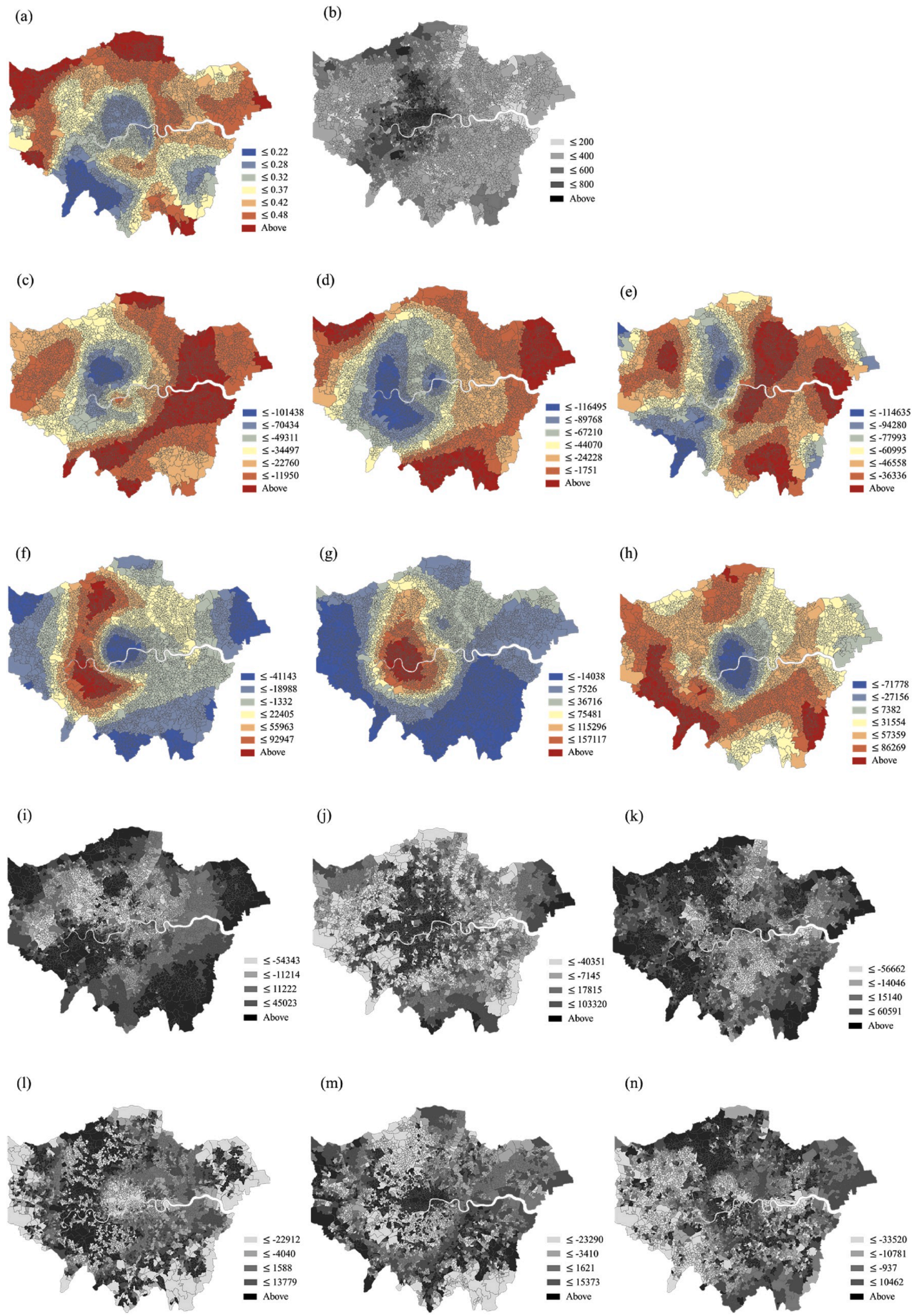

Fig. 5. Maps with results of the geographically weighted regression of the principal component (PC) model: (a) local R^2 values; (b) estimated values of the dependent variable; (c)-(h) coefficient values of the explanatory variables of PC1-PC6, respectively; (i)-(n) contributing values of each explanatory variable (PC1-PC6, respectively) to estimated values of the dependent variable.

has the lowest value of Akaike Information Criterion (AIC) (134,185) and the highest value of adjusted R^2 (0.219).

