Spatial Ordered Probit Model with Heteroscedastic Skewed Errors and Exogenous Spillovers

Hyunjun Hwang
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712, USA
Email: hjhwang@utexas.edu

Chandra R. Bhat (corresponding author)
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712, USA
Tel: 512-471-4535; Email: bhat@mail.utexas.edu

ABSTRACT
Understanding individual travel decisions requires accounting for interdependencies between individuals and locations in close spatial or social proximity, as ignoring these relationships can produce biased results and misleading inferences. This study introduces a novel model (labeled LSLX-LSAE for Local Spatial Lag of X-Local Spatial Asymmetric Error) that incorporates localized spatial processes in both exogenous variable spillovers and error correlations while accommodating heteroscedasticity and flexible error distributions. By constraining the range of spatial effects to better reflect travel choice contexts and employing a maximum approximate composite marginal likelihood (MACML) inference approach, the model resolves identification and computational challenges. Simulation experiments demonstrated satisfactory parameter recovery of the LSLX-LSAE estimator. An empirical application to walking frequency data from the 2023 Puget Sound Regional Travel Study reveals that residential spatial unit characteristics differ in the spatial extent of their influence on walking frequency, with some operating only within the immediate census tract and others exhibiting spillover effects from neighboring areas. Model comparison against four alternative specifications confirmed the statistical superiority of the proposed model in both disaggregate and aggregate data fit measures. Average treatment effect analysis further demonstrated that the assumption of global versus local spatial processes has a considerable influence on estimated effect magnitudes, in some cases altering not only the size but also the sign of treatment effects. These findings suggest that policies targeting the built environment to promote walking should be spatially differentiated, reflecting the distinct spatial scales at which different residential spatial unit characteristics influence pedestrian behavior.

Keywords: Spatial econometrics, discrete choice modeling, composite marginal likelihood, travel behavior, active travel

1. INTRODUCTION

Understanding and predicting the patterns and determinants of individual and collective travel decisions lies at the core of travel behavior research. Such insights are critical for informing transportation planning, urban development policies, and sustainability strategies. Traditional modeling approaches, such as multinomial choice models, ordered-response models, count models, and multiple discrete-continuous models, have been widely used to analyze the factors shaping individual travel behavior. While these models provide valuable insights, they typically rely on the assumption that individual utilities or propensities are independent across decision-makers. This assumption ignores spatial interdependencies, wherein an individual's choices are influenced by the decisions, attributes, or unobserved circumstances of others in close spatial or social proximity (see Giubergia et al., 2025 for an extended discussion on this issue). Failing to account for spatial interdependencies can lead to inconsistent parameter estimates, misleading inferences, and ultimately, suboptimal policy recommendations. Spatial econometric approaches explicitly address these concerns by incorporating spatial structure into behavioral models, thereby offering a more realistic representation of interrelated decision-making processes.

The spatial econometric literature, systematically established by Anselin (1988) and advanced along many different directions in subsequent studies (see Anselin, 2021), identifies three primary mechanisms of spatial interdependencies. First, endogenous spatial interaction occurs when an individual's choices directly influence the choices of nearby individuals. For instance, in a travel behavior context, which is the topical emphasis of the current paper, the intensity of walking of an individual may increase the intensity of walking of another individual in close proximity, through direct observation-based effects. Second, exogenous spatial interaction arises when an individual's personal or locational/social characteristics (that determine the individual's choices) also influence another individual's choices (rather than an individual's choices directly influencing another individual's choices). For example, consider the case of an individual A whose walking behavior is positively influenced by the good land-use mix and pedestrian facilities around individual A's residence. Individual B lives in an adjacent spatial area, but that person's neighborhood lacks the kinds of land-use mixing and pedestrian facilities around individual A's residence. But as individual B ventures beyond their immediate surroundings to individual A's spatial sphere, individual B discovers the built environment there as a conducive place for walking and so may start walking more there or may also start walking more in their own neighborhood because of a “feel good” effect of walking that then leads to a tempering of the negative effect of the relatively poor pedestrian built environment characteristics of the individual's own neighborhood. Third, spatial error interaction emerges when individuals in proximity share similar unobserved factors that influence travel behavior. Residents of the same neighborhood may experience common infrastructure conditions (such as shaded and lighted walkways), local driving norms (such as better yielding behavior to pedestrians), or policy environments (such as better enforcement of traffic laws) that may not be observed by the analyst, and that create correlated error disturbances in walking behavior across individuals in close proximity, even after controlling for observed individual and area characteristics.

Significant progress has been made in extending conventional discrete choice models into increasingly advanced spatial econometric frameworks that account for the interdependencies underlying travel behavior and human mobility. Among the many spatial econometric specifications, the General Nesting Spatial (GNS) model represents a comprehensive framework, accommodating all the three types of spatial dependence discussed above. It serves as a conceptual umbrella from which a wide range of more parsimonious models can be derived by imposing

appropriate parameter restrictions, as shown in the table below (see also Halleck Vega and Elhorst, 2013). For example, the Spatial Durbin Model (SDM), in structural form, incorporates both endogenous and exogenous interaction effects, assuming spatially uncorrelated errors. In contrast, the Spatial Durbin Error Model (SDEM), in structural form, addresses exogenous spatial interactions and spatial error correlation, but does not consider endogenous spatial effects. The Spatial Lag of X (SLX) model, on the other hand, focuses solely on exogenous spatial interaction effects, assuming independence in outcomes and errors.

The GNS model, as enticing as it may seem as an umbrella model, is, at best, weakly identified when applied to practical applications (see Manski, 1993; Halleck Vega and Elhorst, 2013; Bhat, 2014). Worse, in spatial econometrics, since the nature of spatial connections is assumed a priori, incorporating various spatial interactions makes it difficult to distinguish which factors are driving the spatial dependency observed in the data, leading to what Gibbons and Overman (2012) criticized as "Mostly pointless spatial econometrics." Hence, model specification should be guided by the specific spatial effect components anticipated in the empirical application, as no single modeling framework can adequately serve all spatial econometric contexts.

Beyond the inherent identification problems, there are conceptual challenges in incorporating endogenous spatial interaction effects into discrete outcomes for travel behavior, an issue that has received relatively little attention in the spatial econometric literature. This is because, in discrete outcome models, it is not possible to have individual A's actual observed discrete choice affect another individual B's actual observed discrete choice, while also having individual B's observed discrete choice affect individual A's observed discrete choice, as would be typically needed in a spatial dependence context. Such a specification would result in a logical inconsistency problem wherein the sum of the different possible combinations of discrete probabilities across the two individuals would not sum to one (see Maddala, 1983; Bhat et al., 2016).[1] As a result, when considering endogenous spatial interaction, all spatial econometric models for discrete choice develop the endogenous interaction effect based on the underlying latent propensities that determine the observed discrete outcomes. The net result is that discrete choice models with endogenous spatial interaction effects (GNS, SAC, SDM, SAR) assume, in structural form, that individuals at one spatial location base their decisions on the latent propensities of other individuals located elsewhere. But, by definition, the latent propensities of other individuals are not fully observed (not just by the analyst, but also other decision makers). Of course, one may yet defend the use of the discrete choice models listed above with endogenous spatial interaction on the basis that their reduced forms boil down to exogenous spatial interaction effects and spatial error interactions. But such a reduced form imposes the condition that the spatial spillover effects are global in nature (spatial spillover refers to the situation when a change in an exogenous variable for one individual in a specific location affects the outcome of another individual in another location; a global spillover implies that a change in an exogenous variable of any individual affects every other individual everywhere else). Such a global spillover effect is difficult to maintain in most travel behavior contexts (where local spillover effects are more likely to exist), since there is little reason to believe that an intervention or change in an exogenous variable at one location will

[1] As discussed by Anselin (2021), technically speaking, in cross-sectional settings and even for linear regression models with spatial dependence, there is a need for caution in the interpretation of endogenous interaction effects as the effects of the outcome of one individual on another individual's outcome. Effectively, what is observed here is the simultaneous combination of observed outcomes across individuals. When the observed outcome is a continuous variable, such a simultaneous system is estimable and so one can use an endogenous spatial interaction specification. However, when the observed outcome is discrete (or, more generally, limited-dependent outcomes), such a simultaneous system breaks down because of the logical inconsistency problem.

universally affect the decisions of agents at all other locations, regardless of spatial distance. Besides, as may be deduced from the reduced form, the extent of spatial dependence in the SAC and SAR specifications is reduced to a single autocorrelation coefficient, and the ratio of the spatial spillover to direct effects is a priori assumed to be the same across all exogenous variables (Elhorst, 2010; Halleck Vega and Elhorst, 2015).

As with the endogenous spatial interactions just discussed, the models in Table 1 that accommodate spatial error interactions (GNS, SAC, SDEM, SEM) also specify a global error correlation effect, where unobserved factors are correlated over the entire domain of space being considered. Again, in reality, such correlations in travel behavior investigations are likely to be local in nature, with little reason for distant observations (spatial units) to be correlated due to unobserved factors. Further, an issue that rarely has been raised in the context of these spatial error specifications is that unobserved correlation effects across observations and error heteroscedasticity across space (that is, the variance in outcome observations located at different points in space) are both assumed to originate from the single source of spatial dependence (as specified through a proximity-based weight matrix and an autocorrelation structure). In reality, heteroscedasticity may be a function of unique individual characteristics specific to pre-defined "neighborhoods" or to other groupings of observational units/individual related factors that is completely independent of spatial correlation effects. For example, individuals with high formal education attainment may be uniformly more likely to recognize the health benefits of walking, resulting in smaller interpersonal variations compared to other educational groups. Both the global error correlation problem as well as the strong tie between spatial correlation and heteroscedasticity can at once be addressed by directly specifying a correlation structure on the error terms as well as allowing for separate heteroscedasticity across observations based on exogenous variables (see Bhat and Sener, 2009 for an example of introducing such observation-specific heteroscedasticity into spatial models).

Table 1 Structural Components of Spatial Interactions

Model Type	Endogenous Spatial Interactions	Exogenous Spatial Interactions	Spatial Error Interactions
General Nesting Spatial (GNS) (Anselin, 1988)	×	×	×
Spatial Autoregressive Combined (SAC) (Kelejian and Prucha, 1998)	×		×
Spatial Durbin Model (SDM) (Anselin, 1988)	×	×	
Spatial Durbin Error Model (SDEM) (LeSage and Pace, 2009)		×	×
Spatial Autoregressive (SAR) (Cliff and Ord, 1973)	×		
Spatial Lag of X (SLX) (LeSage and Pace, 2009; Elhorst, 2010; Halleck Vega and Elhorst, 2015)		×	
Spatial Error Model (SEM) (Cliff and Ord, 1973)			×

Different from all the models discussed thus far and identified in Table 1, the SLX model offers an alternative approach without global spillover effects, and admits more flexibility in the exogenous spatial interaction effects by allowing the extent of spatial spillover (relative to the direct effects) to vary across exogenous variables. However, it does not consider correlation effects due to unobserved factors. As such, the SLX model also does not consider error heteroscedasticity across observations. However, in most outcomes, we expect some level of spatial correlation due to unobserved factors too. As importantly, across observations, we expect heteroscedasticity because of unobserved spatial unit characteristics. For a continuous outcome, usual estimation methods (such as ordinary least squares) will still remain as an unbiased and consistent estimator (though will be an inefficient estimator) even under heteroscedasticity; but, for discrete choice outcomes (and more broadly for limited-dependent outcomes), ignoring heteroscedasticity in the errors will, in general, result in an inconsistent maximum likelihood estimator (in addition to the estimator being inefficient) (see Greene, 2000 in the context of general econometrics, and Elhorst, 2024 for a discussion in the context of spatial econometrics). For these reasons, in this paper, we consider an SLX setup for the exogenous spatial interaction effects and supplement that with a direct error correlation/heteroscedasticity structure, which also avoids latent-propensity and global dependence inherent in most other spatial models for discrete choice analysis.

Another key aspect in spatial econometrics concerns the distributional assumption underlying the error terms. Spatial models typically presuppose that individual error terms follow a symmetric distribution (such as a normal or logistic distribution), though such assumptions are rarely substantiated by empirical evidence. At the same time, it is widely recognized that mis-specification of the error distribution will, in general, lead to inconsistent estimates of the choice probabilities as well as the effects of exogenous variables (Caffo et al., 2007; Bhat et al., 2017). Such mis-specifications of error distribution lead to severe problems even in aspatial models, but get magnified in spatial models. For example, McMillen (2010, 2012) suggests that distributional form mis-specification of errors can themselves lead to spurious spatial correlation in residuals, and the study by Pinkse and Slade (2010) identifies the normality assumption almost universally used in spatial models as being "implausible". In this context, allowing for skewed and long-tailed error distributions is not merely an esoteric econometric consideration, but can be motivated on decision-theoretic grounds for travel behavior outcomes. For example, consider the case of the intensity of monthly walking frequency, as collected on an ordinal scale (ordinal scales are typically used in such cases, because it is easier for respondents to recollect their past behavior in grouped categories rather than the exact count; see Quinio and Lam, 2021; Thigpen, 2019). Even after conditioning on relevant exogenous factors, some individuals are likely to have a high intensity of walking due to unobserved individual and other location factors. In addition, clustering at the lower end of the spectrum is likely to prevail along with a bounding at zero. The net result would be a right-skewed error distribution with a long thin right tail, as was observed by Bhat (2024) in an analysis of walk frequency. More generally, there could be a variety of reasons why the unobserved error term distribution may be skewed with thin/fat tails on one side (see Calabrese et al. (2017) who identify different reasons for such skewed error distributions, including omitted variables with asymmetric distributions, or a misspecification in the functional form of an introduced exogenous variable).

