A mixed Bayesian spatial logit model for networked multivariate binary choice data

Yuki Oyama[1] ・ Daisuke Murakami[2] ・ Rico Krueger[3]

[1]Member of JSCE, Shibaira Institute of Technology
E-mail: oyama@shibaura-it.ac.jp
[2]Member of JSCE, The Institute of Statistical Mathematics
E-mail: dmuraka@ism.ac.jp
[3]Non-member, Technical University of Denmark
E-mail: rickr@dtu.dk

This study deals with spatially correlated multivariate binary choice data, which is a set of individuals' perceptions of multiple spatial units. This type of data includes an individual's spatial perception, visit experience, or choice set formation. To analyze such data, we propose a mixed spatial autoregressive logit model, which lies in an interface between spatial statistics and discrete choice modeling, accounting for inter-individual heterogeneity and spatial dependence among utilities. To estimate the model, we rely on Bayesian estimation aided by Pólya-Gamma data augmentation that addresses the non-conjugacy of the logit model and allows for fast posterior inference. The proposed methods are applied to the data of individuals perceiving streets within a specific area, and we analyze the formative factors of spatial city image.

Key Words: *Spatial logit; Mixed logit; Bayesian inference; Pólya-Gamma data augmentation; Spatial perception*

1. Introduction

This study aims to analyze multivariate binary responses regarding individuals' spatial perception, typically represented by a two-dimensional vector $\boldsymbol{y} = [y_{ni}]$ where $n \in \{1, \ldots, N\}$ is an individual and $i \in \{1, \ldots, S\}$ is a spatial unit. This type of data includes the responses representing spatial perception (if an individual perceives a spatial unit), visit experience (if she has been to a place during a period), or choice set formation (if she considers a destination as an alternative to travel). Such responses are likely to be different among individuals. Also, the responses of the same individual may often reveal a spatial pattern; for example, an individual has been to places within a specific region. Therefore, the analysis of such spatial data requires the consideration of (1) inter-individual heterogeneity and (2) spatial dependence.

To deal with such multivariate spatial choice data, this study proposes a mixed spatial autoregressive logit model, which combines a mixed logit (MXL) model in the discrete choice modeling field (e.g., Train, 2009[1)]) and a spatial autoregressive (SAR) logit model in the spatial statistics (e.g., LeSage, 2000[2)]). Although this model simultaneously accounts for inter-individual heterogeneity and spatial dependence among utilities, it remains a difficulty in the model estimation. In particular, the capture of spatial dependence in the logit framework is difficult due to the inter-dependence of the utility functions. The incorporation of random parameters requires more computational effort.

Among estimation methods for SAR choice models, such as maximum likelihood[3)], generalized method of moments (GMM)[4)], and linearized GMM[5)], Bayesian estimation[2)] has shown outperformance in terms of accuracy and stability to small sample sizes (e.g., Calabresse and Elkink, 2014[6)]). However, Bayesian Markov chain Monte Carlo (MCMC) estimation is also known to be computationally expensive especially for the spatial logit model due to the non-conjugacy of the likelihood function. Recently, Polson et al. (2013)[7)] presented Pólya-Gamma data augmentation (PG-DA) technique to address the non-conjugacy in MCMC estimation of binary logit models. PG-DA augments the Gibbs sampler by introducing an additional Pólya-Gamma distributed latent variable, which ensures conjugate updates and thus does not need an inefficient Metropolis algorithm within MCMC estimation. Krueger et al. (2020)[8)] and Krisztin and Piribauer (2021)[9)] successfully applied this technique to spatial data analysis.

In this study, we also rely on Bayesian estimation

aided by PG-DA. Focusing on its compatibility with the Bayesian MXL model[10], we present its extension for the mixed spatial logit model. We show that the conjugate posterior distribution for individual-specific (random) parameters can be derived, as well as the fixed parameters. The efficiency and stability of this hierarchical Bayesian MCMC with PG-DA will be demonstrated with numerical experiments.

Furthermore, by applying the proposed model, we empirically analyze the spatial perception of individuals. The data indicates if an individual perceives a street within a specific area (Oyama, 2023[11]). As well as the main effects of street attributes, we capture the observed and unobserved heterogeneity of individuals. Moreover, given that the streets compose a network, we consider the spatial dependence using network-based spatial weight matrices. This application reveals the formative factors of the subjective spatial city image of residents and visitors and can help policies related to the image building of the area.

The contributions of the paper are summarized as follows:

- We analyze spatial data with individuals' subjective perceptions. This is an exploration of the interface between spatial statistics and discrete choice modeling.
- We propose a mixed Bayesian spatial logit model to analyze such multivariate spatially-correlated binary choice data, by utilizing the PG-DA strategy.
- We apply the proposed method and empirically analyze individuals' spatial perceptions. The analysis reveals the formative factors of the subjective spatial city image.

The remainder of the paper is structured as follows. Section 2. provides a literature review. Section 3. presents a mixed spatial autoregressive logit model (SAR-MXL). Section 4. describes the Bayesian estimation strategy. Section 5. shows a numerical experiment, and Section 6. presents the empirical application to spatial perception analysis. Finally, Section 7. concludes the study.

2. Literature review

(1) Spatial choice model

Spatial binary choice models have been actively developed in spatial econoemtrics [12]. McMillen (1992) [3] is one of the pioneering studies proposing a probit model considering spatial dependence in residuals and response variables. Pinkse and Slade (1998)[4] developed a one-step generalized method of moments (GMM) estimation for spatial error probit and logit models. The GMM has been linearized for fast computation [5] while extended to the two-step GMM to improve estimation accuracy[12]. The GMM-based approaches have drawbacks including the unbounded parameter space for spatial autocorrelation coefficients and the need for instrumental variables[9].

Bayesian estimation has also been used to estimate spatial logit models [2]. Although its high accuracy has been validated[6], the MCMC can be inefficient due to the non-conjugacy and the need for the Metropolis step. To address this problem, the PG-DA strategy[7] has been introduced for binary and multinomial spatial logit regressions models [9),13], and confirmed better accuracy and computational efficiency.

However, aforementioned spatial econometric studies have ignored inter-individual heterogeneity despite its importance in modeling location choice as well as individual perceptions. Bhat et al. (2004)[14] and Bhat (2011) [15] developed maximum likelihood estimation considering the heterogeneity in spatial logit regression. Yet, to the best of the author's knowledge, the PG-DA has never been extended for spatial mixed logit modeling.

(2) Mixed choice models

Mixed choice models (e.g., Train, 2009[1]) extend standard choice models (such as binary or multinomial logit or probit models) hierarchically by allowing model parameters to vary randomly across individuals. This makes mixed choice models highly flexible. Mixed choice models can accommodate unobserved heterogeneity, flexible substitution patterns and correlation in unobserved factors over time (Train, 2009[1]). Due to this flexibility, mixed choice models are widely used in practical work.