4.2. Regression results of the PC model

As previously mentioned, the five models in Table 4 were selected based on exploratory regression analysis. However, in this analysis, the number of variables to be considered for analysis is so large that it is also important to adopt an additional approach to check and compare the models. In the approach, a PCA is additionally conducted for the variables to build an additional model (PC model) that considers all the variables. The PC score results and analysis based on the PC model are shown in Table 5.

The immigrant, de-urbanization, and racial polarization PCs have negative impacts with statistical significance, while the mixed-land use and city access PCs have positive impacts with statistical significance. When comparing the PC model with the previous five models, there are consistencies wherein the immigrant and racial polarization factors, including non-UK nationality and non-Caucasian variables, and the de-urbanization factor, including unemployment rate and building occupancy rate, have negative impacts, and the mixed-land use and city access factors, including employment- and education-related variables, have positive impacts on housing price.

Although the entrepreneurial variables have a positive and statistically significant impact in the previous five models, in the PC model, the entrepreneurial factor does not have a statistically significant effect. This may be because, although the entrepreneurial factor can have an impact in some analytical models based on statistical rules or specific areas in a city, the existence or amount of its impact is still vague as various factors must be considered. The entrepreneurial impact on city evaluation may thus depend on other factors such as the city access component.

However, there is still room to improve the analysis of the models. Estimating a spatial lag model or GWR can be considered alternative analytical methodologies.

4.3. Spatial lag models and results

The key idea of spatial econometrics is expressed as the spatial lag. The motivation of spatial econometrics is the idea that regions are not independent but are interdependent. Therefore, spatial econometrics aims to measure impacts arising from a spatially dependent structure using spatially lagged variables. Computing spatially lagged variables requires a spatial weight matrix. Drukker et al. [44] and Kondo [45] provide the specific command of statistical software to construct spatial weight matrix and calculate spatially lagged variables. The spatial lag is defined as analogous to the time lag in time series analysis, and spatial econometrics incorporates spatial lag into cross-section analysis to consider spatial dependence between own region and neighboring regions [45]. The two-dimensional spatial information is mathematically expressed by the spatial weight matrix.

According to Kondo [45], the spatial weight matrix plays an important role in spatial analysis and the spatial weight matrix is generally row-standardized, which indicates that the sum of each row is equal to 1, in the context of spatial econometrics. Therefore, this study uses row-standardized weight matrix.

A simple extension of the previous analytical models is to examine spatial spillover effects on housing price across regions. Neighboring housing price might affect own housing price. The spatial structure on housing price is a crucial aspect, and the spatial lag provides an important insight. The results of the spatial lag models of housing price are presented in Table 6. Table 6 shows that the spatial lag of housing price has a significant positive impact in all spatial lag models, indicating that an increase in neighboring housing price also leads to an increase in own housing price. Furthermore, the values of Adj R^2 of all spatial lag models are larger and those of AIC of all spatial lag models are smaller than those of previous models.

Among the six models including PC lag model, the results in Table 6 indicate that PC lag has the lowest value of Akaike Information Criterion (AIC) (133,286) and the highest value of adjusted R^2 (0.354). Regarding changes in coefficient values between the previous models (Tables 4 and 5) and the spatial models (Table 6), large changes indicate that results were biased and inefficient in previous regressions (Tables 4 and 5). Hence, variables are not independent and should not be treated as such. That is why the spatial lag models were introduced in this section and the results provide important and further insights for this study.

