A POISSON-LOGNORMAL CONDITIONAL-AUTOREGRESSIVE MODEL FOR MULTIVARIATE SPATIAL ANALYSIS OF PEDESTRIAN CRASH COUNTS ACROSS NEIGHBORHOODS

Yiyi Wang
Assistant Professor
Civil Engineering Department
Montana State University
yiyiwang.ut@gmail.com

Kara M. Kockelman
(Corresponding author)
Professor and William J. Murray Jr. Fellow
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
kkockelm@mail.utexas.edu
Phone: 512-471-0210

The following is a pre-print, the final publication can be found in *Accident Analysis and Prevention*; 60: 71-84, 2013.

ABSTRACT

This work examines the relationship between 3-year pedestrian crash counts across Census tracts in Austin, Texas, and various land use, network, and demographic attributes, such as land use balance, residents' access to commercial land uses, sidewalk density, lane-mile densities (by roadway class), and population and employment densities (by type). The model specification allows for region-specific heterogeneity, correlation across response types, and spatial autocorrelation via a Poisson-based multivariate conditional auto-regressive (CAR) framework and is estimated using Bayesian Markov chain Monte Carlo methods. Least-squares regression estimates of walk-miles traveled per zone serve as the exposure measure. Here, the Poisson-lognormal multivariate CAR model outperforms an aspatial Poisson-lognormal multivariate model and a spatial model (without cross-severity correlation), both in terms of fit and inference.

Positive spatial autocorrelation emerges across neighborhoods, as expected (due to latent heterogeneity or missing variables that trend in space, resulting in spatial clustering of crash counts). In comparison, the positive aspatial, bivariate cross correlation of severe (fatal or incapacitating) and non-severe crash rates reflects latent covariates that have impacts across severity levels but are more local in nature (like lighting conditions and local sight obstructions), along with spatially-lagged cross correlation. Results also suggest greater mixing of residences and commercial land uses is associated with greater pedestrian crash risk across different severity levels, ceteris paribus, presumably since such access produces more potential conflicts between pedestrian and vehicle movements. Interestingly, network densities show variable effects, and sidewalk provision is associated with lower severe-crash rates.

Key Words: pedestrian crashes, crash modeling, count models, multivariate conditional autoregressive models, spatial data

MOTIVATION

Pedestrian-vehicle crashes kill close to 5,000 Americans a year, accounting for over 10 percent of total roadway fatalities (NHTSA 2011). Motor vehicle data are regularly tabulated and their crashes receive significant research attention, including emphasis of data-modeling techniques (Abdel-Aty and Essam-Radwan 2000, Miaou et al. 2003, Lord 2006, Caliendo et al. 2007, Ma et al. 2008), as well as more straightforward studies (Davies et al. 2005). In contrast, relatively little analytical research has tackled the question of pedestrian-vehicle crash rates (especially area-level count data), although pedestrians arguably represent the most vulnerable of road users.

Focusing on zone-level pedestrian crash counts offers several benefits. This macro-level approach complements more focused pedestrian safety investigations, such as those emphasizing intersections (e.g., Weir et al. 2009, Naderan and Shahi 2009, Cottrill and Thakuriah 2010). Zone systems do not neglect any (reported) crashes: for example, almost two thirds of all U.S. pedestrian-related crashes and 76% of all pedestrian *fatalities* occur away from intersections (NHTSA 2011, FHWA 2007). So an intersection focus may miss over half the population of interest. Focused analyses also miss the spatial autocorrelation present in such data, due largely to missing variables (such as shoulder widths, use of planning strips, land use setting, and other variables typically uncontrolled for). Spatial models work well for zone-based data and seek to identify such patterns (Morency and Cloutier, 2006).

To this end, this paper analyzes zone-based pedestrian crash count totals across two severity levels (severe [i.e., fatal and incapacitating injury] and non-severe [i.e., incapacitating, light injury, and no injury cases]) over a 3-year period in Austin, Texas, using a multivariate conditional autoregressive (CAR) count model, which accounts for correlations across severity levels, unobserved heterogeneity (in zones), and spatial autocorrelation (from error terms). The capital of Texas, Austin is a medium-sized urban region (with a county population just over 1 million), with a fair amount of pedestrian activity, thanks to generally sunny conditions, a large college student population, and walk-friendly culture. The county's 3-year pedestrian crash counts (by severity level, from year 2007 to 2009, as reported by police) were aggregated using ArcGIS's *spatial join* function over Thiessen polygons built around each census tract's centroid, as described later. The paper begins with an overview of related research and methodological details of the multivariate CAR approach, followed by results and conclusions.

PEDESTRIAN CRASH STUDIES

Recent years offer a rising number of research studies on pedestrian safety. For example, Weir et al. (2009) studied vehicle-pedestrian injury collisions across 176 San Francisco census tracts, while controlling for local traffic volumes, shares of arterial streets with and without transit service, some land use attributes, population, employment, and residents' income levels. Their log-linear OLS results suggest that pedestrian injury/fatality counts rise with traffic volumes, shares of arterial streets lacking transit, share of land *zoned* for neighborhood commercial and mixed residential/neighborhood commercial uses, numbers of residents and (resident) workers, and share of persons living in poverty. They did not normalize crash counts by exposure, which is fundamental to count prediction, so many of their modeled effects are size effects (proxying for exposure). Miranda-Moreno et al. (2011) simultaneously modeled pedestrian activity (in log-linear form, as an exposure variable) and crash counts (using a standard negative binomial specification) at signalized intersections in Montreal, Canada. They concluded that many built environment, transport system, and traveler attributes (such as land use types, network intensity,

transit supply, and demographic characteristics) in the vicinity of intersections are strong predictors of pedestrian activity but have rather small effects on collision frequency (after controlling for exposure). Using unsupervised learning methods, Prato et al. (2012) identified fatal pedestrian-crash hot spots or clusters in Israel, as a function of lighting conditions, local demographics, share of two-wheel vehicles in the traffic flow, and roadway attributes.

Using Poisson regression (with heterogeneity and under-reporting components), Cottrill and Thakuriah (2010) evaluated the influence of transit accessibility, crime rates, and general demographics (like income and children population) on pedestrian crash rates in the Chicago area. Their results suggest that safety improvements targeting transit may moderate high pedestrian rates in various areas. Statistically significant spatial clustering of crash counts was reflected in model residuals, via local indicators of spatial association (LISA); however, such dependence was not captured by their model specification.

Spatial dependence across observational units is prevalent in transportation data sets, including traffic volumes (e.g., Wang and Kockelman [2009]), land development decisions (Wang et al. 2012), and crash prediction (e.g., Levine et al. 1995a, Levine et al. 1995b, Wang et al. 2009). Miaou et al. (2003) showed the existence of spatial autocorrelation among adjacent roadway segments in their analysis of vehicle crashes along rural two-lane highways in Texas, using several variations of a conditional autoregressive (CAR) count model. Wang et al. (2009) examined traffic congestion's influence over vehicle crashes along 70 homogenous segments of a British expressway, while accounting for both heterogeneity and spatial autocorrelation using a series of Poisson-based CAR models. While they found that congestion did not play a significant role along their case corridor, other covariates' roles were consistent with existing work (e.g., higher grades are associated with higher crash rates).

Many studies have explored multivariate count models for aspatial settings and confirm that significant interactions across crash types (e.g., severity levels) exist, due to omitted variables and other unobserved (latent) factors (Park and Lord 2006, Ma et al. 2008, El-Basyouny and Sayed 2009, Valverde and Jovanis 2007). However, only Miaou and Song (2005) and Wang et al. (2011) appear to offer a multivariate approach to crash count modeling (producing site rankings for safety improvements) in a spatial setting. Wang et al. (2011) used a two-stage mixed multivariate framework, where the first stage uses an "intrinsic" spatial model for a total (univariate) crash count (with no spatial autocorrelation coefficient, which is restrictive [Spiegelhalter et al. 2003]), and the second stage is a multinomial logit model (for count splits, after conditioning on the total).

This paper proposes and develops a multivariate Poisson log-normal CAR model to reveal spatial autocorrelation, zone-specific heterogeneity, and correlation across pedestrian crash counts and severity levels. The model is built upon the multivariate CAR model proposed by Jin et al. (2005) and estimated using Bayesian Markov chain Monte Carlo methods with a sampling scheme described in the Appendix. Spatial multivariate analysis of *pedestrian* crash counts is a novelty. Covariates include zone-level land use, transit access, network intensity, sidewalk density, and demographic attributes. The City of Austin's map layers were used to derive land use covariates for neighborhoods across Travis County, offering more realistic land use information than zoning maps (as used in Weir et al.'s [2009] analysis).

METHODOLOGY

Two spatial model specifications are common for relating neighboring sites' responses: the spatial autoregressive model (SAR), as discussed in Elhorst (2009) and Anselin (1988), and the conditional autoregressive model (CAR), as appears in Besag (1975). Cressie (1995) has shown that the SAR specification is a special type of CAR model, at least in a continuous-response setting. CAR models are more commonly used in spatial analysis of count data, thanks to faster computation (see, e.g., Czado et al. 2010, Song et al. 2006, Gelfand and Vounatsou 2003, Miaou et al. 2003, and Pettitt et al. 2002). The SAR approach tends to be difficult to employ for limited-response frameworks, especially with large data sets (as discussed in Wang et al. [2012] and Wang et al. [2012]) and yields parameter estimates similar to those estimated from the CAR model (Miaou et al. 2003). For all these reasons, this work relies on the CAR specification for analyzing crash counts.

Conditional autoregressive (CAR) specifications appear to begin with Besag (1975), and are mostly estimated using Bayesian methods. Conditional distributions of (univariate) CAR-model response variables are, in most cases, defined by a series of conditional distributions, as shown in Equation 1 (Cressie 1995).

$$\lambda_i|\lambda_{-i} \sim N\left[\mu_i + \sum_{j=1}^{n} c_{ij}\left(\lambda_j - \mu_j\right),\ \sigma_i^2\right] \tag{1}$$

where λ_i indicates the spatially autocorrelated variable (typically a response variable, like mean crash rates, average traffic flows, or household incomes), λ_{-i} denotes such variables at neighboring locations, μ_i is the expected response value (such that $E(\lambda_i) = \mu_i$), σ_i^2 is the conditional variance, and c_{ij} are known or unknown weights (with $c_{ii} = 0$), describing the proximity or closeness between locations *i* and *j*.

These conditional distributions lead to a multivariate normal (MVN) joint distribution of the spatially correlated variables, based on the factorization theorem (Besag 1975, Wall 2004):

$$\boldsymbol{\lambda} \sim MVN_n[\boldsymbol{\mu}, (\boldsymbol{I} - \boldsymbol{C})^{-1}\boldsymbol{M}] \tag{2}$$

where the column vector $\boldsymbol{\lambda}$ is a stacked version of the *n* λ_i's (as is the vector $\boldsymbol{\mu}$), $\boldsymbol{I}$ is an identity matrix, $\boldsymbol{C}$ is an *n* by *n* weight matrix defined by contiguity or distance and $\boldsymbol{C} = \left[c_{ij}\right]$, and $\boldsymbol{M}$ is a diagonal matrix with $\boldsymbol{M}_{\text{ii}} = \sigma_i^2$. This joint distribution is used, along with the likelihood function of the data set, to implement the Gibbs sampler for estimating the posterior distributions of all unknown parameters.

