Autologistic models for benchmark risk/vulnerability assessment of urban terrorism outcomes

Jingyu Liu

*Interdisciplinary Program in Statistics
University of Arizona, Tucson, AZ, USA*

Walter W. Piegorsch*

*Interdisciplinary Program in Statistics and Department of Mathematics
University of Arizona, Tucson, AZ, USA*

A. Grant Schissler

*Department of Mathematics and Statistics
University of Nevada, Reno, NV, USA*

and

Susan L. Cutter

*Hazards & Vulnerability Research Institute and Department of Geography
University of South Carolina, Columbia, SC, USA*

Summary

We develop a quantitative methodology to characterize vulnerability among 132 U.S. urban centers (‘cities’) to terrorist events, applying a place-based vulnerability index to a database of terrorist incidents and related human casualties. A centered autologistic regression model is employed to relate urban vulnerability to terrorist outcomes and also to adjust for autocorrelation in the geospatial data. Risk-analytic ‘benchmark’ techniques are then incorporated into the modeling framework, wherein levels of high and low urban vulnerability to terrorism are identified. This new, translational adaptation of the risk-benchmark approach, including its ability to account for geospatial autocorrelation, is seen to operate quite flexibly in this socio-geographic setting.

*Address for correspondence: Walter W. Piegorsch, University of Arizona, Keating Bldg., Tucson, AZ 85721, USA. E-mail: piegorsch@math.arizona.edu

Keywords: Benchmark dose; Centered autologistic model; Geospatial analysis; Maximum pseudo-likelihood; Quantitative risk analysis; Spatial autocorrelation.

1 Introduction: urban vulnerability assessment

Quantifying vulnerability to natural or human-induced hazardous events is a critical component in modern risk/hazard analysis. For example, multiple factors can affect a city’s or other urban locale’s response to a hazardous impact, and it is possible to quantify and/or index the city’s vulnerabilities associated with that response. Work in this area often characterizes those populations, localities, ecosystems, and other entities at risk to hazardous events, and such efforts have seen substantive development (Peng et al., 2005; Mazzorana et al., 2009; Santini et al., 2010; MacKenzie, 2014). Many localities exhibit vulnerabilities due to their physical and geographic characteristics, others contain populations that are socially vulnerable, while still others possess built-environment factors that make them vulnerable to certain hazards. In some locations, a combination of vulnerabilities affects a community’s disaster resiliency and response. Indeed, urban hazards are by their very nature multidimensional processes, with many factors underlying the occurrence and incidence of each adverse event. While it is difficult to distill such processes down to a few explanatory component(s), efforts to do so can provide important guidance to many interested parties—municipal planners, emergency managers, social services providers, disaster relief organizations, etc.—when assessing their locality’s level of vulnerability to natural or human-induced hazards.

Towards this end, we previously exploited a series of vulnerability indices (Cutter et al., 2003; Borden et al., 2007) to benchmark urban susceptibility to adverse events, focusing on terrorism outcomes (Piegorisch et al., 2007). Data from that study consisted of binary incidence indicators, Y_i , indexed over 132 U.S. metropolitan areas (‘cities’) deemed most susceptible to environmental and human-induced hazards by the U.S. Federal government ($i = 1, \dots, 132$). Of interest with these data was the probability, π_i , that the i th urban center would have experienced casualties due to terrorist incidents during the 35-year period (1970–

2004) over which the terrorism outcomes were collected. Our goal was to study how place-based factors of urban vulnerability related to terrorist attacks, and to use statistical features of those relationships to indicate and characterize the associated levels of urban vulnerability. In particular, we translated a technique from toxicological risk assessment—called benchmark risk analysis (Crump, 2012)—to characterize an urban area’s susceptibility to adverse events. Broadly speaking, the benchmark approach models π_i as a function of some target predictor variable, x_i , quantifying the hazardous input, and builds a risk function $R(x)$ that models how the mean risk/vulnerability of the study subjects—here, the cities—changes over x . It then asks: at which value of x does $R(x)$ achieve a pre-specified *benchmark response* (BMR). The solution is generically called the *benchmark dose*, or in our case the *benchmark index* (BMI) above which cities are at increased vulnerability to the hazard. The particular mathematical expression for the BMI will depend upon the statistical model chosen to represent π and R as functions of x .

Given the particular model, a point estimate, $\widehat{\text{BMI}}_{\text{BMR}}$, for the BMI, and corresponding $1 - \alpha$ confidence bounds can be calculated from the observed data pairs (x_i, Y_i) (see below). In most settings, the adverse outcome is detrimental, and thus only a lower confidence limit is required; this is denoted by BMIL_{BMR} (Crump, 1995). In both cases, where needed for clarity the subscript explicitly reminds us of the dependence on the BMR input value.

We should note that the terms ‘risk’ and ‘vulnerability’ carry very specific meanings in the literature, although at times different writers utilize them in different ways. Our use of vulnerability here is similar to that defined by Haimes (2006), the manifestation of a location’s inherent condition(s) that can be exploited to adverse effect. We use the term risk in a more colloquial sense, however, suggesting some general source or instance of threat, danger, or peril, rather than any technical definition or notion.

In our earlier work (Piegorisch et al., 2007), we manipulated the benchmark paradigm to identify urban BMIs at varying levels of BMR, corresponding to increasing urban vulnerability to terrorism casualties based on our 132 cities data. The result provided a quantification of

vulnerability in these various urban locations, based on their own particular socioeconomic, physical, and built-environment factors. We did so by constructing a single place-based vulnerability index (PVI) as a weighted average of these individual components; the PVI served as our x variable (see our 2007 article for full details). The consequent estimates of BMI provided benchmark points on this PVI scale at which urban vulnerability to terrorist casualties grew large, relative to lower values on the same scale.

A critical feature missing in Piegorsch et al. (2007), however, was recognition that one city’s terrorism vulnerability could be affected by its proximity to another city, i.e., potential spatial autocorrelation between the Y_i s can affect calculation of $\widehat{\text{BMI}}_{\text{BMR}}$. Indeed, to our knowledge no approach currently exists that allows for calculation of benchmark doses/indices/etc. when potential autocorrelation exists in the data. Here, we propose an extension of the benchmark dose approach for use with spatially autocorrelated observations, and apply it to our 132 cities data. To do so, we expand our linear predictor to include a spatial autocovariate alongside the target benchmark predictor, a tactic often seen to assuage concerns with the effects of geospatial autocorrelation (Koutsias, 2003; de Frutos et al., 2007; Vaughn et al., 2015). Our goal is to develop a valid risk-analytic technology that can relate urban vulnerability to terrorist activity while taking into account possible spatial autocorrelation after adjusting for place-based factors of the urban vulnerability. We then manipulate statistical features of the relationship to indicate and characterize the different levels of urban vulnerability to terrorist events in our data set.

In Sec. 2, below, we reintroduce the 132 cities data from Piegorsch et al. (2007), including an updated version of the PVI more amenable to benchmark calculations with autocorrelated observations. Sec. 3 then describes a series of autologistic models we consider to analyze the data, along with details on how the risk-benchmark calculations can be incorporated into them. Sec. 4 follows with a short Monte Carlo simulation to study operating characteristics of the consequent BMI and BMIL estimates. Sec. 5 applies the new methods to the 132 cities data, while Sec. 6 ends with a brief discussion. All calculations we present below are

performed in the **R** programming environment (R Core Team, 2016).

2 Example: urban vulnerability data

To build a foundation for use of our autocorrelated risk-benchmark methodology, we employ data on urban vulnerability to terrorism events from our earlier work (Borden et al., 2007; Piegorsch et al., 2007). Therein, we examined 132 U.S. metropolitan locations (‘cities’) deemed at greatest risk to terrorist hazards; see the display in Table 1. The cities are also listed in Piegorsch et al. (2007), where full details of the larger data set can be found. Summarizing from that source, the binary outcomes, Y_i , indicate whether or not the i th city experienced a terrorist-related event during the 35-year period 1970–2004, in which one or more human casualties (injuries or deaths) were recorded. These values also appear in Table 1. Notice that of the 132 cities in the table, $\sum Y_i = 36$, or 27.3%, reported a casualty event.

Associated with these adverse outcomes, each i th city is assigned a PVI that acts as our target predictor variable x_i . This PVI is a shifted version of the original place-based index built from separate socioeconomic, physical, and built-environment measures in Piegorsch et al. (2007): our construction requires the target predictor to be non-negative while the Piegorsch et al. vulnerability index was manipulated to have roughly zero mean. We discuss our shifted PVI in more detail below.

Of interest is the probability, π , that a city experiences terrorist-related casualties during the time span (1970–2004) under study. Our goal is to relate π to the city’s broader place-based vulnerability, as quantified by the PVI.

The responses in Table 1 are dichotomous and can be viewed as binomial: $Y_i \sim \text{Bin.}(1, \pi_i)$. To model the relationship between π and the PVI, Piegorsch et al. (2007) employed a binary generalized linear model (GLiM):

$$\pi_i = P[Y_i = 1] = g^{-1}(\eta_i),$$

where the linear predictor η_i was taken as the simple linear form $\beta_0 + \beta_1 x_i$ and the single covariate x_i was the PVI. The β -values were unknown regression parameters, and $g(\cdot)$ was the link function of the GLiM. They then translated this model into a form amenable to the sort of benchmark risk analysis seen in toxicological risk assessment (Crump, 1984): after fitting the GLiM to the data, point estimates of the unknown β -parameters were manipulated into estimates of the BMI for some fixed BMR. Associated lower confidence limits, BMILs, were then constructed under the model, to serve as benchmark points marking various levels of urban vulnerability to terrorist casualties. As noted above, however, missing from this analysis was the recognition that responses among cities closely located to each other could exhibit spatial autocorrelation. How to adjust the benchmark analysis to account for such autocorrelation is the focus of the methods presented in the next section.

3 Autologistic modeling

Motivated by the vulnerability data in Table 1, we consider how to model a set of binary observations as a function of a target regressor variable while adjusting for extant spatial autocorrelation. We assume the relationship between the variables may be affected by the observations' proximities to each other (Chun & Griffith, 2013, Chapter 2) leading to the development of local-area spatial statistics. We appeal to the well-known logistic regression model, which is widely applied with binary data (Hosmer et al., 2013). A simple and direct way to account for spatial autocorrelation within a logistic model allows the response probability, π_i , at the i th spatial location—here, the i th city—to depend upon the other observed binary responses in some pre-defined neighborhood, N_i , around that city. Besag (1972) first proposed such a model to incorporate spatial autocorrelation, the conditional autologistic model

$$\pi_i = P[Y_i = 1 | x_i, y_j, j \neq i] = \frac{1}{1 + \exp\{-\beta_0 - \beta_1 x_i - \beta_2 \sum_{j \in N_i} a_{ij} y_j\}} . \quad (1)$$

In our notation, Y_i is the outcome indicator for casualties, x_i is the PVI, the a_{ij} s stipulate the i th neighborhood N_i (see below), and $\sum_{j \in N_i} a_{ij}y_j$ is a spatial autocovariate constructed around each Y_i . We refer to β_2 as the spatial autocorrelation parameter. When $\beta_2 = 0$, no spatial autocorrelation exists in the data, and the autologistic model then reduces to a standard logistic GLiM (which we call the ‘independence model’).

To define the neighborhoods, N_i , for our 132 cities, we consider a simple adjacency characterization: set $a_{ij} = 1$ if the j th metropolitan area is physically adjacent to the i th metropolitan area, i.e., it shares any edge boundary—a ‘rook’ connection, as opposed to ‘queen’ edge-and-corner connections—as delineated by the border definitions used in Borden et al. (2007). Otherwise set $a_{ij} = 0$. The corresponding adjacency matrix $\mathbf{A} = \{a_{ij}\}$ defines a simple neighborhood structure among the 132 cities in the data set. Fig. S1 and Table S1 in a supplemental document specify this adjacency structure across our 132 cities; Fig. S1 also maps the casualty incidence.

3.1 Centered autologistic model

The autologistic model in (1) is a popular choice for modeling spatially dependent binary data. Unfortunately, Caragea & Kaiser (2009) identified some irregularities with the traditional autologistic form. In (1), the autocovariate modeling the dependence between Y_i and the other Y_j s is non-negative for any i . Consequently, the odds that a response equals 1 under the traditional model relative to the odds that the response equals 1 under the independence model (the standard logistic GLiM) can change only in one direction for any nonzero neighbors. A simple argument shows that this leads to an anomaly: assume for convenience that there is positive autocorrelation in the data under a simple adjacency neighborhood structure. (Similar arguments will hold under negative autocorrelation.) Then the regression coefficient for the autocovariate should be positive. In fact, the autocovariate under the model in (1) is the number of nonzero neighbors, and if the majority of a response’s neighbors are 0, the odds that the response equals 1 under that model relative to the in-

dependence model is expected to decrease. Yet, as the autocovariate under the traditional model is always non-negative, that odds ratio can only increase given a positive regression coefficient for the autocovariate. Any meaningful autocovariate should be negative in this situation (Hughes et al., 2011).

To account for this anomaly, Caragea & Kaiser (2009) offered a correction that extends the model into a centered autologistic form, by redefining the spatial autocovariate (essentially, by centering it):

$$\pi_i = P[Y_i = 1 | x_i, y_j, j \neq i] = \frac{1}{1 + \exp\{-\beta_0 - \beta_1 x_i - \beta_2 \sum_{j \in N_i} a_{ij}(y_j - \mu_j)\}} \quad (2)$$

where the new quantity $\mu_j = E[Y_j | \beta_2 = 0] = \{1 + \exp(-\beta_0 - \beta_1 x_j)\}^{-1}$ is the expected value of Y_j under an independence model with no spatial autocorrelation. Centering makes the estimates more stable and interpretable, and as such the centered autologistic construction in (2) provides a direct and simple way to build spatial dependency into a binary regression GLiM (Hughes et al., 2011; Kolaczyk & Csárdi, 2014, §8.3). We therefore turn to (2) for constructing quantitative risk/vulnerability assessments with such data.

3.2 Maximum pseudo-likelihood estimation

Allowing for spatial autocorrelation induces dependencies among the observations, complicating standard likelihood analysis with our centered autologistic model. Instead, we estimate the unknown β -parameters in (2) via maximization of the pseudo-likelihood function proposed by Besag (1975). Besag found that by multiplying together the conditional probability distributions of the Y_i s given their neighbors, the resulting pseudo-likelihood

$$L_P(\beta) = \prod_{i=1}^n P[Y_i = 1 | x_i, y_j, j \neq i] = \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i}$$

possesses many of the same features as the usual likelihood function, despite any existing dependencies among the observations. In particular, maximum pseudo-likelihood estimates

(MPLEs) of the β -parameters are consistent and asymptotically normal under typical regularity conditions (Arnold & Strauss, 1991).