Mixed choice models are formulated by convolving a choice kernel (such as logit or probit) with a mixing distribution, which controls the flexibility of the model. This mixing distribution can be parametric (like normal or log-normal), semi-parametric (like a mixture of normal) or nonparametric (Vij & Krueger, 2017[16]). In practice, the normal distribution is the most widely adopted mixing distribution. Parameter estimation in discrete choice models can be performed using either frequentist maximum simulated likelihood methods or Bayesian MCMC methods (Train, 2009[1]).

(3) Spatial perception analysis

How people perceive and form the spatial image of a city has been a major topic of interest in various research fields[17),18)]. To this end, the seminal work of Linch (1960)[19)] presented a framework for analyzing the relationship between the spatial cognition of the city and the external properties of the environment. Although Lynch's theory has led and inspired research in spatial cognition[20)], behavioral geography[21)], and urban studies[22)], the extent to which urban elements affect the spatial cognition of individuals is still unclear.

Coulton et al. (2001)[23)] and Smith (2010)[24)] analyzed the formation of walking neighborhood areas perceived by residents, on the basis of cognitive maps drawn by them. Guimpert et al. (2018)[25)] also collected cognitive maps in Santiago, Chile. They converted the maps to mesh-based perception data, and modeled the mesh perception probability of a resident with mixed binary choice modeling. Oyama (2023)[11)] presented a street-based neighborhood perception model to analyze the effect of streetscape features extracted from street images. These quantitative approaches successfully measured the effects of the spatial attributes of the spatial units and demographic characteristics of respondents on the perception of neighborhoods. However, their analyses treated the spatial units independently and ignored the spatial dependence among the utilities. Because the perception units are small enough to be closely related to each other, the consideration of spatial dependence can significantly improve the prediction performance of the model.

3. Model formulation

(1) Mixed spatial autoregressive logit model

Let $y_{ni} \in \{0, 1\}$ be a binary response of an individual $n \in \{1, \ldots, N\}$ for a spatial unit $i \in \{1, \ldots, S\}$. We aim to model the probability of y_{ni} being one based on the binary logit framework:

$$P(y_{ni} = 1) = \frac{\exp(u_{ni})}{1 + \exp(u_{ni})} \tag{1}$$

where u_{ni} is the utility function of space i perceived by individual n.

To reasonably analyze such subjectively perceived spatial data, we introduce the assumptions of spatial dependence and inter-individual heterogeneity. This results in the formulation of a **mixed spatial autoregressive logit** (SAR-MXL) model. The utility u_{ni} is then defined as

$$u_{ni} = \boldsymbol{\alpha}^\top \boldsymbol{x}_{ni,F} + \boldsymbol{\beta}_n^\top \boldsymbol{x}_{ni,R} + \rho \sum_{j \neq i} w_{ij} u_{nj} + \epsilon_{ni}. \tag{2}$$

This can be written more compactly in the matrix form:

$$\boldsymbol{u}_n = \mathbf{A}^{-1}(\mathbf{X}_n^F \boldsymbol{\alpha} + \mathbf{X}_n^R \boldsymbol{\beta}_n + \boldsymbol{\epsilon}_n), \tag{3}$$

where

$$\mathbf{A} = \mathbf{I} - \rho \mathbf{W}, \tag{4}$$

and $\mathbf{I}$ is an $S \times S$ identity matrix.

The covariates $\boldsymbol{x}_{ni}$ include the observed attributes of individual n and spatial unit i, which are decomposed into $\boldsymbol{x}_{ni,F}$ that pertain to the fixed parameters $\boldsymbol{\alpha}$ and $\boldsymbol{x}_{ni,R}$ with the random parameters $\boldsymbol{\beta}_n$. The latter captures the inter-personal heterogeneity in the perception.

A nonnegative and row-stochastic spatial weight matrix $\mathbf{W} = [w_{ij}]$ contains information on the $S \times S$ linkages between the spatial units in the sample, and ρ is a scalar parameter measuring the strength of spatial autocorrelation. In this SAR model specification, the utility u_{ni} of individual n for spatial unit i depends on the characteristics of other spatial units $j \neq i$, and the spatial multiplier $\mathbf{A}^{-1} = (\mathbf{I} - \rho \mathbf{W})^{-1} = \sum_{r=0}^{\infty} \rho^r \mathbf{W}^r$ captures the spatial dependence.

This model also implicitly considers two random components: the spatially independent errors associated with the logit model formulation (1), and a normally distributed random shock ϵ_{ni} that captures the spatial dependence in the errors.

The model parameters to be estimated are $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho)$. The likelihood function of the model is

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \rho) = \prod_{n \in N} \prod_{i \in S} \frac{\exp(u_{ni})^{y_{ni}}}{1 + \exp(u_{ni})} \tag{5}$$

(2) Baysian specification

Following the standard MXL model formulation, we assume the random (individual-specific) parameter $\boldsymbol{\beta}_n$ to be distributed according to a multivariate normal, i.e., $\boldsymbol{\beta}_n \sim \mathcal{N}(\boldsymbol{\zeta}, \boldsymbol{\Sigma})$, where $\boldsymbol{\zeta}$ is a K_R-dimensional mean vector and $\boldsymbol{\Sigma}$ is a $K_R \times K_R$ covariance matrix. The remaining parameters, $\boldsymbol{\alpha}, \boldsymbol{\zeta}, \boldsymbol{\Sigma}$ and ρ are treated in a fully-Bayesian manner. The fixed parameters $\boldsymbol{\alpha}$ and the mean vector $\boldsymbol{\zeta}$ are assumed to follow multivariate normal distributions. As for the covariance matrix $\boldsymbol{\Sigma}$, we employ Huang's half-t prior (Huang and Wand, 2013[26)]), which Akinc and Vandebroek (2018)[27)] showed outperforms the inverse Wishart prior that is often employed in fully Bayesian MMNL models[1)] in terms of parameter recovery. We also employ a beta distribution for the prior of spatial parameter ρ.

The generative process of the fully Bayesian SAR-MXL model is summarized as follows:

1. $\rho \sim \text{Beta}(a_\rho, b_\rho)$
2. $\boldsymbol{\alpha} \sim \mathcal{N}(\boldsymbol{\lambda}_0, \Xi_0)$
3. $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$
4. $a_k \sim \text{Gamma}(\frac{1}{2}, \frac{1}{A_k^2}), k = 1, \ldots, K_R$
5. $\boldsymbol{\Sigma} \sim \text{IW}(\nu + K_R - 1, 2\nu\text{diag}(\boldsymbol{a}))$
6. $\boldsymbol{\beta}_n \sim \mathcal{N}(\boldsymbol{\zeta}, \boldsymbol{\Sigma})$
7. $y_{ni} \sim \text{Logit}(\boldsymbol{\alpha}, \boldsymbol{\beta}_n, \rho, \mathbf{X}_n)$

Note that (4) and (5) induce Huang's half-t prior. We have known hyperparameters $\{\boldsymbol{\lambda}_0, \Xi_0, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \nu, A_{1:K}\}$ and a collection of model parameters $\theta = (\boldsymbol{\alpha}, \boldsymbol{\zeta}, \boldsymbol{\Sigma}, \boldsymbol{a}, \boldsymbol{\beta}_{1:N}, \rho)$ whose posterior distribution to be estimated.