4.4. GWR results of model 4

As mentioned earlier, GWR is a useful regression model to analyze relationships in which the influences of independent variables on the dependent variable do not remain constant for all locations. The GWR analysis runs a regression for each location and accounts for different responses in different parts of the focus area. Therefore, the GWR analysis is adopted for further analysis. The GWR analysis was conducted in the ArcGIS software using these explanatory variables with Model 4.

In the GWR analysis, a kernel, or bandwidth, that moves over the focus area and seeks to fit the best results for each subarea was used. The ArcGIS software provides an adjusted bandwidth that changes its size as it moves throughout the area under analysis. The bandwidth size defines the rate at which the influence of the coefficients decreases as the distance increases. In the ArcGIS software, the bandwidth can be selected and determined with the Golden search method that is minimize the value of the Akaike Information Criterion (AIC).

The overall adjusted R^2 of the GWR analysis in Model 4 was 0.52 (Table 7), which was larger than that of both of the normal exploratory regression (0.219 in Model 4) in Table 4 and the spatial lag model (0.331 in Lag model 4) in Table 6. Furthermore, the overall AIC of the GWR analysis (131,959 in Model 4) in Table 7 was smaller than that of the normal exploratory regression (134,185 in Model 4) in Table 4 and the spatial lag model (133,452 in Lag model 4) in Table 6.

Here, the bandwidth used for the estimation of each area was 17,481. Further, in the GWR analysis, the local R^2 was calculated by estimating the model for every local area (Fig. 4 and Table 7). Fig. 4 shows that the R^2 is higher in the southwestern areas than the eastern and central areas.

The explanatory variables in the data used in Model 4 are related to unemployment rate, time to city center, and entrepreneurs. Fig. 4 also shows the estimated coefficient value. The coefficient values of variables related to Caucasian and time to city center are similar. The characteristic feature of unemployment rate is that the coefficient value north (south) of the Thames River in the central areas is high (low). Regarding the number of entrepreneurs, it was found that the coefficient values in the central areas are low, and those in the outer areas are high.

The GWR model analysis runs a regression for each location, instead of a sole regression for the entire study area. Therefore, a quantitative contributing value of each variable for the estimated dependent variable in each location can be calculated using the estimated results of the coefficient parameters of each variable. The contributing values can be calculated by the estimated coefficient value multiplied by the observed value of each variable. The contributing values of explanatory variables are also shown in Fig. 4.

The southwest and northwest ends generally tend to have a high number of entrepreneurs, and the central areas tend to have a low number. However, some exceptions in central areas near the Tech City show a markedly high value.

4.5. GWR results for the PC model

Results of the PC model in Table 5 reveal that the adjusted R^2 value of the PC model, in which all variables are considered based on the PCA, is higher than that in the previous five models. Therefore, in this study, the

GWR analysis was conducted in the ArcGIS software using these explanatory variables with the PC model.

The overall adjusted R^2 and AIC of the GWR analysis in the PC model were 0.51 and 132,019 respectively (Table 8). Although the values of the adjusted R^2 and AIC are larger and smaller respectively than those of the normal regression analyses (Tables 4 and 5) and the PC lag model in Table 6, they are slightly worse results than those of the GWR analysis for Model 4 shown in Table 7.

The local R^2 was as shown in Fig. 5 and Table 8, and the bandwidth used for the estimation of each area was 8526. Fig. 5 shows that the local R^2 of outer areas is higher than that of central areas. Fig. 5 also shows the estimated coefficient values of PC1-6. The coefficient value of PC5, the city access component, is similar to the coefficient value of the Caucasian and time to city center variables in Model 4. Moreover, the coefficient value of PC6, the entrepreneurial component, is similar to the coefficient value of the entrepreneur variable in Model 4. These results show that consistency and robustness exist between the analyses of Model 4 and the PC model.

The contributing values of explanatory variables were also calculated, and the results are shown in Fig. 5. With regard to PC4, the mixed-land use component, the downtown or city center areas have low contributing values for housing price, but the value is high around the outside central areas, while PC5, the city access component, has the opposite tendency. PC6, the entrepreneurial component, tends to have a high contributing value for the northwest end and a low value for the central areas; however, some exceptions for the central areas near Tech City show a markedly high value. The result has a tendency similar to the GWR results for Model 4.