3.3 Autologistic benchmark risk analysis

To our knowledge, no previous work has attempted to incorporate spatial correlation—or any other form of autocorrelation—into a benchmark risk analysis. As we show below, doing so with the autologistic model requires nontrivial manipulation of the standard benchmark approach. For the binary-data setting considered here, we continue to relate a pre-specified level of outcome response—the BMR—to the response probability $\pi(x)$ viewed as a function of $x = \text{PVI}$. The goal is to find the smallest positive x at which some background-adjusted function of $\pi(x)$ equals the BMR. The solution is the benchmark index, past which a city’s vulnerability to terrorist casualties is elevated to unacceptable levels for that BMR. The ‘background’ adjustment is included to account for extemporaneous or spontaneous factors out of the risk assessor’s control; a prototypical example is correction for spontaneous (zero-exposure) tumor incidence when assessing the risk of a carcinogenic exposure. We discuss this adjustment further below.

An added complexity with the centered autologistic model in (2) is the presence of the centered autocovariate $\sum_{j \in N_i} a_{ij}(y_j - \mu_j)$. The covariate is designed to improve precision in the model by accounting for the recognized source of spatial autocorrelation; however, it also introduces additional terms that prevent us from simply inverting the estimated regression model to find the $\widehat{\text{BMI}}_{\text{BMR}}$. A similar challenge occurs when calculating benchmark doses in epidemiological studies, where along with the exposure/dose variable of interest, x , investigators often record secondary covariates felt to describe known sources of variation for the disease outcome under study. To overcome this, Budtz-Jørgensen et al. (2001) proposed a clever definition for the benchmark response: view BMR as a specified proportional increase over zero-level background in the odds of an adverse event, evaluated at the same constellation of secondary covariates. At least for logistic-type models, the functional dependency on

secondary covariates then cancels out in the various ratio operations. Budtz-Jørgensen et al. provided no guidance on selection of the BMR here, but we might imagine in our terrorism vulnerability application taking $\text{BMR} = 10$ or 25 , say, since an increase of this magnitude in the odds could be a useful marker for practical use. In any case, we will follow standard practice and avoid selecting a value for BMR that deviates drastically from the range of corresponding responses seen in the data.

Employed with the centered autologistic regression model in (2), we therefore define the BMI as the dose resulting in a pre-specified increase (of BMR multiples) in the odds for an abnormal response. That is, solve for x in

$$\text{BMR} = \frac{\pi(x|\boldsymbol{\beta})/\{1 - \pi(x|\boldsymbol{\beta})\}}{\pi(0|\boldsymbol{\beta})/\{1 - \pi(0|\boldsymbol{\beta})\}}, \quad (3)$$

where $\pi(\cdot|\boldsymbol{\beta})$ is the centered autologistic response probability based on (2) and $\boldsymbol{\beta} = [\beta_0 \ \beta_1 \ \beta_2]^T$ is the vector of unknown autologistic coefficients. Note for adverse outcomes that we anticipate $\beta_1 > 0$, representing an underlying increase in the probability of an adverse event as x increases. This typically defines a proportional rise in odds, thus we operate with $\text{BMR} > 1$.

One must be careful, however, when employing the relationship in (3) within our spatial setting. The target predictor x is not an isolated value: for any x , definitive or abstract, there is some corresponding location, say, s , with some neighborhood, N_s , underlying it. Write this value of x as x_s . The odds ratio in (3) is not calculated across locations, rather, it is calculated with regard to the location, s , associated with the chosen x_s . The numerator odds of observing an adverse event for the arbitrary location s are $\pi(x_s|\boldsymbol{\beta})/\{1 - \pi(x_s|\boldsymbol{\beta})\}$. The denominator odds against which the numerator is compared are calculated at $x_s = 0$, but as if they occur at the same spatial location s with the same neighborhood N_s . Thus the odds ratio in (3) simplifies to

$$\frac{\pi(x_s|\boldsymbol{\beta})/\{1 - \pi(x_s|\boldsymbol{\beta})\}}{\pi(x_s = 0|\boldsymbol{\beta})/\{1 - \pi(x_s = 0|\boldsymbol{\beta})\}} = \frac{\exp(\beta_0 + \beta_1 x_s + \beta_2 \sum_{t \in N_s} a_t(y_t - \mu_t))}{\exp(\beta_0 + \beta_2 \sum_{t \in N_s} a_t(y_t - \mu_t))} = \exp\{\beta_1 x_s\}.$$

Setting this equal to the BMR then yields the unique BMI at the given BMR:

$$\text{BMI}_{\text{BMR}} = \frac{\log(\text{BMR})}{\beta_1}. \quad (4)$$

This is the vulnerability index resulting in a $(\text{BMR}-1)$ -fold increase in odds of an adverse event, relative to a zero value of the index.

Before continuing, we should clarify what exactly is meant by a zero-valued PVI. In order to give it a standardized flavor, but for no other compelling reason, the placed-based index originally presented by Piegorsch et al. (2007), say, PVI07, had roughly zero mean and unit variance. For use in (3), however, setting $x = 0$ has a specific interpretation: a PVI of zero is the baseline index against which a city’s PVIs is compared. That is, a city whose PVI is zero exhibits the smallest, ‘spontaneous’ level of vulnerability imaginable for the population of cities under consideration. A natural tactic then would take the original placed-based values in Piegorsch et al. (2007) and add $|\min\{\text{PVI07}\}|$ to each. In the 132-city data set, $\min\{\text{PVI07}\} = -2.4208$ belongs to Juneau, AK. We felt this would be unwise, however, since it gives Juneau a PVI of exactly zero, and this has the effect of comparing all other cities in the study to Juneau’s particular vulnerability status. Instead we shifted the original Piegorsch et al. index by adding a suitably large positive constant to it; we chose $\text{PVI} = \text{PVI07} + |\Phi^{-1}(0.0001)|$, where $\Phi^{-1}(0.0001) = -3.719$ from the standard normal c.d.f. In effect, one might say that this marks an abstract baseline for odds ratio comparisons in (3) as a similarly located ‘city’ whose vulnerability characteristics place it in the lower 0.01% of all urban centers of interest. Although this simple shift has its own arbitrary aspects, we found that it produced a reasonable pattern of spread and separation for purposes of our analysis. Note that we do not attempt to assign any formal stochastic model to our PVI; we view the index simply as a fixed constant derived from each city’s underlying socioeconomic, physical, and built-environment characteristics, as outlined by Piegorsch et al. (2007).

We explored a number of other transformations for the original PVI07 index, but found

that none gave a more conveniently spaced and sufficiently separated pattern of positive values as these simple, shifted PVIs. It is the shifted values that appear in Table 1.

3.4 BMI estimation and inference

Given data pairs (x_i, Y_i) as the PVI and binary outcome for each i th city ($i = 1, \dots, n$), we fit the centered autologistic regression model from (2) via maximum pseudo-likelihood, as in Sec. 3.2. This produces MPLEs $\hat{\beta} = [\hat{\beta}_0 \ \hat{\beta}_1 \ \hat{\beta}_2]^T$. Unfortunately, the calculations are not trivial, and require computer iteration. We employ the **R** package *ngspatial*; see <https://CRAN.R-project.org/package=ngspatial>. Given the MPLE $\hat{\beta}_1$ (and a predetermined level for the BMR), from (4) the corresponding point estimate for the BMI is then simply

$$\widehat{\text{BMI}}_{\text{BMR}} = \frac{\log(\text{BMR})}{\hat{\beta}_1}. \quad (5)$$

As discussed above, we further desire a lower $1 - \alpha$ confidence limit on the BMI, denoted as BMIL_{BMR} . The straightforward form for the BMI in (4) allows for a particularly simple construction, at least conceptually: if b_{1U} is an upper $1 - \alpha$ confidence limit on β_1 such that $P[\beta_1 < b_{1U}] = 1 - \alpha$ at least approximately in large samples, then clearly

$$P\left[\frac{\log(\text{BMR})}{b_{1U}} < \frac{\log(\text{BMR})}{\beta_1}\right] = P\left[\frac{\log(\text{BMR})}{b_{1U}} < \text{BMI}_{\text{BMR}}\right] = 1 - \alpha \quad (6)$$

if $\text{BMR} > 1$. This defines a $1 - \alpha$ lower confidence limit on the BMI, thus we take $\text{BMIL}_{\text{BMR}} = \log(\text{BMR})/b_{1U}$. We note in passing that (6) can be written as $P[b_{1U}^{-1} \log\{\text{BMR}\} < \text{BMI}_{\text{BMR}}, \forall \text{BMR} > 1] = 1 - \alpha$, which provides a simultaneous lower confidence statement on the BMI, in the sense of Piegorsch & West (2005).

Actually finding such a b_{1U} confidence limit is a more difficult task, however, since it is inappropriate under the centered autologistic model to imitate standard practice and estimate the standard error of $\hat{\beta}_1$ by inverting the MPL information matrix (Varin et al., 2011). Instead, we follow a suggestion by Hughes (2014), who calls for a computer-intensive,

parallel, parametric bootstrap approach. The method returns a bootstrap distribution of $B > 0$ resampled β_1 values, based on the original data. An approximate upper $1 - \alpha$ confidence limit can then be taken as the $(1 - \alpha)B$ quantile from this bootstrap distribution, denoted as b_{1B} . We found that at least $B = 5000$ bootstrap replicates were required for the bootstrap procedure to stabilize and produce viable confidence limits, at least for sample sizes at or above $n = 100$ as seen with the 132-cities data.

From this, a bootstrap-based lower confidence limit for the BMI becomes

$$\text{BMIL}_{\text{BMR}} = \frac{\log(\text{BMR})}{b_{1B}} . \quad (7)$$

4 Monte Carlo evaluations

The BMIL constructed in Sec. 3.4 relies on large-sample approximating arguments to be valid. To investigate the performance of the confidence limits in practice, we conducted a series of Monte Carlo evaluations. We examined the coverage properties of the BMIL, along with selected characteristics of the underlying $\widehat{\text{BMI}}$ estimator.

4.1 Simulation settings

We began by selecting a series of three different scenarios under which a set of PVIs would be constructed: (i) $n = 100$ values individually sampled from a uniform distribution whose lower limit is the minimum of the PVIs in Table 1 and whose upper limit is the maximum of those values, and distributed these 100 values randomly and independently—i.e., the PVIs had no embedded spatial correlation—onto a 10×10 lattice, (ii) $n = 100$ values individually sampled from the empirical c.d.f. of the actual PVI values in Table 1 again distributed randomly and independently onto a 10×10 lattice, or (iii) a 10×10 rectangular lattice of PVIs from Table 1 configured by manipulating $n = 100$ of the 132 cities so that their positions on the lattice were roughly equivalent to their relative positions on the U.S. map; see Table S4 in the supplemental document. For the first two (‘uniform’ and ‘empirical’)

settings, we also expanded the sample size to $n = 400$ and $n = 900$ ‘cities’ by creating 20×20 and 30×30 lattices, respectively.

Using each constructed lattice of ‘cities,’ we defined their neighborhood structure and corresponding adjacency matrix $\mathbf{A}$ for the autologistic model using the same simple adjacency described earlier: $a_{ij} = 1$ if city j is physically adjacent (i.e., connects directly on the lattice) to city i , and $a_{ij} = 0$ otherwise.

Given values for the n PVIs and their adjacency/neighborhood structure on the simulation lattice, we generated binary responses Y_i with response probabilities defined via Equation (2), using the perfect sampler given by Hughes (2014). We selected a series of different configurations for the three unknown parameters in β ; we began with $\beta = [-2.5, 0.5, -1]^T$ which are roughly the values of the MPLEs from fitting the centered autologistic model to the 132 cities data; see Sec. 5, below. We then manipulated the value of each parameter to vary somewhat from this base such that the response probability under the independence model at PVIs near zero took small values between about 5% and 20%, and where the corresponding BMIs at $\text{BMR} = 10$ and $\text{BMR} = 25$ were between 2.5 and 6.5. These latter values bound roughly 95% of the calculated PVIs from Table 1. This led to: $\beta_0 = -2.5, -1.5$; $\beta_1 = 0.5, 1$; and $\beta_2 = -1, 0, 1$. Coupled with the seven different sample size/lattice combinations above, this produced a total of 84 different design/parameter configurations for study.

For each simulation configuration, 2000 simulated data sets were generated. In each simulated data set, we calculated the consequent MPLEs and from these, the corresponding $\widehat{\text{BMI}}$ from (5) and bootstrap-based BMIL from (7). We then assessed whether or not the BMIL correctly covered (i.e., remained below) the true value of the BMI for that parameter configuration. We set our nominal confidence level to $1 - \alpha = 0.95$ and recorded the empirical proportion of cases where the BMIL did in fact express correct coverage. Notice that with 2000 simulations per configuration, the approximate standard error of our empirical coverage rates at the nominal 95% level is $\sqrt{(0.05)(0.95)/2000} = 0.005$ and it never exceeds $\sqrt{(0.5)(0.5)/2000} = 0.011$.

Note that under our construction in (7), the BMIL will correctly cover the true BMI from below if and only if the upper bootstrap limit b_{1B} correctly covers β_1 from above. For simplicity, we only assessed if the latter condition held. In fact, the operation is independent of the BMR, so that our coverage assessments hold for any choice of $BMR > 1$.

We also collected summary statistics on the performance of the various point estimators under the centered autologistic model. For the $\widehat{BMI}$, we found the average relative bias $\frac{1}{2000} \sum_{t=1}^{2000} (\widehat{BMI} - BMI)/BMI$. Notice from the form of the BMI in (4) that any terms involving BMR cancel in this relative bias calculation, so that the result will again be independent of the numerical value for the BMR.

Similarly, for the autologistic parameters β_0 and β_1 we found the average relative biases $\frac{1}{2000} \sum_{t=1}^{2000} (\hat{\beta}_j - \beta_j)/|\beta_j|$ ($j = 0, 1$). For the autocovariate parameter we simply calculated the average bias $\frac{1}{2000} \sum_{t=1}^{2000} (\hat{\beta}_2 - \beta_2)$, since the true values of $|\beta_2|$ were set to either 0 or 1 under our configurations.

4.2 Simulation results

4.2.1 Coverage rates

Empirical coverage rates of our bootstrap-based lower 95% confidence limit on BMI under our centered autologistic model appear in Table 2. As noted above, these also represent empirical coverage rates of the bootstrap-based 95% upper confidence limit on β_1 . As can be seen, all values rest above the nominal 95% confidence coefficient; in a number of cases the excess is significantly above the target level. As sample size increases, the rates drop towards the nominal level, on average. These results suggest that the bootstrap-based limits appear to operate in a reasonable, if slightly conservative manner.