4. Estimation strategy

(1) Pólya-Gamma data augmentation

The main challenge of the standard Gibbs sampler for posterior inference with the logit model is the unavailability of a conjugate prior to the kernel, which requires a computationally inefficient Metropolis method. To address this challenge, we use a Pólya-Gamma data augmentation (PG-DA) estimation strategy proposed by Polson et al. (2013)[7]. PG-DA is the state-of-the-art technique to handle non-conjugacy in MCMC estimation of binary logit models by introducing a PG-distributed latent variable.

Following Polson et al. (2013)[7], the Laplace transform of the PG density is given by

$$\frac{\exp(u_{ni})^a}{(1+\exp(u_{ni}))^b} = 2^{-b}\exp(\kappa u_{ni})\int_0^\infty \exp\left(-\frac{\omega_{ni}u_{ni}^2}{2}\right)p(\omega_{ni})d\omega_{ni} \tag{6}$$

where $\kappa = a - b/2$, and ω_{ni} is a PG distributed random variable with scale b and location parameter zero, i.e., $p(\omega_{ni}) \sim \mathcal{PG}(b, 0)$.

Note that this integral identity does not rely on numerical approximation and can be estimated by sampling from the conditional posterior PG distribution. Using the result in Polson et al. (2013)[7], the conditional posterior for $\boldsymbol{\omega}$ also takes the form of a PG distribution:

$$p(\boldsymbol{\omega}_n|\boldsymbol{\alpha}, \boldsymbol{\beta}_n, \rho, \boldsymbol{y}_n) = \mathcal{PG}\left(1, \mathbf{A}^{-1}(\mathbf{X}_n^F\boldsymbol{\alpha} + \mathbf{X}_n^R\boldsymbol{\beta}_n + \boldsymbol{\epsilon}_n)\right) \tag{7}$$

(2) Conditional distributions

We notice that when $a = y_{ni}$ and $b = 1$, Eq.(6) coincides with the likelihood of the logit model (1):

$$\frac{\exp(u_{ni})^{y_{ni}}}{1+\exp(u_{ni})} = \frac{1}{2}\exp(\kappa_{ni}u_{ni})\int_0^\infty \exp\left(-\frac{\omega_{ni}u_{ni}^2}{2}\right)p(\omega_{ni})d\omega_{ni}. \tag{8}$$

Using this result, the conditional posterior density of $\boldsymbol{\alpha}$ is given by

$$\begin{aligned} &p(\boldsymbol{\alpha}|\boldsymbol{y}_{1:N}, \boldsymbol{\beta}_{1:N}, \rho, \boldsymbol{\omega}_{1:N}) \\ &\propto p(\boldsymbol{\alpha}|\boldsymbol{\lambda}_0, \boldsymbol{\Xi}_0)\prod_{n=1}^N\prod_{i=1}^I p(\boldsymbol{y}_{ni}|\boldsymbol{\beta}_n, \boldsymbol{\alpha}, \rho, \boldsymbol{\omega}_n) \\ &\propto \exp\left(-\frac{1}{2}(\boldsymbol{\alpha}-\boldsymbol{\lambda}_0)'\boldsymbol{\Xi}_0(\boldsymbol{\alpha}-\boldsymbol{\lambda}_0)\right)\cdot \\ &\quad \prod_{n=1}^N \exp\left(\boldsymbol{\kappa}_n\mathbf{A}^{-1}\mathbf{X}_n^F\boldsymbol{\alpha} - \frac{\boldsymbol{\omega}_n\boldsymbol{u}_n^2}{2}\right), \end{aligned} \tag{9}$$

and similarly, the conditional posterior density of $\boldsymbol{\beta}_n$ is given by

$$\begin{aligned} &p(\boldsymbol{\beta}_n|\boldsymbol{y}_n, \boldsymbol{\alpha}, \rho, \boldsymbol{\omega}_n) \\ &\propto p(\boldsymbol{\beta}_n|\boldsymbol{\zeta}, \boldsymbol{\Sigma})\prod_{i=1}^I p(y_{ni}|\boldsymbol{\beta}_n, \boldsymbol{\alpha}, \rho, \boldsymbol{\omega}_n) \qquad (10) \\ &\propto \exp\left(-\frac{1}{2}(\boldsymbol{\beta}_n-\boldsymbol{\zeta})'\boldsymbol{\Sigma}(\boldsymbol{\beta}_n-\boldsymbol{\zeta})\right)\cdot \\ &\quad \exp\left(\kappa_n\mathbf{A}^{-1}\mathbf{X}_n^R\boldsymbol{\beta}_n - \frac{\boldsymbol{\omega}_n\boldsymbol{u}_n^2}{2}\right). \qquad (11) \end{aligned}$$

As a result, the conditional posterior distributions of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_n$ reduce to the following Gaussian distributions:

$$\boldsymbol{\alpha}|\boldsymbol{y}_{1:N}, \boldsymbol{\beta}_{1:N}, \rho, \boldsymbol{\omega}_{1:N} \sim \mathcal{N}(\boldsymbol{\mu}_\alpha, \boldsymbol{\Sigma}_\alpha) \tag{12}$$

$$\boldsymbol{\beta}_n|\boldsymbol{y}_{1:N}, \boldsymbol{\alpha}, \rho, \boldsymbol{\omega}_n \sim \mathcal{N}(\boldsymbol{\mu}_{\beta_n}, \boldsymbol{\Sigma}_{\beta_n}) \tag{13}$$

where

$$\boldsymbol{\mu}_\alpha = \boldsymbol{\Sigma}_\alpha(\Xi^{-1}\boldsymbol{\lambda}_0 + \sum_{n=1}^N(\mathbf{A}^{-1}\mathbf{X}_n^F)^\top\boldsymbol{\Omega}_n(z_n - \mathbf{A}^{-1}\mathbf{X}_n^R\boldsymbol{\beta}_n - \mathbf{A}^{-1}\epsilon_n)) \tag{14}$$

$$\boldsymbol{\Sigma}_\alpha = (\boldsymbol{\Sigma}^{-1} + \sum_{n=1}^N(\mathbf{A}^{-1}\mathbf{X}_n^F)^\top\boldsymbol{\Omega}_n(\mathbf{A}^{-1}\mathbf{X}_n^F))^{-1} \tag{15}$$