5. Conclusions

The following findings were obtained based on the analysis and results presented in this study. First, the social, economic, and environmental characteristics of the area tend to be related; that is, for example, people of a particular religious belief or race tend to live together in a specific urban or de-urbanized area. In this study, these aspects are categorized into several items, which have been narrowed down to six items based on the correlations among variables. Among them, this study focused on the employment situation in the area in terms of unemployed people and employment rate as well as the entrepreneurial environment of the area in terms of the number of entrepreneurs.

As a result, the second finding revealed a sharp contrast between the central and outer regions, and differences were seen in metropolitan areas such as London with regard to regional employment situations and entrepreneurial environments. This impacts the valuation of the area. This is closely related to the infrastructure environment, in which public transportation and buildings are mainly maintained and integrated at the center. This does not mean that employment in all industries and trades is caused by deliberate regional industrial/commercial policies. Although unemployment is not intense in downtown areas, many entrepreneurs exist in the central areas, and they tend to concentrate in the metropolitan areas where housing prices are high. Thus, a sharp contrast is observed in metropolitan areas such as London, and the so-called disruptions can be observed mainly in the central areas. This largely contributes to the evaluation of areas represented by housing prices.

In the metropolitan areas, the characteristics of each area are evident, as each element is closely related to the social, economic, and environmental background. Furthermore, these are not merely characteristics of the employment situation, entrepreneurial environment, and housing prices, and they can also be known as disruptions and disparities. Therefore, to resolve this problem and enhance the areas, it is necessary to enforce policies that consider whether the negative impact on areas affected by unemployment is large or small; whether the positive impact on areas influenced by employment situations in existing companies is large or small; and which type of areas will produce a better effect through support measures for entrepreneurship in each region. The analysis results in this study reveal the importance and effectiveness of providing entrepreneurial support while resolving social issues such as unemployment and empty houses in the surrounding areas by taking advantage of the existing corporate environment in the cities, as in the Tech City concept in London.

Furthermore, the study results prompt further discussion about support actions. For example, measures against unemployment must first be implemented in areas that have been negatively impacted by unemployment, and aggressive support measures must be implemented in areas where entrepreneurial support is effective, regardless of the area being in central London or the suburbs. The findings for London can be used to develop the concept of environmental employment policies as well as support measures for entrepreneurship throughout the world in the future.

Funding

This work was supported by a Grant-in-Aid for Early-Career Scientist (JSPS) (17K12982 (B) and 19K13808) and a Chuo University Grant for Special Research for 2018–2019.

Declaration of competing interest

None.

CRediT authorship contribution statement

Hiroki Nakamura: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization, Project administration, Funding acquisition.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.seps.2020.100820.

