Received: 6 December 2021 | Revised: 11 May 2022 | Accepted: 29 May 2022

DOI: 10.1111/ssqu.13190

ORIGINAL ARTICLE

Can a measurement error perspective improve estimation in neighborhood effects research? A hierarchical Bayesian methodology

Duncan J. Mayer | **Robert L. Fischer**

Jack Joseph and Morton Mandel School of Applied Social Sciences, Case Western Reserve University, Cleveland, Ohio, USA

Correspondence
Duncan J. Mayer, Jack Joseph and Morton Mandel School of Applied Social Sciences, Case Western Reserve University, Cleveland, OH, USA.
Email: DuncanMayer77@gmail.com, Djm276@case.edu

Abstract

Objective: Neighborhood effects research often employs aggregate data at small geographic areas to understand neighborhood processes. This article investigates whether empirical applications of neighborhood effects research benefit from a measurement error perspective.

Methods: The article situates neighborhood effects research in a measurement error framework and then details a Bayesian methodology capable of addressing measurement concerns. We compare the proposed model to conventional linear models on crime data from Detroit, Michigan, as well as two simulated examples that closely mirror the sampling process.

Results: The Detroit data example shows that the proposed model makes substantial differences to parameters of interest and reduces the mean squared error. The simulations confirm the benefit of the proposed model, regularly recovering parameters and conveying uncertainty where conventional linear models fail.

Conclusion: A measurement error perspective can improve estimation for data aggregated at small geographic areas.

KEYWORDS
measurement error, Bayesian inference, census, neighborhood effects, crime, child welfare

Quantitative social scientists have shown great interest in linking administrative and/or publicly available data sources with other sources to effectively evaluate programs and policies, as well as to identify neighborhoods with a high risk of crime or child maltreatment. Incorporating these data has many benefits for researchers and consumers of research, including but not limited to reduced cost, improved timeliness, and increased access to long-term outcomes (Fischer et al. 2019). Neighborhood effects research, which seeks to map the conditions of a neighborhood to consequential outcomes, has been particularly active in this area (Graif, Gladfelter, and Matthews 2014), often pairing a variety of data sources at small geographic areas such as zip codes, census blocks, or tracts. However, public and administrative data are often not directly observed at small geographic units, resulting in aggregate data with observed uncertainty. The uncertainty in these data is ignored when researchers employ conventional models to address research questions. This is important, as data analyses desire to accurately communicate uncertainty, something essential for policymakers (Imbens 2021), and ignoring existing uncertainty in variables may result in overstating the precision of estimates or arriving at incorrect conclusions (Fewell, Davey Smith, and Sterne 2007; Gustafson 2004).

The unique circumstances of neighborhood effects research have led scholars to call for and develop methodological improvements (Coulton et al. 2007; Savitz and Raudenbush 2009). The extant literature, however, has not incorporated the sampling process unique to neighborhood effects research. This article contributes to the literature related to quantitative methodology in neighborhood effects research by investigating the benefit of a measurement error perspective on aggregate data at small geographic areas, focusing on areas of crime and child maltreatment. After describing the circumstances encountered in neighborhood effects research, the article describes a hierarchical Bayesian methodology capable of addressing measurement concerns by incorporating existing uncertainty as auxiliary information. We then compare the proposed model to conventional linear models in real and simulated examples. The results show substantial differences between adjusted and unadjusted estimates; however, models that do not adjust for the observed uncertainty do not recover parameters and often greatly overstate the precision of parameter estimates.

NEIGHBORHOOD EFFECTS RESEARCH

Researchers of child maltreatment and crime have considered structural contributors to disorder, motivated primarily by theories of social disorganization. Social disorganization is used to describe the relationship between the physical and social structure of a neighborhood, with a focus on social ties and the ability of residents to exercise control over their surroundings, leading to the consideration of structural contributors to patterns of criminal behavior and maltreatment rates (Coulton et al. 2007; Shaw and McKay 1942). Different approaches have been taken to incorporate neighborhood conditions; many have used administrative or public data on a level of aggregation designated as a proxy for a neighborhood (e.g., census tract or zip code) and used the proxy as the unit of analysis, where neighborhood-level information may be paired with person-level information (e.g., Coulton, Korbin, and Su 1999; Elo et al. 2009).

In the United States, neighborhood-level information is often drawn from the U.S. Census Bureau (e.g., the American Community Survey [ACS]). Child welfare researchers employ a variety of neighborhood-level variables, including financial information such as property values, the percent of the neighborhood under the federal poverty line, and the percent employed or participating in welfare programs. Demographic information has been of interest as well, including the percent of the neighborhood identifying as white or African American, the total population of the neighborhood, the percent of a neighborhood over the age of 65, and the percent of the neighborhood that identifies as Hispanic or Latino (Coulton, Korbin, and Su 1999; Drake and Pandey 1996; Slack et al. 2017). Similarly, crime researchers commonly use a variety of neighborhood-level variables from the same source, including the percent of homes that are vacant, female-headed households, and the percent of residents that identify as black, are under the federal poverty line, or divorced. Additionally, they have used indications of mobility, divorce rates, and many other variables in an effort to map neighborhood characteristics or changes in various types of crime (Baumer, Wolff, and Arnio 2012; Disha 2019; Wheeler, Kim, and Phillips 2018; Wo 2019).