The validity of the MVN distribution (shown in Equation 2) requires its covariance matrix to be symmetric and positive definite, conditions that can be satisfied by imposing certain constraints on the forms of $\boldsymbol{C}$ and $\boldsymbol{M}$. For example, one may let $\boldsymbol{C} = \rho\boldsymbol{W}$ and $\sigma_i^2 = \frac{\sigma^2}{w_{i+}}$, where ρ is referred to as the spatial autocorrelation coefficient, $\boldsymbol{W}$ is a row-standardized weight matrix (i.e., $\boldsymbol{W} = \left[w_{ij}^*\right]$ and $w_{ij}^* = \frac{w_{ij}}{w_{i+}}$), and w_{i+} is the i^{th} row sum of *W*. By construction, $\boldsymbol{W}$'s diagonal elements are all zeros (Cressie 1995, Wall 2004). The CAR specifications permit contiguity and distance-

based weight matrices but preclude the K^{th}-nearest neighbor weighting scheme because such weights violate the symmetry condition. First-order contiguity weights are defined such that $w_{ij} = 1$ if i and j share a common border, $w_{ij} = 0$ otherwise, and $\boldsymbol{W}$'s diagonal elements (w_{ii}) are all zeros by construction (Cressie 1995). When ρ is fixed at 1, the CAR specification becomes an "intrinsic" CAR model (prevalent in empirical studies), and requires less computing time but presents theoretical and conceptual issues that undermine its validity. Specifically, when the precision parameter (σ^2) is unknown (as is always the case), the functional form of the joint distribution of the spatial random effects ($\boldsymbol{\lambda}$), shown in Equation 2, and σ^2 are not identified under the "intrinsic" CAR specification (Cressie 1995, Gelfand and Vounatsou 2003). Thus, one cannot be as confident in his/her estimates, nor convergence of the parameter draws, due to potentially improper distributional assumptions. Conceptually, the spatial autocorrelation coefficient, ρ, measures the overall spatial relationship of the data, whereas the precision parameter σ^2 accounts for the variation of the spatial dependence. Omitting ρ blurs one's estimates and can lead to counterintuitive interpretations (Spiegelhalter 2003).

A Poisson Log-Normal Multivariate CAR Model

To successfully specify a multivariate CAR (MCAR) structure, an important consideration is the validity of the joint covariance matrix. Rather than focusing on the covariance matrix, it is typical to start from its inverse, or the precision matrix, because the latter is faster to compute and the computation can be implemented using several full-blown methods, see, e.g., the decomposition methods employed by Carlin and Banerjee (2003) and Gelfand and Vounatsou (2003). However, working directly with the precision matrix, instead of the covariance matrix, often obscures interpretation of the correlation structure of the phenomenon under study. In contrast, a judiciously designed covariance matrix allows one to incorporate more behavioral realism, while ensuring the estimability of the resulting model.

This work extends the multivariate CAR model proposed by Jin et al. (2005) to allow for region-specific heterogeneity and a non-Gaussian first stage, leading to what we called a Poisson log-normal MCAR model. Rather than having to transform the aggregated counts to continuous response (e.g., standardized mortality ratio [SMR], Jin et al. 2005), the proposed model allows one to directly analyze spatial count data, such as area-level pedestrian crash count data.

The first stage is expressed as a Poisson process:

$$y_{ik} \sim \text{Poisson}(\lambda_{ik}) \tag{3}$$

where y_{ik} is the observed pedestrian crash counts by severity levels (k=1, 2) for the i^{th} polygon of Travis County, and the mean crash rates, λ_{ik}, in the second stage, are expressed as:

$$\lambda_{ik} = E_{ik}^{\alpha} \cdot \exp\left(x_i'\beta_k + \phi_{ik} + u_i\right) \tag{4}$$

where E_{ik} is an exposure measure of pedestrian crashes (e.g., Walk-Miles traveled for each zone) with the unknown parameter α describing any non-linear relationship between the risk and mean rates, x_i indicates a column vector of covariates including a constant term, β_k denotes a column vector of parameter coefficients specific to each observation type k, and ϕ_{ik} represents the spatial random effect defined by a MCAR structure discussed later in this section. The

heterogeneity error term, u_i, captures zone-specific heterogeneity that is not explained by spatial effects and is assumed to follow a normal distribution, $u_i \sim N(0, \sigma_u^2)$, leading to the Poisson-lognormal spatial model. Alternatively, its exponential term may take on a gamma distribution, $\exp(u_i) \sim \text{Gamma}(\theta, \theta)$, leading to a negative binomial model (Miaou et al. 2003).

Spatial random effects, ϕ_{ik}, follow the multivariate conditional autoregressive model proposed by Jin et al. (2005):

$$\begin{pmatrix} \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}'_{12} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \right) \tag{5}$$

where $\boldsymbol{\phi}_k$ contains the spatial random effects across n locations for a given response type k, k =1, 2, and Σ_{kl} represents $n \times n$ covariance matrices (k, l =1, 2). Multivariate normal theory leads to the following conditional distributions that jointly determine Equation (5).

$$\begin{aligned} &\boldsymbol{\phi}_1 | \boldsymbol{\phi}_2 \sim N(\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\phi}_2, \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}'_{12}) \\ &\boldsymbol{\phi}_2 \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{22}) \end{aligned} \tag{6}$$

For ease of presentation, let $\mathbf{A} = \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}$, $\boldsymbol{\Sigma}_{11\cdot 2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}'_{12}$. Therefore, the joint distribution of $\boldsymbol{\phi}$ is written as:

$$\begin{pmatrix} \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{11\cdot 2} + \mathbf{A} \cdot \boldsymbol{\Sigma}_{22}\mathbf{A}' & \mathbf{A} \cdot \boldsymbol{\Sigma}_{22} \\ (\mathbf{A}\boldsymbol{\Sigma}_{22})' & \boldsymbol{\Sigma}_{22} \end{pmatrix} \right) \tag{7}$$

This joint distribution always exists as long as its covariance matrix is symmetric and positive-definite. The conditions that ensure such property are that $\boldsymbol{\Sigma}_{11\cdot 2}$ and $\boldsymbol{\Sigma}_{22}$ are positive definite (Harville 1997 cited in Jin et al. 2005), as discussed later in this section.

The crux of the problem is then to specify the matrices $\mathbf{A}$, $\boldsymbol{\Sigma}_{11\cdot 2}$, and $\boldsymbol{\Sigma}_{22}$, which will uniquely determine the functional form of the covariance matrix of the joint distribution, as shown in Equation (7). Assume that $\text{var}(\boldsymbol{\phi}_1 | \boldsymbol{\phi}_2) = \boldsymbol{\Sigma}_{11\cdot 2} = [(\mathbf{D} - \rho_1 \mathbf{W})\tau_1]^{-1}$ and $\text{var}(\boldsymbol{\phi}_2) = \boldsymbol{\Sigma}_{22} = [(\mathbf{D} - \rho_2 \mathbf{W})\tau_2]^{-1}$, with scalars τ_1 and τ_2 being the scale parameters. The remaining undetermined quantity (in order to uniquely identify the joint distribution's covariance matrix in Equation [5]) is the transformation matrix $\mathbf{A}$. One may consider $\mathbf{A} = \eta_0 \mathbf{I} + \eta_1 \mathbf{W}$ and thus $E(\phi_1 | \phi_2) = (\eta_0 \mathbf{I} + \eta_1 \mathbf{W})\phi_2$. This parameterization suggests that the conditional mean of ϕ_{i1} at a given location i equals to a scaled ϕ_{i2} value (at the same location) plus a weighted average of neighboring ϕ_{j2} values.

The MCAR model developed through Equations (5) to (7) implies that the covariance matrices of ϕ_2 and ϕ_1 are independent, following the CAR structure with different spatial autocorrelation coefficients respectively. Correlations across different response types are captured by the transformation matrix, $\mathbf{A}$, with the parameter η_0 representing aspatial cross correlation and η_1 describing spatially-lagged cross correlation. Jin et al. (2005) prove that this MCAR model is more general and encompasses the model proposed by Carlin and Banerjee (2003) and Gelfand and Vounatsou (2003).

The spatial autocorrelation coefficients ρ_1 and ρ_2 describe the spatial dependence for the two crash types respectively and should lie within the range $[\frac{1}{\min}, \frac{1}{\max}]$ for the covariance matrices $[\tau_2(\mathbf{D} - \rho_2\mathbf{W})]^{-1}$ and $[\tau_1(\mathbf{D} - \rho_1\mathbf{W})]^{-1}$ to be positive definite and thus invertible, where max and min denote the maximum and minimum eigenvalues of the standardized weight matrix, $\mathbf{D}^{-1}\mathbf{W}$. Note that the matrix, $\mathbf{D}^{-1}\mathbf{W}$, is row-standardized by construction. Negative spatial dependence is rare and thus the lower bound is often set to 0, while for the row-standardized weight matrix, its maximum eigenvalue is guaranteed to be 1 (Cressie 1995). **D** is a diagonal matrix with the i^{th} diagonal element representing the i^{th} row sum of **W**. The precision parameters τ_1 and τ_2 scale the covariance structures.

Note that trivariate and higher-order specifications (for data sets involving four or more response types) can be derived using the above lines of reasoning, as described in this paper's Appendix.

DATA SETS

The new multivariate CAR model developed here was estimated using a Bayesian sampling scheme coded in WinBUGS and is used here to analyze a three-year total of pedestrian crash counts in Travis County from 2007 to 2009. The model controls for transit stop density, land use balance (measured by entropy), residential access to commercial land, school access, network intensity (computed as lane-mile densities by roadway classes), and sidewalk densities.

The study zones rely on Thiessen polygons, built around Austin's census tracts, to ensure that high-crash locations, regularly along tract edges (important roadways) and often at tract corners (important intersections), are uniquely assigned to a polygon zone rather than missing or arbitrarily assigned to adjacent tracts. By default, ArcGIS creates Thiessen polygons based on a given set of polygons (or their centroids) within a rectangular area that covers the given geographic area, resulting in several unreasonably large polygons at the periphery of the county. Thus, Travis County's boundary file was used to cut the "redundant" portion of the polygons to yield a new boundary that follows closely with the original Travis County's shape, as shown in Figure 1. Table 1 summarizes the covariates for the 218 zones derived from census tracts in Travis County.