Existing implementations of non-normal error structures in spatial modeling, notably by Bhat et al. (2017) and Farzammehr et al. (2021), have relied on multivariate skew normal (MSN) distributions across observations. One issue with the use of the MSN distribution in Bhat et al. (2017) is that the construction of the skew necessarily introduces correlation across observations,

so independence across the MSN error terms cannot be tested. In Farzammehr, they relax the restricted version of the MSN distribution of Bhat by using a more generalized version, and estimate the resulting model using a Bayesian hierarchical representation of the model system. However, these MSN specifications can still be restrictive in their distributions, because of the extent of thin/fat tails they can accommodate. Of course, one can consider other skew distributions, such as the multivariate skew-t distribution or even more general skew-elliptical distributions. But all these skew approaches manifest a singularity in Fisher's information matrix for the skew parameter, and can become difficult to estimate even in an aspatial context, leave alone a spatial context (see Arellano-Valle and Azzalini, 2006; Bhat and Sidharthan, 2012; Gómez-Déniz et al., 2021). Calabrese et al. (2017), who also consider non-normal errors, instead adopt a multivariate generalized extreme value (MGEV) distribution for the non-normality, and estimate the model using a Cholesky decomposition applied to the inverse of the covariance matrix of observations, exploiting the sparse nature of this inverse matrix. Again, while a definite advance over the use of normally distributed error terms, this study too adopts a rather restrictive non-normal distribution. All of the three studies above that have considered non-normal error terms also confine themselves to spatial error dependence, without considering spatial exogenous interaction (that is, spillover) effects (in the rest of this paper, we will use the terms spatial exogenous interaction effects and spillover effects interchangeably).

1.1 The Current Paper

This study contributes to the spatial analysis field through two novel modeling specifications. First, we construct a spatial model that captures both spatial dependence in observed covariates and spatial interactions in the unobserved error structure without relying on the traditional spatial autoregressive specifications. When applying such traditional specifications to discrete outcomes, several computational challenges arise beyond the conceptual and theoretical limitations associated with global spillover effects. In particular, estimation of the traditional non-SLX models involves computing the inverse of large N×N matrices, where N denotes the number of decision makers. This computational burden not only severely reduces estimation speed but also generates accumulated rounding errors when inverting sparse matrices (see Silveira Santos and Proença, 2019). Although Taylor series approximation methods have been proposed to mitigate these challenges, they require imposing certain constraints on the spatial weight matrix structure. Our modeling approach combines the SLX framework with Bhat and Sener's (2009) direct specification method for error correlation and heteroscedasticity, resulting in a spatial model that simultaneously incorporates spatial spillover effects and spatial error correlation without the restrictive initial assumption of global spillover effects or the computational challenges previously discussed (our approach does not entail the inversion of large matrices, which enables the approach to be applied to even big data applications).

Second, this study, to the best of our knowledge, represents the first attempt to develop a unified spatial analysis model (which we label as the LSLX-LSAE model for Local Spatial Lag on X-Local Spatial Asymmetric Error model) that brings together spatial spillover effects and spatial correlation/heteroscedasticity (as discussed in the previous section), along with a more flexible, asymmetric, and skewed distribution specification for the error terms, all within the context of a limited-dependent outcome (specifically, an ordered-response outcome). Unlike the Spatial Durbin Error Model (SDEM), our proposed model does not rely on an autoregressive disturbance structure. Instead, spatial dependence in the error component is introduced through a directly specified covariance structure with distance-decaying correlation and observation-specific

heteroscedasticity.[2] Simulation studies by Agiakloglou and Tsimpanos (2023) and Mur and Angulo (2009) have shown that estimating spatial models ignoring the error considerations identified above can lead to substantial instability in model performance and result in a complete failure to detect spatial spillovers. The studies listed earlier of Bhat et al. (2017), Calabrese et al. (2017), and Farzammehr et al. (2021) also demonstrate, through simulation and real world applications, the superior performance of spatial models that consider non-normal errors (relative to traditional normal-error models) in terms of better data fit, more accurate model estimates and elasticity effects, and improved point predictions. Pace and Calabrese (2022) demonstrate that, even if ignoring error non-normality may not always substantially affect the accuracy of parameter estimates and point predictions, explicitly considering the possibility of non-normal errors will, in general, allow for more robust inference as well as more accurate prediction intervals (because of the dependence of prediction intervals on the distribution of error terms). Accordingly, in this paper, to better accommodate asymmetric and skewed error distributions, we adopt Bhat's (2024) transformation-based flexible error structure approach in which the unobserved error terms are assumed to follow inverse Yeo-Johnson (YJ) transformed distributions of underlying normal distributions. The inverse YJ transformation is performed element-wise using a single parameter embedded in the transformation, allowing each observation's error distribution to closely approximate a variety of unimodal distributions such as the extreme value, skew-normal, and skew-t distributions, all depending on a single parameter value (see Bhat, 2024). Some earlier studies have also shown the efficacy and the parsimony of this YJ- transformation approach, showing that it is comparable to even mixtures of skew-normal distributions (let alone single skew-normal distributions) in the flexibility in distribution it provides (see Gallaugher et al., 2020; Watthanacheewakul, 2021; Marimuthu et al., 2022; Jadhav et al., 2023). The inverse YJ-transformed observation-specific non-normal error terms are then brought together across observations using an implicit Gaussian copula to develop a flexible non-normal multivariate error dependence structure across observations. The estimation of the unified framework is undertaken using Bhat's (2011) maximum approximate composite marginal likelihood (MACML) estimation method, achieving both computational stability and scalability to large datasets.

The rest of this paper is structured as follows. In the next section, we present the LSLX-LSAE model structure in the context of an ordered-response framework (though the model itself can be applied to other types of outcomes too). This section also discusses the estimation procedure for the proposed spatial ordered probit model. Section 3 undertakes simulation experiments to demonstrate the effectiveness of the proposed estimation procedure to recover parameters of an underlying spatial ordered response model with non-normally distributed error terms. The section also identifies the inaccuracy resulting from ignoring non-normality. Section 4 shows empirical results as a demonstration of the model and suggests policy implications based on the findings, and the final section concludes the paper.

[2] The LSLX-LSAE model proposed here is similar to the SDEM model, but the SDEM imposes the assumption of global dependence in spatial error correlation and, in earlier applications, uses normally distributed error terms (see Halleck Vega and Elhorst, 2015). Our LSLX-LSAE model considers a non-normally distributed error specification and allows for local interaction effects in the spatial error correlations, while also delineating the process generating error correlation effects from the process leading to error heteroscedasticity. Besides, earlier SDEM models have been estimated primarily for continuous outcome variables, not limited-dependent outcome models as in this paper.

2. MODELING FRAMEWORK

2.1 The Basic Structure

Let q $(q = 1, 2, ..., Q)$ be an index for decision agents (or also labeled as individuals in the following presentation), k $(k = 1, 2, ..., K)$ be the index of the category of ordinal response y_q, and y_q^* be the continuous latent propensity that is mapped to the ordered response y_q through thresholds. Starting from the SLX model for individual outcomes,

$$y_q^* = \mathbf{X}_q\boldsymbol{\beta} + \sum_{q'=1}^{Q} w_{qq'}\mathbf{X}_{q'}\boldsymbol{\gamma} + \varepsilon_q, \;\; y_q = k \text{ if } \psi_{k-1} < y_q^* < \psi_k \tag{1}$$

where $\mathbf{X}_q$ is a $(1 \times A)$ row vector of exogenous variables corresponding to individual q (excluding a constant), and $\boldsymbol{\beta}$ is a corresponding $(A \times 1)$ column vector of coefficients representing the direct effect of the exogenous variables. $w_{qq'}$ represents the weight of the spillover effect between individuals q and q'. Of course, because we already include the direct effect of X_q in the $\boldsymbol{\beta}$ vector, $w_{qq} = 0$. Also, let C_q represent the set of all individuals q' who have an exogenous interaction effect on individual q (note that global spillover effects assume that, for each and every individual, C_q includes all other individuals, while our local spillover specification allows the interaction effects to be limited to a select group of other individuals). With this definition of set C_q, $w_{qq'} > 0$ for $q' \in C_q$, and $w_{qq'} = 0$ for $q' \notin C_q$. Further, as is typical for SLX models (and most other spatial analysis forms), we row-normalize the weight matrix so that $\sum_{q'} w_{qq'} = 1$.[3] Thus, $\sum_{q'=1}^{Q} w_{qq'}X_{q'}$ in Equation (1) represents the weighted average of the covariates of individuals q' that influence the choice of individual q, and the $\boldsymbol{\gamma}$ vector captures the indirect (spillover) effects from others (the elements of $\boldsymbol{\gamma}$ measures the average effect of a one-unit increase in the covariates of others). The parameter ψ_k in Equation (1) is the upper bound threshold for ordinal level k, which, in the usual ordered-response form, satisfies $-\infty = \psi_0 < \psi_1 < ... < \psi_{K-1} < \psi_K = \infty$. ε_q is a non-normal asymmetric error term which is assumed to be an inverse-YJ transform of another normally

[3] There is no uniform consensus in the spatial analysis literature about the use of a row-normalized weight matrix or a scalar maximum eigenvalue-based normalization of the original weight matrix (see Tan et al., 2025). But, in the empirical context of our study that uses an SLX model to capture spillover effects in walking frequency, we argue that the row-normalization approach makes more behavioral sense. Essentially, row-normalization preserves the notion that it is the average socioeconomic characteristics and an averaged perception of built environment (BE) factors in neighboring spatial pockets that would impact walking of an individual, rather than the cumulative of these characteristics and factors across all neighboring spatial pockets. For example, it is likely that a general (average) notion of land-use mixing and walking infrastructure facilities at neighboring locations would affect an individual's own walk frequency, rather than a cumulative sum of land-use mixing opportunities and total number of walking facilities at neighboring locations, as overall individual perception of the built environment affects walking behavior (see Saelens and Handy, 2008 for review). In any case, it is important to note that the methodology we propose here is universally applicable to any form of the exogenous interaction weight matrix adopted, even though we row-normalize the original weights as follows: $w_{qq'} = w_{qq'} / \sum_{q'} w_{qq'}$. We discuss the form of the elements of the non-normalized matrix $w_{qq'}$ in the next section.

distributed random variable η_q with mean zero and variance σ_q. That is, $\varepsilon_q = t_\lambda^{-1}(\eta_q)$, where $t_\lambda^{-1}(\cdot)$ is the inverse-YJ transform with parameter $0 < \lambda < 2$ defined as follows:

$$\varepsilon_q = t_\lambda^{-1}(\eta_q) = \begin{cases} (\lambda\eta_q + 1)^{\left(\frac{1}{\lambda}\right)} - 1 & \text{if } \eta_q \geq 0 \\ 1 - [(\lambda - 2)\eta_q + 1]^{\left(\frac{1}{2-\lambda}\right)} & \text{if } \eta_q < 0 \end{cases} \tag{2}$$

If $0 < \lambda < 1$, $t_\lambda^{-1}(\cdot)$ is a function of η_q that shifts all values to the right while preserving their signs. Therefore, the positive region of ε_q (where η_q is positive) spreads out more to the right, while the negative region of ε_q (where η_q is negative) shrinks closer to zero and accumulates there. As a result, ε_q ends up having a right-skewed distribution. If $1 < \lambda < 2$, the opposite happens - $t_\lambda^{-1}(\cdot)$ shifts η_q values to the left while preserving their signs. Therefore, ε_q has a left-skewed distribution. If $\lambda = 1$, $t_\lambda^{-1}(\cdot)$ does not change the value function, so ε_q has the same distribution as η_q (see Bhat, 2024 for visual plots of the effects of the inverse YJ-transformation). For later use, we will also write the direct YJ transformation, which is as follows:

$$\eta_q = t_\lambda(\varepsilon_q) = \begin{cases} \dfrac{(\varepsilon_q + 1)^\lambda - 1}{\lambda} & \text{if } \varepsilon_q \geq 0 \\ -\dfrac{(-\varepsilon_q + 1)^{2-\lambda} - 1}{2 - \lambda} & \text{if } \varepsilon_q < 0 \end{cases} \tag{3}$$

Next define $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2', ...\mathbf{X}_Q')'$ ($Q \times A$ matrix), $\mathbf{y}^* = (y_1^*, y_2^*, ..., y_Q^*)'$ ($Q \times 1$ vector), $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_Q)'$ ($Q \times 1$ vector), $\boldsymbol{\eta} = (\eta_1, \eta_2, ..., \eta_Q)'$ ($Q \times 1$ vector), and $\boldsymbol{\psi} = (\psi_1, ..., \psi_{K-1})'$ $[(K-1) \times 1]$ vector. For future use, we also define $\boldsymbol{\psi} = (-\infty, \boldsymbol{\psi}', \infty)'$ $[(K+1) \times 1]$ vector. Now, stack the weights into a matrix $\mathbf{W}$ of size $Q \times Q$, with the weights in each row corresponding to the weights of other individuals corresponding to the individual in that row. Let $\mathbf{t}_\lambda^{-1}(.)$ represent a vector function that applies the inverse YJ-transform element-wise to each component of a column vector and returns the transformed column vector, so that the non-normal error vector may be written as $\boldsymbol{\varepsilon} = \mathbf{t}_\lambda^{-1}(\boldsymbol{\eta})$. Correspondingly, $\boldsymbol{\eta} = \mathbf{t}_\lambda(\boldsymbol{\varepsilon})$. Then, we may write the latent propensity of the proposed LSLX-LSAE model in Equation (1) in matrix notation as:

$$\mathbf{y}^* = \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} = \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\gamma} + \mathbf{t}_\lambda^{-1}(\boldsymbol{\eta}) \tag{4}$$

2.2 Specification of the Exogenous Interaction Weight Matrix and Extent of Influence

An important consideration in the current LSLX-LSAE model is the form of the row-normalized weight matrix elements $w_{qq'}$ in Equation (1) to get $\mathbf{W}$ in Equation (4) and the determination of the extent of spatial exogenous interaction (that is, spillover) effects (i.e., the construction of the set C_q). Some earlier studies suggest that the nature, and extent, of spatial dependence should be determined a priori based on theoretical notions (see, for example, Neumayer and Plümper, 2016).