Considerations for aggregate data

The neighborhood-level variables listed above are often incorporated into analyses as if the point estimates are observed from a random sample of neighborhoods without consideration of the existing uncertainty. This permits the use of classical methods of analysis (viz., the generalized linear model). Although insights may be gained from behaving as if the samples are generated in this manner, typically, the measurements are constructed differently. Generally, a sample is drawn from residents, the distribution of a variable of interest conditioned on a neighborhood or other level of aggregation, and then a point estimate, often a mean, median is computed. The variance of the chosen estimator depends on the sample and the chosen quantity. The ACS data are based on surveys of approximately 3 million households, rather than theoretically every household as is the case with the decennial census. The ACS regularly publishes a margin of error associated with a point estimate. The information contained in measures of sampling variability is vital to the interpretation of the point estimate, even when the variables are not the focus of the study, as variables measured with error treated as controls threaten the validity of inference for focal variables through residual confounding (Fewell, Davey Smith, and Sterne 2007).

Sampling variability

The Census Bureau's ACS regularly publishes margins of error, which relates to the classical view of confidence intervals (e.g., a statement about the estimation process), as it is half the width of a confidence interval for a given confidence level. The Census Bureau's standard confidence level is 0.9, corresponding to a z-score of approximately 1.645. Researchers may prefer a standard error, which is the standard error multiplied by a z score of 1.645. To recover the standard error, the margin of error is divided by the appropriate z-score. Researchers may also be interested in the variance, which is the square of the standard error (Fuller 2018).

MATERIALS AND METHODS

This section considers the Bayesian methodology and motivates a flexible class of prior distributions employed in the examples to follow. This is followed by a discussion of the approach to propagating uncertainty and approaching measurement error in neighborhood effects through hierarchical models. The proposed model is then illustrated and compared to conventional linear models on "real" and simulated data, first on publicly accessible assault data from Detroit, Michigan. Finally, two simulations show the impact and flexibility of the model and confirm the benefits of the proposed adjustments.

Bayesian inference

The familiar unstandardized form of the Bayes theorem can be written as $P(\theta|y) \propto P(\theta)P(y|\theta)$. For θ, a parameter (or vector of parameters) and y, the data. $P(\theta|y)$ is called the posterior, $P(\theta)$ the prior, and $P(y|\theta)$ the likelihood. Hierarchical models are a natural extension in this framework, where a joint prior distribution can be incorporated as $p(\phi, \theta|y) \propto p(\phi, \theta)p(y|\phi, \theta)$. Hierarchical models have many motivations, such as avoiding over- or underfitting or improving point estimates for individuals by borrowing strength (Congdon 2019; Gelman and Hill 2006); however, we focus on the ability of the hierarchical model to reflect the structure of the problem by allowing realizations and parameters to be drawn from a population distribution.

Prior distributions

In applied work, there has often been a desire to specify "uninformative" or "objective" prior distributions, including matching, flat, regularizing, or entropy maximizing prior distributions (Consonni et al. 2018). Yet, applied researchers should generally avoid diffuse or flat priors, as they can lead to degeneracies in the posterior in situations with finite data and complex models, as well as make understanding the joint behavior challenging, which is vital in Bayesian inference (Gelman, Simpson, and Betancourt 2017). In the examples that follow, we often employ diffuse regularizing priors. There are several decisions to be made about regularization, particularly the type and amount of regularization, including the location and scale of the prior distribution, as well as its analytic form. Common forms include Gaussian, which shrinks estimates toward the prior mean; Laplace, which shifts estimates by a constant amount; and Cauchy, which impact estimates more as they get farther from the prior mean (Gelman et al. 2003), and stronger regularization can be achieved with spike-slab or "horseshoe" priors (Congdon 2019). Intuitively, as the unstandardized posterior is the product of the likelihood and the prior, a more concentrated prior will apply stronger regularization. Regularization can be controlled by hyperpriors, which one can enhance further through richer models such as mixtures or regression (Congdon 2019; Gelman et al. 2003; Richardson et al. 2002). Simulating the prior predictive distribution can assist researchers in making reasonable choices (Gabry et al. 2019).

There is often little prior information for variance parameters in hierarchical models; consequently, they have received considerable attention in the literature. Historical examples often advocate the use of inverse-gamma priors for the variance of latent Gaussian parameters (a conjugate choice). Gustafson (1998) recommends a Cauchy distribution folded at the origin when little else is known, suggesting that it corresponds to a diffuse prior for variance between units. Other authors have emphasized penalizing priors, such that the high density lies on zero, effectively hedging toward simpler models (Simpson et al. 2017), or locally uniform priors corresponding to little prior knowledge (Gelman 2006).

Adjusting for measurement error

Although the Bayesian methodology does not maintain a monopoly on adjustments for measurement error, it retains several distinct advantages. In a Bayesian methodology, all unknowns are treated as parameters, which can be specified in the joint distribution over data and parameters such that samples can be drawn from the posterior distributions for the measurements. The incorporation of prior information is encoded into different points in the model, and finally, the framework naturally propagates the uncertainty from measurement to outcome models. For example, uncertainty in interval estimates is adjusted naturally once the joint distribution is specified (Gelman et al. 2003; Gustafson 2004; Richardson and Gilks 1993a).

Fundamentally, the goal of modeling measurement error is to appropriately condition the parameters and represent the measurements of the process. Broadly, there are three ways of gaining traction on this conditional distribution: (i) observe some number of cases for the true distribution, often called the validation study design or substudy, (ii) observe multiple realizations or with multiple instruments at a specific occasion, or (iii) design with ancillary information (Buzas, Stefanski, and Tosteson 2014; Richardson and Gilks 1993a). The first two suggestions for approaching measurement error are advantageous for research design stages and permit incorporating information when available. However, the recommendations of repeated draws or other validation study designs are often of limited applicability to those using public and administrative data in neighborhood effects studies. There is often auxiliary information available to use in adjusting for measurement error when building measurement models.