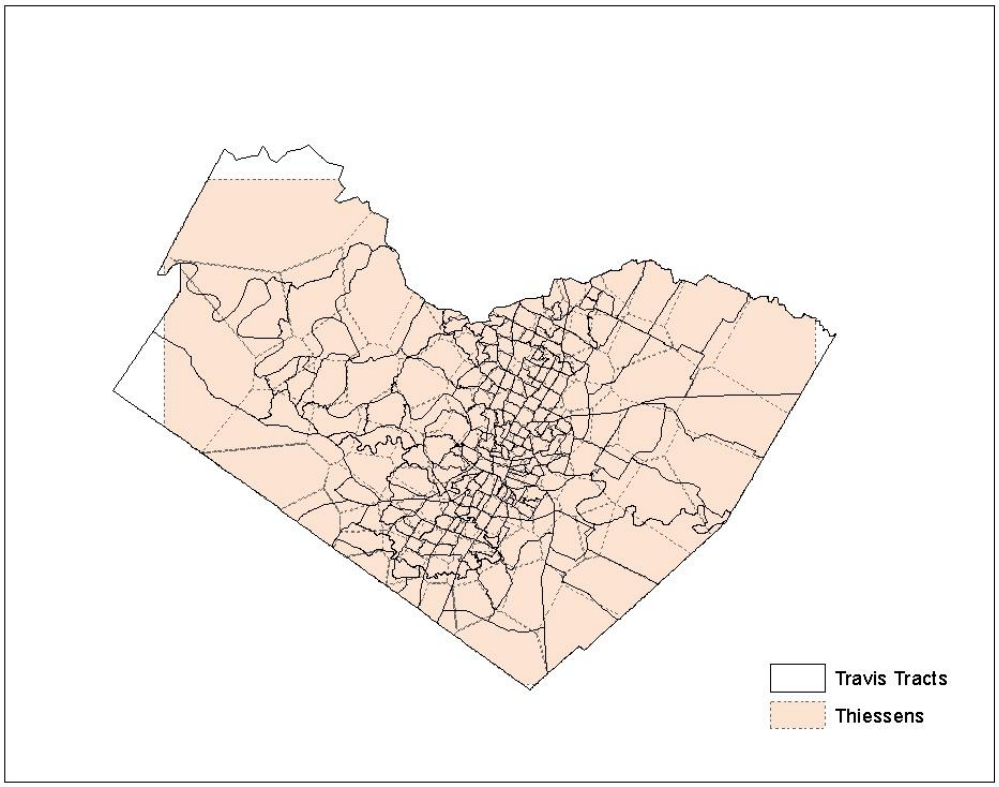

Figure 1. Thiessen Polygons based on Travis County Census Tract Centroids

Table 1. Summary Statistics of Covariates and Response Variables across Thiessen Polygons (n=218)

	Mean	*Std Dev*	*Min*	*Max*
Transit Access				
% SFDU[a] near Transit in zone (within 1/2 mi.)	0.628	0.433	0	1
% APT[b] near Transit (1/2 mi.)	0.655	0.432	0	1
Transit Density (# of bus stops per sq. mile)	13.66	17.57	0	98.6
Land Use				
Land Use Entropy	0.647	0.229	0.037	0.989
% Resid. Parcels near Commercial (1/2 mi.)	0.759	0.304	0	1
Network Intensity				
LnMiDenFWY	4.228	6.435	0.000	44.43
LnMiDenART	8.836	6.783	0.104	51.20
LnMiDenLOC	2.435	3.770	0.000	18.93

Sidewalk Density	6.718	6.076	0.000	28.85
Vehicle Miles Traveled *(per day in 2010)*				
VMTFWY	1.59E+05	2.93E+05	0	1.52E+06
VMTART	3.22E+05	3.32E+05	937	3.61E+06
VMTLOC	1.37E+04	2.93E+04	0	2.45E+05
Demographics & Employment *(2007)* [c]				
Population Density	2,470	2,611	5	1.563E4
Basic Emp. Density	356	653	0	5,137
Retail Emp. Density	235	279	0	1,842
Service Emp. Density	598	762	1	5,308
Access to School				
% SFDU near school (within 1/2 mi.)	0.514	0.352	0	1
% APT near school (within 1/2 mi.)	0.487	0.386	0	1
Exposure Measure				
Walk-Miles Traveled (WMT[d]) (in miles over a two-weekday period)	68.80	41.26	4.79	291.3
Response Variable				
Severe Crash Counts (Fatal & Incapacitating Crashes, 2007–2009)	0.89	1.53	0	15
Non-Severe Crash Counts (Incapacitating, Possible Injury, & No Injury Crash Counts, 2007–2009)	3.23	7.4	0	100

Notes: [a.] SFDU stands for single-family dwelling units, including single family and large-lot single family dwelling units; [b.] APT denotes apartments (e.g., group quarter, duplex, apartment/condo defined by City of Austin's land use archive); [c.] .population and employment densities are computed as the estimated counts (by overlaying traffic-analysis-zone-level count information obtained from CAMPO) divided by polygon size; [d] WMT is the crash exposure measure, estimated using household travel survey data and least squares regression (with details provided in this paper's Results section).

The influence of land use and the built environment on pedestrian safety has been documented by Dumbaugh (2005) and Dumbaugh and Rae (2009). Land use attributes affect pedestrian crash counts, by influencing both walking frequency (and thus pedestrian exposure), traffic volumes or vehicle exposure, and situational complexity for travelers (which influences collision likelihood) (Clifton et al. 2008, Miranda-Moreno et al. 2011). Clifton et al. (2004) showed how areas with high transit access are associated with much higher pedestrian crash rates, and with crashes involving children. Hence, several land use variables are explored here for each zone, including land use entropy, the share residential dwelling units that are close to transit stops, and the share that are close to commercial activities, as summarized in Table 1.

Land use information comes from the City of Austin's 2006 land use maps. Year 2006 immediately precedes the 2007-2009 crash-count period while covering all of Travis County. In contrast, the next available (year 2008) map covers only Austin and its extraterritorial jurisdiction. Few sites experienced development over the 2-year period (2006 to 2008), as described in Wang et al. (2012), and the 2006 map allows for more data points (approximately 65,000 parcels) in this analysis. Land use entropy or balance is formulated using Cervero and Kockelman's (1997) approach and has been used in a variety of settings (see, e.g., Frank and Pivo 1994 and Brown et al. 2009):

$$LU\ Balance = -\sum_{k=1}^{4} p_k \ln(p_k) / \ln(4)$$

where p_k is the proportion of a particular land use k, representing residential, commercial, industry, and office uses. An evenly balanced situation (i.e., each of the four uses take up 25 percent of land area, a situation that rarely exists in practice) delivers an entropy value of 1, whereas smaller entropy values imply less balanced land use patterns.

Another relevant variable is the proximity of residential dwelling units to transit service, which is likely to correlate with pedestrian exposure and possibly the presence of more driver sight obstructions, as discussed in Clifton et al. (2008). This study controls for the number of dwelling units (both single-family and multi-family units) that are within one-half mile of transit stops.

A positive association between the presence of children (under 14 years of age) and pedestrian crashes has been established in the literature (see, e.g., Clifton et al. 2004 and NHTSA 2011). Plausible causes include children's shorter stature (making them harder for motorists to see), their (often) less-developed sense of motion, and an inexpert ability to judge traffic conditions and signal lights. This study computes the percentage of residential parcels (within each polygon) that are within one-half mile of schools, which may proxy this particularly vulnerable population. School information was obtained from. The only available data year for the Texas Education Agency's school locator (http://wgisprd.tea.state.tx.us/sdl/) is 2010, so we assume that school locations remain the same over the 4-year period (from 2007 to 2010). Access to schools is represented as the share of residential units (including SFDUs and apartments) that lie within the half-mile buffers created around school points, using ArcGIS's proximity routine.

Lane-miles by functional classifications can reflect route availability and network connectivity, and thereby affect people's propensity to walk. CAMPO's coded 2005 network and Census Tigerline files were used to extract roadway information. The former provides numbers of lanes, operating speeds, and vehicle counts (imputed from travel demand model results, rather than from observed traffic counts, which are relatively few and far between [in space and time]) for most of the county's roadways. As shown in Figure 2, local streets complement this CAMPO network, but they do not come with traffic-volume and number-of-lane attributes.

Figure 2. A Snapshot of Austin's TigerLine Network (2008) Overlaid on Thiessen Polygons

RESULTS AND ANALYSIS

This section summarizes results for the Poisson log-normal MCAR model applied to the three-year pedestrian-crash data set. Presented first are the estimation results of walk-miles traveled (WMT) per zone using the 2005/2006 Austin Travel Survey (ATS) data set, which provides a glimpse of traffic analysis zone- (TAZ-) level walk miles based on the 569 walking trips (out of a total of 14,113 trips surveyed over a 1-weekday period). These walking trips occur across 217 TAZs, among which 154 zones lie within Travis County and can be linked to this county's 2005 map. The surveyed walk trips were scaled up by the ratio of zone population to zone sample size (to reflect the zone's population share) and then used as the response variable in the TAZ-based WMT model, described below. Parameters estimated from the TAZ-level WMT model were then used to impute walk miles for each Thiessen polygon.

A Model for Walk-Miles Traveled

Covariates that may influence walk miles include zone size (in square miles), population, employment by types (i.e., basic, retail, and service), land use, and coded lane miles by road classes (freeway, arterial, and local streets). These covariates' summary statistics are shown in Table 2.

Table 2 Summary Statistics of Covariates for the Walk-Miles Traveled (WMT) Model (n=154 zones)

	Mean	*Std Dev*	*Min*	*Max*
Response-Related Variables				

WMT (miles per zone per two-weekday period)		2,753	8,124	0	71,531
WMT per capita (miles per zone per two-weekday period divided by zone population)		0.686	0.992	0	7.200
Network					
LnMiDenFWY		4.228	6.435	0.000	44.430
LnMiDenART		8.836	6.783	0.104	51.207
LnMiDenLOC		2.435	3.770	0.000	18.932
Sidewalk (total length, in miles)		13.511	12.328	0.000	67.397
Land Use					
Entropy		0.399	0.243	0.000	0.918
# Resid. parcels near Bus Stops		304	389	0	2,255
Zone Size					
Area (sq. mi)		1.67	7.53	0.04	87
Demographics					
Population (of zone)		2,652	2,473	5	12,532
Employment Counts	Base	377	851	0	7,084
	Retail	250	270	0	1,493
	Service	791	1,252	0	8,891
N_{obs} = 154 TAZs					

Notes: WMT = total walk-miles traveled = $wmt \cdot \frac{\text{Zone Population}}{\text{Num of Residents Sampled}}$, where wmt indicates walk-miles traveled by the ATS sample population. LnMile = lane-miles, FWY = Freeway, ART = Arterial streets, & LOCAL = Local streets.