References

[1] Startup Genome LLC. Global startup ecosystem report 2017. San Francisco, CA: Startup Genome; 2017.
[2] Urbaniec K, Mikulčić H, Duić N, Lozano R. SDEWES 2014—sustainable development of energy, water and environment systems. J Clean Prod 2016;130:1–11.
[3] Urbaniec K, Mikulčić H, Rosen MA, Duić N. A holistic approach to sustainable development of energy, water and environment systems. J Clean Prod 2017;155:1–11.
[4] Mack E, Mayer H. The evolutionary dynamics of entrepreneurial ecosystems. Urban Stud 2015;53:2118–33.
[5] Nakamura H. Relationship among land price, entrepreneurship, the environment, economics, and social factors in the value assessment of Japanese cities. J Clean Prod 2019;217:144–52.
[6] Koronakos G, Smirlis Y, Sotiros D, Despotis DK. Assessment of OECD better life index by incorporating public opinion. Soc Econ Plann Sci 2019. https://doi.org/10.1016/j.seps.2019.03.005.
[7] Van Ryzin GG, Immerwahr S. Importance-performance analysis of citizen satisfaction surveys. Publ Adm 2007;85(1):215–26.
[8] James O. Evaluating the expectations disconfirmation and expectations anchoring approaches to citizen satisfaction with local public services. J Publ Adm Res Theor 2007;19:107–23.
[9] Berardi U. Clarifying the new interpretations of the concept of sustainable building. Sustain Cities Soc 2013;8(1):72–8.
[10] Häkkinen T. Assessment of indicators for sustainable urban construction. Civ Eng Environ Syst 2007;24(4):247–59.
[11] Sharifi A, Murayama A. A critical review of seven selected neighbourhood sustainability assessment tools. Environ Impact Assess Rev 2013;38:73–87.
[12] Scerri A, James P. Accounting for sustainability: combining qualitative and quantitative research in developing 'indicators' of sustainability. Int J Soc Res Methodol 2010;13(1):41–53.
[13] Hagerty MR. Unifying livability and comparison theory: cross-national time-series analysis of life-satisfaction. Soc Indicat Res 1999;47:343–56.
[14] Hagerty MR. Social comparisons of income in one's community: evidence from national surveys of incomes and happiness. J Pers Soc Psychol 2000;78:764–71.

[15] Kawakubo S, Ikaga T, Murakami S. Questionnaire-based validation of environmental performance assessment tool for municipalities. J Environ Eng 2013;78(693):883–92.
[16] Borseková K, Vaňová A, Vitálišová K. Smart specialization for smart spatial development: innovative strategies for building competitive advantages in tourism in Slovakia. Soc Econ Plann Sci 2017;58:39–50. https://doi.org/10.1016/j.seps.2016.10.004.
[17] Nakamura H, Managi S. Effects of subjective and objective city evaluation on life satisfaction in Japan. J Clean Prod 2020:120523. https://doi.org/10.1016/j.jclepro.2020.120523.
[18] Liao PS. Parallels between objective indicators and subjective perceptions of quality of life: a study of metropolitan and county areas in Taiwan. Soc Indicat Res 2009;91:99–114.
[19] Rosen S. Hedonic prices and implicit markets: product differentiation in pure competition. J Polit Econ 1974;82(1):34–55.
[20] Anselin L, Le Gallo J. Interpolation of air quality measures in hedonic house price models: spatial aspects. Spatial Econ Anal 2006;1(1):31–52.
[21] Anselin L, Lozano-Gracia N. Errors in variables and spatial effects in hedonic house price models of ambient air quality. Empir Econ 2008;34(1):5–34.
[22] Can A. The measurement of neighbourhood dynamics in urban house prices. Econ Geogr 1990;66:254–72.
[23] Can A. Specification and estimation of hedonic housing price models. Reg Sci Urban Econ 1992;22:453–74.
[24] Dubin RA. Estimation of regression coefficient in the presence of spatially autocorrelated error terms. Rev Econ Stat 1998;70:466–74.
[25] Dubin RA. Spatial autocorrelation and neighbourhood quality. Reg Sci Urban Econ 1992;22:432–52.
[26] Kim CW, Phipps TT, Anselin L. Measuring the benefits of air quality improvement: a spatial hedonic approach. J Environ Econ Manag 2003;45:24–39.
[27] Valente J, Wu SS, Gelfand A, Sirmans CF. Apartment rent prediction using spatial modelling. J R Estate Res 2005;27:105–36.
[28] Baltagi BH, Fingleton B, Pirotte A. Spatial lag models with nested random effects: an instrumental variable procedure with an application to English house prices. J Urban Econ 2014;80:76–86.
[29] Elhorst JP. Dynamic models in space and time. Geogr Anal 2001;33:119–40.
[30] Anselin L, Griffith DA. Do spatial effects really matter in regression analysis? Pap Reg Sci 1988;65(1):11–34.
[31] Zhang L, Xiong L, Cheng B, Yu C. How does foreign trade influence China's carbon productivity? Based on panel spatial lag model analysis. Struct Change Econ Dynam 2018;47:171–9.
[32] Fotheringham AS, Brunsdon C, Charlton ME. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. Environ Plann A 1998;30:1905–27.
[33] Kestens Y, Thériault M, Des Rosiers F. Heterogeneity in hedonic modelling of house prices: looking at buyers' household profiles. J Geogr Syst 2006;8(1):61–96.
[34] Zhou Q, Wang C, Fang S. Application of geographically weighted regression (GWR) in the analysis of the cause of haze pollution in China. Atmos Pollut Res 2019;10:835–46.
[35] Fotheringham AS, Crespo R, Yao J. Geographical and temporal weighted regression (GTWR). Geogr Anal 2015;47(4):431–52.
[36] Liang X, Liu Y, Qiu T, Jing Y, Fang F. The effects of locational factors on the housing prices of residential communities: the case of Ningbo, China. Habitat Int 2018;81:1–11.
[37] Dziauddin MF, Powe NA, Alvanides S. Estimating the effects of light rail transit (LRT) system on residential property values using geographically weighted regression (GWR). Appl Spat Anal Pol 2015;8(1):1–25.
[38] Lelo K, Monni S, Tomassi F. Socio-spatial inequalities and urban transformation. The case of Rome districts. Soc Econ Plann Sci 2019;68:100696. https://doi.org/10.1016/j.seps.2019.03.002.
[39] Zhai L, Li S, Zou B, Sang H, Fang X, Xu S. An improved geographically weighted regression model for PM2.5 concentration estimation in large areas. Atmos Environ 2018;181:145–54.
[40] Qu M, Wang Y, Huang B, Zhao Y. Source apportionment of soil heavy metals using robust absolute principal component scores-robust geographically weighted regression (RAPCS-RGWR) receptor model. Sci Total Environ 2018;626:203–10.
[41] Sabokbar HF, Roodposhti MS, Tazik E. Landslide susceptibility mapping using geographically-weighted principal component analysis. Geomorphology 2014;226:15–24.
[42] Pontarollo N, Serpieri C. A composite policy tool to measure territorial resilience capacity. Soc Econ Plann Sci 2018;100669. https://doi.org/10.1016/j.seps.2018.11.006.
[43] Foord J. The new boomtown? Creative city to Tech City in east London. Cities 2013;33:51–60.
[44] Drukker DM, Peng H, Prucha IR, Raciborski R. Creating and managing spatial weighting matrices with the spmat command. STATA J 2013;13(2):242–86.
[45] Kondo K. Introduction to spatial econometric analysis: creating spatially lagged variables in Stata. RIETI technical paper series 16-T-001 2016.