The simplest setting can be described using the conventional notation, labeling x^* the observations with error, x the unobserved true variable, and z a vector of covariates. Researchers can specify the distribution x^*, as conditional on x or z (or both), treating x as a parameter (more typically, a vector of parameters), and then conditioning the outcome on the parameters x of this model. In epidemiology, under a factorization that makes conditioning the variable with error explicit, the term that expresses the conditioning, $f(x^* \mid z, x)$, is called the measurement model and describes the relationship between the measures (surrogates) and true covariate. The factorization leads naturally into $f(x \mid z)$, the exposure model, which describes the distribution of the covariate, and $f(y \mid x)$, the response model, which specifies the relationship between the variate y and covariate(s) x (Richardson and Gilks 1993b). Generally, f is taken to be parametric, where the measurement model encodes assumptions about the amount of error generated, as well as the form of the measurement error mechanism.

While not the focus of this article, the specification of the exposure distribution plays an important role in these models. In the hierarchical extension Bayes theorem, the exposure model specifies the distribution of ϕ, and many of the assumptions encoded by the exposure model are the same as those found in other hierarchical models for second-stage parameters. These include the assumption that the vector of parameters, θ_j, can be drawn from ϕ, as well as the reasonability of the specified population distribution, including uni- or multimodality, and the thickness of the tails (Gelman et al. 2003). Misspecification of the exposure can impair inference for the population parameters as well as the response model. Misspecification of the exposure can impair inference for the population parameters as well as the response model. However, aspects of the response model are somewhat robust to misspecification of the exposure, particularly in the regression setting for well-measured covariates (Richardson and Leblond 1997). Despite the cases of robustness, challenges remain, and suggested refinements for more flexible exposure models include finite mixtures and classes of heavy-tailed distributions (Hossain and Gustafson 2009; Richardson et al. 2002).

Variables measured with the error may produce reasonable estimates if the amount of error is not too great. Gustafson (2004) shows a reparameterization of the normal variance in terms of the measurement error mechanism and the unobserved variable, emphasizing the impact of measurement error measured by $\mathrm{Var}(X^*|X)$ and suggesting that measurement error below 0.1 is not a serious threat to inference in continuous features (the categorical case is similar, although contingent on the base rate). This offers a way researchers may motivate adjustments in neighborhood effects research, where measurement error is often observed. There is some caution with the employment of such a threshold, however, as the implications of measurement error are often difficult to predict in practice, particularly in cases of multivariate measurement error or when the error is present in the dependent variable (Fewell, Davey Smith, and Sterne 2007; Gustafson 2004).

Addressing uncertainty in neighborhood effects with hierarchical models

Many of the neighborhood effect studies employ ACS data as measures of neighborhood characteristics. These data provide two values, a point estimate and a margin of error. For an outcome y_j, exposure of interest x_j, covariate z_j, which is not observed directly but for which point estimates z_j^* are drawn with standard errors e_j, all for neighborhood j, the unobserved variable can be made a function of the two available pieces of information. Assuming normally distributed measurement error, a measurement model can be written in the sampling form as $Z_j^* \sim f(Z_j, e_j)$. There is then great flexibility in the model for the population parameters of neighborhoods $Z_j \sim f(\gamma_z, \tau_z)$. Note that z_j is a vector of length j, with j equal to the number of neighborhoods, and represents the measurements without error. γ_z, τ_z, are then assigned prior distributions to model the variance and mean of neighborhoods, which may include elaboration with covariates (Gustafson 2004; Richardson and Gilks 1993a). The measurement error mechanism can be encoded further in the exposure and response models, particularly whether it is classical and/or nondifferential measurement error.

DATA EXAMPLES

In this section, we apply the proposed models to public and simulated data to compare their performance to that of conventional linear models. We first use ACS data to model publicly available assault data from Detroit in 2017. We then present simulated examples, in which uncertainty is introduced through additive error. The focus on additive error is a natural choice; additive measurement error can be motivated by multiplicative error through calibration or logarithmic transformation. For simplicity, the presentations focus on the case of continuous parameters of interest, which are also arguably the most common in neighborhood effects research. Furthermore, much of the neighborhood effects literature (e.g., the articles in the Neighborhood Effects Research section) tests theories through linear models; as a result, the examples that follow do not model the discreetness of the outcome, without loss of generality. As is typical in the measurement error literature, we describe an approach to these problems that ignores the uncertainty

TABLE 1 Select descriptive statistics for census tracts in Detroit, MI ($N = 282$)

	Mean	Median	Standard deviation	Min	Max
Total population	2354.05	2218.00	1119.38	293.00	5765.00
Margin of error	351.52	328.00	153.17	82.00	1008.00
African American	1902.68	1785.50	1073.11	94.00	5162.00
Margin of error	322.38	296.50	154.98	63.00	1037.00
Median income	34,975.02	32,083.00	16,686.61	10,045.00	134,300.00
Margin of error	11,640.09	9970.00	7477.86	1139.00	72,213.00
Vacancies	1251.99	1240.50	477.70	311.00	3367.00
Margin of error	55.43	52.00	22.39	12.00	122.00
Crime (assault offenses)	74.88	67.50	41.35	4.00	322.00

in the estimates as a “naïve approach” (Gustafson 2004; Hossain and Gustafson 2009); we compare the proposed model with linear models estimated by ordinary least squares (OLS). In the following examples, comparisons are made between credibility and confidence intervals, as well as the related standard errors and posterior standard deviations. Although these are theoretically distinct, the hope is that the comparison can illuminate the flexibility of our proposed methodology.