A weighted least squares (WLS) regression model was adopted for predicting total WMT in the zone yielded the best fit (R^2_{adj} = 0.51) and parameter estimates among the four model specifications attempted (ordinary least squares [OLS], WLS, Tobit, and Heckit). Table 3 provides a summary of these WLS results (with statistically insignificant covariates removed). Weights were set at $\sqrt{n_i/N_i^2}$ to assure homoscedasticity of the error term, where n_i represents the number of respondents who were sampled for the i^{th} TAZ and N_i denotes the population counts for the corresponding TAZ. This weight was used because the total WMT (i.e., the response variable) is computed as the sample average times total population, with variance $N_i^2\sigma^2/n_i$.

Table 3 Weighted Least Squares Regression Results for Walk-Miles Traveled (WMT) Model, with Y= ln(Total WMT per zone).

Parameters	*Coef.*	*Std. Error*	*T-Statistic*
Constant	1.887	0.272	6.94
LnMiLOC	0.068	0.027	2.51
Area (sq. mi.)	0.344	0.123	2.80
Population	1.61E-03	2.58E-4	6.25
Sidewalk (mi.)	0.062	0.040	1.55
R^2	0.53		
Adj. R^2	0.51		
n_{obs}	154 zones		

Note: The weights are set at $\sqrt{n_i}/N_i$ for each zone.

Lane-miles by road class are shown to be a significant factor in explaining walk distances per zone. So are zone size (in square miles), population counts, and sidewalk lengths. The response variable here is an estimate of zone-level WMT, imputed using the average WMT per ATS respondent from that zone, multiplied by zone's population. A WLS scheme applies because WMT values are imputed using the population scaling factor described earlier, which introduces heteroskedasticity (in error term variances). The weights are set at $\sqrt{n_i}/N_i$ for each zone. Parameter estimates from Table 3 were then used to estimate WMT for each Thiessen zone, which served as the exposure measure in the MCAR model for pedestrian crash counts, as discussed in the next section.

A Poisson Log-Normal Multivariate CAR Model for Pedestrian Crashes

Recall the model specification in the Methodology section of this paper. The parameter ϕ_{ik} is defined such that: $\boldsymbol{\phi}_2 \sim \mathrm{N}(\mathbf{0}, [(\mathbf{D} - \rho_2\mathbf{W})\tau_2]^{-1})$ and $\boldsymbol{\phi}_1|\boldsymbol{\phi}_2 \sim \mathrm{N}(\mathbf{A}\boldsymbol{\phi}_2,\ [(\mathbf{D} - \rho_1\mathbf{W})\tau_1]^{-1})$. Analogous to the spatial random effects ϕ_{ik}, which are zero-centered, the logarithmic mean crash rates $\ln(\lambda_{ik})$ can also be expressed by a MCAR structure. The only difference between ϕ_{ik} and $\ln(\lambda_{ik})$ is that the latter variable's mean value is no longer centered at zero, but rather $[\ln(E_{ik}^{\alpha}) + x_i'\beta_k + u_i]$. Let column vectors $\mathbf{Z_1}$ and $\mathbf{Z_2}$ substitute for $\ln(\boldsymbol{\lambda}_1) = [\ln(\lambda_{11}), \ln(\lambda_{21}), \ldots, \ln(\lambda_{n1})]'$ and $\ln(\boldsymbol{\lambda}_2) = [\ln(\lambda_{12}),\ \ln(\lambda_{22}), \ldots, \ln(\lambda_{n2})]'$. The conditional distributions for $\mathbf{Z_1}$ and $\mathbf{Z_2}$ are multivariate normal:

$$\mathbf{Z}_1|\mathbf{Z}_2 \sim \mathrm{N}(\boldsymbol{\mu} = \ln(\mathbf{E}^{\alpha}) + \mathbf{X}\cdot\boldsymbol{\beta}_1 + \mathbf{u} + (\eta_0\mathbf{I} + \eta_1\mathbf{W})(\mathbf{Z}_2 - \ln(\mathbf{E}^{\alpha}) - \mathbf{X}\cdot\boldsymbol{\beta}_2 - \mathbf{u}),\ \Omega = [\tau_1(\mathbf{D} - \alpha_1\mathbf{W})]^{-1})$$

$$\mathbf{Z}_2 \sim \mathrm{N}(\ln(\mathbf{E}^{\alpha}) + \mathbf{X}\cdot\boldsymbol{\beta}_2 + \mathbf{u}, [\tau_2(\mathbf{D} - \alpha_2\mathbf{W})]^{-1}) \quad (8)$$

where $\mathbf{E}$ is an *n* by 1 vector of walk-miles traveled (with unknown parameter α helping explain any non-linear relationship between exposure levels and average rates), $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are two column vectors specific to each of the two crash types (with 1 denoting severe crashes [including fatal and incapacitating injuries] and 2 denoting non-severe crashes [including non-

incapacitating, light, possible, and no injury crashes]), **X** is the covariance matrix with the i^{th} row being the observed explanatory variables (including a constant term) for region *i*, and **u** a vector of the *n* heterogeneity error terms, $\mathbf{u} = (u_1, \ldots, u_n)'$.

As noted earlier, the model was estimated using Bayesian Markov chain Monte Carlo techniques (with the sampling scheme as described in the Appendix). Convergence was checked using Geweke's diagnostic tests, as coded in R's "coda" package (Plummer et al. 2012). A bivariate specification is used here because it generated a better fit (DIC value) than the three trivariate models tested[1]. Table 4 summarizes parameter estimates and inference after removing several markedly insignificant covariates (defined as those with pseudo t statistics lower than 1). These variables include transit density and residential parcel shares near schools for severe crashes, and land use entropy for non-severe crashes, as shown as blank rows in Table 4.

Table 4. Parameter Estimates and Inference of the Area-Level Pedestrian Crash Model.

[1] The first of these two models had fatalities and incapacitating injury crashes in two separate classes, and non-incapacitating injury, light-injury, and no-injury crashes in the third class, while the second had fatalities in a single class, incapacitating injury and non-incapacitating injury crashes in the second class, and light-injury and no injury crashes in the third class. The third model had fatalities and incapacitating injury crashes in a class, non-incapacitating and light injury crashes in the second class, and no injury crashes in the third class.

	Severe or Not?	Mean Estimate	Std Dev	Pseudo T-stat.	MC error	2.5% Estimate	Median	97.5% Estimate	Elasticity
Constant	1 (yes)	-0.652 (2.31)*	0.169	-3.87	0.002	-0.844	-0.570	-0.463	
	2 (no)	0.462 (3.42)	0.142	3.25	0.002	0.300	0.458	0.621	
Transit Density	1								
	2	0.482 (1.53)	0.137	3.51	0.002	0.325	0.393	0.635	0.03
Land Use Entropy	1	-0.595 (1.98)	0.278	-2.14	0.001	-0.912	-0.432	-0.284	-0.05
	2								
% Resi. Parcels near Commercial	1	1.158 (1.86)	0.480	2.41	0.003	0.611	0.689	1.696	0.04
	2	0.950 (2.02)	0.497	1.91	0.002	0.383	0.581	1.507	0.06
LnMileDen FWY	1	0.225 (1.86)	0.089	2.52	0.001	0.123	0.223	0.326	0.03
	2	0.111 (1.92)	0.070	1.59	0.001	0.031	0.123	0.189	0.04
LnMileDen ART	1	0.607 (1.63)	0.189	3.21	0.002	0.392	0.574	0.819	0.47
	2	0.830 (1.74)	0.306	2.71	0.002	0.481	0.565	1.173	0.52
LnMileDen LOC	1	-0.259 (1.82)	0.089	-2.91	0.003	-0.360	-0.092	-0.159	-0.41
	2	-0.033 (1.69)	0.014	-2.31	0.000	-0.050	0.155	-0.017	-0.21
Population Density	1	0.208 (1.77)	0.136	1.54	0.001	0.054	0.203	0.360	0.04
	2	0.213 (2.31)	0.190	1.12	0.002	-0.004	0.234	0.425	0.08
% Resi. Parcels near Schools	1	-0.323 (1.42)	0.107	-3.01	0.001	-0.445	-0.270	-0.203	-0.03
	2								
Sidewalk Density	1	-0.374 (1.67)	0.104	-3.61	0.003	-0.492	-0.238	-0.258	-0.13
	2	-0.571 (1.47)	0.164	-3.48	0.003	-0.756	-0.569	-0.382	-0.22
ln(VMTART)	1	0.008 (2.14)	0.004	1.87	0.006	0.003	1.010	0.013	0.01
	2	0.024 (1.73)	0.010	2.51	0.008	0.013	1.515	0.035	0.05
ρ_1		0.728 (1.16)	0.127	5.71	0.002	0.575	0.724	0.873	
ρ_2		0.612 (1.21)	0.102	5.99	0.002	0.496	0.612	0.728	
α		0.131 (1.85)	0.057	2.31	0.001	0.051	0.123	0.196	
η_0		0.712 (1.17)	0.134	5.31	0.001	0.563	0.714	0.865	
η_1		0.312 (1.48)	0.076	4.13	0.002	0.226	0.312	0.398	

τ_{v1}	1.352 (1.69)	0.348	3.886	0.009	0.788	1.310	2.138	
τ_{v2}	2.677 (1.85)	0.476	5.623	0.007	1.863	2.640	3.716	
τ_1	1.653 (1.14)	0.495	3.342	0.007	0.852	1.615	2.707	
τ_2	2.113 (1.42)	0.261	8.083	0.004	1.635	2.115	2.655	
DIC	3200.5							
Mean of LogLik	-2568.1							
RMSE	2.41							
Run times = 59 mins; # of Iteration=15,000; Burn-in period=5,000; # of chains = 3;								

Note: "1" rows denote values for fatal and incapacitating injury crash count prediction, and "2" rows denote parameter values for predicting other (non-severe) crash counts. Geweke's diagnostic statistics are provided in the Mean Estimate column, in parentheses.

Table 4's elasticities were computed as the average (over the entire sample) percentage change in the mean crash rate (or expected value, λ_i) per one percent change in the kth covariate (for each zone, i). These mean crash rates incorporate Eq. 7's unknown/latent error terms, as simulated for the region-specific errors, spatial autocorrelation, and correlations across various response types.

Interpretation of Model Results

Table 4's results reveal noticeable spatial clustering patterns of zone-based crash counts. Severe (i.e., fatal and incapacitating) counts are estimated to have a statistically (and practically) significant spatial autocorrelation coefficient of 0.73, whereas non-severe (i.e., non-incapacitating, light, possible, or no injury) counts yield a slightly lower, but still significant coefficient of 0.61. Apart from these within-category spatial autocorrelations, statistically (and practically) significant spatial dependence emerges across the two crash-type categories: η_1 is estimated to be 0.31 and measures spatially lagged effects of cross-correlation across the two categories.