We should mention that we experienced some minor instabilities with the MPL fitting algorithm. Cases occurred where the algorithm failed to converge with some bootstrapped resamples, so that *ngspatial* reported only ‘NA’ for $\hat{\beta}_1$. The resulting bootstrap distribution therefore contained fewer than the desired $B = 5000$ resampled $\hat{\beta}_1$ values. Table 3 gives

the rates of how often this occurred; each value is the number of reported NAs out of 10^7 fitting attempts (5000 bootstrap samples $\times$ 2000 simulated data sets) at each simulation configuration. As can be seen, the failure rates are quite low, and only occur at the smallest sample size of $n = 100$. When this occurs in practice, users may simply increase B to perhaps 7500 or 10000, in order to produce sufficient bootstrap resamples for the BMIL. We view this as a slight inconvenience and a tolerable consequence of employing such a complex model/fitting procedure.

4.2.2 Estimator bias

Empirical biases of the MPLEs under our centered logistic model appear in Table 4 (relative biases for β_0 and β_1). Some general patterns are clear: $\hat{\beta}_0$ and $\hat{\beta}_2$ typically exhibit small negative bias, while $\hat{\beta}_1$ exhibits small positive bias. The absolute relative biases are usually on the order of 10^{-1} or smaller, and all biases trend towards zero as sample size increases.

Similar values for $\widehat{\text{BMI}}$ appear in Table 5. We see a more variable pattern among these relative bias terms, however, with a mix of both positive and negative values. Most remain small, although a few relative biases past ± 2 do appear with $n = 100$ using the Empirical generator for $x = \text{PVI}$. As sample size increases, however, the relative biases do mitigate and move towards zero on average.

We speculate that the mixed bias patterns for $\widehat{\text{BMI}}$ may be a function of the reciprocal operation used in (5) when defining the BMI estimator. The relative bias for $\hat{\beta}_1$ appears stable; however, using $1/\hat{\beta}_1$ in $\widehat{\text{BMI}}$ induces greater instability. This is perhaps not surprising, as it is not uncommon to find that reciprocal random variables are often less stable than their original counterparts.

5 Urban vulnerability example, revisited

Returning to our example data on placed-based urban vulnerability to terrorism, we apply the centered autologistic model to the data in Table 1. Our goal is to update the earlier study

in Piegorsch et al. (2007) by incorporating potential spatial autocorrelation when studying the relationship between our PVI measure and the terrorist casualty data, and also to explore how this autocorrelated model operates in a real, risk-analytic, benchmarking application.

Regressing Y on $x = \text{PVI}$ via the centered-autologistic model in (2) produces the MPLEs $\hat{\beta}_0 = -2.4460$, $\hat{\beta}_1 = 0.4111$, and $\hat{\beta}_2 = -1.0990$. Fig. 1 presents histograms (using Scott’s rule for bin selection) and overlaid kernel density estimates of the corresponding bootstrap distributions for $\hat{\beta}_1$ and $\hat{\beta}_2$, based on the 5000 bootstrap values from *ngspatial*. (No ‘NA’ instabilities were observed with this bootstrap resample.) The former appears roughly bell-shaped with perhaps a slight right skew, while the latter displays a left skew towards a few extreme negative values. For this sample size, at least, the large-sample features of the MPLEs may not be fully realized with this data set.

The *ngspatial* program provides pointwise 95% bootstrap confidence intervals for the non-intercept parameters as $0.0122 < \beta_1 < 0.8574$ and $-3.2750 < \beta_2 < -0.0436$. Since the first interval fails to contain $\beta_1 = 0$, we can conclude that our nonnegative PVI measure significantly affects the response probability π ; from the strictly positive values in the interval it appears to do so via an increasing relationship. Similarly, since the second interval fails to contain $\beta_2 = 0$, we can conclude that the autocovariate is important for describing π . And, since the β_2 interval presents only negative values, it indicates that the spatial correlation appears negative. This latter result is intriguing: it suggests that when a city exhibits a positive terrorist casualty result, an adjacent city would expect *not* to experience such a result, and *vice versa*.

5.1 Benchmark analysis

Moving to the benchmark calculations, we refer to Equation (5) for a point estimate of BMI_{BMR} . As noted above, the choice for the BMR here is an open issue, since the benchmark paradigm has not previously been applied with models of this sort. Thus a certain portion of this analysis must be viewed as exploratory in nature. After examining a variety of possible

BMRs, we settled on $\text{BMR} = 10$ as a first choice for the relative increase in odds over background. Using (5), this led to $\widehat{\text{BMI}}_{10} = \log(10)/\hat{\beta}_1 = 5.6011$. The upper, one-sided, 95% bootstrap confidence bound on β_1 from the MPL fit (above) is $b_{1B} = 0.7882$, producing a 95% lower bootstrap limit of $\text{BMIL}_{10} = \log(10)/b_{1B} = 2.9213$.

For comparison purposes, we also considered a higher benchmark response at $\text{BMR} = 25$, anticipating that an increase of this magnitude in the odds could be an informative marker for practical use. This gave $\widehat{\text{BMI}}_{25} = \log(25)/\hat{\beta}_1 = 7.8300$, with 95% lower bootstrap limit of $\text{BMIL}_{25} = \log(25)/b_{1B} = 4.0838$. Urban localities with PVIs larger than this value might be viewed as exhibiting excess vulnerability to terrorism casualties, at least based on our analysis.

To illustrate the benchmark delineations geospatially, Fig. 2 maps the 132 urban centers from Table 1, distinguishing those whose PVIs exceed our two benchmarks above: cities whose PVIs exceed BMIL_{25} are marked in black, while those whose PVIs exceed only BMIL_{10} are marked in dark grey. Locations with PVIs below both benchmarks are marked in light grey. The figure shows that urban locations exhibiting high relative benchmark vulnerability are situated primarily in the eastern half of the United States. There, the high-vulnerability locations are generally concentrated along shorelines (eastern seaboard, Gulf Coast, Great Lakes, various rivers), and in the interior south. A few western locations also exceed BMIL_{25} , but not to the extent seen east of the mountain states. Lesser vulnerability urban centers appear almost exclusively in the West and through the northern tier of the map; no such lesser-vulnerability cities appear in the deep South. This latter phenomenon might be partly explained by patterns of urban development, where Eastern urban centers are more dense, less separated (and often geographically continuous—the “megapolis” effect), more constrained by geophysical borders such as rivers, and older than those in the West.

6 Discussion

We have described a statistical methodology that can characterize the vulnerability of U.S. urban centers (‘cities’) to terrorist attack, using a place-based vulnerability index (PVI) and a database of terrorist incidents and related human casualties. We show how to incorporate potential autocorrelation in the geospatial data via a centered autologistic model, and enhance a risk-analytic ‘benchmark’ approach for identifying urban locations at increased risk/vulnerability to terrorist events. In fact, the methods are extensible-enough to apply in a variety of data scenarios where spatial autocorrelation may confound more simplistic models; the possibilities far exceed the specific urban-vulnerability risk analysis studied here.

One can imagine an alternative construction to our autologistic strategy: fit a generalized linear mixed model (GLMM) with a logit link and spatially varying random intercepts. Similar approaches are common in disease mapping with count data, and they can account for spatial correlation across the mapped units in an effective manner (Waller & Carlin, 2010). Indeed, the two strategies could produce roughly similar results when the spatial correlation is positive. We are unsure how to implement a GLMM with random intercepts to produce negative spatial correlation, however, as would be required with the data in our urban vulnerability analysis. Nonetheless, further manipulation of GLMMs with spatially varying coefficients for the risk-analytic problems introduced here could prove quite fruitful for future research.

A critical component required in our analysis is the definition of the neighborhood structure surrounding each metropolitan area. We approached this using a simple, strict definition of adjacency, where one location’s boundary—under the definitions we employed in Borden et al. (2007)—must literally contact another’s for them to be defined as neighbors; see Fig. 2 and also Fig. S1 in the supplemental document. One can alternatively use these models to study spatial autocorrelation as a function of distance between cities, however. For example, we took paired longitudes and latitudes within each of the 132 Euclidean urban region to define ‘centers’ for each city (see Table S3 in the supplemental document). From these, we

calculated the distance, say d_{ij} , between the center points for the i th and j th city pairing. Then to define the neighborhood, N_i , for city i in (2) we populated that N_i with those cities j whose distance from city i rests below some fixed positive bound on the distance, say, $d_{ij} < D$. Fig. S2 in the supplemental document gives a visualization of the consequent neighborhood structure at $D = 100$.

To illustrate, we returned to the data in Table 1 and calculated the point estimate and lower confidence limit for BMI_{25} . We applied a range of values for D , from 50 to 300 miles of city separation. The results appear in Fig. 3. We see that the point estimate oscillates somewhat, but in particular BMIL_{25} stays within in a narrow band between about 2.25 and 3.65. Compare these to $\widehat{\text{BMI}}_{25} = 7.83$ and $\text{BMIL}_{25} = 4.06$ for the strict adjacency neighborhoods in Sec. 5. Moving to different separation distances appears to lower the benchmark points to within a fairly stable range.

Some explanation for this phenomenon may be available by conducting a similar distance analysis on the autocorrelation parameter estimate $\hat{\beta}_2$, as plotted in Fig. 4. There, we see that $\hat{\beta}_2$ also oscillates somewhat, but it generally stays near zero. In particular, all of its corresponding, pointwise, 95% bootstrap intervals contain $\beta_2 = 0$. The particularly wide limits at $D = 50$ are likely due to the small number of pairings that occur for this very tight neighborhood definition: only 26 city pairings report as neighbors when D is set to 50. By contrast, 82 pairings report as neighbors under the adjacency structure. Thus by moving to a distance-based neighborhood definition, we find no significant indication of spatial autocorrelation under our model. The stricter adjacency definition does provide such an indication, however, and by accounting for this autocorrelation in the model the corresponding estimate and lower confidence limit on BMI_{25} become more sensitive and increase above those seen in Fig. 4. A similar phenomenon is exhibited at $\text{BMR} = 10$ (results not shown).

If desired, one can alternatively incorporate other metrics to define connections between the cities, beyond those of simple adjacency or Euclidian distance employed here. Many forms

are possible, including Manhattan (‘city-block’) or Tchebychev (‘chessboard’) distance, Mahalanobis distance to incorporate non-spherical alignments in the separation dimensions, or other constructs that define how two locations ‘connect’ to form the pertinent neighborhoods (Liu, 2017, Sec. 3.2). The choice between these will depend on the particular needs of the risk analyst and on the geo-spatial nature of the incumbent data.

Our application to the U.S. data on urban vulnerability allows for interesting knowledge discovery: the indication of negative, adjacency-based, spatial correlation along the 132 cities in the database may seem counterintuitive, and leads to intriguing speculation on its possible cause. Perhaps the occurrence of terrorist events in one city tends to increase emergency preparedness and urban response planning in adjacent cities, leading to fewer terrorism (or at least lowered casualty) events. Or, perhaps putative terrorists may not target nearby cities in order to maximize their desired impact across a wider geographic space.

Also, well below a quarter (20/132) of the cities in Fig. 2 are labeled in lesser-vulnerability light grey, suggesting a rather disquieting vulnerability landscape: one might be concerned that most any large- and medium-sized urban center in the 50 United States is likely to exceed this low-level benchmark at $BMR = 10$ and exhibit moderate-to-high vulnerability to terrorist casualties. We posed a similar question in our previous study (Piegorsch et al., 2007), although we caution against making any direct comparison between our earlier results and those seen herein. While both presentations are rooted in geospatial comparisons among the same 132 U.S. urban centers, the underlying models and forms of benchmark analyses constructed here differ in a substantive fashion from those earlier, more-simplistic results. Nonetheless, it is possible to attempt some basic comparisons between the two risk-analytic outcomes. We give a visualization-based juxtaposition in the supplemental document—see supplemental Sec. S3. Therein, our autologistic approach offers a more-sensitive/less-conservative indication of the vulnerability palette across these 132 cities, at least under the settings we employed.

It is worth reemphasizing the risk-analytic nature of our modeling approach. On its own,

the PVI stands as a perfectly useful metric for quantifying and distinguishing multi-source urban vulnerability to hazardous events. By employing PVI as a benchmarking predictor and then relating it to data on terrorist casualties via the autologistic model in (2), however, we build a specific focus into the risk assessment. Our results help identify those values of PVI which indicate targeted, place-based vulnerability to terrorists impacts. For example, Boise, ID exhibits a high predicted vulnerability to general hazards, based on its large PVI of 5.415. This is largely due to its past experiences with natural and human-induced catastrophic events (Borden et al., 2007). In effect, our model employs Boise’s particular mix of socio-economic, physical, and built-environment features, as encapsulated in its PVI, to provide risk-analytic guidance on that city’s specific vulnerability to casualties that might occur from a terrorist attack. Even though Boise did not experience such an attack over the 35-year period in our study’s database, our analysis suggests that its terrorist casualty risk exists at a heightened level. (See supplemental Sec. S3 for more on this particular example.)

From a policymaking perspective, our approach could be employed for potential allocation of national and regional funding to support urban preparedness and response to terrorism events. Suppose officials in, e.g., Norfolk, VA or Charleston, SC were considering new or updated forms of coastal anti-terrorist protection. According to our analysis, their relatively high PVIs indicate heightened vulnerability to terrorist events at the community level, even though, as with Boise, above, neither city experienced terrorism casualties during the 35-year period represented by the data in Table 1. This could motivate increased funding allocation(s) or other enhanced efforts to connect urban risk management programs with terrorism vulnerability assessments in these communities (Kunreuther, 2002).

As in Piegorsch et al. (2007), we find that not all urban areas are equally at risk, nor do they have similar underlying vulnerabilities. A location’s capacity to adequately prepare for and respond to hazard(s) varies widely across the U.S., especially in urban areas. We see that “place matters” and that policymakers cannot ignore geographic differences in socio-economic, built-, and natural environments when allocating anti-terrorism resources and

training. Doing so will correspondingly limit urban areas’ abilities to prepare for and respond to terrorist events. Hence the query remains: do our findings affect future urban planning in vulnerability screening, disaster management, and allocation of homeland security resources, and are these questions of valid public concern? We think so, and we believe much more work is required to quantify and understand urban vulnerability to terrorist events and to broader-scale hazards. We hope our results will stimulate further research on these various issues.

Supplemental materials

A supplemental document contains additional materials associated with the results presented here. These include further details and visualizations with the neighborhood structures for the 132 cities in Table 1, the specific PVI values used for the 10×10 ‘map’ lattice in our simulation study from Sec. 4, a comparison of our results in Figure 2 with those seen in Piegorsch et al. (2007), additional study of the centered autologistic model, and sample **R** code to fit the centered autologistic model.