But, in many cases, theoretical notions, while providing a general sense of interactions, do not offer enough practical guidance regarding how to specify the fading nature of the (social or spatial) distance-based influence between individuals, and the threshold beyond which each exogenous variable's influence all but ceases to exist (see Halleck Vega and Elhorst, 2013 for a discussion). Thus, many empirical spatial studies adopt one of four popular specifications of non-normalized weight matrices (with elements $w_{qq'}$): (1) a contiguity specification (in the form of a Boolean representing whether two individuals share a common border, or a shared border length specification), (2) a certain predetermined number of nearest neighbors, (3) an inverse distance decay specification; $w_{qq'} = d_{qq'}^{-\alpha}$, or (4) a negative exponential distance decay function $w_{qq'} = \exp(-\alpha d_{qq'})$. The last of these two need to maintain the condition $\alpha > 0$, and can be specified with a threshold cut-off (as we have done for the error dependence above). Earlier studies using the last two types of specifications, however, do not typically estimate the intensity of distance-based fading; rather, for the sake of estimation simplicity, they prespecify the values of the distance decay parameter (thus, for example, setting α to be one).

In this paper, in terms of functional form, we adopt the negative exponential distance decay form because we expect the spillover effects to not be confined simply to a form of contiguity (note also that the negative exponential distance decay form can closely mimic strict contiguity patterns if the decay parameter has a high magnitude). Also, the negative exponential distance decay function displays much faster spillover fading than the inverse distance decay function (in the latter, spillovers can extend over long ranges, with the characteristics of distant neighbors and their neighborhoods still exerting influence, risking unrealistically strong long-distance effects). In travel behavior contexts, the exogenous spillovers operate through rather localized environment attributes, which is consistent with the use of the negative exponential decay function. Besides, from an estimation standpoint, the negative exponential decay function provides more smoothness by way of defining a cutoff threshold (because of the quick progression toward a zero value), while the inverse distance function tends to have a long right tail of influence making cut-off thresholds feel like a cliff. However, rather than prespecifying the α parameter as in most earlier studies, we explicitly estimate this parameter. Besides, unlike almost all earlier studies, in the current paper, we adopt different decay functions for the spillover effects and for the spatial error correlation effects (that is, we do not constrain the α parameter for spillover effects to be the same as the ρ parameter for error correlation effects).[4]

One additional issue in our specification of the weights relates to the determination of the set C_q. We achieve this by estimating the model without setting a threshold cut-off for the spillover effect (this is tantamount to a global spillover effect to begin with), and then determining a cutoff point where the weight values effectively become very close to zero. Of course, a question may be raised regarding the need to have a cutoff point once a model is already estimated with a specific decay parameter. There are two reasons for this. From a conceptual standpoint, it is useful to obtain a behavioral sense of the extent of spatial spillovers by explicitly identifying the cutoff point from the estimated α parameter. The second is that, to facilitate estimation of the entire model with many parameters, we actually pursue the effort in three steps as discussed in the Section 2.4. In doing so, it becomes very efficient to identify the cutoff value to determine the elements of set C_q.

[4] To be fair, we are able to do so because of the use of the SLX specification. As indicated by Halleck Vega and Elhorst (2015), allowing different parameterizations of the spillover and error decays can be problematic from an identification point of view in other traditional spatial dependence models such as the SDM.

2.3 Specification of the Local Spatial Multivariate Error

To complete the model formulation, we need to specify a multivariate distribution specification for the normal error vector $\boldsymbol{\eta}$, which allows correlation and heteroscedasticity in the error vector $\boldsymbol{\varepsilon}$. In essence, specifying $\boldsymbol{\varepsilon} = \mathbf{t}_{\lambda}^{-1}(\boldsymbol{\eta})$ and a multivariate normal distribution for $\boldsymbol{\eta}$ represents an implicit copula technique (see Bhat, 2024). In the current paper, we construct the $(Q \times Q)$ covariance matrix $\boldsymbol{\Omega}$ of $\boldsymbol{\eta}$ by partitioning this covariance matrix into a diagonal $(Q \times Q)$ matrix $\boldsymbol{\sigma}$ of standard deviations and a $(Q \times Q)$ correlation matrix $\boldsymbol{\Omega}^*$: $\boldsymbol{\Omega} = \boldsymbol{\sigma}\boldsymbol{\Omega}^*\boldsymbol{\sigma}$. Each diagonal element σ_q of the diagonal matrix $\boldsymbol{\sigma}$ is specified as $\sigma_q = \exp(\mathbf{Z}_q\boldsymbol{\theta})$, where $\mathbf{Z}_q$ is a $(1 \times I)$ row vector exogenous variables, and $\boldsymbol{\theta}$ is a corresponding $(I \times 1)$ column vector of parameters to be estimated. To avoid identification problems, $\mathbf{Z}$ should not include a constant term. Next, we specify the off-diagonal elements of the correlation matrix $\boldsymbol{\Omega}^*$ as follows:

$$\omega_{qq'} = \begin{cases} \exp(-\rho d_{qq'}) & \text{if } d_{qq'} < \kappa \\ 0 & \text{if } d_{qq'} \geq \kappa \end{cases} \tag{5}$$

where $d_{qq'}$ is the distance (in social or geographic space) between two individuals ($q \neq q'$) and $\rho > 0$ (which we will refer to as the spatial error correlation (or SEC) parameter). κ represents a threshold distance beyond which there is no error correlation between individuals. The above specification meets all the requirements of a proper distance-decaying correlation matrix. Specifically, it ensures that $\boldsymbol{\Omega}^*$ (a) is symmetric with all off-diagonal entries being less than one, (b) has error correlation elements that fade with distance, and, most importantly, (c) is positive-definite (see Schoenberg, 1938; Cressie, 1993)[5]. The specification, as in most spatial analysis situations, precludes negative correlations because we expect unobserved factors to permeate through space in a positive manner in travel behavior choice contexts. For later use, we will define D_q as the set of all individuals q' who have an error correlation effect with individual q (that is, individuals who are within the threshold κ of individual q), so that $\omega_{qq'} > 0$ for $q' \in D_q$, and $\omega_{qq'} = 0$ for $q' \notin D_q$.

The model specification proposed here (a) is based on local spillover effects, (b) allows error heteroscedasticity to be distinct from that generated through pure error spatial dependency effects (that is, error correlation and error heteroscedasticity are not tied closely together), (c) allows for independent spatial dependency effects in exogenous interaction influences and error correlations, and (d) accommodates flexible non-normal and asymmetric error distributions.

2.4 Estimation Methodology

The parameters to be estimated in the LSLX-LSAE model include: $\boldsymbol{\psi} = \{\psi_1, \psi_2 ..., \psi_{K-1}\}'$ (ordered-response thresholds), $\boldsymbol{\beta}$ (effects of individuals' exogenous variables), $\boldsymbol{\gamma}$ (exogenous spatial interaction spillover effects), $\boldsymbol{\theta}$ (effects of individuals' exogenous variables on the variance of the

[5] Similar to the case of the spillover effects, we adopt this exponential decay function specification rather than an inverse distance (power law) specification for error correlations too because the exponential decay function leads to rapidly decaying correlation with distance. In many travel behavior contexts (such as the influence of unobserved neighborhood demographic and BE factors on walking frequency), we expect localized spatial error dependence effects such that the influence of observations quickly becomes negligible beyond a low distance threshold κ.

YJ-transformed normally distributed error term η_q before transformation), λ (the YJ-transformation parameter), α (the spatial spillover decay parameter) and ρ (the SEC parameter). To respect the constraints $0 < \lambda < 2$, $\alpha > 0$ and $\rho > 0$ during estimation, we parameterize these as follows: $\lambda = \frac{2\exp(\lambda^*)}{1+\exp(\lambda^*)}$, $\alpha = \exp(\alpha^*)$, and $\rho = \exp(\rho^*)$. Once the estimation is completed, in a last iteration, we use the inverse parameterizations to obtain the estimates of λ, α, and ρ. Collect all of these parameters into a single parameter vector $\boldsymbol{\tau} = \{\boldsymbol{\psi}', \boldsymbol{\beta}', \boldsymbol{\gamma}', \boldsymbol{\theta}', \lambda^*, \alpha^*, \rho^*\}$.

Let individual q be observed to choose ordinal category m_q ($m_q \in \{1, 2, ..., K-1, K\}$). Define two $[Q \times (K+1)]$ mask matrices $\mathbf{M}_{low}$ and $\mathbf{M}_{high}$ as follows. First fill both matrices with elements of zero. Then, for each individual q, (a) set the element in the q^{th} row and m_q^{th} column of $\mathbf{M}_{low}$ to one, and (b) set the element in the q^{th} row and $(m_q+1)^{th}$ column of $\mathbf{M}_{high}$ to one. Let $\mathbf{G}_{low} = (\mathbf{M}_{low}\boldsymbol{\psi})$, $\mathbf{G}_{high} = (\mathbf{M}_{high}\boldsymbol{\psi})$, $\mathbf{B} = \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\gamma}$, $\mathbf{S}_{low} = \boldsymbol{\sigma}^{-1}\mathbf{t}_\lambda(\mathbf{G}_{low} - \mathbf{B})$, and $\mathbf{S}_{high} = \boldsymbol{\sigma}^{-1}\mathbf{t}_\lambda(\mathbf{G}_{high} - \mathbf{B})$. Then, the likelihood function may be written as:

$$
\begin{aligned}
L(\boldsymbol{\tau}) &= \text{Prob}\left[\mathbf{G}_{low} < \mathbf{y}^* < \mathbf{G}_{high}\right] = \text{Prob}\left[\mathbf{G}_{low} < \mathbf{B} + \boldsymbol{\varepsilon} < \mathbf{G}_{high}\right] = \text{Prob}\left[\mathbf{G}_{low} < \mathbf{B} + \mathbf{t}_\lambda^{-1}(\boldsymbol{\eta}) < \mathbf{G}_{high}\right] \\
&= \text{Prob}\left[(\mathbf{G}_{low} - \mathbf{B}) < \mathbf{t}_\lambda^{-1}(\boldsymbol{\eta}) < (\mathbf{G}_{high} - \mathbf{B})\right] = \text{Prob}\left[\{\mathbf{t}_\lambda(\mathbf{G}_{low} - \mathbf{B})\} < \boldsymbol{\eta} < \{\mathbf{t}_\lambda(\mathbf{G}_{high} - \mathbf{B})\}\right] \qquad (6) \\
&= \int_{\mathbf{S}_{low}}^{\mathbf{S}_{high}} \boldsymbol{\phi}_Q(\boldsymbol{\Omega}^*) d\boldsymbol{\eta},
\end{aligned}
$$

where $\boldsymbol{\phi}_Q(\boldsymbol{\Omega}^*)$ represents the Q-dimensional standard multivariate normal density function with correlation matrix $\boldsymbol{\Omega}^*$.

The likelihood function in Equation (6) entails the computation of the order of 2^Q multivariate normal cumulative distribution functions of size Q, which is impractical to evaluate using traditional numerical simulation techniques (see Bhat et al., 2016; Müller and Czado, 2005). So, we use the composite marginal likelihood (CML), as introduced by Bhat (2011) for the estimation of models with spatial dependence. The key to the CML approach is the closure property under marginalization of the multivariate normal distribution. Importantly, although the original error term vector $\boldsymbol{\varepsilon}$ is non-normal, applying the inverse YJ transformation conveniently reduces the likelihood function to an integral over a multivariate normal density (see Equation (6)). Thus, in combination with the CML approach, the inverse YJ transform to generate non-normality in spatial dependence models affords particular elegance and simplicity relative to other approaches to generate non-normality. In particular, CML estimation becomes no more difficult than the case of adopting a normal error term for $\boldsymbol{\varepsilon}$. To see this, and following Bhat (2011), we maximize a much easier-to-compute two-dimensional surrogate likelihood function constructed by compounding all combinations of pairwise probabilities of each observation q with observations q' that are contained in set E_q, defined as the union of the sets C_q and D_q. The identification of the set E_q makes the CML approach computationally fast, as only the pairings of each observation q with observations q' that are contained in set E_q need to be considered in the CML estimation rather than all pairings of each observation q with other observations. Given the

local spillovers and local spatial error dependence we expect in travel behavior models, this makes the estimation of the proposed LSLX-LSAE model quick. Our approach to identify the set E_q is discussed in the next section.

The pairwise CML estimator is the one that maximizes the compounded probability of all pairwise individuals. Under usual regularity assumptions, the CML estimator is consistent and asymptotically normal distributed, thanks to the unbiasedness of the CML score function (for a formal proof, see Yi et al., 2011 and Bhat, 2014; for extended discussions of the general CML estimator, see Molenberghs and Verbeke, 2005; Varin and Vidoni, 2008; Xu and Reid, 2011). For the proposed LSLX-LSAE model, the pairwise CML function may be written as:

$$\begin{aligned} L_{CML}(\boldsymbol{\tau}) &= \prod_{q=1}^{Q-1} \prod_{\substack{q'=q+1 \\ q' \in E_q}}^{Q} L_{qq'}, \text{ where } L_{qq'} = P\left(y_q = m_q, y_{q'} = m_{q'}\right) \\ &= \begin{bmatrix} \Phi_2(S_{high,q}, S_{high,q'}, \boldsymbol{\Omega}^*_{qq'}) - \Phi_2(S_{high,q}, S_{low,q'}, \boldsymbol{\Omega}^*_{qq'}) \\ -\Phi_2(S_{low,q}, S_{high,q'}, \boldsymbol{\Omega}^*_{qq'}) + \Phi_2(S_{low,q}, S_{low,q'}, \boldsymbol{\Omega}^*_{qq'}) \end{bmatrix}, \end{aligned} \tag{7}$$

where $S_{high,q}$ and $S_{low,q}$ refer to the q^{th} element of the $(Q \times 1)$ vectors $\mathbf{S}_{high}$ and $\mathbf{S}_{low}$, respectively, $\boldsymbol{\Omega}^*_{qq'}$ represents the qq'^{th} element of the matrix $\boldsymbol{\Omega}^*$, and Φ_2 is the bivariate cumulative standard normal distribution function.