Area-level crash counts also exhibit strong correlation between the two severity levels, as measured by a statistically (and practically) significant η_0 value of 0.71. This value implies that severe and less severe pedestrian crash rates correlate in a very positive way, even after controlling for exposure and various other zone-level attributes. Such aspatial cross-correlation is expected, and attributable to omitted variables shared by crash types within a zone (but not across zones, as reflected via the η_1 term estimate). Examples of such missing-variables correlation (across response types) include presence of unusual site conditions (like heavy industry or entertainment zones), distinctive local lighting conditions (affecting night-time crash rates), and sight obstructions (affecting pedestrian and motorist visibility at all times). In contrast, the spatially lagged effects of cross-correlation capture missing variables that are spatially clustered but wider spread, thus affecting many nearby zones, and are shared across crash-severity levels—such as terrain features, weather conditions, and various socio-economic variables.

The relationship between crash exposure (WMT per zone) and crash rates is estimated to be highly non-linear (with an average exponent, α , of 0.131, rather than 1 [for the linear case]),

with rates (per mile walked) falling off dramatically as walk levels rise, presumably thanks to drivers expecting more pedestrians in high-WMT zones and responding accordingly and/or safer pedestrian environments encouraging more walking. This is a salient result: crash rates fall substantially (per WMT) as pedestrian exposure (WMT) rises, ceteris paribus, as shown in Figure 3. Also, this parameter was assumed to be identical across severity levels, based on DIC values not changing when distinctive exponents were permitted. ln(VMTs) for local streets and freeways failed to show significance and were removed from the model.

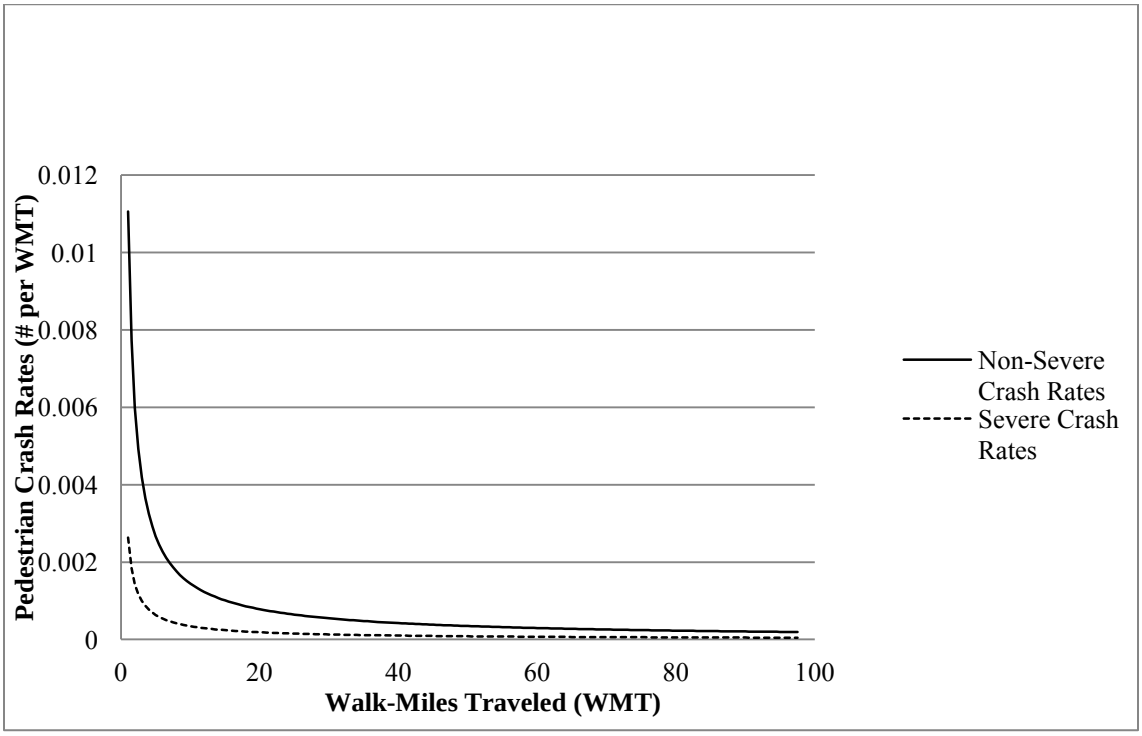

Figure 3 Relationship between Pedestrian Crash Rates (# per WMT) and Walk-Miles Traveled over a Two-Workday Period

After controlling for exposure (WMT), greater land use balance was estimated to lower severe crash rates, as reflected by a negative coefficient estimate on the entropy measure. But entropy's effect on less-injurious crash rates was not statistically (nor practically) significant, so it was not included in that piece of the final specification (as reflected in Table 4). Shares of apartment and single-family parcels near commercial parcels showed very similar parameter estimates and so were merged to form a single covariate (the share of residential parcels within ½ mile of commercial parcels). In contrast, an increase in pedestrian crashes across both crash types is predicted (everything else constant) when a higher share of a zone's residential parcels are near commercial land uses, as reflected by (modest) elasticities of +0.04 and +0.06 (for the two crash types, respectively). The variance term for the non-severe crash rates is estimated to be 2.7, surpassing the variance for the severe crash rates by 1.4, as expected, since non-severe crash rates are generally higher than the rates for severe crash rates and permit greater variation in the error term.

Higher bus-stop density appears to contribute somewhat to less-injurious crash rates (after controlling for pedestrian exposure), but its effect on severe pedestrian crash rates was found to

be minimal (and so was removed from the final model for that crash type). Residents' proximity to schools was found to have almost no practical effect on either crash rate (after controlling for WMT estimates in each zone), but its coefficient was statistically significant in the case of severe crash rates.

Network intensity covariates yielded mixed effects: A higher density of arterial streets is predicted to notably contribute to both severe and less-severe crashes (with elasticities of +0.47 and +0.52, respectively), whereas freeway intensity had little practical effect. Interestingly, a higher local-street density is estimated to significantly ***lower*** severe crash rates, and, to a lesser extent, non-severe crash rates. It would be useful to be able to control for traffic levels, instead of simply centerline miles, to get a better sense of how these network-design effects (arterials vs. locals) play out, in order to better anticipate an optimal balance in serving all travelers while protecting pedestrians.

Residual Spatial Autocorrelation

Residuals were computed as the difference between estimated values and observed values for pedestrian crash counts by severity levels, as shown in Figures 4 and 5. Both maps show negligible, positive spatial dependence, as measured by Moran's I value and a significant measure reflected by the p-value. Moran's I is a spatial version of the common Pearson correlation coefficient: the closer this statistic is towards 1 or -1, the stronger the spatial autocorrelation (with positive values indicating spatial clustering and negative values denoting spatial dispersion).

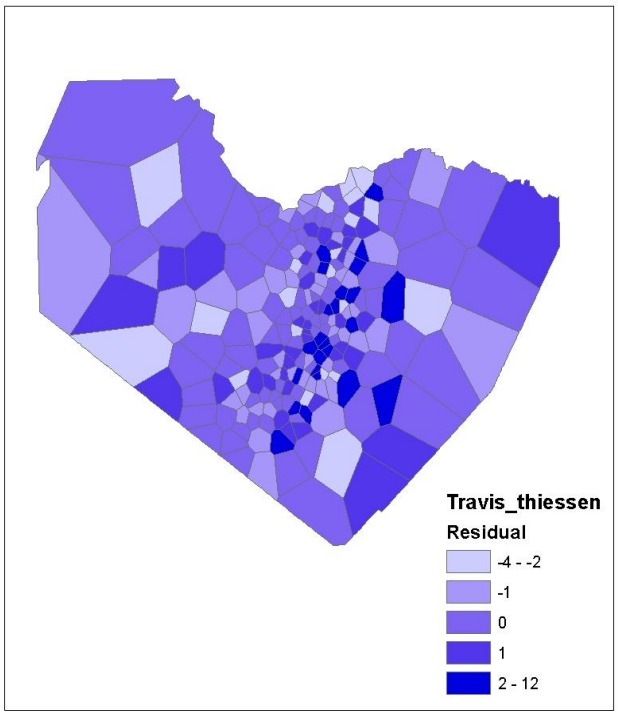

Figure 4 Spatial Distribution of Residuals for Severe Crash Counts.

Note: Moran's I = 0.013 (with p value = 0.70)

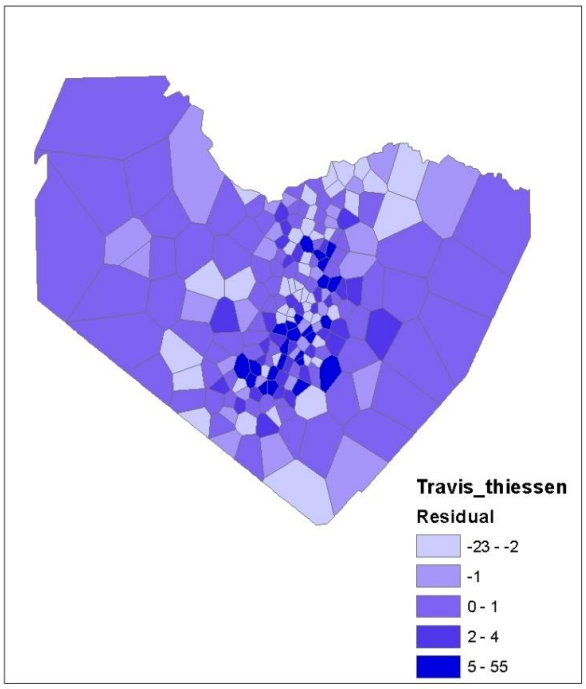

Figure 5. Spatial Distribution of Residuals for Non-Severe Crash Counts.

Note: Moran's I = 0.028 (with p-value = 0.03)

The Poisson log-normal MCAR model is also compared with an aspatial multivariate Poisson-lognormal model and a spatial Poisson-lognormal model (without correlations across different severity levels), with results shown in Table 5. The Poisson log-normal MCAR model yields the lowest DIC value and Moran's I of residuals among the three models tested. Including spatial autocorrelation effect has proved to greatly improve fit statistics, as reflected by the marked increase in the mean log-likelihood (and decrease in DIC values) after convergence in achieved. The Poisson log-normal CAR model (i.e., Model II, which incorporates spatial autocorrelation within each severity level but omit cross-severity correlation) reduces DIC value by 10% from a pure multivariate Poisson log-normal model (Model III). Another decrease of 34% in DIC value results from Model I's incorporating aspatial and spatially-lagged cross-correlation into Model II. Similar observations can also be found for comparing the root mean squared errors (RMSE) across the three models.