Acknowledgements

Thanks are due to Rachel E. Reeves for cartographic assistance; to Grace Chiu, Stephan R. Sain, and Hideyasu Shimadzu for helpful discussions during the development of this work; and to the Editor, Associate Editor, and an anonymous referee for insightful comments on an earlier draft of the manuscript. This material represents a portion of the first author’s Ph.D. dissertation from the University of Arizona Graduate Interdisciplinary Program in Statistics. The research was supported in part by grant #ES027394 from the U.S. National Institutes of Health.

References

- Arnold, B. C., & Strauss, D. J. (1991). Pseudolikelihood estimation: some examples. *Sankhyá, Ser. B*, 53(1), 233–243.
- Besag, J. E. (1972). Nearest-neighbour systems and the auto-logistic model for binary data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(1), 75–83.
- Besag, J. E. (1975). Statistical analysis of non-lattice data. *Journal of the Royal Statistical Society, Series D (The Statistician)*, 24(3), 179–195. doi: 10.2307/2987782
- Borden, K. A., Schmidtlein, M. C., Emrich, C. T., Piegorsch, W. W., & Cutter, S. L. (2007). Vulnerability of US cities to environmental hazards. *Journal of Homeland Security and Emergency Management*, 4(2), Article Number 5. doi: 10.2202/1547-7355.1279
- Budtz-Jørgensen, E., Keiding, N., & Grandjean, P. (2001). Benchmark dose calculation from epidemiological data. *Biometrics*, 57(3), 698–706. doi: 10.1111/j.0006-341X.2001.00698.x
- Caragea, P. C., & Kaiser, M. S. (2009). Autologistic models with interpretable parameters. *Journal of Agricultural, Biological, and Environmental Statistics*, 14(3), 281–300. doi: 10.1118/jabes.2009.07032
- Chun, Y., & Griffith, D. A. (2013). *Spatial Statistics and Geostatistics: Theory and Applications for Geographic Information Science and Technology*. Thousand Oaks, CA: SAGE Publications Inc. doi: 10.1080/13658816.2015.1107074
- Crump, K. S. (1984). A new method for determining allowable daily intakes. *Toxicological Sciences*, 4(5), 854–871. doi: 10.1016/0272-0590(84)90107-6
- Crump, K. S. (1995). Calculation of benchmark doses from continuous data. *Risk Analysis*, 15, 79–89. doi: 10.1111/j.1539-6924.1995.tb00095.x

- Crump, K. S. (2012). Benchmark analysis. In A. H. El-Shaarawi & W. W. Piegorsch (Eds.), *Encyclopedia of Environmetrics* **1** (2nd ed., pp. 191–198). Chichester: John Wiley & Sons. doi: 10.1002/9780470057339.vab009
- Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003). Social vulnerability to environmental hazards. *Social Science Quarterly*, *84*(2), 242–261. doi: 10.1111/1540-6237.8402002
- de Frutos, A., Olea, P. P., & Vera, R. (2007). Analyzing and modelling spatial distribution of summering lesser kestrel: The role of spatial autocorrelation. *Ecological Modelling*, *200*(1–2), 33–44. doi: 10.1016/j.ecolmodel.2006.07.007
- Haimes, Y. Y. (2006). On the definition of vulnerabilities in measuring risks to infrastructures. *Risk Analysis*, *26*(2), 293–296. doi: 10.1111/j.1539-6924.2006.00755.x
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). New York: John Wiley & Sons. doi: 10.1002/9781118548387
- Hughes, J. (2014). *ngspatial*: A package for fitting the centered autologistic and sparse spatial generalized linear mixed models for areal data. *The R Journal*, *6*(2), 81–95.
- Hughes, J., Haran, M., & Caragea, P. C. (2011). Autologistic models for binary data on a lattice. *Environmetrics*, *22*(7), 857–871. doi: 10.1002/env.1102
- Kolaczyk, E. D., & Csárdi, G. (2014). *Statistical Analysis of Network Data with R*. New York: Springer. doi: 10.1007/978-1-4939-0983-4
- Koutsias, N. (2003). An autologistic regression model for increasing the accuracy of burned surface mapping using landsat thematic mapper data. *International Journal of Remote Sensing*, *24*(10), 2199–2204. doi: 10.1080/0143116031000082073
- Kunreuther, H. (2002). Risk analysis and risk management in an uncertain world. *Risk Analysis*, *22*(4), 655–664. doi: 10.1111/0272-4332.00057

- Liu, J. (2017). *Autologistic Modeling in Benchmark Risk Analysis* (Ph.D. thesis). Interdisciplinary Program in Statistics, University of Arizona, Tucson, AZ.
- MacKenzie, C. A. (2014). Summarizing risk using risk measures and risk indices. *Risk Analysis*, *34*(12), 2143–2162. doi: 10.1111/risa.12220
- Mazzorana, B., Zischg, A., Largiadèr, A., & Hübl, J. (2009). Hazard index maps for woody material recruitment and transport in alpine catchments. *Natural Hazards and Earth System Sciences*, *9*, 197–209. doi: 10.5194/nhess-9-197-2009
- Peng, R. D., Schoenberg, F. P., & Woods, J. A. (2005). A space–time conditional intensity model for evaluating a wildfire hazard index. *Journal of the American Statistical Association*, *100*(469), 26–35. doi: 10.1198/016214504000001763
- Piegorsch, W. W., Cutter, S. L., & Hardisty, F. (2007). Benchmark analysis for quantifying urban vulnerability to terrorist incidents. *Risk Analysis*, *27*(6), 1411–1425. doi: 10.1111/j.1539-6924.2007.00977.x
- Piegorsch, W. W., & West, R. W. (2005). Benchmark analysis: shopping with proper confidence. *Risk Analysis*, *25*(4), 913–920. doi: 10.1111/j.1539-6924.2005.00636.x
- R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>
- Santini, M., Caccamo, G., Laurenti, A., Noce, S., & Valentini, R. (2010). A multi-component GIS framework for desertification risk assessment by an integrated index. *Applied Geography*, *30*, 394–415. doi: 10.1016/j.apgeog.2009.11.003
- Varin, C., Reid, N., & Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica*, *21*(1), 5–42.

- Vaughn, N. R., Asner, G. P., Smit, I. P. J., & Riddel, E. S. (2015). Multiple scales of control on the structure and spatial distribution of woody vegetation in African savanna watersheds. *PLOS One*, 10(12), Article No. e0145192. doi: 10.1371/journal.pone.0145192
- Waller, L. A., & Carlin, B. P. (2010). Disease mapping. In A. E. Gelfand, P. J. Diggle, M. Fuentes, & P. Guttorp (Eds.), *Handbook of Spatial Statistics* (pp. 217–243). Boca Raton, FL: CRC Press. doi: 10.1201/9781420072884-c14

Table 1: Terrorism casualty indicator (Y) and place-based vulnerability measure (PVI) for 132 U.S. metropolitan areas

Metropolitan area	Y	PVI	Metropolitan area	Y	PVI
Albany, NY	1	4.199	Little Rock, AR	0	3.701
Albuquerque, NM	0	2.942	Los Angeles, CA ¹	1	4.140
Allentown/Bethlehem, PA	0	3.933	Louisville, KY	0	3.841
Amarillo, TX	0	3.419	Lubbock, TX	0	3.110
Anchorage, AK	0	2.757	Madison, WI	1	3.303
Atlanta, GA	1	5.402	McAllen, TX	0	4.396
Augusta, GA	0	4.450	Memphis, TN	0	4.632
Augusta, ME	0	2.909	Miami/Ft. Lauderdale, FL	1	4.314
Austin, TX	0	3.611	Milwaukee, WI	0	4.080
Bakersfield, CA	0	2.880	Minneapolis/St. Paul, MN	0	4.059
Baltimore/Annapolis, MD	0	4.942	Mission Viejo, CA	1	3.873
Barre/Montpelier, VT	0	2.308	Mobile, AL	0	4.144
Baton Rouge, LA	0	6.735	Modesto, CA	0	2.911
Birmingham, AL	1	3.905	Montgomery, AL	0	3.824
Bismarck, ND	0	3.101	Nashville, TN	0	3.368
Boise, ID	0	5.415	New Haven, CT	1	3.821
Boston, MA	1	4.323	New Orleans, LA	1	6.838
Bridgeport/Stamford, CT	1	2.700	New York/Newark, NY	1	5.873
Buffalo, NY	1	3.998	Norfolk, VA ²	0	6.045
Cape Coral, FL	0	4.262	Oklahoma City, OK	1	3.768
Carson City, NV	0	2.580	Olympia, WA	0	2.638
Charleston, SC	0	6.262	Omaha, NE	1	4.050
Charleston, WV	0	3.557	Orlando, FL	0	4.466
Charlotte, NC	0	4.765	Oxnard, CA	0	2.512
Chattanooga, TN	0	4.183	Palm Bay/Melbourne, FL	0	3.764
Cheyenne, WY	0	2.691	Pensacola, FL	1	4.337
Chicago, IL	1	5.123	Philadelphia, PA	1	5.456
Cincinnati, OH	0	3.941	Phoenix/Glendale/Mesa, AZ	1	2.815
Cleveland/Akron, OH	1	4.911	Pierre, SD	0	2.211
Colorado Springs, CO	0	2.457	Pittsburgh, PA	1	4.386
Columbia, SC	0	4.856	Portland, OR	0	3.730
Columbus, GA	1	3.614	Poughkeepsie/Newburgh, NY	0	3.921
Columbus, OH	0	3.684	Providence, RI	0	3.439
Concord, NH	0	2.235	Raleigh/Durham, NC	0	4.736
Corpus Christi, TX	0	3.949	Reno, NV	0	2.752
Dallas/Ft. Worth, TX	1	4.166	Richmond, VA	0	5.655
Dayton, OH	0	3.498	Riverside/San Bernardino, CA	0	3.544
Denver, CO	1	3.890	Rochester, NY	0	3.456
Des Moines, IA	0	3.877	Sacramento, CA	1	3.465
Detroit/Warren, MI	1	3.907	Salem, OR	0	2.705
Dover, DE	0	3.664	Salt Lake City/Ogden, UT	1	3.910
El Paso, TX	0	3.908	San Antonio, TX	0	4.634
Flint, MI	0	3.373	San Diego, CA	1	3.200
Frankfort, KY	0	2.739	San Francisco, CA ³	1	3.769
Fresno, CA	0	3.322	Santa Fe, NM	0	2.506
Ft. Wayne, IN	0	2.780	Sarasota/Bradenton, FL	0	4.345
Grand Rapids, MI	0	3.083	Scranton, PA	0	4.157
Greensboro/Winston Salem, NC	1	4.533	Seattle/Tacoma, WA	0	3.404
Harrisburg, PA	0	3.937	Shreveport, LA	0	4.249

Continued on next page

Table 1 – *Continued from previous page*

Metropolitan area	Y	PVI	Metropolitan area	Y	PVI
Hartford, CT	0	3.378	Spokane, WA	0	2.939
Helena, MT	0	1.990	Springfield, IL	0	3.012
Honolulu, HI	0	3.482	Springfield, MA	0	3.279
Houston, TX	0	5.563	St. Louis, MO	1	4.663
Huntsville, AL	0	3.932	Stockton, CA	0	3.307
Indianapolis, IN	0	3.340	Syracuse, NY	0	4.130
Jackson, MS	0	4.279	Tallahassee, FL	0	3.722
Jacksonville, FL	0	3.810	Tampa/St. Petersburg, FL	1	4.869
Jefferson City, MO	0	3.029	Toledo, OH	0	4.436
Juneau, AK	0	1.298	Topeka, KS	1	3.003
Kansas City, MO	0	3.979	Trenton, NJ	1	4.129
Knoxville, TN	0	3.690	Tucson, AZ	0	2.951
Lancaster, PA	0	3.604	Tulsa, OK	0	3.695
Lansing, MI	0	3.086	Washington DC	1	5.697
Las Vegas, NV	0	3.329	Wichita, KS	1	3.089
Lexington, KY	0	3.340	Worcester, MA	0	2.867
Lincoln, NE	0	3.174	Youngstown, OH	0	4.143

¹ Los Angeles region also includes Glendale, Long Beach, and Huntington Beach, CA.

² Norfolk region also includes Chesapeake, Newport News, and Virginia Beach, VA.

³ San Francisco region also includes Oakland, San Jose, and Fremont, CA.

Table 2: Empirical coverage rates for centered autologistic benchmark index lower confidence limit (BMIL) based on 2000 simulated data sets, stratified by PVI generation method (top; see text), true autologistic regression parameter configuration (left column), and sample size n . Nominal coverage level is set to 95%.

Coefficients: ($\beta_0, \beta_1, \beta_2$)	Uniform PVI			Empirical PVI			Map PVI
	$n = 100$	$n = 400$	$n = 900$	$n = 100$	$n = 400$	$n = 900$	$n = 100$
(-1.5, 0.5, -1)	0.972	0.957	0.962	0.958	0.958	0.960	0.964
(-1.5, 0.5, 0)	0.974	0.958	0.956	0.966	0.958	0.959	0.963
(-1.5, 0.5, 1)	0.972	0.959	0.967	0.970	0.965	0.961	0.968
(-1.5, 1, -1)	0.984	0.971	0.965	0.974	0.961	0.957	0.951
(-1.5, 1, 0)	0.986	0.969	0.968	0.974	0.960	0.966	0.969
(-1.5, 1, 1)	0.987	0.967	0.973	0.979	0.963	0.963	0.979
(-2.5, 0.5, -1)	0.978	0.960	0.954	0.957	0.959	0.959	0.961
(-2.5, 0.5, 0)	0.979	0.960	0.964	0.953	0.957	0.961	0.963
(-2.5, 0.5, 1)	0.974	0.967	0.963	0.964	0.961	0.960	0.967
(-5, 1, -1)	0.986	0.971	0.960	0.970	0.964	0.967	0.974
(-2.5, 1, 0)	0.991	0.976	0.968	0.969	0.967	0.960	0.980
(-2.5, 1, 1)	0.990	0.973	0.970	0.967	0.968	0.969	0.978
Average	0.981	0.966	0.9640	0.967	0.962	0.962	0.968

Table 3: Empirical rates of convergence failure for the MPL algorithm, each across 10^7 fitting attempts (5000 bootstrap resamples from 2000 simulated data sets), stratified by PVI generation method (top; see text), true autologistic regression parameter configuration (left column), and sample size n . Dashes indicate no convergence failures.