Defining R_q as the cardinality of set E_q for the q^{th} individual, and $R = \sum_{q=1}^{Q-1} R_q$ (R refers to the total number of pairings in the CML function of Equation (7)), the covariance matrix of the CML estimator is given by $\dfrac{\left[\hat{\boldsymbol{H}}^{-1}\right]\left[\hat{\boldsymbol{J}}\right]\left[\hat{\boldsymbol{H}}^{-1}\right]'}{R}$, where

$$\hat{\boldsymbol{H}} = -\frac{1}{R}\left[\sum_{q=1}^{Q-1} \sum_{\substack{q'=q+1 \\ q'' \in E_q}}^{Q} \frac{\partial^2 \log L_{qq'}}{\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}'}\right]_{\hat{\boldsymbol{\tau}}_{CML}}, \text{ or alternatively,} \tag{8}$$

$$\hat{\boldsymbol{H}} = \frac{1}{R}\sum_{q=1}^{Q-1} \sum_{\substack{q'=q+1 \\ q' \in E_q}}^{Q} \left(\left[\frac{\partial \log L_{qq'}}{\partial \boldsymbol{\tau}}\right]\left[\frac{\partial \log L_{qq'}}{\partial \boldsymbol{\tau}'}\right]\right)_{\hat{\boldsymbol{\tau}}_{CML}} \tag{9}$$

However, the estimation of the "vegetable" matrix $\boldsymbol{J}$ is more difficult in this case, but may be estimated using the windows resampling procedure as proposed by Bhat (2011, 2014).

2.5 Estimation Procedure

The CML estimation of the proposed model entails several parameters, and estimating all of these at once can be challenging without good start values. Also, identifying the set E_q prior to estimation makes the process fast, as discussed earlier. But, at the same time, E_q itself is based on the sets C_q and D_q, which are based on the spillover parameter α and SEC parameter ρ. For both of these reasons, we proceed in estimation in a three-step manner. In the first step, we ignore

spatial error correlation, but consider direct exogenous variable effects, spatial spillover effects, error heteroscedasticity, and error non-normality. The estimation of the resulting aspatial model to obtain estimates of $\boldsymbol{\tau} = \{\boldsymbol{\psi}', \boldsymbol{\beta}', \boldsymbol{\gamma}', \boldsymbol{\theta}', \lambda^*, \alpha^*\}$ is straightforward using maximum likelihood techniques. Different specifications of exogenous variables are comprehensively tested at this stage. Important to note is that ignoring spatial correlation as done at this first step should not affect the consistency of the parameters in $\boldsymbol{\tau}$, including the spatial spillover decay parameter estimate α because spillover effects enter through observed covariates rather than through the dependent variable (ignoring spatial error correlation in the first stage affects efficiency but does not generally lead to inconsistency in the estimation of the spillover decay parameter). Using this estimate of α, we can determine the cutoff point where the spillover effect is effectively zero to then determine the set C_q for each individual *q*. Then, in a second step, we fix all the parameters estimated in this first step, construct the weight matrix W based on the spillover cutoff point, and run the CML estimation across all pairings of individuals (note that this second estimation is not confined only to the set C_q for each individual *q*) to obtain an initial estimate of ρ. This allows us to determine the distance threshold κ for spatial error dependence to develop the set D_q for each individual *q*. Finally, in the third step, with initial start values for all parameters obtained in the earlier two steps, and the set E_q developed as the union of the set C_q and D_q, we estimate all parameters at once.

3. SIMULATION STUDY

3.1 Simulation Design

The simulation exercises are conducted to verify the ability of the proposed LSLX-LSAE model to recover parameters from datasets generated from known underlying parameters. For the simulation study, we considered an ordered-response model with five categories. We assigned 1200 observation units or individuals $(Q = 1200)$ evenly over a 20×20 spatial grid system comprising a total of 400 spatial units, placing three individuals per spatial unit. Starting from the origin point, the top right vertex of each of the 400 spatial units $(C = 1, 2, ...400)$ is associated with cartesian coordinates (i_c, j_c), $i_c \in \{1, 20\}$, $j_c \in \{1, 20\}$. Each unit in the cartesian coordinate system corresponds to five miles, so that each edge of the square grid system is 100 miles in length.

Five independent variables are employed. The first two variables (X_{q1}, X_{q2}) are binary dummy variables associated with individuals and do not have any spatial spillover effect. The values for each of these two binary variables are randomly generated across individuals as a realization of a Bernoulli process. The next two (X_{q3}, X_{q4}) are also binary variables, but now associated with spatial units, with individuals located within the same spatial unit *c* assigned the same value for each of the two variables (even though these two variables are spatial-unit specific, for notational ease, we write them as (X_{q3}, X_{q4})). In particular, for X_{q3}, all individuals in spatial units *c* whose top right vertex has coordinates $(i_c \le 5,\ j_c \le 5)$, $(i_c \le 5, 10 < j_c \le 15)$, $(5 < i_c \le 10,\ 5 < j_c \le 10)$, $(5 < i_c \le 10, 15 < j_c \le 20)$, $(10 < i_c \le 15,\ j_c \le 5)$, $(10 < i_c \le 15, 10 < j_c \le 15)$, $(15 < i_c \le 20,\ 5 < j_c \le 10)$, or $(15 < i_c \le 20, 15 < j_c \le 20)$ are assigned a value of $X_{q3} = 1$, while all individuals in the other spatial units *c* are assigned a value of $X_{q3} = 0$. Across the entire grid system, this creates a block-wise checkboard pattern, where each block consists of 5 by 5 array of

spatial units. In contrast to X_{q3} that varies slowly over space, X_{q4} is set to vary more rapidly over space. This is to accommodate different distribution patterns across variables, as undertaken in earlier studies (see, for example, Heagerty and Lumley, 2000). Specifically, all individuals in spatial units whose top right vertex has an odd value for i_c are assigned a value of $X_{q4}=1$, while all individuals in spatial units whose top right vertex has an even value for i_c are assigned a value of $X_{q4}=0$. Across the entire grid system, this results in a vertical striped pattern with a width of one cell, alternating between 1s and 0. A visual pattern of the distribution of (X_{q3}, X_{q4}) is provided in Figure 1, with the spatial units for which these variables take a value of 1 being shaded. Also, different from the effects of (X_{q1}, X_{q2}), we allow both a direct effect for the spatial unit-specific variables (X_{q3}, X_{q4}) on the choices of individuals located within a spatial unit, as well as a local spatial spillover effect for the variables at one spatial unit on the choices of individuals located in other proximal spatial units. The fifth variable (Z_{q1}) is a binary variable that engenders heteroscedasticity across individuals. As with the first two variables (X_{q1}, X_{q2}), the values for (Z_{q1}) are also randomly generated across individuals as a realization of a Bernoulli process. Once the five independent variables have been generated, these are fixed for the rest of the simulations.

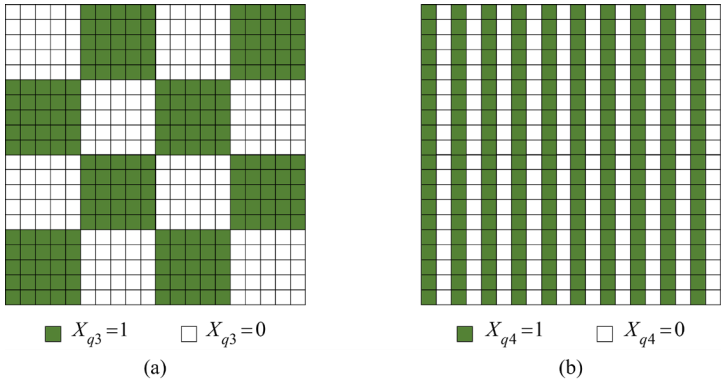

Figure 1 Distribution patterns of independent variables assigned to spatial units – (a) X_{q3} and (b) X_{q4}

Next, define the following vectors: $\mathbf{X}_k = (X_{1k}, X_{2k}, ..., X_{Qk})'$ ($Q\times 1$ vectors) for $k = 1,2,3,4$, and $\mathbf{Z_1} = (Z_{11}, Z_{21}, ..., Z_{Q1})$ ($Q\times 1$ vector). Then, in the context of our simulation, Equation (4) may be written as follows:

$$\begin{aligned}&\mathbf{y}^* = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \mathbf{X}_3\beta_3 + \mathbf{X}_4\beta_4 + \mathbf{W}\mathbf{X}_3\gamma_3 + \mathbf{W}\mathbf{X}_4\gamma_4 + \mathbf{t}_\lambda^{-1}(\boldsymbol{\eta})\\ &\boldsymbol{\eta} \sim MVN(\mathbf{0}, \boldsymbol{\sigma}\boldsymbol{\Omega}^*\boldsymbol{\sigma}), \boldsymbol{\sigma} = \exp(\mathbf{Z}_1\theta)\end{aligned} \tag{10}$$

In our simulations, the parameters are set as follows: $(\beta_1, \beta_2, \beta_3, \beta_4) = (-1.0, 1.0, 1.0, -1.0)$ representing the direct effect of individual characteristics, $(\gamma_3, \gamma_4) = (3.0, -3.0)$ for spatial spillover effects, and $\theta = 0.8$ for the heteroscedasticity effect. This leaves the YJ parameter λ and the parameters determining the exogenous interaction weight matrix $\mathbf{W}$ and the local spatial multivariate error correlation matrix $\boldsymbol{\Omega}^*$. Based on Section 2.4, we specify

$\lambda = \frac{2\exp(-0.5)}{1+\exp(-0.5)} = 0.7551$, which reflects a rightward skew for the error term (as has been observed in an aspatial ordered-response model for walking frequency in Bhat, 2024). The elements of **W** are next obtained by row-normalization (see Section 2.2) by specifying the elements of the non-normalized weight matrix $w_{qq'}$ as $w_{qq'} = \exp(-\alpha d_{qq'})$ with $\alpha = \exp(\alpha^*) = \exp(-0.5) = 0.6065$. In this specification, however, we specify $w_{qq'} = 0$ for all pairs of individuals q and q' within the same spatial unit, so that the spatial unit-specific variables have spillover effects across spatial units, with only a direct effect within spatial units. The distance $d_{qq'}$ for two individuals in different spatial units is specified to be the Euclidean distance between the centroids of the spatial units in which the individuals are located (the average Euclidean distance between spatial units in the entire grid system is 25.9). Also, to maintain a local spillover effect, we specified spillovers only among individuals in spatial units that are within a threshold distance of 15.18 miles of each other (equivalent to $w_{qq'}$ being 0.0001 or higher). In doing so, spatial spillover is confined to 6.1% of the individual pairs in the simulation sample. Finally, in the simulation, we specify the elements $\omega_{qq'} = \exp(-\rho d_{qq'})$ of the local spatial multivariate error correlation with the SEC parameter $\rho = \exp(\rho^*) = \exp(-0.2) = 0.8187$. We employ a higher value for ρ relative to α to recognize that unobserved correlation effects (that include a whole variety of different unobserved characteristics) are likely to span longer distances than spillover effects associated with specific exogenous variables. Importantly, to allow for unobserved error correlations even in the choices of individuals residing within the same spatial unit, we assigned a value of $d_{qq'}$ (for the elements in the spatial error correlation matrix) of the average distance between two randomly located points in space within the spatial unit (Santalo, 2004). Given that all spatial units in the grid system used in this simulation are uniform in shape and size, we assigned a constant distance of 2.65 miles between any two individuals residing in the same unit, regardless of the spatial units' locations. Again, to maintain a local spatial error effect, we specified error correlation effects only among individuals in spatial units that are within a distance threshold of 28.12 miles (equivalent to $\omega_{qq'}$ being 10^{-10} or higher). With this threshold, the spatial error effect is confined to 13.9% of the individual pairs.

Using values drawn from the multivariate distribution of $\boldsymbol{\eta}$, followed by the inverse YJ transformation of the error draws, the multivariate (1200x1)-latent variable vector $\mathbf{y}^*$ is generated for the specific realization of the multivariate vector $\boldsymbol{\eta}$. For this specific instance of the vector $\mathbf{y}^*$, quintile values (that is, 20th, 40th, 60th, and 80th percentile values) are computed across the 1200 elements of the vector $\mathbf{y}^*$. The above procedure is undertaken 1000 times to generate 1000 instances of latent variable vectors and 1000 sets of quintile values, and the average values of the quintiles across the 1000 sets are obtained and specified as the parameter values of the $\boldsymbol{\Psi}$ vector: $\psi_1 = -1.640$, $\psi_2 = 0.291$, $\psi_3 = 1.629$, and $\psi_4 = 3.028$. Using these threshold values, the elements of each of the 1000 instances of generated $\mathbf{y}^*$ vectors are converted into a choice among five ordinal categories to create 1000 data sets.

The CML estimation procedure of Section 2.5 is applied to each data set, based on only those pairings of individuals corresponding to the set E_q developed as the union of the set C_q and

D_q discussed there. Thus, our simulations not only provide the data-specific values of the vector $\{\boldsymbol{\psi}', \boldsymbol{\beta}', \boldsymbol{\gamma}', \boldsymbol{\theta}', \lambda, \alpha, \rho\}$, but also provide the thresholds corresponding to the sets C_q and D_q (which can then be compared with the thresholds of 15.18 miles and 28.12 miles, as specified in the generation of the data sets). The Godambe information-based covariance matrix and the corresponding standard errors of the parameters from each data set are also computed. To evaluate the ability of our proposed model to recover the coefficients and their standard errors, we calculated the following statistics, following Bhat et al. (2015) and Mondal and Bhat (2022):

(1) The mean percentage bias (MPB): The absolute value of the percentage deviation of the mean estimate from the true value, which can be formulated as follows:

$$\text{MPB}(\%) = \left| \frac{\text{Mean Estimate} - \text{True Value}}{\text{True Value}} \right| \times 100$$

(2) The finite sample standard error (FSSE): The standard deviation of parameter estimates across datasets, which serves as the empirical standard error with finite samples.

(3) The asymptotic standard error (ASE): The average of estimated standard errors for each parameter across datasets, computed following the procedure discussed in the previous section.

(4) Relative Efficiency (RE): The ratio of the ASE to the FSSE, computed as RE=ASE/FSSE. This metric evaluates how well the estimated asymptotic standard errors, which are computed using the window resampling procedure, approximate the FSSE.