Detroit assault data

To compare our proposed methodology, we pair publicly accessible crime data from Detroit Michigan (Ashby 2019) with 5-year ACS data, each aggregated at the tract level. We focus on assault offenses that occur in census tracts in Detroit as a function of the characteristics of the tract, all measured in 2017. This is not meant to be a claim about the “real” effect of such characteristics on assaults; rather, it is meant as a demonstration of the methodology. The raw variables are used rather than percentages, and we include the population as a covariate. In addition to the number of residents in a tract, we collect measurements on features frequently seen in the literature, the number of residents who identify as black or African American, the median income, and the number of vacant homes in the tract. Select descriptive statistics and the margins of error are described in Table 1.

We follow the approach discussed above, defining a hierarchical model in which the realizations for a characteristic in a tract are conditional on the process defined by the joint distribution of the population parameters. Although this analysis proceeds with minimal use of domain knowledge, such a specification not only more closely reflects the state of our knowledge than independent distributions but also allows the application of domain knowledge in priors through the parameters of the population (e.g., mean and variance), as well as through the covariance pattern. The impact of the adjustments can be anticipated by studying the ratio of observed variability to the standard errors. However, the summary statistics for sampling variability lend themselves to an interpretation of errors being drawn from a single distribution, which is not likely to be the case in all data. The highest mean and standard deviation of the standard errors relative to the observations are found in median income with a scaled mean and standard deviation of 0.42 and 0.27, respectively, followed by population and race, each with means and standard deviation of approximately 0.19 and 0.09, respectively. This leads us to believe the adjustments will be most impactful for income, race, and population. However, any single observation may have a large amount of uncertainty.

Letting our four observed neighborhood characteristics be labeled x_{ij} with associated standard error e_{ij} for characteristic i in tract j, with the standard error treated as an estimate for the standard deviation of the measurement error, which we assume to be normally distributed and nondifferential, the model is described in Equations (1.1)–(1.4). where (1.1) is the measurement model, (1.2) the exposure model, (1.3)

TABLE 2 Main results for the Detroit assault data example

	Naïve[a]				Hierarchical specification				
Variable	**Estimate**	**SE**	**Confidence interval[b]**		**Mean**	**Median**	**SD**	**Credible interval[c]**	
Constant	4.11***	0.03	4.09	4.21	4.12	4.11	0.02	4.07	4.17
Total population	0.13*	0.05	0.03	0.24	0.20	0.20	0.07	0.09	0.31
AfricanAmerican/black	0.26***	0.04	0.18	0.34	0.29	0.29	0.05	0.21	0.36
Median income	−0.15**	0.03	−0.20	−0.09	−0.36	−0.36	0.06	−0.45	−0.27
Vacancy	0.02	0.05	−0.08	0.12	−0.01	−0.01	0.05	−0.11	0.09
MSE[3]	0.23				0.18				

Note. All variables are mean-centered and scaled to unit variance.

[a]Ordinary least squares on the aggregated data.

[b]95 percent.

[c]Mean squared error.

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

the response model, and (1.4) the priors. The adjusted observations are treated as parameters drawn from a common multivariate normal distribution with a 4 by 1 mean vector Θ and 4 by 4 covariance matrix Σ. We place standard normal priors on the population means, and each element of Θ. We then factor out the individual variances from the covariance matrix (Equation 1.2), as we prefer to assign individual priors to the variances and the correlation matrix. Each of the hierarchical variances is assigned a half-Cauchy prior describing diffuse predata belief, and we assign the correlation matrix the Lewandowski-Kurowicka-Joe (onion/LKJ) prior with parameter of 2 defining a distribution over the space of the correlation matrix (Lewandowski, Kurowicka, and Joe 2009). The parameters are used to define a linear model for the mean of y_j with coefficients that have independent normal priors and variance defined by a gamma distribution.

$$x_{1,j} \sim N\left(\lambda_j, e_{1,j}\right),\ x_{2,j} \sim N\left(\gamma_j, e_{2,j}\right)$$
$$x_{3,j} \sim N\left(\tau_j, e_{3,j}\right),\ x_{4,j} \sim N\left(\alpha_j, e_{4,j}\right), \tag{1.1}$$

$$\begin{gathered}\left[\lambda_j, \gamma_j \tau_j, \alpha_j\right]^t \sim \text{MVNORM}\left[\Theta,\ \Sigma\right]\\ \Theta = \left(\mu_\lambda, \mu_\gamma, \mu_\tau, \mu_\alpha\right)^t\\ \Sigma = \text{diag}\left(\sigma_\lambda, \sigma_\gamma, \sigma_\tau, \sigma_\alpha\right)\ \rho\ \text{diag}\left(\sigma_\lambda, \sigma_\gamma, \sigma_\tau, \sigma_\alpha\right)\\ \rho = \begin{bmatrix} 1 & \rho_{\lambda,\gamma} & \rho_{\lambda,\tau} & \rho_{\lambda,\alpha}\\ \rho_{\lambda,\gamma} & 1 & \rho_{\gamma,\tau} & \rho_{\gamma,\alpha}\\ \rho_{\lambda,\tau} & \rho_{\lambda,\tau} & 1 & \rho_{\tau,\alpha}\\ \rho_{\lambda,\alpha} & \rho_{\gamma,\alpha} & \rho_{\tau,\alpha} & 1\end{bmatrix},\end{gathered} \tag{1.2}$$

$$\ln(y_j) = \beta_0 + \beta_1\lambda_j + \beta_2\gamma_j + \beta_3\tau_j + \beta_4\alpha_j + N\left(0, \sigma_y\right)\ \forall \text{j}, \tag{1.3}$$

$$\begin{gathered}\beta_l \sim N(0,5),\ \Theta_k \sim N(0,1), \sigma_k \sim \text{cauchy}_+(0,1),\ \forall l, k\\ \sigma_y \sim gamma\,(2,1), \rho \sim lkj\,(5).\end{gathered} \tag{1.4}$$