Coefficients: ($\beta_0, \beta_1, \beta_2$)	Uniform PVI			Empirical PVI			Map PVI
	$n = 100$	$n = 400$	$n = 900$	$n = 100$	$n = 400$	$n = 900$	$n = 100$
(-1.5, 0.5, -1)	-	-	-	-	-	-	-
(-1.5, 0.5, 0)	-	-	-	-	-	-	-
(-1.5, 0.5, 1)	0.0001	-	-	0.0002	-	-	0.0001
(-1.5, 1, -1)	0.0027	-	-	0.0311	-	-	0.0185
(-1.5, 1, 0)	0.0019	-	-	0.0280	-	-	0.0145
(-1.5, 1, 1)	0.0019	-	-	0.0133	-	-	0.0067
(-2.5, 0.5, -1)	-	-	-	-	-	-	0.0001
(-2.5, 0.5, 0)	-	-	-	-	-	-	-
(-2.5, 0.5, 1)	-	-	-	0.0003	-	-	0.0001
(-2.5, 1, -1)	0.0001	-	-	0.0003	-	-	0.0002
(-2.5, 1, 0)	-	-	-	0.0002	-	-	0.0001
(-2.5, 1, 1)	0.0001	-	-	0.0001	-	-	0.0003

Table 4: Empirical biases for centered autologistic MPLEs based on 2000 simulated data sets, stratified by PVI generation method (top; see text), true autologistic regression parameter configuration (left column), and sample size n .

Coefficients: ($\beta_0, \beta_1, \beta_2$)	Relative bias of β_0						Map PVI $n = 100$
	Uniform PVI			Empirical PVI			
	$n = 100$	$n = 400$	$n = 900$	$n = 100$	$n = 400$	$n = 900$	
(-1.5, 0.5, -1)	-0.053	-0.011	-0.009	-0.095	-0.020	-0.011	-0.070
(-1.5, 0.5, 0)	-0.070	-0.014	-0.007	-0.092	-0.019	-0.006	-0.083
(-1.5, 0.5, 1)	-0.152	-0.053	-0.024	-0.181	-0.087	-0.075	-0.170
(-1.5, 1, -1)	-0.144	-0.024	-0.012	-0.030	-0.002	-0.015	0.007
(-1.5, 1, 0)	-0.191	-0.030	-0.012	-0.122	0.012	-0.015	-0.093
(-1.5, 1, 1)	-0.207	-0.024	-0.017	-0.225	0.000	-0.014	-0.211
(-2.5, 0.5, -1)	-0.052	-0.013	-0.004	-0.053	-0.010	-0.009	-0.047
(-2.5, 0.5, 0)	-0.058	-0.013	-0.008	-0.071	-0.012	-0.010	-0.063
(-2.5, 0.5, 1)	-0.024	-0.006	0.004	-0.009	0.033	0.031	-0.006
(-2.5, 1, -1)	-0.096	-0.021	-0.006	-0.108	-0.012	-0.007	-0.087
(-2.5, 1, 0)	-0.103	-0.023	-0.009	-0.111	-0.014	-0.007	-0.112
(-2.5, 1, 1)	-0.111	-0.016	-0.008	-0.154	-0.027	-0.014	-0.163
Average	-0.105	-0.021	-0.009	-0.104	-0.013	-0.013	-0.092

Coefficients: ($\beta_0, \beta_1, \beta_2$)	Relative bias of β_1						Map PVI $n = 100$
	Uniform PVI			Empirical PVI			
	$n = 100$	$n = 400$	$n = 900$	$n = 100$	$n = 400$	$n = 900$	
(-1.5, 0.5, -1)	0.050	0.010	0.008	0.083	0.017	0.010	0.061
(-1.5, 0.5, 0)	0.063	0.013	0.006	0.085	0.018	0.007	0.075
(-1.5, 0.5, 1)	0.074	0.016	0.008	0.100	0.027	0.017	0.062
(-1.5, 1, -1)	0.130	0.024	0.011	0.096	0.013	0.012	0.070
(-1.5, 1, 0)	0.153	0.029	0.012	0.123	0.006	0.012	0.103
(-1.5, 1, 1)	0.150	0.024	0.014	0.144	0.012	0.011	0.125
(-2.5, 0.5, -1)	0.051	0.013	0.005	0.059	0.010	0.011	0.047
(-2.5, 0.5, 0)	0.056	0.012	0.008	0.074	0.012	0.012	0.067
(-2.5, 0.5, 1)	0.063	0.020	0.009	0.088	0.028	0.013	0.067
(-2.5, 1, -1)	0.089	0.020	0.006	0.093	0.012	0.007	0.076
(-2.5, 1, 0)	0.093	0.022	0.009	0.094	0.014	0.007	0.096
(-2.5, 1, 1)	0.090	0.019	0.009	0.084	0.018	0.012	0.094
Average	0.089	0.019	0.009	0.094	0.016	0.011	0.079

Coefficients: ($\beta_0, \beta_1, \beta_2$)	Bias of β_2						Map PVI $n = 100$
	Uniform PVI			Empirical PVI			
	$n = 100$	$n = 400$	$n = 900$	$n = 100$	$n = 400$	$n = 900$	
(-1.5, 0.5, -1)	-0.129	-0.034	-0.019	-0.115	-0.032	-0.017	-0.137
(-1.5, 0.5, 0)	-0.068	-0.011	-0.011	-0.046	-0.005	-0.012	-0.074
(-1.5, 0.5, 1)	0.058	0.024	0.004	0.054	0.020	0.006	0.045
(-1.5, 1, -1)	-0.525	-0.111	-0.043	-3.142	-0.140	-0.053	-2.942
(-1.5, 1, 0)	-0.228	-0.076	-0.033	-1.553	-0.061	-0.033	-1.673
(-1.5, 1, 1)	-0.011	-0.024	-0.008	-0.276	-0.006	-0.007	-0.149
(-2.5, 0.5, -1)	-0.125	-0.028	-0.014	-0.128	-0.038	-0.015	-0.139
(-2.5, 0.5, 0)	-0.064	-0.014	-0.006	-0.057	-0.015	-0.009	-0.085
(-2.5, 0.5, 1)	0.047	0.025	0.006	0.048	0.024	0.007	0.054
(-2.5, 1, -1)	-0.207	-0.064	-0.024	-0.382	-0.041	-0.022	-0.216
(-2.5, 1, 0)	-0.123	-0.031	-0.019	-0.101	-0.020	-0.017	-0.111
(-2.5, 1, 1)	0.064	0.010	-0.001	0.033	0.013	0.001	0.040
Average	-0.109	-0.028	-0.014	-0.472	-0.025	-0.014	-0.449

Table 5: Empirical relative bias for centered autologistic BMI estimator in (5) based on 2000 simulated data sets, stratified by PVI generation method (top; see text), true autologistic regression parameter configuration (left column), and sample size n .

Coefficients: ($\beta_0, \beta_1, \beta_2$)	Uniform PVI			Empirical PVI			Map PVI
	$n = 100$	$n = 400$	$n = 900$	$n = 100$	$n = 400$	$n = 900$	$n = 100$
(-1.5, 0.5, -1)	0.072	0.019	0.005	-0.132	0.085	0.022	0.387
(-1.5, 0.5, 0)	0.055	0.011	0.004	-2.334	0.062	0.021	0.512
(-1.5, 0.5, 1)	0.073	0.015	0.004	-0.781	0.063	0.018	-0.125
(-1.5, 1, -1)	-0.011	-0.001	-0.001	0.885	0.053	0.012	1.158
(-1.5, 1, 0)	-0.020	-0.004	-0.002	0.141	0.064	0.011	0.392
(-1.5, 1, 1)	-0.019	0.001	-0.004	2.820	0.047	0.011	0.063
(-2.5, 0.5, -1)	0.024	0.016	0.008	0.373	0.086	0.019	0.158
(-2.5, 0.5, 0)	0.089	0.012	0.001	0.900	0.056	0.013	0.161
(-2.5, 0.5, 1)	0.099	0.009	0.003	-0.023	0.056	0.021	0.499
(-2.5, 1, -1)	-0.022	-0.006	0.000	0.173	0.029	0.008	0.047
(-2.5, 1, 0)	-0.027	-0.008	-0.003	0.149	0.022	0.007	0.059
(-2.5, 1, 1)	-0.020	-0.004	-0.003	0.101	0.016	0.002	0.162
Average	0.024	0.005	0.001	0.189	0.053	0.014	0.289

Figure 1: Histograms and overlaid kernel density estimates (dark curves) of the bootstrap distributions for $\hat{\beta}_1$ and $\hat{\beta}_2$, based on 5000 bootstrap replicates from the centered autologistic fit of the 132 cities data. Histograms employ Scott's rule for bin selection.

Figure 2: Map of 132 cities from Table 1, coded to indicate vulnerability status from our autologistic analysis: light grey indicates city $PVI < BMIL_{10}$, dark grey indicates $BMIL_{10} \leq PVI < BMIL_{25}$, while black indicates $PVI \geq BMIL_{25}$. (Online version includes colorized map.)

Figure 3: Relationship between $\widehat{\text{BMI}}_{25}$ estimates (solid dots) and separation distance D defining a neighborhood within the centered autologistic fit for the 132 cities data. Lower dashes are pointwise 95% BMIL_{25} values.

Figure 4: Relationship between estimated autocovariance parameter $\hat{\beta}_2$ (solid dots) and separation distance D defining a neighborhood within the centered autologistic fit for the 132 cities data. Upper and lower dashes are pointwise 95% confidence limits for β_2 .

Supplementary material to ‘Autologistic models for benchmark risk/vulnerability assessment of urban terrorism outcomes’

Jingyu Liu

Interdisciplinary Program in Statistics
University of Arizona, Tucson, AZ

Walter W. Piegorsch

Interdisciplinary Program in Statistics and Department of Mathematics
University of Arizona, Tucson, AZ

A. Grant Schissler

Department of Mathematics and Statistics
University of Nevada, Reno, NV

and

Susan L. Cutter

Hazards & Vulnerability Research Institute and Department of Geography
University of South Carolina, Columbia, SC

This supplementary document provides supporting material for *Autologistic models for benchmark risk/vulnerability assessment of urban terrorism outcomes* by J. Liu, et al., (the ‘main document’). The various sections below address a variety of supplemental/supporting topics, and are not intended to flow naturally between each other. They are, however, presented in roughly the same order in which their counterpart topics appeared in the main document. As therein, we denote the pre-set benchmark response as BMR, the benchmark vulnerability index as BMI, the point estimator as $\widehat{\text{BMI}}_{\text{BMR}}$, and the $1 - \alpha$ lower confidence limit as BMIL_{BMR} . The original data set of 132 urban locations (“cities”) is given in the main document.

S1. Neighborhood assignments for 132-cities data

The complete list of neighborhood assignments for all 132 urban centers under the simple adjacency definition appears in Table S1. Recall that adjacency is quantified via the scalars a_{ij} where $a_{ij} = 1$ if urban center j is physically adjacent to city i , and $a_{ij} = 0$ otherwise. Also, $a_{ii} = 0$ for all i . Two urban centers are adjacent when they share any boundary, as defined by the border definitions from Borden et al. (2007). Physical adjacency here is more complex than it may at first seem, since the definitions for each of the 132 metropolitan areas are often much larger than the individual ‘city’ they envelop. As a result, two cities

may be rather far apart from each other as the crow flies, but be adjacent neighbors under our construction. For example, the San Francisco area also includes Oakland, San Jose, and Fremont, CA (see main document Table 1), and as a result is defined as adjacent to each of Modesto, Sacramento, and Stockton, CA. Figure S1 visualizes this effect, and displays the neighborhood structure as a function of how many adjacent neighbors each urban area possesses. A five-pointed star ($\star$) is also overlayed on those locales where casualty events were recorded, i.e., $Y_i = 1$ in main Table 1.

A similar list of neighborhood assignments for all 132 urban centers under an alternative neighborhood definition based on a 100-mile radius from the defined longitude/latitude location of each city appears in Table S2. Table S3 follows with the specific numerical values for longitude and latitude pairings with each urban location, as used here and also whenever they were needed in the main document.

Figure S2 gives a display similar to Figure S1 using the alternative neighborhood definitions based on a 100-mile radius.

Table S1: All neighbors under the strict adjacency definition for 132 metropolitan areas studied in the main document. A blank entry under “Neighbors” indicates a standalone metropolitan area with no adjacent neighbors.

Metropolitan area	Neighbors
Albany, NY	
Albuquerque, NM	Santa Fe, NM
Allentown/Bethlehem, PA	Lancaster, PA; Philadelphia, PA; Scranton, PA; Trenton, NJ
Amarillo, TX	
Anchorage, AK	
Atlanta, GA	
Augusta, GA	Columbia, SC
Augusta, ME	
Austin, TX	San Antonio, TX
Bakersfield, CA	Los Angeles, CA; Oxnard, CA; Riverside/San Bernardino, CA
Baltimore/Annapolis, MD	Harrisburg, PA; Lancaster, PA; Philadelphia, PA; Washington DC
Barre/Montpelier, VT	
Baton Rouge, LA	New Orleans, LA
Birmingham, AL	
Bismarck, ND	
Boise, ID	
Boston, MA	Concord, NH; Providence, RI; Worcester, MA
Bridgeport/Stamford, CT	Hartford, CT; New Haven, CT; Poughkeepsie/Newburgh, NY; Springfield, MA; Trenton, NJ
Buffalo, NY	Rochester, NY
Cape Coral, FL	Sarasota/Bradenton, FL
Carson City, NV	Reno, NV; Sacramento, CA
Charleston, SC	
Charleston, WV	
Charlotte, NC	Columbia, SC; Greensboro/Winston Salem, NC
Chattanooga, TN	
Cheyenne, WY	Denver, CO
Chicago, IL	
Cincinnati, OH	Dayton, OH