$$\boldsymbol{\mu}_{\beta_n} = \boldsymbol{\Sigma}_{\beta_n}(\boldsymbol{\Sigma}^{-1}\boldsymbol{\zeta} + (\mathbf{A}^{-1}\mathbf{X}_n^R)^\top\boldsymbol{\Omega}_n(z_n - \mathbf{A}^{-1}\mathbf{X}_n^F\alpha - \mathbf{A}^{-1}\epsilon_n)) \tag{16}$$

$$\boldsymbol{\Sigma}_{\beta_n} = (\boldsymbol{\Sigma}^{-1} + (\mathbf{A}^{-1}\mathbf{X}_n^R)^\top\boldsymbol{\Omega}_n(\mathbf{A}^{-1}\mathbf{X}_n^R))^{-1} \tag{17}$$

$$\boldsymbol{\Omega}_n = \text{diag}(\omega_{n1}, \ldots, \omega_{nI}) \tag{18}$$

$$z_{ni} = \kappa_{ni}/\omega_{ni} \tag{19}$$

Note that in the above derivation, our study extendedly combines the mixed logit case (Bansal et al., 2019[10]) and

the spatial logit case (Krisztin and Piribauer, 2021[9]).

(3) Augmented Gibbs sampler

The joint distribution of the model factorizes as

$$p(\boldsymbol{y}_{1:N}, \boldsymbol{\theta}) = p(\rho)p(\boldsymbol{\alpha}|\boldsymbol{\lambda}_0, \Xi_0)p(\boldsymbol{\zeta}|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)\cdot p(\boldsymbol{\Sigma}|a, \nu)\prod_{k=1}^{K_R} p(a_k|A_k)\prod_{n\in N} p(\boldsymbol{\beta}_n|\boldsymbol{\zeta}, \boldsymbol{\Sigma}) \tag{20}$$

Then, the augmented Gibbs sampler algorithm for the SAR-MXL model can be summarized as follows:

1. Sample ρ using a griddy Gibbs sampler[28]
2. Sample $\boldsymbol{\alpha}$ using Eq.(12)
3. Sample a_k from Gamma$(\frac{\nu+K_R}{2}, \frac{1}{A_k^2} + \nu(\boldsymbol{\Sigma}^{-1})_{kk})$
4. Sample $\boldsymbol{\Sigma}$ from IW$(\nu + K_R - 1, 2\nu\text{diag}(a) + \sum_{n=1}^{N}(\boldsymbol{\beta}_n - \boldsymbol{\zeta})(\boldsymbol{\beta}_n - \boldsymbol{\zeta})^\top)$
5. Sample $\boldsymbol{\zeta}$ from $\mathcal{N}(\frac{1}{N}\sum_{n=1}^{N}\boldsymbol{\beta}_n, \frac{\boldsymbol{\Omega}}{N})$
6. for $n = 1 : N$:
   (a) Sample $\boldsymbol{\omega}_n$ using Eq.(7)
   (b) Sample $\boldsymbol{\beta}_n$ using Eq.(13)

Note that we used a beta distribution for the prior of ρ, following the literature[29]. However, because the conditional posterior for ρ does not have a conjugate distribution, a griddy Gibbs sampler[28] is used to sample from the conditional posterior, based on

$$p(\rho|\boldsymbol{y}_{1:N}, \boldsymbol{\alpha}, \boldsymbol{\beta}_{1:N}) = \frac{p(\rho)p(\boldsymbol{y}_{1:N}|\rho, \boldsymbol{\alpha}, \boldsymbol{\beta}_{1:N})}{\int_{-1}^{1} p(\rho)p(\boldsymbol{y}_{1:N}|\rho, \boldsymbol{\alpha}, \boldsymbol{\beta}_{1:N})d\rho} \approx \frac{p(\rho)p(\boldsymbol{y}_{1:N}|\rho, \boldsymbol{\alpha}, \boldsymbol{\beta}_{1:N})}{\sum_{r\in G} p(r)p(\boldsymbol{y}_{1:N}|r, \boldsymbol{\alpha}, \boldsymbol{\beta}_{1:N})} \tag{21}$$

5. Numerical experiment

This section presents a numerical experiment to demonstrate the performance of the proposed method. The experiment is based on simulated observations and mainly aims to assess the accuracy of the estimates under different strengths of spatial dependence.

Note that the efficiency of PG-DA compared to the traditional MCMC has been already proven in the literaturec[7], in terms of computational time and sampling efficiency. Also, because our data involves two dimensions and is not typical for spatial statistics, meaning that no other estimation methods or existing package is available. We therefore focus on the accuracy in this experiment. Also, we implemented the model by writing our own Python code.

(1) Simulation setup

The observations y_{ni} for the experiment are randomly sampled from the binomial distribution with the probability of the SAR-MXL model:

$$y_{ni} \sim \mathcal{B}\left(1, \frac{\exp(u_{ni})}{1+\exp(u_{ni})}\right) \tag{22}$$

where $\boldsymbol{u}_n$ is defined by (3).

In this experiment, we consider 50 individuals ($N = 50$) and 200 spatial units ($S = 200$). We considered two parameters with fixed effects ($K_F = 2$) and two random parameters ($K_R = 2$), and their true values were set as follows:

- $\boldsymbol{\alpha}^* = (1.0, -1.0)$
- $\boldsymbol{\zeta}^* = (1.0, -1.0)$
- $\boldsymbol{\Sigma}^* = \text{diag}(0.5, 0.5)$

The explanatory variables $\mathbf{X}$ were sampled from the standard normal distribution and then normalized. The spatial weight matrix $\mathbf{W}$ was constructed using 4-nearest neighbors based on a spatial location pattern randomly generated from a uniform distribution $\mathcal{U}(0, 1)$.

We tested different strengths ρ of spatial dependence and examined how the estimation results were affected. During the MCMC estimation, we drew 10,000 times and the initial 5,000 draws were discarded as burn-in. As for the Griddy Gibbs sampler for the conditional posterior for ρ, we used 100 grid points.

(2) Results

Tables 1 and 2 report the results. Overall, the true parameter values were well reproduced, including the random parameters, regardless of the strength of spatial dependence in the generated observations. Even for the extreme cases ($\rho^* = 0.9$ and $\rho^* = -0.9$), the autocorrelation parameter ρ was estimated with small biases. Given that the poor performance of GMM methods with high levels of autocorrelation is known[6], this result clearly shows the stability of our Bayesian method with respect to spatial dependence.

It is worth noting that the estimates were slightly biased when ρ^* is positive and large, compared to when ρ^* is small or negative. A possible explanation for this is that, with a high and positive autocorrelation, the observations can exhibit similar spatial patterns and the valid sample size can be reduced.