Table 5. Comparison of Full Model Results (I) to Aspatial Model (II) and Spatial Model without Cross Correlation (III) Results

	Poisson Log-Normal MCAR	*Poisson Log-Normal CAR*	*Poisson Log-Normal Multivariate*
Model No.	I	II	III

Parameter Constraints	-	η_0 & $\eta_1 = 0$	ρ_1, ρ_2, & $\eta_1 = 0$
DIC	3200.5	4852.31	5061.41
Mean LogLik	-2568.1	-3731.13	-3999.12
RMSE	2.41	4.21	6.70
Moran's I of Residuals for Severe Crash Counts	0.013 (p-value = 0.70)	0.132 (0.06)	0.651 (0.04)
Moran's I of Residuals for Non-Severe Counts	0.028 (p-value = 0.03)	0.192 (0.09)	0.581 (0.01)

CONCLUSIONS

This paper proposed, calibrated, and applied a Poisson log-normal multivariate CAR model, which captures zone-specific heterogeneity, correlation across response types, and spatial dependence ascribed to the latent error term. The use of Thiessen polygons to aggregate area-level crash count data is recommended, rather than using the natural tract boundaries, to ensure that high-crash locations can be uniquely assigned to a polygon zone (rather than arbitrarily assigned to or split across adjacent tracts).

This new spatial multivariate model was applied to analyze the relationship between area-level pedestrian crash counts and various land use, network, and demographic factors, including residents' proximity to schools, land use balance, transit access, network intensity, sidewalk density, and resident demographics. Walk-miles traveled were used as the exposure measure and imputed using 2005/2006 Austin Travel Survey's walk trips. Parameter estimates suggest, for example, that roadway-provision (and no doubt roadway-use) decisions have very important roles to play, as these effectively proxy here for traffic levels.

Pure, positive spatial autocorrelation (indicating clustering patterns) appears present across Austin neighborhoods, as expected (due to measurement errors that trend in space and the spatial clustering patterns of crash counts). The spatially lagged effects of cross-response correlation (estimated to be statistically and practically significant) capture missing variables that are both spatially clustered and shared across crash types, such as socio-economic variables (like ethnicity and poverty). In contrast, the model's aspatial cross-correlation ($\eta_0 = 0.712$) represents omitted variables that are meaningful for both crash-severity levels but apply within zones, more locally (like relatively poor lighting conditions and the presence of unusual sight obstructions).

From a planning and policy perspective, this paper's results reinforced the importance of advocating walking in order to reduce crash rates, as reflected by the drastic decrease in crash rates as walk miles traveled increase. Providing walking facilities (such as sidewalks and other pedestrian paths) and greater local street intensity for all road users may also reduce crash rates, per walk-mile traveled, as suggested by the conspicuous elasticity estimates for sidewalk and local-street provision in the pedestrian crash model's results. In addition, balanced land development offers a mild, positive impact in reducing severe crashes and could serve as a countermeasure to curb pedestrian fatalities. Other countermeasures may include providing

pedestrian signals that count down (to warn walkers of time remaining), pedestrian (and cyclist) overpasses/underpasses, walk beacons at popular mid-block crossings, pedestrian phases that turn on before the green signal for vehicles (crossing in the same direction), and more safety programs for vulnerable road users (like school children and disabled pedestrians), while restricting parking near intersections, as suggested in Zegeer and Bushell (2011).

Incorporating spatial effects has proved to substantially improve inference and fit statistics in analyzing area-level pedestrian crash count data. The model developed here follows Jin et al.'s specification and presents a novel alternative to the Poisson multivariate CAR model proposed by Gelfand and Vounatsou (2003) and Song et al. (2006) thanks to a more intuitive parameterization of the spatial influence (by focusing on the covariance matrix rather than the precision matrix), the ability of teasing out aspatial cross correlation from its spatially lagged counterpart, and faster computation. However, several enhancements shall be pursued. For example, more network variables should be explored, such as at-grade intersection density and link-level traffic flow (which is currently missing for all local streets and many segments with higher functional classifications). Again, from a planning and policy perspective, it is crucial to investigate what types of variables tend to generate what kind of spatial autocorrelation: a pure within-severity-level dependence, a spatially-lagged cross-severity correlation, or an aspatial cross correlation. This opens door to a new chapter of spatial count data analysis, by exploring models with spatially lagged covariate terms (e.g., the spatial Durbin model [LeSage and Pace 2009] and possibly a CAR variation with spatially lagged covariates). A temporal extension shall also be pursued to identify any time trend in the occurrence of pedestrian crash count across neighborhoods.

REFERENCES

Abdel-Aty, M., Essam Radwan, E. (2000) Modeling traffic accident occurrence and involvement. *Accident Analysis and Prevention* 32: 633–642.

Anselin, L. (1988) *Spatial Econometrics: Methods and Models.* Kluwer Academic Publisher, Norwell, MA

Besag, J. (1975) Statistical Analysis of Non-Lattice Data. *Statistician* 24 (3):179–195.

Brown, B., Yamada, I., Smith, K., Zick, C., Kowaleski-Jones, L., Fan, J. (2009) Mixed land use and walkability: Variations in land use measures and relationship with BMI, overweight, and obesity. *Health Place* 15(4): 1130-1141.

Caliendo, C., Guida, M., Parisi, A. (2007) A crash-prediction model for multilane roads. *Accident Analysis and Prevention* 39: 657- 670.

Carlin, B. P., and Banerjee, S. (2003). Hierarchical multivariate CAR models for spatiotemporally correlated survival data. *Bayesian Statistics* 7: 45–63.

Clifton, K., Burnier, C., Schneider, R., Huang, S., and Kang, M. (2008) Pedestrian Demand Model for Evaluating Pedestrian Risk Exposure. Technical Report. URL: http://www.kellyjclifton.com/MoPeD/SHAPedestrianModelingpresentation5_19_2008.pdf

Clifton, K., Burnier, C., and Fults, K. (2004) Women's involvement in pedestrian-vehicle crashes: influence of personal and environmental factors. Women's Issues in Transportation Conference proceedings 2 (35): 155-162.

Cottrill, C., and Thakuriah, P. (2010) Evaluating pedestrian crashes in areas with high low-income or minority populations. *Accident Analysis & Prevention* 42 (6): 1718-1728.

Cressie N. A. (1995) *Statistics for Spatial Data. Revised Edition.* John Wiley & Sons, Inc. New York.

Davies, R.B., Cenek, P.D., Henderson, R.J. (2005) The effect of skid resistance and texture on crash risk, International Surface Friction Conference, 2005, Christchurch, New Zealand, Transit New Zealand, Wellington.

EI-Basyouny, K., and Sayed, T. (2009) Accident prediction models with random corridor parameters. *Accident Analysis and Prevention* 41: 1118-1123.

Ewing, R. (2006) Fatal and Non-fatal Injuries. Understanding the Relationship Between Public Health and the Built Environment: A Report Prepared for the LEED-ND Core Committee. URL: http://www.activeliving.org/files/LEED%20ND%20report.pdf

Elhorst, P. (2009) Spatial Panel Data Models. In Fischer M., and Getis A. (eds.) *Handbook of Applied Spatial Analysis*: 377-407. Springer, Berlin.

FHWA (2007) Safety at unsignalized intersections. Federal Highway Adminsitration, US Department of Transportation, Washington DC. URL: http://safety.fhwa.dot.gov/intersection/unsignalized/presentations/unsig_pps_041409/long.cfm

Frank, L. and Pivo, G. (1994) Impacts of mixed use and density on utilization of three modes of travel: single-occupant vehicle, transit, and walking. *Transportation Research Record* 1466: 37-43.

Gelfand, A.E., and Vounatsou, P. (2003) Proper multivariate conditional autoregressive models

for spatial data analysis. *Biostat* 4 (1): 11-15.

Gschlöÿl, S. and Czado, C. (2008) Does a Gibbs sampler approach to spatial Poisson regression models outperform a single site MH sampler? Computational Statistics and Data Analyis 52:4184-4202.

Jin, X., Carlin, B., and Banerjee, S. (2005) Generalized Hierarchical Multivariate CAR Models for Areal Data. *Biometrics*,61: 950-961.

Leyden, K.M. (2003) Social capital and the built environment: the importance of walkable neighborhoods. *American Journal of Public Health* 93 (9): 1546–1551.

National Population Projections 2009 (2008) U.S. Census Bureau. Washington, D.C. URL: http://www.census.gov/population/projections/data/national/2009.html.

Levine, N., Kim, K., and Nitz, L. (1995) Spatial analysis of Honolulu motor vehicle crashes: I. Spatial patterns. *Accident Analysis and Prevention* 27 (5): 663-674.

Levine, N., Kim, K., and Nitz, L. (1995) Spatial analysis of Honolulu motor vehicle crashes: II. Zonal generators. *Accident Analysis and Prevention* 27 (5): 675-685.

Lord, D. (2006) Modeling motor vehicle crashes using Poisson-gamma models: examining the effects of low sample mean values and small sample size on the Estimation of the fixed dispersion parameter. *Accident Analysis and Prevention* 38 (4): 751-766.

Ma, J., Kockelman, K.M., and Damien, P. (2008) A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. *Accident Analysis and Prevention* 40 (3): 964–975.

Miaou, S-P., Song, J., and Mallick, B. (2003) Roadway traffic crash mapping: a space-time modeling approach, *Journal of Transportation & Statistics* 6 (1): 33-58.

Miaou, S-P., Song, J. (2005) Bayesian ranking of sites for engineering safety improvements: Decision parameter, treatability concept, statistical criterion, and spatial dependence. *Accident Analysis and Prevention* 37 (4): 699–720.

Miranda-Moreno, L., Morency, P., and El-Geneidy, A. (2011) The link between built environment, pedestrian activity and pedestrian-vehicle collision occurrence at signalized intersections. *Accident Analysis and Prevention* 43(5): 1624–1634.

Morency, P., and Cloutier, M.S. (2006) From targeted "black spots" to area-wide pedestrian safety. *Injury Prevention* 12: 360–364.

Naderan, A., and Shahi, J. (2009) Aggregate crash prediction models: Introducing crash generation concept. *Accident Analysis and Prevention* 42 (1): 339-346.

Neider, M. B., McCarley, J.S., Crowell, J.A., Kaczmarski, H., and Kramer, A.F. (2010) Pedestrians, vehicles, and cell phones. *Accident Analysis & Prevention* 42 (1): 589-594

Park, E.S., and Lord, D. (2007) Multivariate Poisson-lognormal models for jointly modeling crsh frequency by severity. *Transportation Research Record* No. 2019: 1-6.

Pettitt, A., Weir, I.S., and Hart, A. G. (2002) A conditional autoregressive Gaussian process for irregularly spaced multivariate data with application to modeling large sets of binary data. *Statistics and Computing* 12(4): 353-367.

Song, J., Ghosh, M., Miaou, S., and Mallick, B. (2006) Bayesian multivariate spatial models for roadway traffic crash mapping. *Journal of Multivariate Analysis* 97: 246-273.

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. (2003) WinBUGS User Manual Version 1.4. URL: http://voteview.org/manual14.pdf.

NHSTA (2011) Traffic Safety Facts 2009 Data, National Highway Traffic Safety Administration, US DOT. URL: http://www-nrd.nhtsa.dot.gov/Pubs/811394.pdf.

Valverde, J. and Jovanis, P. (2007) Identifying road segments with high risk of weather-related crashes using full Bayesian hierarchical models. Proceedings of the 86th Transportation Research Board Annual Meeting Compendium of Papers.