Continued on next page

Table S1 – *Continued from previous page*

Metropolitan area	Neighbors
Cleveland/Akron, OH	Youngstown, OH
Colorado Springs, CO	Denver, CO
Columbia, SC	Augusta, GA; Charlotte, NC
Columbus, GA	
Columbus, OH	Dayton, OH
Concord, NH	Boston, MA
Corpus Christi, TX	
Dallas/Ft. Worth, TX	
Dayton, OH	Cincinnati, OH; Columbus, OH
Denver, CO	Cheyenne, WY; Colorado Springs, CO
Des Moines, IA	
Detroit/Warren, MI	Flint, MI; Lansing, MI; Toledo, OH
Dover, DE	Philadelphia, PA
El Paso, TX	
Flint, MI	Detroit/Warren, MI; Lansing, MI
Frankfort, KY	Lexington, KY; Louisville, KY
Fresno, CA	
Ft. Wayne, IN	
Grand Rapids, MI	Lansing, MI
Greensboro/Winston Salem, NC	Charlotte, NC; Raleigh/Durham, NC
Harrisburg, PA	Baltimore/Annapolis, MD; Lancaster, PA; Philadelphia, PA; Washington DC
Hartford, CT	Bridgeport/Stamford, CT; New Haven, CT; Springfield, MA; Worcester, MA
Helena, MT	
Honolulu, HI	
Houston, TX	
Huntsville, AL	
Indianapolis, IN	
Jackson, MS	
Jacksonville, FL	
Jefferson City, MO	
Juneau, AK	
Kansas City, MO	Topeka, KS
Knoxville, TN	
Lancaster, PA	Allentown/Bethlehem, PA; Baltimore/Annapolis, MD; Harrisburg, PA; Philadelphia, PA
Lansing, MI	Detroit/Warren, MI; Flint, MI; Grand Rapids, MI
Las Vegas, NV	Riverside/San Bernardino, CA
Lexington, KY	Frankfort, KY
Lincoln, NE	Omaha, NE
Little Rock, AR	
Los Angeles, CA ¹	Bakersfield, CA; Mission Viejo, CA; Oxnard, CA; Riverside/San Bernardino, CA
Louisville, KY	Frankfort, KY
Lubbock, TX	
Madison, WI	Milwaukee, WI
McAllen, TX	
Memphis, TN	
Miami/Ft. Lauderdale, FL	
Milwaukee, WI	Madison, WI

Continued on next page

Table S1 – *Continued from previous page*

Metropolitan area	Neighbors
Minneapolis/St. Paul, MN	
Mission Viejo, CA	Los Angeles, CA; Riverside/San Bernardino, CA; San Diego, CA
Mobile, AL	Pensacola, FL
Modesto, CA	San Francisco, CA; Stockton, CA
Montgomery, AL	
Nashville, TN	
New Haven, CT	Bridgeport/Stamford, CT; Hartford, CT
New Orleans, LA	Baton Rouge, LA
New York/Newark, NY	Trenton, NJ
Norfolk, VA ²	Richmond, VA
Oklahoma City, OK	
Olympia, WA	Seattle/Tacoma, WA
Omaha, NE	Lincoln, NE
Orlando, FL	Palm Bay/Melbourne, FL; Tampa/St. Petersburg, FL
Oxnard, CA	Bakersfield, CA; Los Angeles, CA
Palm Bay/Melbourne, FL	Orlando, FL
Pensacola, FL	Mobile, AL
Philadelphia, PA	Allentown/Bethlehem, PA; Baltimore/Annapolis, MD; Dover, DE; Harrisburg, PA; Lancaster, PA; Trenton, NJ
Phoenix/Glendale/Mesa, AZ	Tucson, AZ
Pierre, SD	
Pittsburgh, PA	Youngstown, OH
Portland, OR	Salem, OR
Poughkeepsie/Newburgh, NY	Bridgeport/Stamford, CT; Trenton, NJ
Providence, RI	Boston, MA; Worcester, MA
Raleigh/Durham, NC	Greensboro/Winston Salem, NC
Reno, NV	Carson City, NV; Sacramento, CA
Richmond, VA	Norfolk, VA; Washington DC
Riverside/San Bernardino, CA	Bakersfield, CA; Las Vegas, NV; Los Angeles, CA; Mission Viejo, CA; San Diego, CA
Rochester, NY	Buffalo, NY
Sacramento, CA	Carson City, NV; Reno, NV; San Francisco, CA; Stockton, CA
Salem, OR	Portland, OR
Salt Lake City/Ogden, UT	
San Antonio, TX	Austin, TX
San Diego, CA	Mission Viejo, CA; Riverside/San Bernardino, CA
San Francisco, CA ³	Modesto, CA; Sacramento, CA; Stockton, CA
Santa Fe, NM	Albuquerque, NM
Sarasota/Bradenton, FL	Cape Coral, FL; Tampa/St. Petersburg, FL
Scranton, PA	Allentown/Bethlehem, PA
Seattle/Tacoma, WA	Olympia, WA
Shreveport, LA	
Spokane, WA	
Springfield, IL	
Springfield, MA	Bridgeport/Stamford, CT; Hartford, CT; Worcester, MA
St. Louis, MO	
Stockton, CA	Modesto, CA; Sacramento, CA; San Francisco, CA
Syracuse, NY	
Tallahassee, FL	
Tampa/St. Petersburg, FL	Orlando, FL; Sarasota/Bradenton, FL
Toledo, OH	Detroit/Warren, MI

Continued on next page

Table S1 – *Continued from previous page*

Metropolitan area	Neighbors
Topeka, KS	Kansas City, MO
Trenton, NJ	Allentown/Bethlehem, PA; Bridgeport/Stamford, CT; New York/Newark, NY; Philadelphia, PA; Poughkeepsie/Newburgh, NY
Tucson, AZ	Phoenix/Glendale/Mesa, AZ
Tulsa, OK	Wichita, KS
Washington DC	Baltimore/Annapolis, MD; Harrisburg, PA; Richmond, VA
Wichita, KS	Tulsa, OK
Worcester, MA	Boston, MA; Hartford, CT; Providence, RI; Springfield, MA
Youngstown, OH	Cleveland/Akron, OH; Pittsburgh, PA

¹ Los Angeles region also includes Glendale, Long Beach, and Huntington Beach, CA.

² Norfolk region also includes Chesapeake, Newport News, and Virginia Beach, VA.

³ San Francisco region also includes Oakland, San Jose, and Fremont, CA.

Figure S1: Map of 132 urban centers from main Table 1, coded to indicate neighborhood status based on simple rook adjacency: increasing hues indicate more adjacent neighbors for each location. Also overlaid are five-pointed stars (★) if the locality experienced one or more terrorist-related casualties (i.e., $Y_i = 1$ in main Table 1). © 2017 HVRI. All rights reserved.

Table S2: All neighbors under a 100-mile radius definition for 132 metropolitan areas studied in the main document. A blank entry under “Neighbors” indicates a standalone metropolitan area with no neighbors within 100 miles.

Metropolitan area	Neighbors
Albany, NY	Hartford, CT; Poughkeepsie/Newburgh, NY; Springfield, MA
Albuquerque, NM	Santa Fe, NM
Allentown/Bethlehem, PA	Dover, DE; Harrisburg, PA; Lancaster, PA; New York/Newark, NY; Philadelphia, PA; Scranton, PA; Trenton, NJ
Amarillo, TX	
Anchorage, AK	
Atlanta, GA	Columbus, GA
Augusta, GA	Columbia, SC
Augusta, ME	
Austin, TX	San Antonio, TX
Bakersfield, CA	Los Angeles/Glendale/Long Beach/Huntington Beach; Oxnard, CA
Baltimore/Annapolis, MD	Dover, DE; Harrisburg, PA; Lancaster, PA; Philadelphia, PA; Washington, DC
Barre/Montpelier, VT	Concord, NH
Baton Rouge, LA	New Orleans, LA
Birmingham, AL	Huntsville, AL; Montgomery, AL
Bismarck, ND	
Boise, ID	
Boston, MA	Concord, NH; Hartford, CT; Providence, RI; Springfield, MA; Worcester, MA
Bridgeport/Stamford, CT	Hartford, CT; New Haven, CT; New York/Newark, NY; Poughkeepsie/Newburgh, NY; Springfield, MA
Buffalo, NY	Rochester, NY
Cape Coral, FL	Sarasota/Bradenton, FL; Tampa/St. Petersburg, FL
Carson City, NV	Reno, NV
Charleston, SC	
Charleston, WV	
Charlotte, NC	Columbia, SC; Greensboro/Winston-Salem, NC
Chattanooga, TN	Huntsville, AL; Knoxville, TN
Cheyenne, WY	Denver, CO
Chicago, IL	Milwaukee, WI
Cincinnati, OH	Columbus, OH; Dayton, OH; Frankfort, KY; Indianapolis, IN; Lexington, KY; Louisville, KY
Cleveland/Akron, OH	Detroit/Warren, MI; Toledo, OH; Youngstown, OH
Colorado Springs, CO	Denver, CO
Columbia, SC	Augusta, GA; Charlotte, NC
Columbus, GA	Atlanta, GA; Montgomery, AL
Columbus, OH	Cincinnati, OH; Dayton, OH
Concord, NH	Barre/Montpelier, VT; Boston, MA; Providence, RI; Springfield, MA; Worcester, MA
Corpus Christi, TX	
Dallas/Ft. Worth, TX	
Dayton, OH	Cincinnati, OH; Columbus, OH
Denver, CO	Cheyenne, WY; Colorado Springs, CO
Des Moines, IA	
Detroit/Warren, MI	Cleveland/Akron, OH; Flint, MI; Lansing, MI; Toledo, OH
Dover, DE	Allentown/Bethlehem, PA; Baltimore/Annapolis, MD; Lancaster, PA; Philadelphia, PA; Trenton, NJ; Washington, DC

Continued on next page

Table S2 – *Continued from previous page*

Metropolitan area	Neighbors
El Paso, TX	
Flint, MI	Detroit/Warren, MI; Grand Rapids, MI; Lansing, MI; Toledo, OH
Frankfort, KY	Cincinnati, OH; Lexington, KY; Louisville, KY
Fresno, CA	Modesto, CA
Ft. Wayne, IN	Toledo, OH
Grand Rapids, MI	Flint, MI; Lansing, MI
Greensboro/Winston-Salem, NC	Charlotte, NC; Raleigh/Durham, NC
Harrisburg, PA	Allentown/Bethlehem, PA; Baltimore/Annapolis, MD; Lancaster, PA; Philadelphia, PA; Scranton, PA; Washington, DC
Hartford, CT	Albany, NY; Boston, MA; Bridgeport/Stamford, CT; New Haven, CT; New York/Newark, NY; Poughkeepsie/Newburgh, NY; Providence, RI; Springfield, MA; Worcester, MA
Helena, MT	
Honolulu, HI	
Houston, TX	
Huntsville, AL	Birmingham, AL; Chattanooga, TN
Indianapolis, IN	Cincinnati, OH
Jackson, MS	
Jacksonville, FL	
Jefferson City, MO	
Juneau, AK	
Kansas City, MO	Topeka, KS
Knoxville, TN	Chattanooga, TN
Lancaster, PA	Allentown/Bethlehem, PA; Baltimore/Annapolis, MD; Dover, DE; Harrisburg, PA; Philadelphia, PA; Scranton, PA; Trenton, NJ; Washington, DC
Lansing, MI	Detroit/Warren, MI; Flint, MI; Grand Rapids, MI; Toledo, OH
Las Vegas, NV	
Lexington, KY	Cincinnati, OH; Frankfort, KY; Louisville, KY
Lincoln, NE	Omaha, NE
Little Rock, AR	
Los Angeles, CA ¹	Bakersfield, CA; Mission Viejo, CA; Oxnard, CA; Riverside/San Bernardino, CA
Louisville, KY	Cincinnati, OH; Frankfort, KY; Lexington, KY
Lubbock, TX	
Madison, WI	Milwaukee, WI
McAllen, TX	
Memphis, TN	
Miami/Ft. Lauderdale, FL	
Milwaukee, WI	Chicago, IL; Madison, WI
Minneapolis/St. Paul, MN	
Mission Viejo, CA	Los Angeles/Glendale/Long Beach/Huntington Beach; Oxnard, CA; Riverside/San Bernardino, CA; San Diego, CA
Mobile, AL	Pensacola, FL
Modesto, CA	Fresno, CA; Sacramento, CA; San Francisco/Oakland/San Jose/Fremont, CA; Stockton, CA
Montgomery, AL	Birmingham, AL; Columbus, GA
Nashville, TN	
New Haven, CT	Bridgeport/Stamford, CT; Hartford, CT; New York/Newark, NY; Poughkeepsie/Newburgh, NY; Providence, RI; Springfield, MA; Worcester, MA

Continued on next page

Table S2 – *Continued from previous page*

Metropolitan area	Neighbors
New Orleans, LA	Baton Rouge, LA
New York/Newark, NY	Allentown/Bethlehem, PA; Bridgeport/Stamford, CT; Hartford, CT; New Haven, CT; Philadelphia, PA; Poughkeepsie/Newburgh, NY; Trenton, NJ
Norfolk, VA ²	Richmond, VA
Oklahoma City, OK	Tulsa, OK
Olympia, WA	Seattle/Tacoma, WA
Omaha, NE	Lincoln, NE
Orlando, FL	Palm Bay/Melbourne, FL; Tampa/St. Petersburg, FL
Oxnard, CA	Bakersfield, CA; Los Angeles, CA; Mission Viejo, CA
Palm Bay/Melbourne, FL	Orlando, FL
Pensacola, FL	Mobile, AL
Philadelphia, PA	Allentown/Bethlehem, PA; Baltimore/Annapolis, MD; Dover, DE; Harrisburg, PA; Lancaster, PA; New York/Newark, NY; Scranton, PA; Trenton, NJ
Phoenix/Glendale/Mesa, AZ	
Pierre, SD	
Pittsburgh, PA	Youngstown, OH
Portland, OR	Salem, OR
Poughkeepsie/Newburgh, NY	Albany, NY; Bridgeport/Stamford, CT; Hartford, CT; New Haven, CT; New York/Newark, NY; Scranton, PA; Springfield, MA
Providence, RI	Boston, MA; Concord, NH; Hartford, CT; New Haven, CT; Springfield, MA; Worcester, MA
Raleigh/Durham, NC	Greensboro/Winston Salem, NC
Reno, NV	Carson City, NV
Richmond, VA	Norfolk, VA; Washington, DC
Riverside/San Bernardino, CA	Los Angeles, CA; Mission Viejo, CA; San Diego, CA
Rochester, NY	Buffalo, NY; Syracuse, NY
Sacramento, CA	Modesto, CA; San Francisco/Oakland/San Jose/Fremont, CA; Stockton, CA
Salem, OR	Portland, OR
Salt Lake City/Ogden, UT	
San Antonio, TX	Austin, TX
San Diego, CA	Mission Viejo, CA; Riverside/San Bernardino, CA
San Francisco, CA ³	Modesto, CA; Sacramento, CA; Stockton, CA
Santa Fe, NM	Albuquerque, NM
Sarasota/Bradenton, FL	Cape Coral, FL; Tampa/St. Petersburg, FL
Scranton, PA	Allentown/Bethlehem, PA; Harrisburg, PA; Lancaster, PA; Philadelphia, PA; Poughkeepsie/Newburgh, NY; Trenton, NJ
Seattle/Tacoma, WA	Olympia, WA
Shreveport, LA	
Spokane, WA	
Springfield, IL	St. Louis, MO
Springfield, MA	Albany, NY; Boston, MA; Bridgeport/Stamford, CT; Concord, NH; Hartford, CT; New Haven, CT; Poughkeepsie/Newburgh, NY; Providence, RI; Worcester, MA
St. Louis, MO	Springfield, IL
Stockton, CA	Modesto, CA; Sacramento, CA; San Francisco, CA
Syracuse, NY	Rochester, NY
Tallahassee, FL	
Tampa/St. Petersburg, FL	Cape Coral, FL; Orlando, FL; Sarasota/Bradenton, FL

Continued on next page

Table S2 – Continued from previous page

Metropolitan area	Neighbors
Toledo, OH	Cleveland/Akron, OH; Detroit/Warren, MI; Flint, MI; Ft. Wayne, IN; Lansing, MI
Topeka, KS	Kansas City, MO
Trenton, NJ	Allentown/Bethlehem, PA; Dover, DE; Lancaster, PA; New York/Newark, NY; Philadelphia, PA; Scranton, PA
Tucson, AZ	
Tulsa, OK	Oklahoma City, OK
Washington, DC	Baltimore/Annapolis, MD; Dover, DE; Harrisburg, PA; Lancaster, PA; Richmond, VA
Wichita, KS	
Worcester, MA	Boston, MA; Concord, NH; Hartford, CT; New Haven, CT; Providence, RI; Springfield, MA
Youngstown, OH	Cleveland/Akron, OH; Pittsburgh, PA

¹ Los Angeles region also includes Glendale, Long Beach, and Huntington Beach, CA.
² Norfolk region also includes Chesapeake, Newport News, and Virginia Beach, VA.
³ San Francisco region also includes Oakland, San Jose, and Fremont, CA.