Table–1 Means of posteriors: Case 2 with $\rho^* \in [0.0, 0.9]$.

		$\rho*$									
	TRUE	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\hat{\rho}$		-0.045	0.052	0.155	0.253	0.362	0.458	0.566	0.664	0.763	0.881
$\hat{\alpha}_1$	1	1.054	1.004	0.984	1.010	0.996	0.990	1.047	1.121	1.146	1.160
$\hat{\alpha}_2$	-1	-1.040	-1.026	-0.972	-0.995	-1.105	-1.015	-1.108	-1.117	-1.141	-1.149
$\hat{\beta}_1^{\text{mean}}$	1	0.979	1.139	1.101	0.994	1.157	1.221	1.260	1.067	1.114	1.164
$\hat{\beta}_1^{\text{std}}$	0.5	0.524	0.624	0.564	0.575	0.657	0.537	0.614	0.600	0.729	0.596
$\hat{\beta}_2^{\text{mean}}$	-1	-1.032	-1.133	-0.955	-1.036	-1.169	-1.013	-1.011	-1.079	-1.291	-1.234
$\hat{\beta}_2^{\text{std}}$	0.5	0.675	0.562	0.572	0.526	0.475	0.492	0.596	0.536	0.621	0.598

Table–2 Means of posteriors: Case 2 with $\rho^* \in [-0.9, -0.1]$.

		$\rho*$								
	TRUE	-0.9	-0.8	-0.7	-0.6	-0.5	-0.4	-0.3	-0.2	-0.1
$\hat{\rho}$		-0.914	-0.790	-0.696	-0.600	-0.488	-0.414	-0.311	-0.195	-0.087
$\hat{\alpha}_1$	1	0.996	1.016	1.057	1.040	0.987	1.034	1.047	1.031	1.012
$\hat{\alpha}_2$	-1	-1.014	-1.019	-1.014	-1.048	-1.061	-1.028	-1.033	-0.971	-1.032
$\hat{\beta}_1^{\text{mean}}$	1	0.993	0.946	1.176	1.109	1.061	1.020	0.952	1.002	1.016
$\hat{\beta}_1^{\text{std}}$	0.5	0.580	0.569	0.629	0.580	0.622	0.578	0.535	0.554	0.555
$\hat{\beta}_2^{\text{mean}}$	-1	-1.000	-1.065	-0.974	-1.008	-0.926	-1.044	-1.116	-0.964	-1.039
$\hat{\beta}_2^{\text{std}}$	0.5	0.502	0.552	0.573	0.637	0.545	0.604	0.681	0.553	0.664

6. Empirical results

(1) Data and variables

The data used in this study is the subjective cognitive areas of two neighborhoods in Tokyo, Kiyosumi-Shirakawa and Kiba. The data was collected through an on-street survey, where respondents drew a cognitive map as a polygon on the base map, relying on their experiences or prior information on the neighborhoods. We converted the cognitive maps into street-based perception data $y_{ni} \in \{0, 1\}$, indicating whether the respondent n perceived the street i within the neighborhood area ($y_{ni} = 1$) or not ($y_{ni} = 0$). We collected samples from 25 residents and 25 visitors for each neighborhood, resulting in 100 respondents in total. The target networks of Kiyosumi-Shirakawa and Kiba contain 258 and 473 streets, respectively (Figure 1). Thus, the numbers of observations account for 12,900 and 23,650, giving a total of 36,550.

As for the spatial attributes of the streets, we mainly focus on (1) the distance from the reference point, (2) boundaries, and (3) streetscape features. We show a reference point for each neighborhood as the star-marked point in Figure 1, which was defined as the closest location to the metro station named after the neighborhood (Kiyosumi-Shirakawa Station of the Hanzomon Line, and Kiba Station of the Tozai Line) among the points along the main park (Kiyosumi Garden & Park for Kiyosumi-Shirakawa and Kiba Park for Kiba). The distance is defined as the network shortest path distance. As for boundaries, we used three dummy variables indicating if a street is inside the boundaries of the river, avenue, and administrative district (Figure 1). The landscape features were extracted as pixel values from the street photos we took.

In addition, we consider observed heterogeneity captured by the personal characteristics of the respondents, including age, gender, and familiarity with the neighborhood. Familiarity is measured by the years of residence for residents and by the visit frequency for visitors.

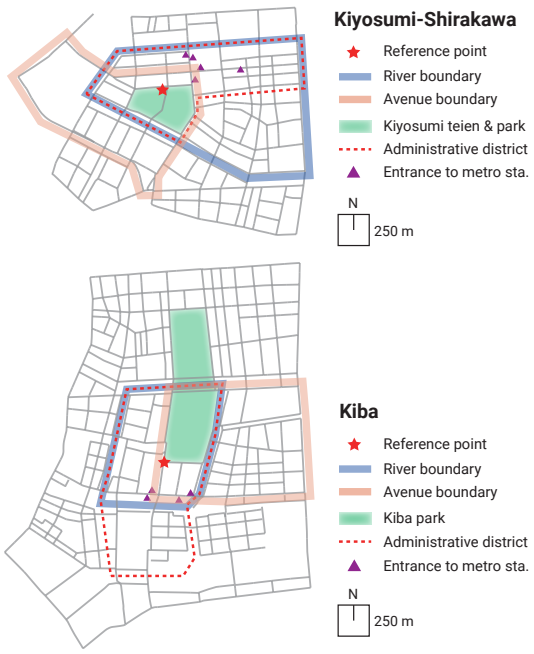
Note that the reader who is interested in the details of the data is referred to Oyama (2023)[11] who provides a detailed description of data and variables with statistics.

(2) Network-based weight matrix

Since the spatial units of our data are streets that compose a network, we defined the spatial weight matrix $\mathbf{W}$ based on the network and the shortest path distance on it. The following weight matrices were tested:

1. Adjacency: $w_{ij} = 1$ only when i is adjacent to j
2. K-nearest: $w_{ij} = 1$ if i is one of k-th nearest neighbors to j
3. Within-D: $w_{ij} = 1$ if i is within distance D from j
4. Angle: $w_{ij} = a_{ij}/180$ where $a_{ij} \in [0, 180)$ is the angle between streets i and j and $a_{ij} = 0$ if i is not adjacent to j

In addition to adjacency and distance-based neighbor matrices, we considered the angle-based matrix based on the hypothesis that two neighboring streets with a small



Figure–1 Networks of Kiyosumi-Shirakawa and Kiba.

angle difference can be recognized as being the same continual street and are positively correlated with each other.

Note that although matrices based on distance, such as power or exponential, are often used in spatial statistics, they were not considered in this study because they performed very poorly for the networks in the preliminary tests. All the tested matrices were normalized so that the sum of each row was equal to one.