Wall, M. (2004) A close look at the spatial structure implied by the CAR and SAR models. *Journal of Statistical Planning & Inference* 121: 311-324.

Wang, X., and Kockelman, K. (2009) Forecasting Network Data: Spatial Interpolation of Traffic Counts Using Texas Data. *Transportation Research Record* No. 2105: 100-108.

Wang, C., Quddus, M.A. , and Ison, S.G. (2009) Impact of traffic congestion on road accidents: a spatial analysis of the M25 motorway in England. *Accident Analysis & Prevention* 41(4): 798-808.

Wang, C., Quddus, M. A., and Ison, S. G. (2011) Predicting accident frequency at their severity levels and its application in site ranking using a two-stage mixed multivariate model. *Accident Analysis & Prevention* 43: 1979-1990.

Wang, X., Kockelman, K., Lemp, J. (2012) The Dynamic Spatial Multinomial Probit Model: Analysis of Land Use Change Using Parcel-Level Data. Forthcoming in the *Journal of Transport Geography*.

Wang, Y., Kockelman, K., and Damien, P. (2012) A Spatial Autoregressive Multinomial Probit Model for Anticipating Land Use Change in Austin, Texas. Proceedings of IATBR's 13th International Conference on Travel Behavior Research Board, in Toronto.

Washington, S. P., Karlaftis, M. G., and Mannering, F. L. (2011) *Statistical and Econometric Methods for Transportation Data Analysis*. CRC Press, Chapman & Hall, Boca Raton, FL.

Weir, M., Weintraub, J., Humphreys, E., Seto, E., and Bhatia, R. (2009) An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. *Accident Analysis & Prevention* 41: 137-145.

Zegeer, C., and Bushell, M. (2011) Pedestrian crash trends and potential countermeasures from around the world. *Accident Analysis & Prevention* 44 (1): 3-11.

APPENDIX

Sampling Scheme for the Bivariate Setting

Having specified the conditional distributions of the mean crash rates, $\boldsymbol{Z} = (\boldsymbol{Z}_1,\ \boldsymbol{Z}_2)$, the focus now is the posterior distribution: $p(\beta, \boldsymbol{u}, Z, \tau, \alpha, \eta|Y) \propto L(Y|\beta, \boldsymbol{u}, Z, \tau, \alpha, \eta) \cdot \pi(\beta) \cdot \pi(\boldsymbol{u}) \cdot \pi(Z) \cdot \pi(\tau) \cdot \pi(\alpha) \cdot \pi(\eta)$, as detailed below.

The posterior distribution $p(\beta, \boldsymbol{u}, Z, \tau, \alpha, \eta|Y)$*:*

$$p(\beta, \boldsymbol{u}, Z, \tau, \alpha, \eta|Y) \propto L(Y|\ Z) \cdot p(Z|\beta, \boldsymbol{u}, \tau, \alpha, \eta) \cdot \pi(\beta) \cdot \pi(\boldsymbol{u}) \cdot \pi(\tau) \cdot \pi(\alpha) \cdot \pi(\eta)$$

$$\propto L(Y|\ Z) \cdot p(Z_1|Z_2, \beta, \boldsymbol{u}, \tau, \alpha, \eta) \cdot p(Z_2|\beta, \boldsymbol{u}, \tau, \alpha, \eta) \cdot \pi(\beta) \cdot \pi(\boldsymbol{u}) \cdot \pi(\tau) \cdot \pi(\alpha) \cdot \pi(\eta)$$

$$\propto \left(\prod_{k=1}^{2}\prod_{i=1}^{n} Z_{ik}^{y_{ik}} \cdot e^{-Z_{ik}}\right) \cdot \tau_1^{\frac{n}{2}} \cdot |D - \alpha_1 W|^{\frac{1}{2}} \cdot \exp\left\{-\frac{\tau_1}{2}[\boldsymbol{Z}_1 - \boldsymbol{m}_1 - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)]' \cdot (D - \alpha_1 W)[\boldsymbol{Z}_1 - \boldsymbol{m}_1 - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)]\right\} \cdot \tau_2^{\frac{n}{2}} \cdot$$

$|D-\alpha_2 W|^{\frac{1}{2}} \cdot \exp\left\{-\frac{\tau_2}{2}(\boldsymbol{Z}_2-\boldsymbol{m}_2)'(D-\alpha_2 W)(\boldsymbol{Z}_2-\boldsymbol{m}_2)\right\} \cdot e^{\Sigma_i u_i} \cdot [\mathrm{Gamma}(\theta,\theta)]^n \cdot$

$[\mathrm{Gamma}(1,0.1)]^2 \cdot [\mathrm{Unif}(0,1)]^2[\mathrm{N}(0,100)]^2$

where $\boldsymbol{m}_1 = X'\boldsymbol{\beta}_1 + \ln(\boldsymbol{E}) + \boldsymbol{u}$ and $\boldsymbol{m}_2 = X'\boldsymbol{\beta}_2 + \ln(\boldsymbol{E}) + \boldsymbol{u}$.

Here, type-specific covariates $\beta = (\beta_1, \beta_2)$ are assumed to follow a flat normal prior, reflected as being centered around zero with a large variance term, e.g., $\beta_1 \sim N(\mathbf{0}, 10^5 I)$ and $\beta_2 \sim N(\mathbf{0}, 10^5 I)$. The precision parameters $\tau = (\tau_1, \tau_2)$ are assumed to follow a rather diffused Gamma distribution, e.g., $\mathrm{Gamma}(1, 0.1)$ with mean 10 and variance 100. Spatial autocorrelation coefficients, $\alpha = (\alpha_1, \alpha_2)$, are assigned a uniform prior over the interval (0, 1), denoted by $\mathrm{Unif}(0, 1)$. The two "bridging" parameters η_0 and η_1 follow a diffused normal prior, $N(0, 10^2)$.

Conditional distributions of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$

$$p(\boldsymbol{\beta}_1|\cdot) \propto \exp\left\{-\frac{\tau_1}{2}[\boldsymbol{Z}_1 - X'\boldsymbol{\beta}_1 - \ln(\boldsymbol{E}) - \boldsymbol{u} - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})]'(D - \alpha_1 W)[\boldsymbol{Z}_1 - X'\boldsymbol{\beta}_1 - \ln(\boldsymbol{E}) - \boldsymbol{u} - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})]\right\}$$

$$\propto \exp\left\{\left[\boldsymbol{\beta}_1 - \frac{1}{2}{A_1}^{-1}\Omega_1\right]' \cdot A_1 \cdot \left[\boldsymbol{\beta}_1 - \frac{1}{2}{A_1}^{-1}\Omega_1\right]\right\}$$ (Using completing the squares technique)

$$\propto N\left(\frac{1}{2}{A_1}^{-1}\Omega_1, {A_1}^{-1}\right)$$

where $A_1 = -\frac{\tau_1}{2}X(D-\alpha_1 W)X'$ and $\Omega_1 = -\frac{\tau_1}{2}[2X(D-\alpha_1 W)(\boldsymbol{Z}_1 - \ln(\boldsymbol{E}) - \boldsymbol{u}) - X(D-\alpha_1 W)(\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})]$.

$$p(\boldsymbol{\beta}_2|\cdot) \propto \exp\Big\{-\frac{\tau_1}{2}\big[-(\boldsymbol{Z}_1 - \ln(\boldsymbol{E}) - \boldsymbol{u})'(D-\alpha_1 W)(\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})$$
$$+ (X'\boldsymbol{\beta}_1)'(D-\alpha_1 W)(\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})$$
$$- (\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})'(\eta_0 I + \eta_1 W)'(D-\alpha_1 W)\big(\boldsymbol{Z}_1 - X'\boldsymbol{\beta}_1 - \ln(\boldsymbol{E}) - \boldsymbol{u}$$
$$- (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})\big)\big]\Big\}$$
$$\cdot \exp\Big\{-\frac{\tau_2}{2}[(\boldsymbol{Z}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})'(D-\alpha_2 W)(-X'\boldsymbol{\beta}_2)$$
$$- (X'\boldsymbol{\beta}_2)'(D-\alpha_2 W)(\boldsymbol{Z}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u}) + (X'\boldsymbol{\beta}_2)'(D-\alpha_2 W)(X'\boldsymbol{\beta}_2)]\Big\}$$

$$\propto \exp\Big\{\Big[\boldsymbol{\beta}_2 - \tfrac{1}{2}{A_2}^{-1}\Omega_2\Big]' \cdot A_2 \cdot \Big[\boldsymbol{\beta}_2 - \tfrac{1}{2}{A_2}^{-1}\Omega_2\Big]\Big\} \propto N\Big(\tfrac{1}{2}{A_2}^{-1}\Omega_2, {A_2}^{-1}\Big)$$

where $A_2 = \frac{\tau_2}{2}X(D-\alpha_2 W)X' - \frac{\tau_1}{2}X(\eta_0 I + \eta_1 W)'(D-\alpha_1 W)(\eta_0 I + \eta_1 W)X'$ and $\Omega_2 =$

$-[\tau_1((\boldsymbol{Z}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})'(\eta_0 I + \eta_1 W)'(D-\alpha_1 W)(\eta_0 I + \eta_1 W)X' - (\boldsymbol{Z}_1 - \ln(\boldsymbol{E}) - \boldsymbol{u})'(D-$

$\alpha_1 W)(\eta_0 I + \eta_1 W)X' + \beta_1' X(D-\alpha_1 W)(\eta_0 I + \eta_1 W)X') + \tau_2(\boldsymbol{Z}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})'(D-\alpha_2 W)X']$.

Conditional distributions of $u_1, u_2, \ldots, u_n$

$p(\boldsymbol{u}|\cdot) \propto \exp\Big\{-\frac{\tau_1}{2}[\boldsymbol{Z}_1 - X'\boldsymbol{\beta}_1 - \ln(\boldsymbol{E}) - \boldsymbol{u} - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})]'(D-$

$\alpha_1 W)[\boldsymbol{Z}_1 - X'\boldsymbol{\beta}_1 - \ln(\boldsymbol{E}) - \boldsymbol{u} - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})] - \frac{\tau_2}{2}(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 -$

$\ln(\boldsymbol{E}) - \boldsymbol{u})'(D-\alpha_2 W)(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})\Big\} \cdot e^{\sum_i u_i}$.

It is difficult to draw $\boldsymbol{u} = u_1, u_2, \ldots, u_n$ simultaneously. Alternatively, one may draw these *n* heterogeneity error terms sequentially, as described below.

$$p(u_i|\cdot) \propto \exp\left\{-\frac{\tau_1}{2}[\boldsymbol{Z}_1 - X'\boldsymbol{\beta}_1 - \ln(\boldsymbol{E}) - \boldsymbol{u} - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})]'(D - \alpha_1 W)[\boldsymbol{Z}_1 - X'\boldsymbol{\beta}_1 - \ln(\boldsymbol{E}) - \boldsymbol{u} - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})] - \frac{\tau_2}{2}(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})'(D - \alpha_2 W)(\boldsymbol{Z}_2 - X'\boldsymbol{\beta}_2 - \ln(\boldsymbol{E}) - \boldsymbol{u})\right\} \cdot e^{u_i} \quad (i=1, 2, \ldots, n)$$

The conditional posterior of u_i does not follow any known distribution and thus cannot be sampled using Gibbs method. Metropolis-Hastings algorithm (Metropolis et al. 1953, Carlin and Louis 2009) and a more recent development, the generalized direct sampling method (Walker et al. 2011), can be utilized in drawing these quantities.