Figure S2: Map of 132 urban centers from main Table 1, coded to indicate neighborhood status based on 100-mile-radius neighborhoods: increasing hues indicate more 100-mile neighbors for each location. Also overlaid are five-pointed stars (★) if the locality experienced one or more terrorist-related casualties (i.e., $Y_i = 1$ in main Table 1). © 2017 HVRI. All rights reserved.

Table S3: The numerical longitudes (as negative degrees, west from prime meridian) and latitudes (as degrees north from equator) used for the 132 U.S. metropolitan areas.

Urban Center	Long.	Lat.	Urban Center	Long.	Lat.
Albany, NY	-73.80	42.67	Little Rock, AR	-92.35	34.72
Albuquerque, NM	-106.62	35.12	Los Angeles, CA ¹	-118.41	34.11
Allentown/Bethlehem, PA	-75.48	40.60	Louisville, KY	-85.74	38.22
Amarillo, TX	-101.82	35.20	Lubbock, TX	-101.88	33.58
Anchorage, AK	-149.19	61.18	Madison, WI	-89.39	43.08
Atlanta, GA	-84.42	33.76	McAllen, TX	-98.24	26.22
Augusta, GA	-81.99	33.46	Memphis, TN	-90.01	35.11
Augusta, ME	-69.77	44.32	Miami/Ft. Lauderdale, FL	-80.21	25.78
Austin, TX	-97.75	30.31	Milwaukee, WI	-87.97	43.06
Bakersfield, CA	-119.00	35.36	Minneapolis/St. Paul, MN	-93.27	44.96
Baltimore/Annapolis, MD	-76.61	39.30	Mission Viejo, CA	-117.65	33.61
Barre/Montpelier, VT	-72.57	44.26	Mobile, AL	-88.09	30.68
Baton Rouge, LA	-91.13	30.45	Modesto, CA	-120.99	37.66
Birmingham, AL	-86.80	33.53	Montgomery, AL	-86.28	32.35
Bismarck, ND	-100.77	46.81	Nashville, TN	-86.78	36.17
Boise, ID	-116.23	43.61	New Haven, CT	-72.92	41.31
Boston, MA	-71.02	42.34	New Orleans, LA	-89.93	30.07
Bridgeport/Stamford, CT	-73.20	41.19	New York/Newark, NY	-73.94	40.67
Buffalo, NY	-78.86	42.89	Norfolk, VA ²	-76.24	36.92
Cape Coral, FL	-82.00	26.64	Oklahoma City, OK	-97.51	35.47
Carson City, NV	-119.74	39.15	Olympia, WA	-122.89	47.04
Charleston, SC	-79.99	32.79	Omaha, NE	-96.01	41.26
Charleston, WV	-81.63	38.35	Orlando, FL	-81.37	28.50
Charlotte, NC	-80.83	35.20	Oxnard, CA	-119.21	34.20
Chattanooga, TN	-85.26	35.07	Palm Bay/Melbourne, FL	-80.66	27.99
Cheyenne, WY	-104.79	41.15	Pensacola, FL	-87.19	30.44
Chicago, IL	-87.68	41.84	Philadelphia, PA	-75.13	40.01
Cincinnati, OH	-84.51	39.14	Phoenix/Glendale/Mesa, AZ	-112.07	33.54
Cleveland/Akron, OH	-81.68	41.48	Pierre, SD	-100.34	44.37
Colorado Springs, CO	-104.76	38.86	Pittsburgh, PA	-79.98	40.44
Columbia, SC	-80.89	34.04	Portland, OR	-122.66	45.54
Columbus, GA	-84.87	32.51	Poughkeepsie/Newburgh, NY	-73.92	41.71
Columbus, OH	-82.99	39.99	Providence, RI	-71.42	41.82
Concord, NH	-71.56	43.23	Raleigh/Durham, NC	-78.66	35.82
Corpus Christi, TX	-97.29	27.71	Reno, NV	-119.82	39.54
Dallas/Ft. Worth, TX	-96.77	32.79	Richmond, VA	-77.47	37.53
Dayton, OH	-84.20	39.78	Riverside/San Bernardino, CA	-117.40	33.94
Denver, CO	-104.87	39.77	Rochester, NY	-77.62	43.17
Des Moines, IA	-93.62	41.58	Sacramento, CA	-121.47	38.57
Detroit/Warren, MI	-83.10	42.38	Salem, OR	-123.02	44.92
Dover, DE	-75.53	39.16	Salt Lake City/Ogden, UT	-111.93	40.78
El Paso, TX	-106.44	31.85	San Antonio, TX	-98.51	29.46
Flint, MI	-83.69	43.02	San Diego, CA	-117.14	32.81
Frankfort, KY	-84.86	38.20	San Francisco, CA ³	-122.45	37.77
Fresno, CA	-119.79	36.78	Santa Fe, NM	-105.95	35.68
Ft. Wayne, IN	-85.14	41.07	Sarasota/Bradenton, FL	-82.55	27.34
Grand Rapids, MI	-85.66	42.96	Scranton, PA	-75.67	41.40
Greensboro/Winston Salem, NC	-79.83	36.08	Seattle/Tacoma, WA	-122.35	47.62
Harrisburg, PA	-76.88	40.28	Shreveport, LA	-93.80	32.47

Continued on next page

Table S3 – *Continued from previous page*

Urban Center	Long.	Lat.	Urban Center	Long.	Lat.
Hartford, CT	−72.68	41.77	Spokane, WA	−117.41	47.67
Helena, MT	−112.03	46.60	Springfield, IL	−89.64	39.78
Honolulu, HI	−157.80	21.32	Springfield, MA	−72.54	42.12
Houston, TX	−95.39	29.77	St. Louis, MO	−90.24	38.64
Huntsville, AL	−86.63	34.71	Stockton, CA	−121.31	37.97
Indianapolis, IN	−86.15	39.78	Syracuse, NY	−76.14	43.04
Jackson, MS	−90.21	32.32	Tallahassee, FL	−84.28	30.46
Jacksonville, FL	−81.66	30.33	Tampa/St. Petersburg, FL	−82.48	27.96
Jefferson City, MO	−92.19	38.57	Toledo, OH	−83.58	41.66
Juneau, AK	−134.42	58.30	Topeka, KS	−95.69	39.04
Kansas City, MO	−94.55	39.12	Trenton, NJ	−74.76	40.22
Knoxville, TN	−83.95	35.97	Tucson, AZ	−110.89	32.20
Lancaster, PA	−76.30	40.04	Tulsa, OK	−95.92	36.13
Lansing, MI	−84.55	42.71	Washington DC	−77.02	38.91
Las Vegas, NV	−115.22	36.21	Wichita, KS	−97.34	37.69
Lexington, KY	−84.46	38.04	Worcester, MA	−71.81	42.27
Lincoln, NE	−96.69	40.82	Youngstown, OH	−80.65	41.10

¹ Los Angeles region also includes Glendale, Long Beach, and Huntington Beach, CA.

² Norfolk region also includes Chesapeake, Newport News, and Virginia Beach, VA.

³ San Francisco region also includes Oakland, San Jose, and Fremont, CA.

S2. 10×10 ‘map’ lattice for PVI generation in simulation study

Table S4 contains an abstraction of the 10×10 ‘map’ lattice used in the main document’s Monte Carlo study on the operating characteristics of $\widehat{\text{BMI}}_{\text{BMR}}$ and BMIL_{BMR} . The lattice was configured by manipulating a selection of $n = 100$ metropolitan areas—as listed in the table—from the original 132-cities data set. The 100 positions on the lattice were chosen roughly equivalent to their cities’ relative positions on the U.S. map. The table also provides the corresponding PVIs used for the final 10×10 lattice.

Table S4: 10×10 ‘map’ lattice of PVIs used in simulation study, based on manipulation of original 132-cities data set.

Anchorage 2.7574	Juneau 1.2982	Bismarck 3.1007	Minneapolis 4.0594	Milwaukee 4.0798	Flint 3.3725	Buffalo 3.9979	Rochester 3.4563	Barre 2.3082	Augusta 2.9088
Seattle 3.4041	Spokane 2.9391	Pierre 2.2114	Madison 3.3032	Grand Rapids 3.0826	Detroit 3.9066	Cleveland 4.9107	Syracuse 4.13	Albany 4.1985	Concord 2.2349
Olympia 2.6381	Helena 1.9902	Cheyenne 2.6906	Des Moines 3.8765	Chicago 5.1226	Ft. Wayne 2.7801	Youngstown 4.1432	Scranton 4.1574	Worcester 2.8671	Boston 4.3228
Portland 3.7296	Boise 5.4146	Denver 3.8896	Omaha 4.0497	Springfield 3.0124	Indianapolis 3.3399	Columbus 3.6835	Allentown 3.9328	Hartford 3.3776	New Haven 3.8205
Sacramento 3.4645	Reno 2.7518	Santa Fe 2.5059	Lincoln 3.1743	Jefferson City 3.0293	Louisville 3.8411	Cincinnati 3.941	Harrisburg 3.9369	Philadelphia 5.4557	New York 5.8728
San Francisco 3.7689	Carson City 2.5796	Amarillo 3.419	Kansas City 3.9785	Tulsa 3.6948	Nashville 3.3678	Frankfort 2.7388	Lancaster 3.6038	Washington DC 5.6974	Dover 3.6643
Stockton 3.3074	Modesto 2.9111	Lubbock 3.1102	Wichita 3.0885	Little Rock 3.701	Memphis 4.6322	Chattanooga 4.1828	Greensboro 4.5331	Raleigh 4.7356	Norfolk 6.0454
Bakersfield 2.8801	Las Vegas 3.3285	Tucson 2.9513	Austin 3.6105	Shreveport 4.2491	Birmingham 3.9051	Atlanta 5.4024	Charlotte 4.7649	Columbia 4.8555	Charleston 6.2617
Oxnard 2.5121	Riverside 3.5439	El Paso 3.9081	Houston 5.5626	Baton Rouge 6.7354	Jackson 4.2785	Montgomery 3.8242	Columbus 3.6138	Augusta 4.4504	Jacksonville 3.8098
Honolulu 3.4816	San Diego 3.2004	San Antonio 4.6336	McAllen 4.3957	New Orleans 6.8381	Mobile 4.1436	Pensacola 4.3368	Tallahassee 3.7221	Cape Coral 4.2622	Orlando 4.4659

S3. Comparisons with the previous 2007 benchmark analysis

As noted in the main Discussion section, we caution against making any direct comparisons between our previous benchmark analysis of these data (Piegorsch et al., 2007) and that seen in the main document. While both presentations are rooted in geospatial comparisons among the same 132 U.S. urban centers, the underlying modeling framework and the nature of the benchmark analysis employed under our autologistic approach differ in a substantive fashion from those implemented in our earlier, more-simplistic article.

It is still worthwhile, at least, to provide a basic, visualization-based comparison of our reported vulnerability profiles with these 132 cities. Figure S3 reproduces a U.S. map with the 132 cities from Piegorsch et al. (2007). It presents the basic red/yellow/green delineation we gave then to distinguish cities felt to be at highest/moderate/lesser vulnerability, respectively, under the simpler models employed in that article. Compare this with the main document's Figure 2, which gives our updated, spatially adjusted, autologistic vulnerability profile and its different high/moderate/lesser vulnerability delineations. (For ease of comparison, we reproduce the figure below as Figure S4, using a similar red/yellow/green color scheme.)

Juxtaposing the two figures, one sees a greater prevalence of red/higher vulnerability cities presented under our autologistic analysis, along with a consequent decrease in yellow/moderate and, in particular, green/lesser vulnerability cities. The general pattern of more red cities in the Southeast and Atlantic states is retained, but many more yellow cities from the 2007 analysis now recharacterize into the red delineation. Important yellow-to-red recharacterizations appear in NY, New England, FL, and TX; Los Angeles has moved from yellow to red as well. Corresponding recharacterizations from green to yellow appear throughout the map. Overall, it appears that by incorporating spatial autocorrelation into the vulnerability analysis for these 132 cities, a somewhat more-severe indication of the vulnerability palette appears across these 132 cities, at least under the benchmark settings we employ.

Also, comparing the patterns in these two figures with the information in Figure S1, above, isolates an interesting fact: each of Boise, ID, Houston, TX, and Norfolk, VA exhibit high PVIs that place them in the upper vulnerability tier in both our earlier and current analyses, but from Figure S1 we see none recorded a terrorist casualty event. Upon closer inspection of the underlying data from Borden et al. (2007), we can report that each locale appears to suffer from certain physical hazards more than most others in our study. In particular, Boise experienced substantial property damage due to a major wind/hail storm in 1972, major flooding and hail damage in 1987, and a huge wildfire in 1992. These events generate large physical-hazard vulnerability, driving up Boise's hazard vulnerability (HazVI) component in our PVI metric; see Piegorsch et al. (2007).