(3) Goodness-of-fit metrics

Given the observations y_{ni} and the retained sample from conditional posterior $\boldsymbol{\theta} = (\boldsymbol{\theta}^1, \ldots, \boldsymbol{\theta}^R)$ where R is the number of the retained realizations, the following three metrics were used to measure goodness-of-fit of the models:

1. LPPD (log pointwise predictive density)[30]:

$$\begin{aligned} \mathrm{LPPD} &= \sum_n \sum_i \log \int P(y_{ni}|\boldsymbol{\theta}_n) p(\boldsymbol{\theta}_n) d\boldsymbol{\theta}_n \\ &\approx \sum_n \sum_i \log \left(\frac{1}{R} \sum_{r=1}^{R} P(y_{ni}|\boldsymbol{\theta}_n^r) \right) \end{aligned} \quad (23)$$

LPPD corresponds to the logarithm of the pointwise likelihood integrated over the posterior distribution of the relevant model parameters and can be approximated with the posterior draws.

Table–3 Model comparison

	Kiyosumi-Shirakawa			
	Logit	SAR	MXL	SAR-MXL
LPPD	-4789.096	-4748.782	-2591.718	-2285.240
RSME	0.341	0.339	0.243	0.227
FPR	0.838	0.840	0.919	0.933

	Kiba			
	Logit	SAR	MXL	SAR-MXL
LPPD	-10293.721	-10188.046	-7437.416	-6611.511
RSME	0.374	0.372	0.316	0.296
FPR	0.802	0.804	0.855	0.876

2. RSME (root mean square error):

$$\mathrm{RSME} = \frac{1}{NS} \sum_n \sum_s \sqrt{(\hat{y}_{ni} - y_{ni})^2} \quad (24)$$

where $\hat{y}_{ni} = E(y_{ni}|\boldsymbol{\theta}_n) \approx \sum_{r=1}^{R} P(y_{ni}|\boldsymbol{\theta}_n^r)/R$. RMSE measures the discrepancy between the predicted and observed street perceptions, $\hat{y}_{ni}$ and y_{ni}.

3. FPR (first preference recovery):

$$\mathrm{FPR} = 1 - \frac{1}{NS} \sum_n \sum_s |I(\hat{y}_{ni} > 0.5) - y_{ni}| \quad (25)$$

where $I(\hat{y}_{ni} > 0.5)$ takes one if $\hat{y}_{ni}$ is greater than 0.5 and zero otherwise. FPR is the percentage of the model predicting the accurate perception on the basis of maximum a posteriori probability (MAP).

(4) Results

We performed the MCMC estimation for each neighborhood with 10,000 iterations, where the first 5,000 iterations were burn-ins, and 100 grid points for the griddy Gibbs sampler.

Model fit We first compared the following four models (where SP stands for spatial dependence and IH for inter-individual heterogeneity):

1. **Logit**: without SP and IH,
2. **SAR**: with SP but without IH,
3. **MXL**: without SP but with IH,
4. **SAR-MXL**: with both SP and IH,

where we used a simple adjacency matrix for $\mathbf{W}$ of SAR and SAR-MXL models. Table 3 reports the result. Because smaller LPPD and RSME and greater FPR indicate better performance, the capture of both spatial dependence and inter-individual heterogeneity improved the model fit. Hereafter we focus on the SAR-MXL model.

Then, we compared the different network-based spatial weight matrices $\mathbf{W}$, as defined in Section 6.(2). Table 4 reports the comparison result. While the adjacency matrix

Table–4 Comparison of different network weight matrices.

	Kiyosumi-Shirakawa			
	LPPD	RMSE	FPR	$\hat{\rho}$
Adjacency	**-2281.270**	**0.227**	**0.934**	0.745
Nearest ($k = 5$)	-2504.171	0.238	0.925	0.422
Nearest ($k = 10$)	-2553.014	0.241	0.922	0.250
Within ($\leq$ 200m)	-2498.425	0.237	0.926	0.394
Within ($\leq$ 400m)	-2523.141	0.239	0.924	0.265
Angle	-2384.135	0.232	0.930	0.660

	Kiba			
	LPPD	RMSE	FPR	$\hat{\rho}$
Adjacency	-6620.576	0.296	0.876	0.756
Nearest ($k = 5$)	-6648.999	0.296	0.877	0.783
Nearest ($k = 10$)	**-6396.574**	**0.289**	**0.882**	0.813
Within ($\leq$ 200m)	-6886.085	0.301	0.873	0.667
Within ($\leq$ 400m)	-7043.031	0.305	0.870	0.619
Angle	-6705.054	0.298	0.874	0.735

performed best for Kiyosumi-Shirakawa, the 10-nearest neighbors based on the shortest path distance performed best for Kiba. This implies that the perception utilities of streets in Kiba more widely depend on each other than those in Kiyosumi-Shirakawa. Yet, the simple adjacency matrix overall performed well. The angle-based adjacency matrix did not outperform the simple one.

Table 4 also reports the estimated posterior mean of spatial parameter $\hat{\rho}$ for each case. Regardless of the definition of $\mathbf{W}$, it was estimated as a positive value, indicating the positive spatial correlation between the utilities of streets spatially close to each other.

Estimation results Table 5 reports the mean, standard deviation, and credible interval of the posterior of each estimated parameter for the SAR-MXL model, and Table 6 reports the effect sizes of the policy variables. The effect sizes, elasticities for continuous variables and marginal effects for dummy variables were calculated using the retained sample from conditional posterior. More specifically, we computed the mean, standard deviation, and credible interval of the effect size of k-th variable for each pair (n, i) and then reported the means of the statistics over the records of $N \times S$. Direct effect sizes evaluate the effects of a change in an attribute of a street on the probability change of the street, while indirect ones evaluate the sum of its effect on the probability changes of the other streets.

The main findings from the estimation results are fourfold. Firstly, the spatial parameter was estimated as a large positive value for both neighborhoods. This indicates a high spatial dependence between the utilities of streets that are spatially close to each other. Since we used a spatial autoregressive (or spatial lag) specification, the utility of a street depends not only on its attributes but also on the attributes of the neighboring streets. The estimated indirect effect sizes also indicate nontrivial spillover effects of attribute changes of a street on its neighboring streets.

Secondly, the marginal effects together with the positive signs of the coefficients for the inner-boundaries variables indicate that the boundaries played as edges of the cognitive area of the individuals. Therefore, mitigating boundaries can be an effective way to expand the cognitive area of a neighborhood. The results suggest that the river boundary is the strongest edge of area cognition in both neighborhoods. It has a particularly large effect in Kiyosumi-Shirakawa, where on average a 7.0% change in perception probability is suggested between inside and outside the river boundary, while the change is comparably trivial as 1.7% in Kiba. It also has a non-trivial indirect marginal effect on surrounding streets, 5.0% change in Kiyosumi-Shirakawa and 1.3% in Kiba, respectively. The avenue boundary has less impact on perception probability than the other boundaries in both neighborhoods.