3.2 Simulation Results

Table 2 presents the performance of our proposed estimation procedure for the LSLX-LSAE model in recovering the true parameters. The MPB measures the difference between the mean estimate (the aggregated simulation result) and the true value, ranging from 0.1% to 8.9%. The elevated MPB for Threshold 2|3 results from its true value being near zero, despite showing a comparable bias magnitude to other thresholds (thus, the small value of the parameter leads to a high percentage bias). In general, the parameters related to the explanatory variables without spatial effects (β_1, β_2, θ) show superior recovery performance compared to the parameters related to the explanatory variables with spatial effects ($\beta_3, \beta_4, \gamma_3, \gamma_4$). This is unsurprising given that the parameters unrelated to spatial effects affect the likelihood function more directly and linearly, while other parameters either exhibit interrelations due to spatial effects or influence the model through more involved, nonlinear relationships. Notably, the computationally-efficient stepwise estimation protocol achieves high recovery accuracy for all parameters governing the spatial dependency and the error structure (λ, α, ρ and threshold distances) with the MPB ranging from 0.2% to 5.2%. The slight overestimation of the spatial spillover threshold distance (15.971 rather than 15.180) increases the proportion of individual pairs with valid spillover effects from 6.1% to 7.7%. However, after normalization, the maximum spatial weight of these added pairs is only 1.8×10^{-4} and their maximum cumulative weight for any single individual is 1.6×10^{-3}. For spatial correlation, estimating a shorter threshold distance (26.659 instead of the true value of 28.120) simply excludes pairs with negligible correlations (between 1.0×10^{-10} and 3.3×10^{-10} calculated from the estimated SEC parameter) during CML likelihood calculation. These proportions are sufficiently small to ensure that the final results remain unaffected. These results demonstrate that

parameters not explicitly estimated in conventional spatial econometric models can be accurately recovered at feasible computational cost using our proposed stepwise procedure.

Table 2 Parameter Recovery Ability of the Estimation Method

Parameters	True value	Mean estimate	MPB	FSSE	ASE	RE
Direct effect parameters						
β_1	-1.000	-1.005	0.5%	0.051	0.047	0.926
β_2	1.000	1.008	0.8%	0.051	0.049	0.949
β_3	1.000	1.029	2.9%	0.087	0.074	0.857
β_4	-1.000	-0.983	1.7%	0.053	0.048	0.905
Spatial spillover effect parameters						
γ_3	3.000	3.019	0.6%	0.082	0.073	0.886
γ_4	-3.000	-3.012	0.4%	0.048	0.042	0.880
Heteroscedasticity parameter						
θ	0.800	0.803	0.3%	0.042	0.039	0.908
YJ parameter						
λ	0.755	0.754	0.2%	0.047	0.056	1.174
Spatial spillover decay parameter embedded in the weight (W) matrix						
α	0.607	0.610	0.5%	0.017	0.014	0.817
Spatial error correlation (SEC) parameter						
ρ	0.819	0.818	0.2%	0.021	0.021	0.987
Thresholds						
Threshold 1|2	-1.640	-1.643	0.2%	0.025	0.023	0.924
Threshold 2|3	0.291	0.317	8.9%	0.033	0.030	0.921
Threshold 3|4	1.629	1.659	1.8%	0.043	0.040	0.933
Threshold 4|5	3.028	3.073	1.5%	0.030	0.026	0.862
Threshold distances for spatial effects						
Spatial Spillover Effect	15.180	15.971	5.2%	3.152	na	na
Spatial Error Dependence	28.120	26.659	5.2%	2.391	na	na

Moving on to the deviation of estimates, the relative efficiency (RE) statistic compares the finite-sample standard deviation (FSSE) of parameter estimates across the 250 simulation replications with the average asymptotic standard error (ASE) computed from the Godambe information matrix using the windows resampling procedure. An RE value of unity indicates exact agreement between asymptotic and empirical dispersion. In finite samples, deviations from unity are expected because (a) the ASE is derived from a composite likelihood approximation rather than the full likelihood, and (b) the J matrix in the sandwich covariance expression is itself estimated using window resampling, which introduces smoothing and truncation effects (Bhat, 2014). In our simulation design with $Q = 1200$ observations, exponential decay in both spillover and spatial error correlation, and local truncation of pairwise interactions (6.1% of pairs for spillovers and 13.9% for spatial error correlation), the RE values range from 0.857 to 0.987 for thirteen of the fourteen parameters, with a single value of 1.174 for the YJ transformation

parameter. The predominance of RE values below unity indicates a modest downward bias in the ASE relative to the empirical FSSE. This pattern is consistent with the known finite-sample behavior of kernel- and window-based covariance estimators under smoothly decaying spatial dependence. Because the window procedure effectively truncates weak long-range correlations, the resulting sandwich variance estimator may slightly compress dispersion in moderate samples (see Kelejian and Prucha, 2007). Importantly, though, the magnitude of these differences is small in absolute terms. For most parameters, the difference between ASE and FSSE is on the order of 0.01–0.02. When evaluated relative to the true parameter magnitudes (which range between approximately 0.6 and 3.0 for the structural parameters), these differences represent only a very small fraction of the underlying effect size. Thus, although the RE statistic indicates mild finite-sample underestimation of dispersion, the absolute deviation between ASE and FSSE is economically and behaviorally negligible. The central criterion—accurate recovery of parameter values—remains strongly satisfied, as evidenced by the very low mean percentage bias across parameters. The single RE value above unity (1.174) occurs for the YJ transformation parameter, which enters the likelihood nonlinearly and is therefore more sensitive to simulation variability across replications. Even in this case, the absolute difference between ASE and FSSE is small and does not materially affect inference.

Overall, the RE results indicate that the window-resampling-based Godambe covariance matrix provides a close approximation to the true finite-sample dispersion in a spatial setting characterized by local spillovers and exponentially decaying spatial error correlation. Given the composite likelihood framework and moderate sample size, the observed RE magnitudes are well within the range typically encountered in spatial composite likelihood and spatial heteroscedastic and autocorrelation consistent (HAC) applications, and they support the reliability of the reported asymptotic standard errors.

4. EMPIRICAL RESULTS

4.1 Data Description

To demonstrate an application of the proposed model, we use data from the 2023 Puget Sound Regional Household Travel Study (Puget Sound Regional Council or PSRC, 2024). The survey was conducted from March to June 2023 in the Puget Sound Region, Washington, United States, and covered the metropolitan area of Seattle. The sampling frame was the list of all households in the four-counties of King, Kitsap, Pierce, and Snohomish counties in the area. Survey sampling and administration details are available in PSRC, 2024. The survey collected information on sociodemographic, employment/schooling, and residential location of all individuals in each sampled household, as well as their travel diaries for seven consecutive days. From the travel diaries, we extract information on the walking frequency of individuals, which is the focus of the current application. Walking behavior is known to be strongly influenced by the built and social environment (see Besser et al., 2021, for example), and hence spatial interdependence is also expected. In addition, because walking trips cover shorter distances than other travel modes, the spatial decay in environmental influences is likely more rapid. Therefore, spatial spillover and correlation effects should be modeled as localized phenomena rather than global effects.

We used the response from the following questions to determine individuals' walking frequency:

- Do you walk outside for 15 minutes or more as a travel mode?
- If you do, how often did you walk outside for 15 minutes or more in the past 30 days?

After excluding respondents who skipped these two questions and individuals who refused to answer key demographic questions, 3,297 individuals from 3,297 households (one individual in each household) were retained in the dataset. The distribution of individual characteristics and the outcome variable are presented in Table 3, and the spatial distribution of respondents is shown in Figure 2. Although respondents are primarily concentrated in the downtown areas of Seattle and Bellevue, the sample also captures significant geographic diversity, including various suburban and rural areas.

Table 3 Socio-Economic and Demographic Characteristics and Outcome Variables of the Sample (N=3,297)

Variable	%	Variable	%
Race		***Vehicle ownership***	
White only	62.5	0	16.7
Asian	15.2	1	45.0
Hispanic	3.0	2	28.1
Others	19.3	3+	10.2
Age		***Employment***	
18-24	4.5	Full time	52.4
25-34	24.2	Self-employed	6.0
35-44	21.2	Part time	7.6
45-54	13.6	Not working now	34.0
55-64	14.4	***Presence of children***	
65-74	15.0	Live with a child	16.8
75+	7.1	No child	83.2
Gender		***Number of adults***	
Woman	48.5	One	47.0
Man	43.7	Two	47.0
Others	7.8	Three or more	6.0
Education		***Income***	
Some college or lower	33.6	Under 25k	10.9
Bachelor's degree	36.0	25-50k	13.6
Graduate degree	30.4	50-75k	14.5
Disability affects ability to travel		75-100k	12.0
Yes	9.1	100-200k	31.3
No	90.9	200k+	17.7

Outcome Variable – Walking frequency in last 30 days (Outside for 15-minutes or more)		
	Count	Percentage
Never in last 30 days	1191	36.1
Two or less days a week	435	13.2
3-4 days a week	547	16.6
5 days a week	332	10.1
6-7 days a week	792	24.0

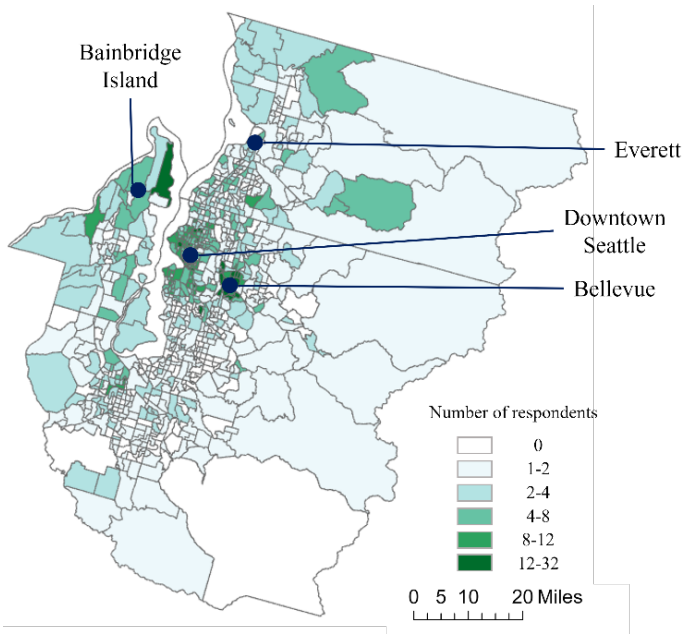

Figure 2 Spatial distribution of respondents (N = 3,297)

In addition to the individual characteristics listed in Table 3, we also considered residential neighborhood built environmental (BE) factors as determinants of walking frequency. We drew four BE variables from the Smart Location Database 3.0 (EPA, 2021). These four variables are part of the National Walkability Index (NWI) at the block group level, each computed as an index ranging from 1 to 20 based on their nationwide ranking.[6] These indices represent: (1) the household-employment mix (measured as an index of the balance between the number of occupied housing units and the number of jobs), (2) the intersection density (measured using only pedestrian-oriented intersections, that is, excluding intersections involving high-speed roads, highway ramps, and parking lot roads), (3) proximity to transit stops (based on the walking distance from the block group centroid to the nearest transit stop), and (4) the employment sector mix (measured as an index of the diversity of jobs across eight employment categories, such as retail, office, and public administration). Finally, for residential neighborhood sociodemographic characteristics, population density and median household income at the census tract level were obtained from the 2020 national census (U.S. Census Bureau, 2021) and appended to each respondent's residential location.

The overall set of explanatory variables includes individual characteristics and residential spatial unit variables (residential neighborhood-level built and social environment factors). In the analysis, only residential spatial unit (RSU) variables were assumed to have spatial spillover effects.[7] Also, for all exogenous variables, we considered alternative functional forms, including

[6] Because the finest level of residential neighborhood geography accessible from the survey is limited to the census tract (due to privacy considerations), block group-level data from the Smart Location Database were aggregated to the tract level. We computed area-weighted averages across block groups within each census tract to obtain tract-level measures, which were then appended to each respondent's residential location at the census tract level.

[7] In other words, we assumed that individual sociodemographic factors and household characteristics contribute only to direct effects and heteroscedasticity, without serving as sources of spatial spillover effects in the model. While this specification is not mandatory in the LSLX-LSAE model, we adopted it owing to the nature of the survey data used

linear and logarithmic continuous forms as well as different non-linear forms (such as a series of dummy variables characterizing different thresholds and piece-wise linear forms). When exogenous variables were already in categorical form (as is the case with individual-level demographic variables), dummy variables were considered in the most disaggregate form and then aggregated as appropriate based on the number of observations within each disaggregate category and statistical tests for non-distinct effects across categories. In the case of RSU variables, the final specification comprised simple binary dummy variables, as presented in Table 4.

Table 4 Binary Dummy Variables of RSU Variables

RSU variable	Threshold in Modeling (Based on RSU-level distribution)				Sample Distribution			
	Split	Threshold value	Low	High	Min	Max	% in Low	% in High
Household-employment mix	Bottom 50% Top 50%	10.6	≤10.6	>10.6	1.0	20.0	52.6	47.4
Intersection density	Bottom 50% Top 50%	14.9	≤14.9	>14.9	1.0	20.0	45.3	54.7
Proximity to transit stops	Bottom 75% Top 25%	18.3	≤18.3	>18.3	1.0	20.0	70.6	29.4
Employment sector mix	Bottom 25% Top 75%	9.0	≤9.0	>9.0	2.7	20.0	23.1	76.9
Population density	Bottom 50% Top 50%	11,000 people/sq mile	≤11,000	>11,000	4.6	89100.4	26.4	73.6
Median household income	Bottom 50% Top 50%	$93,400	≤93,400	>93,400	27,885	250,000	69.8	30.2

The distance between individuals is defined as the Euclidean distance between the centroids of the census tracts in which they reside. Under this definition, the distance between individuals residing in the same tract is zero. This limitation arises from the coarse resolution of locational data and can cause estimation inconsistency in spatial effects. Regarding the spatial spillover effects, zero distance between individuals q and q' in the same census tract implies that the BE characteristics shared by both individuals would affect each person's behavior through both direct and spillover from the other person. To address this issue, we set spillover weights to zero for individuals residing in the same census tract ($w_{qq'} = 0$) regardless of the decay parameter α. This enables us to differentiate between the effects of characteristics of an individual's residential tract and the local spillover effects from average characteristics in neighboring tracts. For spatial error correlations, a zero distance between individuals q and q' in the same tract implies that their error correlation would be set to 1 (as $d_{qq'} = 0$, $\exp(-\rho \times d_{qq'}) = 1$ regardless of the parameter ρ).

in this empirical analysis. Specifically, because the survey consists of a small number of respondents within each spatial unit, the sociodemographic characteristics of respondents in neighboring areas are unlikely to be representative of true population characteristics of those areas. For example, even if a census tract has a high average income, only low-income individuals may have been drawn from that tract in the survey sample. Using individual respondent characteristics from neighboring units as spillover variables would therefore fail to capture the actual neighborhood effects that influence walking behavior. Instead, we used aggregated spatial unit characteristics (e.g., census tract median household income) as sources of spatial spillover effects, as these better represent the true area-level characteristics that capture spatial influences.