The main results are shown in Table 2, we omit the results of the correlation matrix that describes the relationship among the population parameters. Table 2 shows that the estimates from the naïve approach suffer attenuation of varying amounts relative to the results of the hierarchical model. As anticipated, the largest difference between models is found in the estimates for median income. The estimate of the mean of the posterior marginal distribution for median income yields an interpretation of just over twice that of the point estimate from OLS, with estimates for a one standard deviation increase, corresponding to 30 percent and 14 percent decreases in crime. This is also a case when the relative magnitude of the parameters shifts due to different levels of measurement error, as the naïve approach shows the race variable to have the largest point estimate by absolute value. However, after taking the difference in samples from the hierarchical approach, the 95 percent credible interval for the difference in absolute values between was 0.03–0.29, suggesting a high probability (0.98) that the coefficient on median income has a higher absolute value. Differences in interpretation of the point estimates shift for the total population by approximately 8.5 percent and 4 percent for the parameter corresponding to African American or Black. Additionally, as anticipated, the interval estimates from OLS are narrower when compared to the credible intervals, failing to reflect the loss of information occurring in data aggregation. The hierarchical specification has also produced more accurate predictions, on average, as indicated by the lower mean squared error. Often, spatial effects become a concern as geographic units of interest become smaller, as is often the case with neighborhood effects research (Coulton et al. 2007), and a wide variety of approaches are available in the full Bayesian methodology (see, among others, Congdon 2019, chap. 6). In this analysis, we find no evidence of spatial autocorrelation in the residuals with a Moran's I below 0.005. Consequently, we dispense with modeling the spatial effects.

Refinements using domain knowledge

Prior distributions encoding information from previous research, practice experience, or expert opinion are important and can assist in identifying parameters (Gustafson 2004; Richardson and Gilks 1993a). Often, this information can be translated into knowledge of mechanisms available on a given variable, which may be related to the exposure and unrelated to the measurement error (e.g., dependent error). Such dependence can be leveraged within a measurement model or exposure model. In contrast to the Detroit assault example, in the following simulated examples, we explicitly model the exposure using a familiar linear form for lucidity.

Simulated examples

We consider two simplified situations in which we are interested in modeling an outcome Y that is measured without error. In situation one, let I be income, which is measured without error, I^* be the observed, aggregated, measure of income with error, and V be the vacancy rate in the neighborhood measured without error, on which both I and Y depend. Situation two is similar, where we observe I^*, V, and Y; however, we enter a second variable measured with error and aggregated, let P^* and P denote observed and accurately measured property values, respectively. In this setup, Y depends on V, I, and P; however, I depends on P, and P depends on V. In situation one, the measurement errors introduced are independent and identically distributed, while we relax this restriction in situation two. To simulate these data, we set the number of neighborhoods to 1000, with the number of residents in each neighborhood set to 50. In situation one, population-level income and vacancy are drawn from a bivariate normal distribution, which is fully specified by a vector of means and a covariance matrix. For situation one, these are specified according to the mean vector μ_1 and covariance matrix Σ_1.

$$\mu_1 = \begin{bmatrix} 40 \\ 0.3 \end{bmatrix} \; \Sigma_1 = \begin{bmatrix} 20 & 1.3 \\ 1.3 & 0.5 \end{bmatrix},$$

$$\mu_2 = \begin{bmatrix} 40 \\ 60 \\ 0.3 \end{bmatrix} \Sigma_2 = \begin{bmatrix} 20 & 10 & 0.5 \\ 10 & 30 & 3 \\ 0.5 & 3 & 1 \end{bmatrix}. \tag{2}$$

The entries in the first vector of means correspond to income and vacancy. The probability integral transformation is used to transform vacancies to a binomial distribution with n of 1 and p of 0.5; vacancies are measured without error and in these simulations are treated as binary (i.e., high or low vacancies). The hierarchical distribution is created by treating draws from the bivariate normal distribution for income as population incomes for a given neighborhood, such that each resident is assigned a random draw from a normal distribution with a mean equal to the neighborhood's income. Error is added to each resident's income by a draw from a mean zero half-normal distribution with a standard deviation of 30. Through random number generation of a normal density, we simulate an outcome with conditional expectation defined as $E(Y_j|I_j, V_j) = 2 + 2V_j - 3I_j$ for scaled I and V.

In situation two, the setup described above is extended to accommodate property value. In μ_2, the mean vector for the covariates in situation two, the second entry corresponds to property value. Y_j is generated under a normal density with conditional expectation defined as $\mathrm{E}(Y_j|P_j, I_j, V_j) = -2I_j + 1.5V - -3P_j$ for scaled I, V, and P. Measurement error is generated and added to both property value and income. In contrast to the first situation, the variance of the error for each resident's property value and income are sampled from a half-normal population distribution with standard deviations of 30 and 15, respectively. These draws are then used as the variance for a single draw from a normal distribution with mean zero and standard deviation set to the draw from the half-normal, giving heterogeneous error (at the population level) among tracts. We collect results from 10 independent simulations of each situation and average over uncertainty.