Hiroki Nakamura, (Ph.D.), is an associate professor at Faculty of Commerce, Chuo University. He has B.Eng., M.Eng., and Ph.D. in Eng. from the Tokyo Institute of Technology. His research work has been about energy, environment, risk management, and urban management issues. He worked at the Energy and Environment Department, Japan Productivity Center, the 'Multidisciplinary Education and Research Center for Energy Science, Tokyo Institute of Technology Global Center of Excellence (GCOE) Program' as a research assistant, Research Institute for East Asia Environments, Kyushu University as an assistant professor, The University of Kitakyushu as a lecturer, and Urban Institute and Departments of Urban and Environmental Engineering, Kyushu University as an assistant professor.

He has published in peer reviewed journals and books such as *Journal of Transport Geography, Economic Analysis and Policy, Safety Science, Tourism Management,* "Chapter 20 Foundation for the Establishment of Urban and Environmental Plans in Asia and the Required Practical Education," in Takayuki Shimaoka et al. eds. (2016) *Basic Studies in Environmental Knowledge, Technology, Evaluation, and Strategy -Introduction to Asia Environmental Studies*, Springer, and "Climate transport policy in the Asia-Pacific region," in G. Van Calster and W. Vandenberghe eds. (2015) *Research Handbook on Climate Change Mitigation Law*, Edward Elgar Publishing.