However, because error correlation stems from unobserved factors in the model, neither assuming zero correlation between individuals in the same tract (as we do for spatial spillover effects) nor fixing it at unity (by adopting zero distance and fixing the correlation at one) is theoretically justified. Consequently, for the error correlation matrix specifically, we approximate distances for the pairs of individuals residing in the same census tract as the average distance between two randomly placed points within a square having the same area as each tract (Santaló, 2004). With this distance specification, the average distance between individuals in our sample is 18.12 mi, excluding pairs of individuals in the same tract, while the minimum distance is 0.15 mi and the maximum distance is 104.98 mi. Among the total of 5,433,456 pairs of individuals, 13,280 pairs reside in the same tract. The average distance between individuals appears large considering typical walking distances. However, as shown in Figure 3, the distribution of pairwise distances is heavily skewed toward longer distances owing to a sample concentration in urbanized areas, while the mean is inflated by the long right tail arising from the large geographic region covered by the study area.

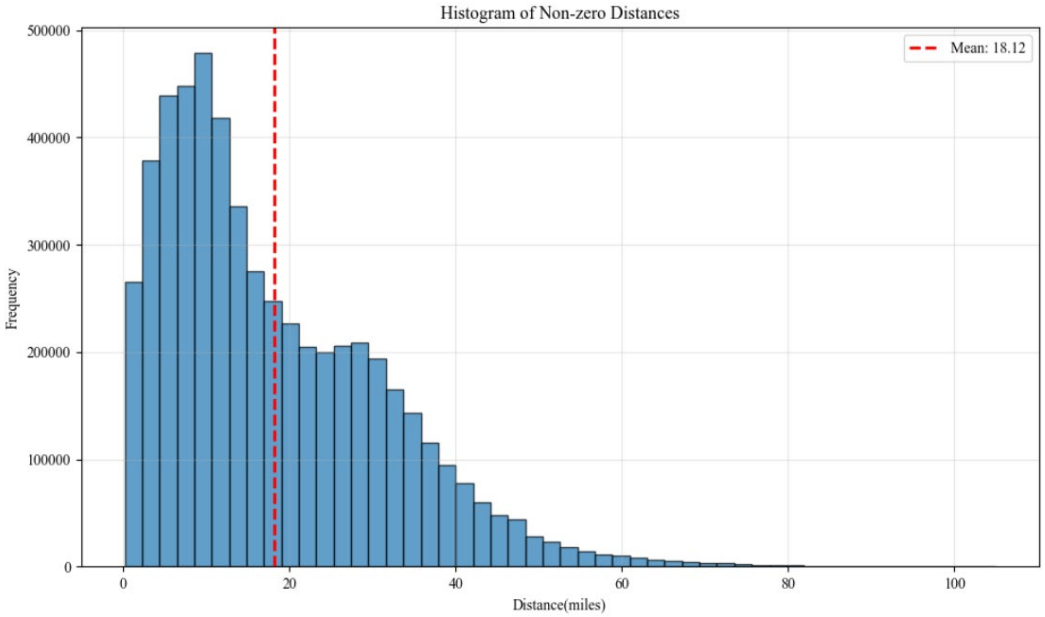

Figure 3 Distribution of distances between individual pairs

We construct the neighborhood sets for local spatial spillover effects (C_q) and local spatial error correlation (D_q) using the same logic and cutoffs as in the simulation. We impose spatial spillover effects only for individual pairs with $w_{qq'} > 0.0001$, thereby excluding pairs with negligible spatial weights. This threshold choice is particularly important given the row-normalization in the LSLX-LSAE model. Because spillover effects are computed as distance-weighted averages across neighbors, lowering the cutoff would produce counterintuitive results: isolated census tracts far from most other units would be estimated to receive substantial spillover effects from distant neighbors simply because those distant units represent a larger share of their (sparse) neighborhoods. Therefore, including more spatially distant pairs does not necessarily improve the model's ability to capture actual spatial processes. For spatial error correlation, we adopt a more lenient threshold of $\omega_{qq'} > 10^{-10}$. Because the error correlation matrix does not

undergo row-normalization, the problematic behavior arising from the cutoffs in spillover effects does not occur, allowing us to include more observation pairs in D_q.

4.2 Estimation Results

4.2.1 Effects of Individual Characteristics

The direct effects of individual characteristics presented in Table 5 show that older individuals (aged 65 or older) have a higher walking propensity than do their younger peers, while those with any form of physical disability tend to have a lower walking propensity. While some earlier studies suggest that older individuals walk less than their younger peers (see Adjaye-Gbewonyo and Briones, 2024 and Hwang et al., 2023), our study may be controlling for the negative age effects found in earlier studies in part through the physical disability variable (in our sample, 15.6% of those aged 65 years or more reported having a physical disability relative to only 5.6% in the younger age group). Besides, older adults may experience reduced time constraints following retirement, which alleviates one of the major barriers to walking participation (Clark and Scott, 2016). Race effects in Table 5 reveal that non-Hispanic white individuals have a higher walk propensity compared to individual of other racial and ethnic groups (perhaps because of their higher propensity to participate in leisure walking; see Adjaye-Gbewonyo and Briones, 2024), while individuals of Asian origin walk the least (technically speaking, a negative effect in an ordered-response model indicates a reduced walk propensity, increasing walk frequency in the lowest frequency category and decreasing walk frequency in the highest frequency category, with the effects for the intermediate frequency category being ambiguous; but for ease in presentation, we will use the verbiage of "walk less" and "walk more" to refer to the more technically rigorous "lower walk propensity" and "higher walk propensity", respectively). Other effects of individual characteristics indicate that walking propensity decreases with higher auto-ownership levels (potentially due to higher car dependence and accessibility; see Sehatzadeh et al., 2011), and in the presence of children in the household (perhaps due to time poverty of those with children; see Bernardo et al., 2015). In contrast, individuals with a graduate degree and high household income (>200,000 annual income) tend to walk more than those with less than a graduate degree and lower household income (≤200,000 annual income), respectively, findings that align with those from previous studies (Berrigan and Troiano, 2002; Cerin et al., 2009). These socioeconomically advantaged groups possess greater social and economic resources, which may enhance their awareness of the physical and mental health benefits of walking and provide greater flexibility to allocate time for walking activities.

Two groups exhibit lower within-group variance in walking propensity: women and individuals with a graduate degree. For women, this reduced variance may reflect more uniform walking behavior due to social companionship effects (that is, women are more likely to pursue physical activity with others, which makes walking behavior more consistent and regular; see Kritz et al., 2021 and Smith et al., 2023). For highly educated individuals, the reduced variance in walking propensity may partially stem from more uniform health literacy levels that translate to more uniform walk frequency levels.

Table 5 Model Estimation Results

Explanatory Variables (base level) / Parameters and Thresholds		Estimates	
		Coef.	t-stat
Individual characteristics			
Direct effects			
Age (≤64 years)	65 or older	0.095	4.15
Disability on mobility (No)	Yes	-0.346	-13.41
Race (Other than Non-Hispanic White and Asian)	Non-Hispanic White	0.194	11.50
	Asian	-0.115	-5.92
Number of vehicles in household (Zero)	One	-0.339	-14.47
	Two or more	-0.557	-16.08
Child in household (No)	Live with a child	-0.068	-2.03
Education level (Bachelor or lower)	Graduate degree	0.098	4.51
Household income (<$200,000)	>$200,000	0.060	2.46
Heteroscedasticity effects			
Gender (Not Women)	Women	-0.102	-6.17
Education level (Bachelor or lower)	Graduate degree	-0.136	-6.81
Residential Spatial Unit (RSU) variables			
Direct effects			
Household-employment mix (Low)	High	0.040	2.61
Intersection density (Low)	High	0.232	6.78
Proximity to transit stops (Low)	High	0.217	11.69
Population density (Low)	High	0.064	3.42
Median household income (Low)	High	0.044	1.99
Spillover effects			
Household-employment mix (Low)	High	0.094	2.02
Median household income (Low)	High	-0.089	-2.87
Employment sector mix (Low)	High	0.354	3.00
Model parameters			
YJ Parameter (λ) (t-stat computed with respect to the value of 1.000)		0.809	4.24
Spatial spillover decay parameter (α)		2.227	2.01
Spatial error correlation parameter (ρ)		4.244	11.31
Spillover effect threshold (miles)		2.215	na
Correlation threshold (miles)		4.547	na
Thresholds			
1|2		-0.184	-2.32
2|3		0.166	1.91
3|4		0.623	4.46
4|5		0.930	6.60

4.2.2 Effects of Residential Spatial Unit (RSU) Variables

The direct effects of the RSU variables in Table 5 show that individuals living in neighborhoods with high household-employment mix, high intersection density, good public transit proximity, high population density, and high median incomes walk more than their peers in other neighborhoods. These findings are consistent with previous research (Ewing and Cervero, 2010, Fonseca et al., 2022, and Tanishita and Van Wee, 2017). But unlike most earlier research on walking behavior, we also investigate spatial spillover effects. Our results reveal that the spatial spillover of high household-employment mix reinforces the direct effect of this variable, while that of median tract household income is the opposite of its direct effect. The latter result reflects the multifaceted influence of neighborhood income levels on the walking environment. While high median neighborhood income may increase walking among those living in that neighborhood (due

to social capital effects, better walking infrastructure, stronger social cohesion, and higher safety perceptions; see Thornton et al., 2016), the high median income of neighboring residential tracts may decrease walking among those living in a specific tract because high-income neighborhoods often coincide with more auto-oriented regional development patterns, larger residential lots, and segregated land uses (see, King and Clarke, 2015). That is, even if a focal tract itself is reasonably walkable, being situated within a higher-income belt may expose residents to a regional mobility culture that prioritizes automobile travel over routine walking. In addition, high-income neighboring areas may influence retail structure, land values, and service distribution at a regional scale (such as, for example, the upscaling of retail outlets and more car parking provided for customers), potentially reducing the prevalence of small, low-cost, walk-accessible, and neighborhood-serving destinations in the focal tract that support everyday walking. Social norms may also play a role, whereby spatial proximity to higher-income populations may correspond to different lifestyle patterns that normalize car-based lifestyles and reduce incidental walking. In short, higher median income has a positive local environmental quality effect but a negative regional structural or normative effect, and the SLX framework is precisely what allows these multi-scale influences to be disentangled.

Finally, our results also indicate that, while employment sector mix does not have a direct effect, it does have a positive spillover effect. Employment entropy across job categories serves as a proxy for non-residential land use diversity, specifically the mix among commercial, service, industrial, and other non-residential land use types (EPA, 2021). A tract surrounded by low-employment sector mix areas is likely situated in a broader area where employment is concentrated in a narrow range of sectors (such as an industrial district or suburban office park), limiting the variety of activities and services accessible on foot to nearby residents. Conversely, a tract neighboring high-employment-mix areas is likely embedded in a broader urban context where diverse employment categories correspond to a wider array of walkable destinations. The absence of a direct effect coupled with a significant spillover effect suggests that the walkability benefits of non-residential land use diversity operate at a spatial scale larger than the census tract, with the overall development pattern of the surrounding area governing residents' walking propensity.

4.2.3 Error Distribution and Spatial Weight Parameters

With respect to the parameters governing the error structure, the YJ parameter is estimated as 0.809 and is statistically significantly different from the value of one, indicating that the error term in the walking ordered-response equation follows a rightward skew, rejecting the commonly adopted assumption of symmetric normality. This is consistent with the finding in Bhat (2024). In terms of spatial weight parameters, both the spatial spillover and spatial error correlation parameters are statistically significantly different from zero (the magnitudes of the two parameters are not directly comparable, because the spatial spillover parameter is embedded within a row-normalized weight matrix, while the spatial error correlation directly governs the decay of inter-individual error dependence). The estimated range of spatial spillover effects is 2.215 miles, while that for spatial correlation is more expansive at 4.547 mi. This difference in spatial reach is theoretically consistent with the distinct roles of the two spatial processes in the model. Spillover effects operate through observed built-environment characteristics, and therefore reflect structured, behaviorally mediated dependencies that fade relatively quickly with distance. In contrast, spatial error correlation captures residual dependence arising from unobserved regional factors that are not tied to a single environmental attribute and may operate at broader spatial scales. Such factors can include regional planning regimes, shared infrastructure systems, macro-level economic

conditions, or geographically persistent cultural patterns that influence mobility behavior beyond the immediate local environment. Because these influences are not confined to a specific tract-level characteristic, it is reasonable that their spatial correlation range exceeds that of observed spillover effects.

Although the estimated spatial threshold distances above exceed the conventional half-mile threshold often associated with short utilitarian walking trips, they remain plausible in the current empirical context that includes recreational and leisure walking, which frequently involves durations exceeding 30–60 minutes and therefore distances well above two miles (Paul et al., 2015).

4.2.4 Measures of Data Fit

The main purpose of the methodology proposed here is to accommodate local spatial exogenous interaction effects along with local spatial error correlation and spatial heteroscedasticity, while also considering a flexible, asymmetric specification for the error terms. Therefore, we provide the performance metrics of the proposed LSLX-LSAE model, but also compare its performance with three restrictive versions: (1) the basic SLX model with global spillover effects (with no spatial correlation dependencies) and the assumption of a normally distributed error (no thresholding of spatial spillover, $\mathbf{\Omega}^*$ is an identity matrix or $\rho \to \infty$, and $\lambda = 1$), (2) the SLX-AE (for SLX with asymmetric error) model without spatial error correlation but with the assumption of skewed error (no thresholding of spatial spillover, $\mathbf{\Omega}^*$ is an identity matrix and λ is free to be estimated), (3) the SLX-SSE (for SLX with Spatial Symmetric Error, both spatial processes being global, and $\mathbf{\Omega}^*$ not restricted to an identity matrix but $\lambda = 1$), and (4) the LSLX-LSSE (the counterpart of the SLX-SSE model with both spatial processes being local).