Application

In these situations, the reasonability of the vector of population parameters for each neighborhood is verifiable; however, the amount of measurement error introduced is substantial. We retain the approach described above with respect to the measurement models; however, we use an explicit prior model for the random vector of means within the measurement model (i.e., the exposure model). The full specifications for situations one and two are given in Equations (3.1)–(3.4), where (3.1) corresponds to the measurement model(s), (3.2) to the exposure model(s), (3.3) to the response model(s), and (3.4) details the priors given for each parameter included.

Situation 1	Situation 2	
$I_j^* \sim N(\zeta_j, e_{I,j})$	$I_j^* \sim N(\xi_j, e_{I,j}),\ P_j^* \sim N(\pi_j, e_{P,j})$	(3.1)
$\zeta_j = \alpha_0 + \alpha_1 V + N(0, \sigma_\zeta)\ \forall j$	$\pi_j = \alpha_0 + \alpha_1 V_j + N(0, \sigma_\pi)\ \forall j\ \ \xi_j = \alpha_2 + \alpha_3 \pi_j + N(0, \sigma_\xi)\ \forall j$	(3.2)
$Y_j = \beta_0 + \beta_1 \zeta_j + \beta_2 V_j + N(0, \sigma_y)\ \forall j$	$Y_j = \beta_0 + \beta_1 \xi_j + \beta_2 V_j + \beta_3 \pi_j + N(0, \sigma_y)\ \forall j$	(3.3)
$\alpha_i \sim N(0,5), \beta_k \sim N(0,5) \forall\ i, k\ \sigma_\zeta, \sigma_y \sim \text{cauchy}_+ (0,1)$	$\alpha_i \sim N(0,5), \beta_k \sim N(0,5) \forall\ i, k\ \sigma_\xi, \sigma_y, \sigma_\pi \sim \text{cauchy}_+ (0,1)$	(3.4)

In (3.1), it is important to note that a model for the measurement of vacancies is omitted, encoding the belief that vacancies are measured without error. The measurement models in (3.1) can be seen as describing multivariate nondifferential measurement error, using the standard errors as estimates. This is seen as we assume the independence of y and the measures with error conditional on the local dependencies specified in (3.3) and (3.4). This also assumes the reasonability of linear exposure and response models, as seen in blocks seven and eight.

Simulation results

In situation one, the naïve estimate shows that the measurement error on income inflates the estimate for vacancy and reduces the estimate for income. Table 3 shows that the coverage procedure excludes many plausible values for the parameters, including the fixed, true effects. The implementation of the Bayesian methodology with a model for the measurement of income has greatly reduced bias for each parameter when compared to the naïve estimate. The posterior mean for the effects of vacancy and income are within one standard deviation of the true parameters. Also of note, the credibility interval is wider than the confidence interval, as the Bayesian methodology has propagated the uncertainty in the data and reflects the loss of information in the measurement process.

In situation two, the naïve model shows reduced estimates for the variables measured with error, income, and property value; however, the attenuation on income is small, and coverage for income does include the true value. The estimate for vacancy is, again, inflated, with more substantial bias than appears in situation one. As with situation one, the coverage procedure for the naïve approach excludes the true value for coefficients on property value and vacancy and provides false confidence in a narrow interval estimate. The Bayesian methodology has reduced bias for the effects, and the true values are within a single standard deviation from the posterior mean. The standard deviations of the posterior distributions show greater variability than the standard error, as the naïve approach again underestimated the uncertainty in the effects following aggregation.

DISCUSSION

This article has contributed to the methodological literature focused on neighborhood effects by investigating the benefits of a measurement error perspective. The results show that treating existing uncertainty as measurement error in aggregate data for small geographic units has the capacity to greatly improve estimation: reducing bias, recovering parameters, and appropriately conveying uncertainty. In contrast, the results of this article show that ignoring existing uncertainty results in models that may fail to recover parameters, including cases when the focal variable is well measured. This is important, as we found no study to date has adjusted for existing uncertainty in aggregate data. In addition to these substantive contributions, the article has detailed a flexible hierarchical Bayesian methodology that has not been introduced to neighborhood effects research. We have shown that the model is flexible enough to fit a variety of research settings, including different dependence structures found in population parameters, as well as multivariate measurement error. Rather than assuming perfect measurements, researchers may consider whether the construct of interest is accurately measured by the data at hand as well as the magnitude of the error. This article has also demonstrated the utility of differentiating between local models, particularly in clarifying assumptions about measurements, population distributions, and outcome models. It is of paramount importance to model the features that are most important to a given social problem. Our examples show how challenging it can be to predict the impact of measurement error, particularly in the setting with multivariate measurement error with correlated independent variables.

Consumers of research, such as policymakers, require optimal information when engaged in decision making, and researchers may endeavor to accurately communicate uncertainty. For example, when designing interventions, the expected returns may be catalyzed by the estimated uncertainty, and ignoring the observed uncertainty can lead to narrower interval estimates, mis-calibrating the expectations of policymakers (Imbens 2021). Incorrect estimates may lead policymakers to focus on intervention points of less importance or to miss features entirely. In the context of the situations presented through simulation, policymakers viewing the results of the naïve approach may feel certain that reducing vacancy may be the most advantageous intervention for a given neighborhood, leading to recommendations such as demolitions, when direct cash transfers may be more beneficial, *ceteris paribus* (of course, this depends on the distribution of income in the neighborhood). Although this article has focused on empirical results,

