# Analyzing the Relationship between Poverty and Child Maltreatment: Investigating the Relative Performance of Four Levels of Geographic Aggregation

*Sarah B. Aron, Jean McCrowell, Alyson Moon, Ryoichi Yamano,
Duston A. Roark, Monica Simmons, Zurab Tatanashvili, and Brett Drake*

The purpose of this article is to compare four different levels of aggregation to assess their utility as areal units in child maltreatment research. The units examined are county, zip code, tract, and block group levels. Each of the four levels is analyzed to determine which show the strongest effects in modeling the correlation between poverty and child maltreatment report rates. Tract-level aggregation appears to be the most generally robust level, with other levels of aggregation being more vulnerable to different kinds of threats. Some zip codes contain very few people, raising reliability issues, but if weighting or minimum population cutoffs are used, this problem is minimized, and zip codes become an attractive choice. County-level data are less homogeneous than other levels, introducing validity concerns. The smaller populations commonly present in block groups also invoke reliability problems, reducing their utility, especially when rare events are examined.

KEY WORDS: *child maltreatment; geography; neighborhood; poverty*

Neighborhood effects on child maltreatment are an important area of study. Many studies of child maltreatment use geographically aggregated data to represent neighborhood-level constructs in statistical models. Levels of aggregation can range from the state or county level to zip codes, neighborhoods, tracts, block groups, or even smaller geographically defined areas (Sampson, Morenoff, & Gannon-Rowley, 2002). Contextual measures of poverty are also often used as control variables in studies using individual-level data. For the current research, *aggregate data* is defined as means, percentages, or similar values derived from and representing a geographic area.

The modifiable areal unit problem (MAUP) is well known in many fields (King, 1997) and has recently received attention in the child maltreatment literature (Lery, 2008, 2009). The MAUP is an issue encountered when arbitrary geographic areas are established. Should the area be small, events within the area may be rare, and reliability can suffer because of poor signal-to-noise ratios. Should the area be large and not homogeneous with regard to key factors, then any aggregate measure of those factors will not

represent the area well (Nakaya, 2000). An ideal areal unit would be one in which the area is large enough to create stable (reliable) counts of the variables of interest and also sufficiently homogenous on all key variables to minimize error due to conflation of dissimilar subareas. For child welfare research, it would be best to use geographic boundaries that are large enough to provide stable counts of maltreatment but small enough to encompass families that are generally similar to each other on key factors, especially poverty, a key construct relative to child maltreatment (Drake & Zuravin, 1998).

An obvious practical question confronting child welfare researchers and agencies involves which levels of aggregation to use for different purposes. This article presents data to assist academic and agency researchers in answering that question. For each of four geographic levels, data are presented showing observed correlations between poverty and child maltreatment reporting rates. Correlations obtained at each level of analysis are presented side by side so that the relative strengths and significance of the correlations can be observed. To the degree that each level of analysis fosters reliability (stability) and

validity (homogeneity within each unit of analysis), error will be reduced, and the observed correlations will be correspondingly higher (Nakaya, 2000). This is, perhaps, the most straightforward way of demonstrating the relative utility of different levels of aggregation, at least relative to poverty and its relationship to maltreatment. To maximize utility and generalizability, a simple and important construct (poverty), a very basic statistical operation (correlation), and the four most universally available levels of aggregation (county, zip code, tract, and block group) are used.

## BACKGROUND
### Contextual and Compositional Variables
Seldom encountered in the child maltreatment literature, descriptions of variables as contextual or compositional are commonly found in the health and social capital literatures (Subramanian, Lochner, & Kawachi, 2002; Veenstra, 2005). For example, income can be measured at the individual level, generally using survey questions. This is termed a *compositional* variable, one that directly describes an individual-level characteristic. Aggregate measures of income (median, mean, percentage below poverty level in a given area) are *contextual* variables representing the economics or other features of a geographic area in which individuals live. A random sampling of individuals in a given area will provide a set of compositional measures that can be used to derive a contextual measure. Full or near-full sampling of individuals within areas (for example, census data) provide another source for contextual data. Compositional and contextual variables are therefore computationally related but remain theoretically quite distinct. Compositional variables provide measures of individual subjects, whereas contextual variables describe the communities in which subjects live. If a geographic area is relatively homogeneous, there will be correspondingly little variation between contextual and compositional measures. Theoretically, if an area is entirely homogenous, then the contextual and compositional measures will be identical.