The restricted models are tested with the proposed model using both disaggregate measures as well as aggregate measures. At the disaggregate level, we use the adjusted composite likelihood ratio test (ADCLRT) statistic (Pace et al., 2011; Bhat, 2014) as well as the predictive log-likelihood (PLL) value and average probability of correct prediction (APCP). The first ADCLRT statistic adjusts the composite log-likelihood difference to follow a chi-square distribution by incorporating information from the $\boldsymbol{J}$ and $\boldsymbol{H}$ matrices obtained during the estimation procedure (see Pace et al., 2011 and Bhat, 2014). Since the SLX and SLX-AE models are estimated using full likelihood, the ADCLRT is not applicable to these models. Additionally, because the SLX-SSE model assumes a global spatial process and therefore utilizes all individual pairs in the CML calculation, comparing its composite log-likelihood value against the localized models is not meaningful. For the two localized models (LSLX-LSSE and LSLX-LSAE), whose threshold distances are similar, we use the composite log-likelihood computed with the shorter threshold distance to ensure a common basis for the ADCLRT comparison. The second predictive log-likelihood (PLL) value uses the coefficient estimates from the CML estimation to compute the probability of each individual's observed walk frequency ordinal category, regardless of the estimation procedure used. The procedure to get PLL is similar to that in the simulation. We generate the random error vector according to the correlation and error structure based on the parameters for 5,000 times to estimate the probability of each individual to fall into the same level as the sample. Specifically, for each individual, we draw 5,000 realizations of the multivariate error vector from the estimated correlation and covariance structure, and compute the simulated probability of the observed ordinal outcome as the proportion of draws that result in that outcome. The PLL is then obtained as the sum of the log of these simulated probabilities across all individuals in the sample. The third statistic of average probability of correct prediction (APCP) is derived in a straightforward manner

once the probabilities of the observed walk frequency are predicted based on the above procedure. At the aggregate level, we computed the probabilities for each individual and each ordered walk frequency category to obtain the predicted percentages in each frequency bin, and computed a weighted absolute percentage error (WAPE) measure (the weights being based on the true percentages in each walk frequency category).

Table 6 presents the goodness of fit measures. The results indicate that the proposed LSLX-LSAE model provides a much better statistical fit based on the ADCLRT statistic at any reasonable level of significance (in fact, almost at the 10^{-10} level) relative to the LSLX-LSSE model. Further, the PLL value for the proposed LSLX-LSAE model is the highest of all the models, as is the APCP. At the aggregate level, the proposed model most closely replicates the true split percentages of the walk frequency ordinal categories, and also exhibits the lowest WAPE measure. Overall, the proposed model clearly outperforms all the other models.

Table 6 Goodness of Fit

		SLX	SLX-AE	SLX-SSE	LSLX-LSSE	LSLX-LSAE
Composite log-likelihood at convergence		na	na	-15,502,300	-2,368,610	-2,045,130
Number of pairs in composite marginal likelihood		na	na	5,433,456	803,846	695,729
ADCLRT statistic (LSLX-LSAE versus LSLX-LSSE)		na	na	na	54.74	na
Predictive log-likelihood (PLL)		-4,706.70	-4,699.11	-4,702.33	-4,686.73	-4,672.77
Average probability of correct prediction (APCP)		0.2802	0.2827	0.2887	0.2889	0.2892
Walking Frequency	Actual Percentage	Predicted Percentage				
Never in last 30 days	36.12	32.25%	32.85%	34.23%	36.52%	36.25%
Two or less days a week	13.20	13.63%	14.17%	13.38%	13.27%	13.14%
3-4 days a week	16.59	17.70%	17.94%	17.54%	16.55%	16.62%
5 days a week	10.07	10.27%	10.24%	9.88%	9.88%	10.15%
6-7 days a week	24.02	26.15%	24.80%	24.97%	23.78%	23.95%
WAPE		7.74%	6.54%	4.16%	0.84%	0.37%

4.2.5 Average Treatment Effects (ATEs)

Walking frequency is captured in ordinal form, so the coefficients in Table 5 indicate only the direction of effects on reaching the highest and lowest outcome levels. These coefficients cannot be directly interpreted in terms of magnitude. Furthermore, as the model includes three pathways of the effect of explanatory variables (direct effects, spatial spillover effects, and heteroscedasticity effects), comparing the contribution of each effect would be useful for deriving policy implications. An average treatment effect (ATE) analysis is undertaken to measure the effect of each explanatory variable on the actual outcome variable. In this study, we used the expected value of walking days per week as the indicator of the distribution of walking frequency, which is calculated using the proportion and the mid-point value of each level (0 for never, 1.5 for 1-2 days per week, 3.5 for 3-

4 days per week, 5 for 5 days per week, and 6.5 for 6-7 days per week). ATE is quantified as the percentage change in expected walking days per week between two statuses, i.e., when all respondents are at the base level (BL) and when they are at the treatment level (TL). To examine how the estimated treatment effects differ across model specifications, ATE results from the proposed model are compared against those from the same four alternative models as in the preceding section (SLX, SLX-AE, SLX-SSE, and LSLX-LSSE). As the primary interest lies in how spatial structure assumptions affect estimated effects, and also for presentation conciseness, our comparison focuses on RSU characteristics for which statistically significant spatial spillover effects are identified in the LSLX-LSAE model.

Table 7 presents the ATE analysis results, where each row shows the percentage change in expected walking days per week when all respondents are shifted from a low (base) to a high (treatment) level of the corresponding RSU variable. For household-employment mix, while the SLX, SLX-AE, and SLX-SSE models predict an increase in expected walking days of 5.42%, 5.05%, and 5.35%, respectively, the two models incorporating local spatial interactions (LSLX-LSSE and LSLX-LSAE) predict larger increases of 10.98% and 10.34%. Employment sector mix shows a comparable pattern, with global spatial models consistently predicting smaller ATE values than the local spatial models. These results suggest that the assumption of global versus local spatial processes has a considerable influence on the estimated magnitude of built environment effects. While the global spatial models predict a positive ATE for median household income (reflecting a positive direct effect that outweighs the negative spatial spillover), the localized models predict a negative ATE (reflecting a negative spatial spillover that dominates the positive direct effect). Comparing across the three global spatial models (SLX, SLX-AE, and SLX-SSE) as well as between the two local spatial models (LSLX-LSSE and LSLX-LSAE), the introduction of error distribution flexibility produces moderate shifts in ATE estimates, while the shift from global to local spatial processes yields substantially larger differences.

Table 7 Average Treatment Effect of RSU Variables with Spatial Spillover Effects

RSU Variable (BL: Low, TL: High)	ATE(%)				
	SLX	SLX-AE	SLX-SSE	LSLX-LSSE	LSLX-LSAE
Household-employment mix	5.42	5.05	5.35	10.98	10.34
Median household income	12.36	8.82	12.68	-2.77	-4.75
Employment sector mix	22.90	21.74	22.96	33.02	31.95

4.3 Policy Implications

One key contribution of applying the LSLX-LSAE model to walking behavior is that it enables us to identify the different spatial scales at which built and social environment factors of RSU influence individuals' behavior, providing insights for policy design, walkability planning, and how walkability is measured.

From a policy design standpoint, several observations may be drawn from our results. First, certain variables exhibit direct effects on walking frequency within a census tract but no spillover effects from neighboring tracts: intersection density, transit proximity and population density. Accordingly, policies targeting these features, such as adding pedestrian connections within existing street networks, positioning transit stops for better accessibility, and increasing residential density with infill development, should be distributed equitably across neighborhoods rather than concentrated in a few areas, with priority given to underserved communities that currently lack

adequate infrastructure facilities. Second, household-employment mix and employment sector mix, which serve as proxies for land-use mix measures, differ in the spatial scale at which they influence walking behavior. Household-employment mix exhibits both a direct effect and a spillover effect, indicating that residents' walking behavior is shaped by the integration of residential and non-residential uses both within their own tract and in surrounding tracts. Employment sector mix, by contrast, exhibits only a spillover effect with no direct effect, indicating that residents respond to the diversity of employment in surrounding tracts rather than within their own. This distinction has practical implications for land-use policy: policies that integrate residential and non-residential uses, such as mixed-use zoning that co-locates housing with commercial and service activities, can generate walkability benefits regardless of the scale at which they are implemented. Moreover, coordinated implementation across adjacent neighborhoods can amplify these benefits through cumulative spillover effects. Policies aimed at enhancing employment sector diversity, by contrast, should be designed at the corridor or regional level, as their walkability benefits are realized only through the characteristics of surrounding areas. Such policies may include encouraging a varied mix of commercial, service, and institutional activities rather than concentrating single-sector employment. Third, the opposing directions of the direct and spillover effects of median tract income call for context-sensitive walkability interventions. In neighborhoods situated within higher-income regional belts, where auto-oriented development patterns may undermine walking, counteracting these tendencies through reducing minimum parking requirements, incentivizing small-scale neighborhood-serving retail, and preserving walk-accessible destinations becomes particularly important. Conversely, in lower-income neighborhoods where the local environment may be less conducive to walking but the surrounding regional context does not impose these auto-oriented pressures, investments focusing on local walkability such as sidewalk quality, lighting, and safety may yield comparatively larger returns.

Beyond policy design implications, the differential spatial scales of walkability components also give rise to broader structural considerations for walkability planning. The combination of the preceding findings reveals the potential for a "walkability trap." When a tract is surrounded by neighborhoods with high land-use mix but itself lacks adequate local walkability conditions (such as poorly connected pedestrian networks, low accessibility of transit, and low population density), its residents may be structurally unable to benefit from the favorable regional land-use environment. This implies that local pedestrian infrastructure serves as a necessary precondition for residents to realize the walkability benefits generated by the land-use characteristics of surrounding areas. Policymakers should therefore ensure that basic pedestrian network quality is established across all neighborhoods before expecting regional land-use mix strategies to yield tangible walkability benefits.

Finally, our findings have implications for how walkability is measured. Conventional walkability indices, such as the National Walkability Index, measure all components within each spatial unit and combine them using predetermined weights (EPA, 2021), without accounting for the influence that surrounding areas' characteristics may exert on residents' walking behavior. To our knowledge, no existing walkability index incorporates the relevant characteristics of neighboring areas. Our results indicate that residents respond not only to the conditions within their own tract but also to certain characteristics of surrounding tracts, and that the extent of this cross-boundary influence varies across components. This suggests that walkability indices incorporating the relevant characteristics of neighboring areas may more accurately represent the walking environment as experienced by residents.

5. CONCLUSION

The current study develops and applies the Local Spatial Lag on X–Local Spatial Asymmetric Error (LSLX-LSAE) model, a unified spatial econometric framework that captures localized spatial spillover effects of observed covariates and localized spatial correlation in unobserved error structures, without autoregressive specifications, while accommodating heteroscedasticity and flexible skewed error distributions. Simulation experiments demonstrated satisfactory parameter recovery of the LSLX-LSAE estimator. We then applied the model to the analysis of pedestrian walking frequency using data from the 2023 Puget Sound Regional Travel Study, with residential spatial unit characteristics drawn from the Smart Location Database 3.0 and the 2020 Decennial Census. The empirical analysis revealed significant heteroscedasticity across social groups, asymmetric error distributions, and that the effects of residential spatial unit characteristics differ in spatial extent, with some confined to the immediate census tract and others extending to neighboring areas. Model comparison against four alternative specifications (SLX, SLX-AE, SLX-SSE, and LSLX-LSSE) confirmed the statistical superiority of the proposed model in both disaggregate and aggregate data fit measures. Importantly, the average treatment effect analysis demonstrated that the assumption of global versus local spatial processes has a considerable influence on estimated effect magnitudes, in some cases altering not only the size but also the sign of treatment effects, underscoring the importance of correctly specifying the spatial structure when drawing policy conclusions. These findings suggest that effective walking promotion requires spatially differentiated strategies, with localized infrastructure investments distributed equitably across neighborhoods, while land-use policies are designed at the regional level to maximize spatial spillover benefits.

The research in this study may be extended in several directions. First, implementing variable-specific spatial spillover decay parameters would accommodate differences in both the spatial reach and magnitude of spillover effects across variables, as identified in our empirical analysis. Second, the current LSLX-LSAE framework incorporates spillover effects only from exogenous variables of sampled observations, precluding the estimation of influences from unsampled spatial units. Incorporating exogenous variables from all spatial units throughout the study area, regardless of sampling, would enhance the model's applicability to individual-level behavioral analyses. We hope that the proposed LSLX-LSAE framework, by offering a computationally tractable approach to modeling localized spatial processes with flexible error structures, will encourage further methodological and empirical investigations of spatial dependence in travel behavior analysis.

ACKNOWLEDGMENTS

This research was partially supported by the National Center for Understanding Future Travel Behavior and Demand (TBD), which is a University Transportation Center sponsored by the U.S. Department of Transportation under grant numbers 69A3552344815 and 69A3552348320. The authors are grateful to Lisa Macias for help in formatting this document.

REFERENCES

Adjaye-Gbewonyo, D., Briones, E. M. (2024). Walking for Leisure and Transportation Among Adults: United States, 2022 (NCHS Data Brief No. 504). Centers for Disease Control and Prevention. https://doi.org/10.15620/cdc/158783

Agiakloglou, C., Tsimpanos, A. (2023). Evaluating the performance of AIC and BIC for selecting spatial econometric models. *Journal of Spatial Econometrics*, *4*(1), 2.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models* (Vol. 4). Springer Science & Business Media.

Anselin, L. (2021). Spatial models in econometric research. In *Oxford Research Encyclopedia of Economics and Finance*.

Arellano-Valle, R. B., Azzalini, A. (2006). On the unification of families of skew-normal distributions. *Scandinavian Journal of Statistics*, *33*(3), 561–574.