TABLE 3 Main results for simulated Situations 1 and 2

Variable	Naïve estimate[a]				Hierarchical specification					True value[b]
	Estimate	SE	Confidence[c]		Mean	Median	SD	Credible[c]		
Situation 1										
Income	−1.55	0.08	−1.71	−1.40	−3.14	−3.13	0.21	−3.50	−2.80	−3
	(0.09)	(0.00)	(0.05)	(0.06)	(0.22)	(0.22)	(0.02)	(0.26)	(0.19)	
Vacancy	2.92	0.23	2.48	3.37	1.84	1.84	0.21	1.34	2.33	2
	(0.19)	(0.03)	(0.19)	(0.20)	(0.27)	(0.27)	(0.02)	(0.28)	(0.26)	
Situation 2										
Income	−1.88	0.10	−2.08	−1.68	−2.03	−2.03	0.14	−2.26	−1.79	−2
	(0.12)	(0.00)	(0.12)	(0.11)	(0.15)	(0.15)	(0.01)	(0.16)	(0.15)	
Vacancy	3.12	0.24	2.64	3.60	1.91>	1.91	0.26	1.47	2.34	2
	(0.29)	(0.01)	(0.29)	(0.29)	(0.27)	(0.27)	(0.01)	(0.28)	(0.27)	
Property value	−1.68	0.10	−1.87	−1.49	−3.00	−3.00	0.16	−3.27	−2.73	−3
	(0.15)	(0.00)	(0.15)	(15)	(0.20)	(0.10)	(0.01)	(0.20)	(0.20)	

Note. The results of 10 independent simulations. Standard deviations of the 10 simulations are presented in parentheses, with the mean of the indicated quantity above.
[a]Ordinary least squares on the aggregated data.
[b]Population parameters set during simulation.
[c]95 percent interval estimates.

there are important implications for theory building. The cumulative effects of measurement error may have deleterious effects on the development of theory in a field that relies heavily on empirical results; propitious theories may be left behind when not empirically supported. The effect of underestimating uncertainty may be most clear in the use of meta-analysis, as they integrate and estimate over effects that have underestimated uncertainty, creating substantive problems for further theoretical development. Future research may investigate the effects accounting for uncertainty in transformations of aggregated data, including indices for concentrated disadvantage and poverty (see, among others, Hannon 2005).

CONFLICTS OF INTEREST

The authors have no relevant financial or nonfinancial interests to disclose.

ORCID

Duncan J. Mayer https://orcid.org/0000-0002-4448-3353
Robert L. Fischer https://orcid.org/0000-0001-5642-6746

REFERENCES

Ashby, M. P. J. 2019. "Studying Crime and Place with the Crime Open Database: Social and Behavioral Sciences." *Research Data Journal for the Humanities and Social Sciences* 4(1):65–80. https://doi.org/10.1163/24523666-00401007

Baumer, Eric P., Kevin T. Wolff, and Ashley N. Arnio. 2012. "A Multicity Neighborhood Analysis of Foreclosure and Crime: A Multicity Neighborhood Analysis of Foreclosure and Crime." *Social Science Quarterly* 93(3):577–601. https://doi.org/10.1111/j.1540-6237.2012.00888.x

Buzas, Jeffrey, Leonard Stefanski, and Tor Tosteson. 2014. "Measurement Error." In *Handbook of Epidemiology*, 2nd ed., edited by W. Ahrens and I. Pigeot, 729–65. New York: Springer Reference.

Congdon, Peter D. 2019. *Bayesian Hierarchical Models: With Applications Using R, Second Edition*, 2nd ed. Boca Raton: Chapman and Hall/CRC.

Consonni, Guido, Dimitris Fouskakis, Brunero Liseo, and Ioannis Ntzoufras. 2018. "Prior Distributions for Objective Bayesian Analysis." *Bayesian Analysis* 13(2):627–79. https://doi.org/10.1214/18-BA1103

Coulton, Claudia J., David S. Crampton, Molly Irwin, James C. Spilsbury, and Jill E. Korbin. 2007. "How Neighborhoods Influence Child Maltreatment: A Review of the Literature and Alternative Pathways." *Child Abuse & Neglect* 31(11–12):1117–42. https://doi.org/10.1016/j.chiabu.2007.03.023

Coulton, Claudia J., Jill E. Korbin, and Marilyn Su. 1999. "Neighborhoods and Child Maltreatment: A Multi-Level Study." *Child Abuse & Neglect* 23(11):1019–40. https://doi.org/10.1016/S0145-2134(99)00076-9.

Disha, Ilir. 2019. "Different Paths: The Role of Immigrant Assimilation on Neighborhood Crime." *Social Science Quarterly* 100(4):1129–53. https://doi.org/10.1111/ssqu.12618

Drake, Brett, and Shanta Pandey. 1996. "Understanding the Relationship between Neighborhood Poverty and Specific Types of Child Maltreatment." *Child Abuse & Neglect* 20(11):1003–18. https://doi.org/10.1016/0145-2134(96)00091-9

Elo, Irma T., Laryssa Mykyta, Rachel Margolis, and Jennifer F. Culhane. 2009. "Perceptions of Neighborhood Disorder: The Role of Individual and Neighborhood Characteristics." *Social Science Quarterly* 90(5):23. https://doi.org/10.1111/j.1540-6237.2009.00657.x

Fewell, Z., G. Davey Smith, and J. A. C. Sterne. 2007. "The Impact of Residual and Unmeasured Confounding in Epidemiologic Studies: A Simulation Study." *American Journal of Epidemiology* 166(6):646–55. https://doi.org/10.1093/aje/kwm165

Fischer, Robert L., Francisca García Richter, E. Anthony, N. Lalich, and C. Coulton. 2019. "Leveraging Administrative Data to Better Serve Children and Families." *Public Administration Review* 79(5):675–83. https://doi.org/10.1111/puar.13047

Fuller, Sirius. 2018. *Using American Community Survey Estimates and Margins of Error*. http://www.census.gov/programs-surveys/acs/guidance/training-presentations/acs-moe.html

Gabry, Jonah, Daniel Simpson, A. Aki Vehtari, Michael Betancourt, and Andrew Gelman. 2019. "Visualization in Bayesian Workflow." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 182(2):389–402. https://doi.org/10.1111/rssa.12378

Gelman, Andrew. 2006. "Prior Distributions for Variance Parameters in Hierarchical Models." *Bayesian Analysis* 1(3):19.