Thirdly, as expected, the distance is negatively associated with the perception probability. A 1% increase in distance from the reference point decreases on average the perception probability of a street by 0.49% in Kiyosumi-Shirakawa and 0.36% in Kiba, respectively. We also captured observed heterogeneity by the interaction terms of distance with personal characteristics and the direction in which a street is located from the reference point. In Kiba, women and people 40 years or older are in particular more sensitive to distance, meaning that their cognitive areas tend to be smaller than those of men and young people. Residents who have lived in the neighborhoods for less than 10 years also have larger cognitive areas than those who have lived longer, possibly reflecting the ambiguity of their cognitive areas.

The sensitivity to distance also highly depends on the direction. A positive value of the coefficient for a direction interaction term mitigates the resistance to distance. In Kiyosumi-Shirakawa, the streets in the south are the least sensitive to distance, followed by the east, north, and west. In Kiba, the streets in the north are the least sensitive to distance, followed by the south, east, and west. This result captures the effect of the extent of Kiba Park; indeed, most respondents drew their cognitive maps so that the entire area of the park is included (Figure 1).

Finally, we evaluated the extent to which the visual landscape features of the streets affect their perception proba-

Table–5 Estimation result of the SAR-MXL model.

	Kiyosumi-Shirakawa				Kiba			
	mean	std. dev.	2.50%	97.50%	mean	std. dev.	2.50%	97.50%
Spatial parameter								
ρ	0.739	0.032	0.663	0.782	0.760	0.009	0.743	0.782
Fixed parameter								
Intercept	-0.303	0.169	-0.639	-0.011	0.057	0.096	-0.144	0.242
Boundaries								
Inner river	1.225	0.134	0.994	1.508	0.195	0.031	0.134	0.256
Inner avenue	0.089	0.057	-0.022	0.201	0.016	0.025	-0.034	0.065
Inner admin. district	0.223	0.045	0.133	0.308	0.107	0.033	0.044	0.173
Distance × demographic chars.								
Live less than 10 yrs	0.315	0.141	0.064	0.584	0.223	0.113	0.006	0.467
Visit once a month or less	-0.092	0.173	-0.456	0.230	0.158	0.117	-0.032	0.401
Female	0.079	0.153	-0.268	0.341	-0.048	0.092	-0.227	0.126
40 yrs or older	-0.065	0.155	-0.334	0.256	-0.241	0.098	-0.469	-0.087
Distance × direction								
East	1.234	0.113	1.029	1.468	0.499	0.040	0.422	0.579
North	0.221	0.161	-0.092	0.536	1.295	0.046	1.209	1.394
South	1.690	0.112	1.477	1.916	0.765	0.050	0.669	0.864
Fixed parameter								
Mean								
Distance	-4.017	0.239	-4.516	-3.612	-2.252	0.083	-2.416	-2.092
Green pixel	0.026	0.005	0.017	0.036	0.026	0.002	0.022	0.030
Building pixel	0.000	0.003	-0.005	0.006	0.003	0.001	0.000	0.005
Road pixel	0.014	0.004	0.007	0.022	0.005	0.002	0.002	0.009
Sky pixel	0.006	0.005	-0.004	0.016	0.016	0.003	0.011	0.021
Std. dev.								
Distance	1.491	0.194	1.166	1.907	1.111	0.054	1.010	1.220
Green pixel	0.030	0.006	0.017	0.040	0.025	0.002	0.021	0.029
Building pixel	0.030	0.003	0.025	0.036	0.021	0.001	0.019	0.023
Road pixel	0.041	0.003	0.035	0.047	0.027	0.002	0.023	0.031
Sky pixel	0.043	0.005	0.034	0.053	0.041	0.003	0.035	0.047

bility in each neighborhood. In both neighborhoods, the pixel values of Green, Building, Road, and Sky are positively associated with street perception probability, but the credible intervals of their effect sizes included negative values, indicating their random policy effects among individuals. Only the Green pixel value in Kiba clearly had a (small but) positive effect on the street perception probability. This result indicates that street greenery forms the spatial image of the Kiba neighborhood and increasing it can reinforce such an image. However, the spillover effect of increasing greenery has a very limited effect.

It should be noted that the effect sizes were the means across all observations with the size of $N \times S$, which might have made the discussion on effect sizes unclear. We will therefore focus on several representative streets in terms of policy or location and analyze the effect sizes more specifically.

7. Concluding remarks

This study proposed a Bayesian SAR-MXL model to analyze multivariate binary choice data. Based on the seminal work of Polson et al. (2013)[7], we extend it to simultaneously capture inter-individual heterogeneity and spatial dependence. The high performance of the framework was demonstrated with both the numerical experiment and empirical application.

In the application, we successfully captured the observed and unobserved heterogeneity and spatial dependence on to what extent the spatial attributes affect the perception probability of streets. As such, the Bayesian SAR-MXL model is a flexible framework to analyze subjective spatial perception and can be applied to other contexts. One may use the SAR-MXL model as a choice set formation model for a destination choice model. Since the choice set definition is crucial in destination choice modeling, the integration of our model may improve its predictability.

Note that our framework is for binary responses, which may limit its applicability. Although the extension to mixed MNL specification is not straightforward (Bansal et al., 2019[10]), in the spatial statistics context, Krisztin et al. (2022)[13] successfully applied the PG-DA strategy to a spatial MNL model and efficiently analyzed urban expansion. Therefore, it would be appealing to seek the possibility of extending our framework to an MNL case.

Table–6 Effect size: Direct effect of a variable of the street of interest, and its indirect effect on the neighboring streets.