Bernardo, C., Paleti, R., Hoklas, M., Bhat, C. (2015). An empirical investigation into the time-use and activity patterns of dual-earner couples with and without young children. *Transportation Research Part A: Policy and Practice*, *76*, 71–91. https://doi.org/10.1016/j.tra.2014.12.006

Berrigan, D., Troiano, R. P. (2002). The association between urban form and physical activity in U.S. adults. *American Journal of Preventive Medicine*, *23*(2), Article 2. https://doi.org/10.1016/S0749-3797(02)00476-2

Besser, L. M., Chang, L.-C., Kluttz, J. (2021). Individual and neighborhood characteristics associated with neighborhood walking among US older adults. *Preventive Medicine Reports*, *21*, 101291. https://doi.org/10.1016/j.pmedr.2020.101291

Bhat, C. R. (2011). The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. *Transportation Research Part B: Methodological*, *45*(7), 923–939. https://doi.org/10.1016/j.trb.2011.04.005

Bhat, C. R. (2014). The Composite Marginal Likelihood (CML) inference approach with applications to discrete and mixed dependent variable models. *Foundations and Trends® in Econometrics*, *7*(1), 1–117. https://doi.org/10.1561/0800000022

Bhat, C. R. (2024). Transformation-based flexible error structures for choice modeling. *Journal of Choice Modelling*, *53*, 100522. https://doi.org/10.1016/j.jocm.2024.100522

Bhat, C. R., Astroza, S., Hamdi, A. S. (2017). A spatial generalized ordered-response model with skew normal kernel error terms with an application to bicycling frequency. *Transportation Research Part B: Methodological*, *95*, 126–148. https://doi.org/10.1016/j.trb.2016.10.014

Bhat, C. R., Dubey, S. K., Alam, M. J. B., Khushefati, W. H. (2015). A new spatial multiple discrete-continuous modeling approach to land use change analysis. *Journal of Regional Science*, *55*(5), 801–841. https://doi.org/10.1111/jors.12201

Bhat, C. R., Pinjari, A. R., Dubey, S. K., Hamdi, A. S. (2016). On accommodating spatial interactions in a Generalized Heterogeneous Data Model (GHDM) of mixed types of dependent variables. *Transportation Research Part B: Methodological*, *94*, 240–263. https://doi.org/10.1016/j.trb.2016.09.002

Bhat, C. R., Sener, I. N. (2009). A copula-based closed-form binary logit choice model for accommodating spatial correlation across observational units. *Journal of Geographical Systems*, *11*(3), 243–272. https://doi.org/10.1007/s10109-009-0077-9

Bhat, C. R., Sidharthan, R. (2012). A new approach to specify and estimate non-normally mixed multinomial probit models. *Transportation Research Part B: Methodological*, *46*(7), 817–833.

Caffo, B., An, M.-W., Rohde, C. (2007). Flexible random intercept models for binary outcomes using mixtures of normals. *Computational Statistics & Data Analysis*, *51*(11), 5220–5235. https://doi.org/10.1016/j.csda.2006.09.031

Calabrese, R., Degl'Innocenti, M., Osmetti, S. A. (2017). The effectiveness of TARP-CPP on the US banking industry: A new copula-based approach. *European Journal of Operational Research*, *256*(3), 1029–1037. https://doi.org/10.1016/j.ejor.2016.07.046

Cerin, E., Leslie, E., Owen, N. (2009). Explaining socio-economic status differences in walking for transport: An ecological analysis of individual, social and environmental factors. *Social Science & Medicine*, *68*(6), Article 6. https://doi.org/10.1016/j.socscimed.2009.01.008

Clark, A., Scott, D. (2016). Barriers to walking: An investigation of adults in Hamilton (Ontario, Canada). *International Journal of Environmental Research and Public Health*, *13*(2), 179. https://doi.org/10.3390/ijerph13020179

Cliff, A. D., Ord, J. K. (1973). *Spatial Autocorrelation*. Pion.

Cressie, N. (1993). Aggregation in geostatistical problems. In *Geostatistics Tróia'92: Volume 1* (pp. 25–36). Springer.

Elhorst, J. P. (2010). Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis*, *5*(1), 9–28.

Elhorst, J. P. (2024). Raising the bar in spatial economic analysis: Two laws of spatial economic modelling. *Spatial Economic Analysis*, *19*(2), 115–132. https://doi.org/10.1080/17421772.2024.2334845

EPA. (2021). *Smart Location Database Version 3.0* [Dataset]. https://www.epa.gov/smartgrowth/smart-location-mapping

Ewing, R., Cervero, R. (2010). Travel and the built environment: A meta-analysis. *Journal of the American Planning Association*, *76*(3), Article 3. https://doi.org/10.1080/01944361003766766

Farzammehr, M. A., Zadkarami, M. R., McLachlan, G. J. (2021). Skew-normal generalized spatial panel data model. *Communications in Statistics - Simulation and Computation*, *50*(11), 3286–3314. https://doi.org/10.1080/03610918.2019.1622718

Fonseca, F., Ribeiro, P. J. G., Conticelli, E., Jabbari, M., Papageorgiou, G., Tondelli, S., Ramos, R. A. R. (2022). Built environment attributes and their influence on walkability. *International Journal of Sustainable Transportation*, *16*(7), 660–679. https://doi.org/10.1080/15568318.2021.1914793

Gallaugher, M. P. B., McNicholas, P. D., Melnykov, V., Zhu, X. (2020). *Skewed Distributions or Transformations? Modelling Skewness for a Cluster Analysis* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2011.09152

Gibbons, S., Overman, H. G. (2012). Mostly pointless spatial econometrics? *Journal of Regional Science*, *52*(2), 172–191.

Giubergia, D., Haddad, A. J., Piras, F., Bhat, C. R., Meloni, I. (2025). Modeling spatial and social interdependency effects on commuting mode choice. *Transportation Research Part A: Policy and Practice*, *196*, 104474.

Gómez-Déniz, E., Arnold, B. C., Sarabia, J. M., Gómez, H. W. (2021). Properties and applications of a new family of skew distributions. *Mathematics*, *9*(1), 87.

Greene, W. H. (2000). *Econometric Analysis*, 4th edition. International Edition, New Jersey: Prentice Hall, 201–215.

Halleck Vega, S., Elhorst, J. P. (2013). *On spatial econometric models, spillover effects, and W*. 53rd ERSA Congress, Palermo, Italy.

Halleck Vega, S., Elhorst, J. P. (2015). The SLX model. *Journal of Regional Science*, *55*(3), 339–363. https://doi.org/10.1111/jors.12188

Heagerty, P. J., Lumley, T. (2000). Window subsampling of estimating functions with application to regression models. *Journal of the American Statistical Association*, *95*(449), 197–211. https://doi.org/10.1080/01621459.2000.10473914

Hwang, H., Haddad, A., Batur, I., Saxena, S., Pendyala, R. M., Bhat, C. R. (2023). An analysis of walking frequency before and after the pandemic. *Technical Paper, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin*.

Jadhav, A., Dhaulakhandi, D., Shandilya, S. K., Malviya, L., Mewada, A. (2023). Data transformation: A preprocessing stage in machine learning regression problems. In *Artificial Intelligence Techniques in Power Systems Operations and Analysis* (pp. 183–194). Auerbach Publications.

Kelejian, H. H., Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, *17*(1), 99–121.

Kelejian, H. H., Prucha, I. R. (2007). The relative efficiencies of various predictors in spatial econometric models containing spatial lags. *Regional Science and Urban Economics*, *37*(3), 363–374. https://doi.org/10.1016/j.regsciurbeco.2006.11.005

King, K. E., Clarke, P. J. (2015). A disadvantaged advantage in walkability: Findings from socioeconomic and geographical analysis of national built environment data in the United States. *American Journal of Epidemiology*, *181*(1), 17–25. https://doi.org/10.1093/aje/kwu310

Kritz, M., Thøgersen-Ntoumani, C., Mullan, B., Stathi, A., Ntoumanis, N. (2021). "It's better together": A nested longitudinal study examining the benefits of walking regularly with peers versus primarily alone in older adults. *Journal of Aging and Physical Activity*, *29*(3), 455–465. https://doi.org/10.1123/japa.2020-0091

LeSage, J., Pace, R. K. (2009). *Introduction to Spatial Econometrics*. Chapman and Hall/CRC.

Maddala, G. S. (1983). *Limited-dependent and Qualitative Variables in Econometrics*. Cambridge University Press.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, *60*(3), 531. https://doi.org/10.2307/2298123

Marimuthu, S., Mani, T., Sudarsanam, T. D., George, S., Jeyaseelan, L. (2022). Preferring Box-Cox transformation, instead of log transformation to convert skewed distribution of outcomes to normal in medical research. *Clinical Epidemiology and Global Health*, *15*, 101043. https://doi.org/10.1016/j.cegh.2022.101043

McMillen, D. P. (2010). Issues in spatial data analysis. *Journal of Regional Science*, *50*(1), 119–141. https://doi.org/10.1111/j.1467-9787.2009.00656.x

McMillen, D. P. (2012). *Quantile Regression for Spatial Data*. Springer Science & Business Media.

Molenberghs, G., Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. Springer.

Mondal, A., Bhat, C. R. (2022). A spatial rank-ordered probit model with an application to travel mode choice. *Transportation Research Part B: Methodological*, *155*, 374–393. https://doi.org/10.1016/j.trb.2021.12.008

Müller, G., Czado, C. (2005). An autoregressive ordered probit model with application to high-frequency financial data. *Journal of Computational and Graphical Statistics*, *14*(2), 320–338.

Mur, J., Angulo, A. (2009). Model selection strategies in a spatial setting: Some additional results. *Regional Science and Urban Economics*, *39*(2), 200–213. https://doi.org/10.1016/j.regsciurbeco.2008.05.018

Neumayer, E., Plümper, T. (2016). Spatial spill-overs from terrorism on tourism: Western victims in Islamic destination countries. *Public Choice*, *169*(3), 195–206.

Pace, L., Salvan, A., Sartori, N. (2011). Adjusting composite likelihood ratio statistics. *Statistica Sinica*, 129–148.

Pace, R. K., Calabrese, R. (2022). Ignoring spatial and spatiotemporal dependence in the disturbances can make black swans appear grey. *The Journal of Real Estate Finance and Economics*, *65*(1), 1–21.

Pinkse, J., Slade, M. E. (2010). The future of spatial econometrics. *Journal of Regional Science*, *50*(1), 103–117. https://doi.org/10.1111/j.1467-9787.2009.00645.x

Puget Sound Regional Council. (2024). 2023 Puget Sound Regional Travel Study: Final Report [Research Report]. Puget Sound Regional Council. https://www.psrc.org/sites/default/files/2024-05/2023-Puget-Sound-Regional-Travel-Study-Final-Report.pdf

Quinio, A. E., Lam, T. (2021). Methods and biases in measuring change with self-reports. *Basic Elements of Survey Research in Education: Addressing the Problems Your Advisor Never Told You About*, 193–217.

Saelens, B. E., Handy, S. L. (2008). Built environment correlates of walking: A review. *Medicine and Science in Sports and Exercise*, *40*(7 Suppl), S550.

Santaló, L. A. (2004). *Integral Geometry and Geometric Probability* (2nd ed.). Cambridge University Press.

Schoenberg, I. J. (1938). Metric spaces and positive definite functions. *Transactions of the American Mathematical Society*, *44*(3), 522–536.

Sehatzadeh, B., Noland, R. B., Weiner, M. D. (2011). Walking frequency, cars, dogs, and the built environment. *Transportation Research Part A: Policy and Practice*, *45*(8), Article 8. https://doi.org/10.1016/j.tra.2011.06.001

Silveira Santos, L., Proença, I. (2019). The inversion of the spatial lag operator in binary choice models: Fast computation and a closed formula approximation. *Regional Science and Urban Economics*, *76*, 74–102. https://doi.org/10.1016/j.regsciurbeco.2019.01.003

Smith, G. S. E., Moyle, W., Burton, N. W. (2023). Frequency of physical activity done with a companion: Changes over seven years in adults aged 60+ living in an Australian capital city. *Journal of Aging and Health*, *35*(9), 736–748. https://doi.org/10.1177/08982643231158424

Tan, C., Kesina, M., Elhorst, J. P. (2025). Parameterizing spatial weight matrices in spatial econometric models. *Political Analysis*, *33*(1), 49–63. https://doi.org/10.1017/pan.2024.16

Tanishita, M., Van Wee, B. (2017). Impact of regional population density on walking behavior. *Transportation Planning and Technology*, *40*(6), 661–678. https://doi.org/10.1080/03081060.2017.1325137

Thigpen, C. (2019). Measurement validity of retrospective survey questions of bicycling use, attitude, and skill. *Transportation Research Part F: Traffic Psychology and Behaviour*, *60*, 453–461. https://doi.org/10.1016/j.trf.2018.11.002

Thornton, C. M., Conway, T. L., Cain, K. L., Gavand, K. A., Saelens, B. E., Frank, L. D., Geremia, C. M., Glanz, K., King, A. C., Sallis, J. F. (2016). Disparities in pedestrian streetscape environments by income and race/ethnicity. *SSM - Population Health*, *2*, 206–216. https://doi.org/10.1016/j.ssmph.2016.03.004

U.S. Census Bureau. (2021). 2020 Census. https://www.census.gov/programs-surveys/decennial-census/2020-census.html

Varin, C., Vidoni, P. (2008). Pairwise likelihood inference for general state space models. *Econometric Reviews*, *28*(1–3), 170–185. https://doi.org/10.1080/07474930802388009

Watthanacheewakul, L. (2021). Transformations for left skewed data. *Proceedings of the World Congress on Engineering*, 7–9.

Xu, X., Reid, N. (2011). On the robustness of maximum composite likelihood estimate. *Journal of Statistical Planning and Inference*, *141*(9), 3047–3054. https://doi.org/10.1016/j.jspi.2011.03.026

Yi, G. Y., Zeng, L., Cook, R. J. (2011). A robust pairwise likelihood method for incomplete longitudinal binary data arising in clusters. *Canadian Journal of Statistics*, *39*(1), 34–51. https://doi.org/10.1002/cjs.10089