Conditional distribution of $\boldsymbol{Z}_1$

$$p(\boldsymbol{Z}_1|\cdot) \propto \left(\prod_{i=1}^{n} Z_{i1}^{y_{i1}} \cdot e^{-Z_{i1}}\right) \cdot \exp\left\{-\frac{\tau_1}{2}[\boldsymbol{Z}_1 - \boldsymbol{m}_1 - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)]' \cdot (D - \alpha_1 W)[\boldsymbol{Z}_1 - \boldsymbol{m}_1 - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)]\right\}$$

where $\boldsymbol{m}_1 = X'\boldsymbol{\beta}_1 + \ln(\boldsymbol{E}) + \boldsymbol{u}$ and $\boldsymbol{m}_2 = X'\boldsymbol{\beta}_2 + \ln(\boldsymbol{E}) + \boldsymbol{u}$. Due to a non-Gaussian first stage, the conditional posterior of $\boldsymbol{Z}_1$ does not follow a known form.

Conditional distribution of $\boldsymbol{Z}_2$

$$p(\boldsymbol{Z}_2|\cdot) \propto \left(\prod_{i=1}^{n} Z_{i2}^{y_{i2}} \cdot e^{-Z_{i2}}\right) \cdot \exp\left\{-\frac{\tau_1}{2}[\boldsymbol{Z}_1 - \boldsymbol{m}_1 - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)]' \cdot (D - \alpha_1 W)[\boldsymbol{Z}_1 - \boldsymbol{m}_1 - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)] - \frac{\tau_2}{2}(\boldsymbol{Z}_2 - \boldsymbol{m}_2)'(D - \alpha_2 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)\right\}$$

Similar to the conditional posterior for $\boldsymbol{Z}_1$, the conditional posterior $p(\boldsymbol{Z}_2|\cdot)$ does not follow a standard distribution either.

Conditional distribution of τ_1

$$p(\tau_1|\cdot) \propto \tau_1^{n/2}\exp\left(-\frac{\tau_1}{2}\cdot T_1\right) \propto \text{Gamma}\left(\frac{n}{2}+1, \frac{T_1}{2}\right)$$

where $T_1 = [\boldsymbol{Z}_1 - \boldsymbol{m}_1 - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)]' \cdot (D - \alpha_1 W)[\boldsymbol{Z}_1 - \boldsymbol{m}_1 - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)]$.

Conditional distribution of τ_2

$$p(\tau_2|\cdot) \propto \tau_2^{n/2}\exp\left(-\frac{\tau_2}{2}\cdot T_2\right) \propto \text{Gamma}\left(\frac{n}{2}+1, \frac{T_2}{2}\right)$$

where $T_2 = (\boldsymbol{Z}_2 - \boldsymbol{m}_2)'(D - \alpha_2 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)$.

Conditional distributin of α_1

$$p(\alpha_1|\cdot) \propto |D - \alpha_1 W|^{\frac{1}{2}}$$
$$\cdot \exp\left\{-\frac{\tau_1}{2}[\boldsymbol{Z}_1 - \boldsymbol{m}_1 - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)]'\right.$$
$$\left.\cdot (D - \alpha_1 W)[\boldsymbol{Z}_1 - \boldsymbol{m}_1 - (\eta_0 I + \eta_1 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)]\right\}$$

Conditional distribution of α_2

$$p(\alpha_2|\cdot) \propto |D - \alpha_2 W|^{\frac{1}{2}} \cdot \exp\left\{-\frac{\tau_2}{2}(\boldsymbol{Z}_2 - \boldsymbol{m}_2)'(D - \alpha_2 W)(\boldsymbol{Z}_2 - \boldsymbol{m}_2)\right\}$$

Conditional distribution of η_0

$$p(\eta_0|\cdot) \propto \exp\left\{-\frac{\tau_1}{2}[(\boldsymbol{Z}_2-\boldsymbol{m}_2)'(\eta_0 I+\eta_1 W)'(D-\alpha_1 W)(\eta_0 I+\eta_1 W)(\boldsymbol{Z}_2-\boldsymbol{m}_2)\right.$$

$$\left.-2(\boldsymbol{Z}_1-\boldsymbol{m}_1)'(D-\alpha_1 W)(\eta_0 I+\eta_1 W)(\boldsymbol{Z}_2-\boldsymbol{m}_2)]\right\}$$

Assume $\eta_1 = 0$, $p(\eta_0|\cdot)$ is then written as:

$$p(\eta_0|\cdot) \propto \exp\left\{-\frac{\tau_1}{2}[{\eta_0}^2(\boldsymbol{Z}_2-\boldsymbol{m}_2)'(D-\alpha_1 W)(\boldsymbol{Z}_2-\boldsymbol{m}_2)\right.$$

$$\left.-2\eta_0(\boldsymbol{Z}_1-\boldsymbol{m}_1)'(D-\alpha_1 W)(\boldsymbol{Z}_2-\boldsymbol{m}_2)]\right\}$$

$$\propto \exp\left\{-\frac{\tau_1}{2}(\boldsymbol{Z}_2-\boldsymbol{m}_2)'(D-\alpha_1 W)(\boldsymbol{Z}_2-\boldsymbol{m}_2)\left(\eta_0-\frac{(\boldsymbol{Z}_1-\boldsymbol{m}_1)'(D-\alpha_1 W)(\boldsymbol{Z}_2-\boldsymbol{m}_2)}{(\boldsymbol{Z}_2-\boldsymbol{m}_2)'(D-\alpha_1 W)(\boldsymbol{Z}_2-\boldsymbol{m}_2)}\right)^2\right\}$$

$$\propto N\left(\frac{(\boldsymbol{Z}_1-\boldsymbol{m}_1)'(D-\alpha_1 W)(\boldsymbol{Z}_2-\boldsymbol{m}_2)}{(\boldsymbol{Z}_2-\boldsymbol{m}_2)'(D-\alpha_1 W)(\boldsymbol{Z}_2-\boldsymbol{m}_2)}, [\tau_1(\boldsymbol{Z}_2-\boldsymbol{m}_2)'(D-\alpha_1 W)(\boldsymbol{Z}_2-\boldsymbol{m}_2)]^{-1}\right)$$

An aspatial model (with cross-type correlations) assumes $\eta_1 = 0$, $\alpha_1 = 0$, and $\alpha_2 = 0$.

The Trivariate Poisson-Lognormal CAR Model

A trivariate MCAR model assumes that the spatial random effects are represented as $\boldsymbol{\phi} = (\boldsymbol{\phi'}_1, \boldsymbol{\phi'}_2, \boldsymbol{\phi'}_3)$, where $\boldsymbol{\phi}_1$ is an *n* by 1 vector of spatial random effects for the latent rates of crash type 1 (or fatal and incapacitating injury), as is the case for crash type 2 (or non-incapacitating injury) and type 3 (possible and no injury)**.** A question emerges as to the sequence of these conditional distributions. A way to determine such question is to try all possible 6 combinations and choose the model with best goodness-of-fit.

For ease of exposition, assume the sequence of conditional distributions as such: $p(\boldsymbol{\phi}) = p(\boldsymbol{\phi}_1|\ \boldsymbol{\phi}_2,\ \boldsymbol{\phi}_3)\cdot p(\boldsymbol{\phi}_2|\ \boldsymbol{\phi}_3)\cdot p(\boldsymbol{\phi}_3)$. Based on multivariate normal theory, the joint distribution of $\boldsymbol{\phi}$ takes the form: $\begin{pmatrix}\boldsymbol{\phi}_1\\ \boldsymbol{\phi}_2\\ \boldsymbol{\phi}_3\end{pmatrix} \sim N\left(\begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\\ \boldsymbol{\mu}_3\end{pmatrix}, \begin{bmatrix}\Sigma_{11} & \Sigma_{12} & \Sigma_{13}\\ \Sigma_{12} & \Sigma_{22} & \Sigma_{23}\\ \Sigma_{13} & \Sigma_{23} & \Sigma_{33}\end{bmatrix}\right)$, where the *n* by 1 vector $\boldsymbol{\mu}_p$ indicates the mean for response type *p* (*p*=1, 2,3), Σ_{pl} is an *n* by *n* matrix describing the covariance structure between response type *p* and *l*. The marginal distribution of $\boldsymbol{\phi}_3$ can be written as: $p(\boldsymbol{\phi}_3) \sim \boldsymbol{N}(\boldsymbol{\mu}_3, \Sigma_{33})$, with $\boldsymbol{\mu}_3 = \mathbf{0}$ and $\boldsymbol{\Sigma}_{33} = [\tau_3(\mathbf{D}-\rho_3\mathbf{W})]^{-1}$. The marginal distribution of $(\boldsymbol{\phi}_2, \boldsymbol{\phi}_3)$ can be obtained by removing irrelevant elements (with respect to $\boldsymbol{\phi}_2$ and

$\boldsymbol{\phi}_3$) from the full distribution, leading to a multivariate normal distribution:
$\begin{pmatrix}\boldsymbol{\phi}_2\\ \boldsymbol{\phi}_3\end{pmatrix} \sim N\left(\begin{pmatrix}\boldsymbol{\mu}_2\\ \boldsymbol{\mu}_3\end{pmatrix}, \begin{bmatrix} _{22} & _{23}\\ _{23} & _{33}\end{bmatrix}\right)$.

$\boldsymbol{\phi}_2|\boldsymbol{\phi}_3 \sim N(\mathrm{A}_{23}\boldsymbol{\phi}_3,\ [(D-\rho_2 W)\ _2]^{-1})$, where A_{23} describes the aspatial correlation between response types 2 and 3, as well as the spatially-lagged correlation between the two response types, formally: $\mathrm{A}_{23} = \eta_{0,23} I + \eta_{1,23} W$.

$\boldsymbol{\phi}_1|\boldsymbol{\phi}_2, \boldsymbol{\phi}_3 \sim N(\mathrm{A}_{13}\boldsymbol{\phi}_3 + \mathrm{A}_{12}\boldsymbol{\phi}_2,\ [(D-\rho_1 W)\ _1]^{-1})$, where A_{13} and A_{12} capture the aspatial and spatially-lagged correlation across response types 1 and 3, and response types 1 and 2, formally: $\mathrm{A}_{13} = \eta_{0,13} I + \eta_{1,13} W$ and $\mathrm{A}_{12} = \eta_{0,12} I + \eta_{1,12} W$