	Kiyosumi-Shirakawa				Kiba			
	mean	std. dev.	2.50%	97.50%	mean	std. dev.	2.50%	97.50%
Marginal effects								
Inner river								
Direct	0.070	0.021	0.035	0.118	0.017	0.004	0.010	0.027
Indirect	0.052	0.015	0.027	0.083	0.013	0.003	0.008	0.020
Inner avenue								
Direct	0.005	0.004	-0.001	0.013	0.001	0.002	-0.003	0.006
Indirect	0.004	0.003	-0.001	0.010	0.001	0.002	-0.002	0.005
Inner admin. district								
Direct	0.013	0.005	0.005	0.023	0.010	0.004	0.003	0.017
Indirect	0.010	0.003	0.004	0.017	0.007	0.003	0.003	0.013
Elasticities								
Distance								
Direct	-0.490	0.102	-0.714	-0.317	-0.358	0.064	-0.497	-0.246
Indirect	-0.162	0.029	-0.224	-0.110	-0.117	0.019	-0.157	-0.084
Green pixel								
Direct	0.061	0.049	-0.027	0.165	0.069	0.033	0.009	0.139
Indirect	0.005	0.005	-0.003	0.015	0.004	0.002	0.0004	0.009
Building pixel								
Direct	0.111	0.141	-0.153	0.398	0.087	0.055	-0.017	0.196
Indirect	0.002	0.002	-0.003	0.006	0.0004	0.0003	-0.0002	0.001
Road pixel								
Direct	0.159	0.186	-0.197	0.532	0.054	0.085	-0.110	0.225
Indirect	0.001	0.001	-0.001	0.004	0.000	0.0005	-0.001	0.001
Sky pixel								
Direct	0.034	0.087	-0.130	0.211	0.047	0.039	-0.025	0.128
Indirect	0.0002	0.002	-0.003	0.004	0.0009	0.001	-0.0003	0.002

References

1) Train, K. E.: *Discrete choice methods with simulation*, Cambridge university press, 2009.
2) LeSage, J. P.: Bayesian estimation of limited dependent variable spatial autoregressive models, *Geographical Analysis*, Vol.32, No.1, pp.19–35, 2000.
3) McMillen, D. P.: Probit with spatial autocorrelation, *Journal of Regional Science*, Vol.32, No.3, pp.335–348, 1992.
4) Pinkse, J. and Slade, M. E.: Contracting in space: An application of spatial statistics to discrete-choice models, *Journal of Econometrics*, Vol.85, No.1, pp.125–154, 1998.
5) Klier, T. and McMillen, D. P.: Clustering of auto supplier plants in the united states: generalized method of moments spatial logit for large samples, *Journal of Business & Economic Statistics*, Vol.26, No.4, pp.460–471, 2008.
6) Calabrese, R. and Elkink, J. A.: Estimators of binary spatial autoregressive models: A monte carlo study, *Journal of Regional Science*, Vol.54, No.4, pp.664–687, 2014.
7) Polson, N. G., Scott, J. G., and Windle, J.: Bayesian inference for logistic models using pólya–gamma latent variables, *Journal of the American statistical Association*, Vol.108, No.504, pp.1339–1349, 2013.
8) Krueger, R., Bansal, P., and Buddhavarapu, P.: A new spatial count data model with bayesian additive regression trees for accident hot spot identification, *Accident Analysis & Prevention*, Vol.144, pp.105623, 2020.
9) Krisztin, T. and Piribauer, P.: A bayesian spatial autoregressive logit model with an empirical application to european regional fdi flows, *Empirical Economics*, Vol.61, No.1, pp.231–257, 2021.
10) Bansal, P., Krueger, R., Bierlaire, M., Daziano, R. A., and Rashidi, T. H.: Pólygamma data augmentation to address non-conjugacy in the bayesian estimation of mixed multinomial logit models, *arXiv preprint arXiv:1904.07688*, 2019.
11) Oyama, Y.: Spatial city image and its formative factors: A street-based area cognition analysis, *SSRN preprint: https://ssrn.com/abstract=4374014*, 2023.
12) Piras, G. and Sarrias, M.: One or two-step? evaluating gmm efficiency for spatial binary probit models, *Journal of choice modelling*, Vol.48, pp.100432, 2023.
13) Krisztin, T., Piribauer, P., and Wögerer, M.: A spatial multinomial logit model for analysing urban expansion, *Spatial Economic Analysis*, Vol.17, No.2, pp.223–244, 2022.
14) Bhat, C. R. and Guo, J.: A mixed spatially correlated logit model: formulation and application to residential choice modeling, *Transportation Research Part B: Methodological*, Vol.38, No.2, pp.147–168, 2004.
15) Bhat, C. R.: The maximum approximate composite marginal likelihood (macml) estimation of multinomial probit-based unordered response choice models, *Transportation Research Part B: Methodological*, Vol.45, No.7, pp.923–939, 2011.
16) Vij, A. and Krueger, R.: Random taste heterogeneity in discrete choice models: Flexible nonparametric finite mixture distributions, *Transportation Research Part B: Methodological*, Vol.106, pp.76–101, 2017.
17) Siegel, A. W. and White, S. H.: The development of spatial representations of large-scale environments, *Advances in child development and behavior*, Vol.10, pp.9–55, 1975.
18) Sadalla, E. K., Burroughs, W. J., and Staplin, L. J.: Reference points in spatial cognition., *Journal of experimental psychology: human learning and memory*, Vol.6, No.5, pp.516, 1980.
19) Lynch, K.: *The image of the city*, MIT press, 1960.
20) Golledge, R. G., Jacobson, R. D., Kitchin, R., and Blades, M.: Cognitive maps, spatial abilities, and human wayfinding, *Geographical review of Japan, Series B.*, Vol.73, No.2, pp.93–104, 2000.
21) McCunn, L. J. and Gifford, R.: Spatial navigation and place imageability in sense of place, *Cities*, Vol.74, pp.208–218,

2018.
22) Filomena, G., Verstegen, J. A., and Manley, E.: A computational approach to 'the image of the city', *Cities*, Vol.89, pp.14–25, 2019.
23) Coulton, C. J., Korbin, J., Chan, T., and Su, M.: Mapping residents' perceptions of neighborhood boundaries: a methodological note, *American journal of community psychology*, Vol.29, No.2, pp.371–383, 2001.
24) Smith, G., Gidlow, C., Davey, R., and Foster, C.: What is my walking neighbourhood? a pilot study of english adults' definitions of their local walking neighbourhoods, *International Journal of behavioral nutrition and physical activity*, Vol.7, No.1, pp.1–8, 2010.
25) Guimpert, I. and Hurtubia, R.: Measuring, understanding and modelling the walking neighborhood as a function of built environment and socioeconomic variables, *Journal of transport geography*, Vol.71, pp.32–44, 2018.
26) Huang, A. and Wand, M. P.: Simple marginally noninformative prior distributions for covariance matrices, 2013.
27) Akinc, D. and Vandebroek, M.: Bayesian estimation of mixed logit models: Selecting an appropriate prior for the covariance matrix, *Journal of choice modelling*, Vol.29, pp.133–151, 2018.
28) Ritter, C. and Tanner, M. A.: Facilitating the gibbs sampler: the gibbs stopper and the griddy-gibbs sampler, *Journal of the American Statistical Association*, Vol.87, No.419, pp.861–868, 1992.
29) LeSage, J. P. and Pace, R. K.: Spatial econometric models, *Handbook of applied spatial analysis: Software tools, methods and applications*, pp. 355–376, Springer, 2009.
30) Gelman, A., Hwang, J., and Vehtari, A.: Understanding predictive information criteria for bayesian models, *Statistics and computing*, Vol.24, No.6, pp.997–1016, 2014.

(2023. 10. 13 Received)