Gelman, Andrew, John Carlin, Hal Stern, and Donald Rubin. 2003. *Bayesian Data Analysis*, 3rd ed. New York: CRC Press.

Gelman, Andrew, and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511790942

Gelman, Andrew, Daniel Simpson, and Michael Betancourt. 2017. "The Prior Can Often Only Be Understood in the Context of the Likelihood." *Entropy* 19(10):555. https://doi.org/10.3390/e19100555

Graif, Corina, Andrew S. Gladfelter, and Stephen A. Matthews. 2014. "Urban Poverty and Neighborhood Effects on Crime: Incorporating Spatial and Network Perspectives." *Sociology Compass* 8(9):1140–55. https://doi.org/10.1111/soc4.12199

Gustafson, Paul. 1998. "Flexible Bayesian Modelling for Survival Data." *Lifetime Data Analysis* 4:19.

Gustafson, Paul. 2004. *Measurement Error and Misclassification in Statistics and Epidemiology: Impacts and Bayesian Adjustments* (Vol. 159). New York: CRC Press.

Hannon, Lance E. 2005. "Extremely Poor Neighborhoods and Homicide." *Social Science Quarterly* 86(s1):1418–34. https://doi.org/10.1111/j.0038-4941.2005.00353.x

Hossain, Shahadut, and Paul Gustafson. 2009. "Bayesian Adjustment for Covariate Measurement Errors: A Flexible Parametric Approach." *Statistics in Medicine* 28:21.

Imbens, Guido W. 2021. "Statistical Significance, P Values, and the Reporting of Uncertainty." *Journal of Economic Perspectives* 35(3):157–74. https://doi.org/10.1257/jep.35.3.157

Lewandowski, Daniel, Dorota Kurowicka, and Harry Joe. 2009. "Generating Random Correlation Matrices Based on Vines and Extended Onion Method." *Journal of Multivariate Analysis* 100(9):1989–2001. https://doi.org/10.1016/j.jmva.2009.04.008

Richardson, Sylvia, and Walter R. Gilks. 1993a. "A Bayesian Approach to Measurement Error Problems in Epidemiology Using Conditional Independence Models." *American Journal of Epidemiology* 138(6):430–42. https://doi.org/10.1093/oxfordjournals.aje.a116875

Richardson, Sylvia, and Walter R. Gilks. 1993b. "Conditional Independence Models for Epidemiological Studies with Covariate Measurement Error." *Statistics in Medicine* 12(18):1703–22. https://doi.org/10.1002/sim.4780121806

Richardson, Sylvia, and Laurent Leblond. 1997. "Some Comment on Misspecification If Priors in Bayesian Modeling of Measurement Error Problems." *Statistics in Medicine* 16:11.

Richardson, Sylvia, Laurent Leblond, Isabelle Jaussent, and Peter J. Green. 2002. "Mixture Models in Measurement Error Problems, with Reference to Epidemiological Studies." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 165(3):549–66. https://doi.org/10.1111/1467-985X.00252

Savitz, Natalya Verbitsky, and Stephen W. Raudenbush. 2009. "Exploiting Spatial Dependence to Improve Measurement of Neighborhood Social Processes." *Sociological Methodology* 39(1):151–83. https://doi.org/10.1111/j.1467-9531.2009.01221.x

Shaw, C. R., and H. D. McKay. 1942. *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.

Simpson, Daniel, Håvar Rue, Andrea Riebler, Thiago G. Martins, and Sigrunn H. Sørbye. 2017. "Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors." *Statistical Science* 32(1):1–28.

Slack, Kristen, Sarah Font, Kathryn Maguire-Jack, and Lawrence Berger. 2017. "Predicting Child Protective Services (CPS) Involvement among Low-Income U.S. Families with Young Children Receiving Nutritional Assistance." *International Journal of Environmental Research and Public Health* 14(10):1197. https://doi.org/10.3390/ijerph14101197

Wheeler, Andrew P., Dae-Y. Kim, and Scott W. Phillips. 2018. "The Effect of Housing Demolitions on Crime in Buffalo, New York." *Journal of Research in Crime and Delinquency* 55(3):390–424. https://doi.org/10.1177/0022427818757283

Wo, James C. 2019. "Mixed Land Use and Neighborhood Crime." *Social Science Research* 78:170–86. https://doi.org/10.1016/j.ssresearch.2018.12.010

AUTHOR BIOGRAPHIES

Duncan Mayer is a PhD candidate in social welfare at the Mandel School of Applied Social Sciences, Case Western Reserve University.

Robert L. Fischer is an associate professor in the Jack, Joseph and Morton Mandel School of Applied Social Sciences, Case Western Reserve University, and codirector of the Center on Urban Poverty and Community Development. His research focuses on nonprofit accountability and the application of evaluation research methodologies in applied and nonprofit settings.

How to cite this article: Mayer, Duncan J., and Robert L. Fischer. 2022. "Can a measurement error perspective improve estimation in neighborhood effects research? A hierarchical Bayesian methodology." *Social Science Quarterly* 103:1260–1272. https://doi.org/10.1111/ssqu.13190