Similarly, Houston experienced three separate, severe storm/flood events: a 1979 tropical storm, 1983's Hurricane Alicia, and 2001's tropical storm Allison. Again, these drove up its HazVI component. The Houston area also possesses a substantial built-environment infrastructure—one of the largest in the U.S.—which raises its built-environment vulnerability (BEVI) component. BEVI is another fundamental constituent of the PVI (Piegorsch et al., 2007).

Comparable factors also affect Norfolk. The region experienced three large hurricane events: Hurricane Bonnie in 1998, Hurricane Floyd in 1999, and Hurricane Isabel in 2003,

Figure S3: Map of 132 cities from main Table 1, coded to indicate vulnerability status from our original risk/vulnerability analysis (Piegorsch et al., 2007): green indicates lesser city vulnerability, yellow indicates moderate city vulnerability, while red indicates high city vulnerability. Note that the map's reference to PVI is the unshifted PVI07 described in the main document. © 2007 Wiley-Blackwell Publishers. All rights reserved.

Figure S4: Map of 132 cities from main Table 1, coded to indicate vulnerability status from our proposed autologistic risk/vulnerability analysis: green indicates lesser city vulnerability, yellow indicates moderate city vulnerability, while red indicates high city vulnerability. © 2016 HVRI. All rights reserved.

leading to substantial flooding/wind/hail damage and a resultantly large HazVI component. And, similar to Houston, Norfolk’s infrastructure is quite complex: it harbors a major naval complex and other heavy port activity that requires many more miles of supporting roads, bridges, tunnels, and rail lines than a typical region of its size. This inflates its BEVI component, even larger than Houston’s. Synthesized into a single index, Norfolk’s vulnerabilities produce one of the highest PVIs in our study.

Taken together, these various factors led to substantive PVI scores for each of these three metropolitan areas. While none of these localities may have recorded a terrorist casualty event during the period of our original study, their vulnerability landscapes as summarized by the PVI nonetheless place them at the high end of our spectrum for heightened vulnerability to casualties from a terrorist event, at least based on the 35-year period in our study’s database.

S4. Non-benchmark risk rankings

Past its use as a tool for benchmark analysis, our autologistic model possesses features that allow for further examination of urban terrorism vulnerability. For example, it is natural to ask how the various cities in our database rank with respect to terrorist casualty risk. To answer this question, one could appeal directly to the PVI as a standalone metric, since by its very design it provides an ordered measure of place-based vulnerability to hazardous outcomes. The PVI is a broad-scale measure, however, and does not take into account the specific information available in our database regarding each city’s terrorism casualty experience. For a more dedicated ranking, and to adjust for potential spatial autocorrelation among the responses, we can turn to the centered autologistic model in main document Equation (2). A variety of possible ordered metrics from the model in (2) could be imagined; certainly one of the most convenient, easily understood, and readily accessible is the predicted probability of response based on (2):

$$\hat{\pi}_i = \frac{1}{1 + \exp\{-\hat{\beta}_0 - \hat{\beta}_1 x_i - \hat{\beta}_2 \sum_{j \in N_i} a_{ij}(y_j - \hat{\mu}_j)\}},$$

where x_i is the i th city’s PVI, each $\hat{\beta}_j$ is the MPLE for the corresponding autologistic parameter, and $\hat{\mu}_j = \{1 + \exp(-\hat{\beta}_0 - \hat{\beta}_1 x_j)\}^{-1}$ is the MPLE for the mean response under the independence model. Using *ngspatial*’s `autologistic()` function, the $\hat{\pi}_i$ s are available as the `fitted.values` attribute from the model fit. Ranking them from largest to smallest gives a quantitative expression of terrorism casualty risk based on the outcomes in our database. This will differ from the broader PVI rankings whenever $\hat{\beta}_2 \neq 0$.

For example, Table S5 compares cities with the 10 largest values of PVI to those with the 10 largest values of $\hat{\pi}_i$ for our 132-cities data. Some intriguing patterns emerge. First and most obvious from a statistical perspective: very small differences between either set of values likely do not represent meaningful differences, so two closely ranked cities may be essentially equivalent in terms of their vulnerability status. With this caveat in mind, however, when going from PVI-based to $\hat{\pi}_i$ -based ranks Washington DC’s rise using $\hat{\pi}_i$ is notable, as is the removal of Baton Rouge, LA; New York, NY; Richmond, VA; Houston,

TX; and Boise, ID from the top-ten. In particular, New York and Houston are two of the largest metropolitan centers in the U.S., so their demotion is worthy of notice. And, Boise’s large PVI had made it particularly noticeable in all of our benchmark-based vulnerability analyses—cf. the discussion in main document Section 6—but it now drops lower in the $\hat{\pi}_i$ -based rankings (to 16th; see Table S6). Conversely, the new $\hat{\pi}_i$ -based top-ten list contains only cities located on or east of the Mississippi River, focusing perhaps even greater attention on the eastern U.S. For completeness, S6 gives the rankings based on $\hat{\pi}_i$ for all 132 cities.

Table S5: Top-ten U.S. metropolitan areas ranked by highest PVI (left columns) or by highest centered autologistic predicted probabilities of response, $\hat{\pi}_i$ (right columns), based on 132-cities data in main document Table 1.

Ranked by PVI		Ranked by $\hat{\pi}_i$	
Metropolitan area	PVI	Metropolitan area	$\hat{\pi}_i$
New Orleans, LA	6.8381	Washington DC	0.7656
Baton Rouge, LA	6.7354	New Orleans, LA	0.7316
Charleston, SC	6.2617	Philadelphia, PA	0.6831
Norfolk, VA ¹	6.0454	Norfolk, VA ¹	0.6354
New York/Newark, NY	5.8728	Columbia, SC	0.5875
Washington DC	5.6974	Tampa/St. Petersburg, FL	0.5786
Richmond, VA	5.6545	Greensboro/Winston Salem, NC	0.5624
Houston, TX	5.5626	Charleston, SC	0.5320
Philadelphia, PA	5.4557	Detroit/Warren, MI	0.5214
Boise, ID	5.4146	Boston, MA	0.5143

¹ Norfolk region also includes Chesapeake, Newport News, and Virginia Beach, VA.

Table S6: Full set of 132 U.S. metropolitan areas ranked from highest to lowest predicted probability of response, $\hat{\pi}_i$, under centered autologistic model. Original data from main document Table 1.

Metropolitan area	$\hat{\pi}_i$
Washington DC	0.7656
New Orleans, LA	0.7316
Philadelphia, PA	0.6831
Norfolk/Chesapeake/Newport News/Virginia Beach, VA	0.6354
Columbia, SC	0.5875
Tampa/St. Petersburg, FL	0.5786
Greensboro/Winston Salem, NC	0.5624
Charleston, SC	0.5320
Detroit/Warren, MI	0.5214
Boston, MA	0.5143
Cleveland/Akron, OH	0.4818
Baton Rouge, LA	0.4682
Richmond, VA	0.4652
Houston, TX	0.4603
Augusta, GA	0.4530
Boise, ID	0.4452
Atlanta, GA	0.4440
San Antonio, TX	0.4410
Pittsburgh, PA	0.4284
Pensacola, FL	0.4234

Continued on next page

Table S6 – *Continued from previous page*

Metropolitan area	$\hat{\pi}_i$
Cape Coral, FL	0.4208
Chicago, IL	0.4158
Dayton, OH	0.4103
Scranton, PA	0.4006
Denver, CO	0.3994
Lancaster, PA	0.3860
Palm Bay/Melbourne, FL	0.3748
Buffalo, NY	0.3746
Omaha, NE	0.3740
St. Louis, MO	0.3707
Cincinnati, OH	0.3700
Memphis, TN	0.3678
Austin, TX	0.3641
Louisville, KY	0.3463
Columbus, OH	0.3457
McAllen, TX	0.3455
Miami/Ft. Lauderdale, FL	0.3379
Portland, OR	0.3354
Jackson, MS	0.3347
Shreveport, LA	0.3320
Frankfort, KY	0.3285
Albany, NY	0.3274
Chattanooga, TN	0.3260
Dallas/Ft. Worth, TX	0.3245
Madison, WI	0.3230
Syracuse, NY	0.3212
Los Angeles/Glendale/Long Beach/Huntington Beach, CA	0.3188
Charlotte, NC	0.3177
Minneapolis/St. Paul, MN	0.3149
New York/Newark, NY	0.3148
Las Vegas, NV	0.3144
Seattle/Tacoma, WA	0.3052
Corpus Christi, TX	0.3052
Huntsville, AL	0.3037
Salt Lake City/Ogden, UT	0.3018
El Paso, TX	0.3017
Birmingham, AL	0.3014
Lexington, KY	0.3013
Des Moines, IA	0.2989
Wichita, KS	0.2963
Topeka, KS	0.2946
Montgomery, AL	0.2944
Jacksonville, FL	0.2932
Oklahoma City, OK	0.2896
Tallahassee, FL	0.2858
Grand Rapids, MI	0.2850
Little Rock, AR	0.2840
Knoxville, TN	0.2831
Baltimore/Annapolis, MD	0.2769
Columbus, GA	0.2768
Orlando, FL	0.2765

Continued on next page

Table S6 – *Continued from previous page*

Metropolitan area	$\hat{\pi}_i$
Sarasota/Bradenton, FL	0.2762
Charleston, WV	0.2721
Honolulu, HI	0.2660
Salem, OR	0.2651
Albuquerque, NM	0.2646
Amarillo, TX	0.2611
Phoenix/Glendale/Mesa, AZ	0.2610
Nashville, TN	0.2570
Sacramento, CA	0.2549
Indianapolis, IN	0.2548
Olympia, WA	0.2543
Fresno, CA	0.2534
Worcester, MA	0.2410
Lubbock, TX	0.2373
Santa Fe, NM	0.2371
Bismarck, ND	0.2366
San Francisco/Oakland/San Jose/Fremont, CA	0.2346
Jefferson City, MO	0.2313
Raleigh/Durham, NC	0.2307
Springfield, IL	0.2301
Spokane, WA	0.2248
Augusta, ME	0.2227
Harrisburg, PA	0.2192
Ft. Wayne, IN	0.2136
Anchorage, AK	0.2121
Toledo, OH	0.1994
Lansing, MI	0.1973
Springfield, MA	0.1910
New Haven, CT	0.1881
Mobile, AL	0.1873
Bakersfield, CA	0.1833
Barre/Montpelier, VT	0.1829
Providence, RI	0.1798
Allentown/Bethlehem, PA	0.1792
Pierre, SD	0.1770
Dover, DE	0.1758
Flint, MI	0.1724
Milwaukee, WI	0.1693
San Diego, CA	0.1675
Helena, MT	0.1641
Kansas City, MO	0.1601
Modesto, CA	0.1477
Tulsa, OK	0.1459
Rochester, NY	0.1438
Lincoln, NE	0.1307
Reno, NV	0.1298
Juneau, AK	0.1287
Oxnard, CA	0.1283
Carson City, NV	0.1234
Bridgeport/Stamford, CT	0.1227
Tucson, AZ	0.1097

Continued on next page

Table S6 – *Continued from previous page*

Metropolitan area	$\hat{\pi}_i$
Cheyenne, WY	0.1082
Trenton, NJ	0.1076
Youngstown, OH	0.1064
Mission Viejo, CA	0.1061
Hartford, CT	0.1009
Colorado Springs, CO	0.0993
Concord, NH	0.0950
Stockton, CA	0.0809
Poughkeepsie/Newburgh, NY	0.0794
Riverside/San Bernardino, CA	0.0566

For further visualization, Figure S5 plots $\hat{\pi}_i$ vs. PVI. Included in the scatterplot are the red/yellow/green vulnerability delineations for each city from Figure S4. As expected, a banded pattern appears: PVIs below BMIL_{10} are green, those above BMIL_{25} are red, and those in between are yellow. The larger pattern in the scatterplot shows a positive, increasing relationship between $\hat{\pi}_i$ and PVI, again as might be expected. The unmistakable string of points curving through the center of the trend is simply the set of cities where the spatial covariate equals zero—e.g., cities with no neighbors—so that their predicted probabilities fall along the simple logistic curve $1/(1 + \exp\{-\hat{\beta}_0 - \hat{\beta}_1 x_i\})$.

Figure S5: Plot of predicted probabilities, under the centered autologistic model, against PVI for 132-cities data. Colors indicate vulnerability status from our proposed autologistic risk/vulnerability analysis in Figure S4: green indicates lesser city vulnerability, yellow indicates moderate city vulnerability, while red indicates high city vulnerability.

S5. Sample R code

The following sample **R** function illustrates implementation of the external *ngspatial* package (Hughes, 2014) and its `autologistic()` function to

- (i) fit the centered autologistic model,
- (ii) call for $B = 5000$ bootstrap replicates to find an upper $1 - \alpha$ confidence limit on the β_1 parameter at $\alpha = 0.05$, and from these
- (iii) calculate $\widehat{\text{BMI}}_{\text{BMR}}$ and a (pointwise) $1 - \alpha$ bootstrap-based BMIL_{BMR} for a given BMR input value.

The required input data are the PVIs $\mathbf{x}$, their corresponding binary observations $\mathbf{Y}$, and the square adjacency matrix $\mathbf{A} = \{a_{ij}\}$. See the *ngspatial* help file (<https://CRAN.R-project.org/package=ngspatial>) for additional information.

```
library(ngspatial)
getBMI <- function(x, Y, A, B=5000, alpha=0.05, BMR=25, seed=123456, ...){
  set.seed(seed)
  X <- cbind(1, x)
  fit <- autologistic(Y ~ X - 1, A = A,
                     control = list(confint = "bootstrap", bootit = B))
  b1 <- fit$coefficients[2]
  b2 <- fit$coefficients[3]
  b1U <- quantile(fit$sample$X2, c(alpha,1-alpha), na.rm = T)[2]
  BMI <- log(BMR)/b1
  BMIL <- log(BMR)/b1U
  result <- list(beta1=b1, beta2=b2, beta1U=b1U, BMIhat=BMI, BMIL=BMIL)
  return(result)
} #end function
```

References

- Borden, K. A., Schmidtlein, M. C., Emrich, C. T., Piegorsch, W. W., & Cutter, S. L. (2007). Vulnerability of US cities to environmental hazards. *Journal of Homeland Security and Emergency Management*, 4(2), Article Number 5. doi: 10.2202/1547-7355.1279
- Hughes, J. (2014). *ngspatial*: A package for fitting the centered autologistic and sparse spatial generalized linear mixed models for areal data. *The R Journal*, 6(2), 81–95.
- Piegorsch, W. W., Cutter, S. L., & Hardisty, F. (2007). Benchmark analysis for quantifying urban vulnerability to terrorist incidents. *Risk Analysis*, 27(6), 1411–1425. doi: 10.1111/j.1539-6924.2007.00977.x