### Neighborhood Context as a Subject of Inquiry
There is a long tradition of empirical exploration in this area (Coulton, Korbin, & Su, 1999; Drake & Pandey, 1996; Garbarino & Sherman, 1980). Various neighborhood characteristics, from cohesion to mobility to density to income (Freisthler, 2004, Lery,

2009), have been identified as contributing to child maltreatment. Much of the work in this area is essentially deductive, testing for anticipated relationships between theoretically derived neighborhood characteristics and observed maltreatment rates. The unit of analysis is generally the geographic area, for example, tracts (Coulton, Korbin, Su, & Chow, 1995). These studies often do not use weighted data, as their purpose is to test theory, not to apply findings back to (generalize to) state or national populations. Only contextual variables are used in studies of this type, as analyses are performed using areas as the unit of analysis.

### Neighborhood Context as a Statistical Control
Both individual and community income have been shown to be strongly associated with child maltreatment (Drake & Zuravin, 1998; Sedlak & Broadhurst, 1996). Models and analyses that do not include some measure of income or poverty invite massive threats to their content validity. Content validity is the methodological requirement that a study attend to all variables that may powerfully influence the model being tested (Drake & Jonson-Reid, 2008). For example, if income is not controlled for, African American race will generally appear to be a powerful predictor of maltreatment, when this relationship is, in fact, largely a spurious result of the uneven distribution of income across racial groups (Ards, Myers, Chung, Malkis, & Hagerty, 2003; Drake, Lee, & Jonson-Reid, 2009). To address this problem, multivariate models attempting to explain child maltreatment must include income or poverty variables as controls. These variables can be either compositional (for example, Sedlak & Broadhurst, 1996) or contextual (for example, Drake, Jonson-Reid, Way, & Chung, 2003).

Many child maltreatment data sets do not include individual (compositional) measures of income. Some data sets, particularly administrative data sets, do include addresses that can be geocoded and then fixed within neighborhoods. In conjunction with census data, this allows for a (contextual) measure of the relative wealth or poverty of areas in which individuals live (Drake et al., 2003; Freisthler, 2004). When such steps are taken, the researcher must be clear that neighborhood-level indicators of poverty do not stand as representative of the individual subject's income level, as this would constitute an ecological fallacy. It is valid,

however, to use such measures as an indicator of the neighborhood context in which the child lives. This is often the best available way to attempt to control for the powerful effects of socioeconomic status in models of child maltreatment. The same lack of individual-level income data confronts medical researchers. One study directly addressed this problem by comparing individual measures of income with neighborhood (block group) income data (Krieger, 1992), with findings indicating that "census-level and individual-level socioeconomic measures were similarly associated with the selected health outcomes" (p. 703).

## Levels of Aggregation in Child Welfare Research

Prior work looking at neighborhood context and maltreatment has been done at block group (Friesthler, Needell, & Gruenewald, 2005), tract (Coulton et al., 1995), zip code (Drake & Pandey, 1996), and county (Garbarino, 1976) levels. Beyond child welfare, a recent overview of studies examining neighborhood effects in social research in general (Sampson et al., 2002) included 40 studies, of which 17 were at the tract level, four were at the zip code level, four were at the neighborhood (clusters of tracts) level, three were at the block group level, and 12 were at other levels, such as police beats. This work found that "an important take-away of our assessment is that these and other neighborhood-level mechanisms can be measured reliably with survey, observational, and *archival* approaches [italics added]" (p. 473). Sampson et al. (2002) also cited difficulties in estimating strengths of associations between neighborhood predictor and outcome variables, given the "many differences in research design and measurement across studies" (p. 473). In the present article, these issues are addressed directly, as key aspects of the design (level of analysis, rarity, weighting) are varied and observed correlations compared.

No corresponding work has been done that includes only child maltreatment reports. In the area of foster placement, recent work by Lery (2008, 2009) attempted to compare the utility of zip codes, tracts, and block groups within Alameda County, California. This was done through exploratory spatial data analysis (Lery, 2008) and by using census-derived indicators of neighborhood risk (high mobility, poverty) as predictors of foster care entries at zip code, county, and tract levels (Lery,

2009). Lery found relative robustness at all three spatial levels in her modeling. Other researchers have compared the utility of tract-level aggregation and neighborhood (clusters of tracts) aggregation in predicting delinquency or crime (Ouimet, 2000; Wooldredge, 2002). In these studies, tracts were found to be equal to or superior to neighborhoods in their predictive power.

Another issue in agency research has to do with definitions of neighborhoods. For some purposes it may be best to define neighborhoods at the individual level (that is, the neighborhood as a particular individual sees it). This method may have high face validity when generating neighborhood data relative to individuals, but it cannot be used to create neighborhood boundaries applicable to all individuals within those boundaries. With regard to the MAUP, this study examines areal units that are both publically available from the census and being used by other researchers. For example, in the Sampson et al. (2002) overview cited earlier, 28 of 40 studies used one of these four areal units, or used areal units that were clusters of these units. Finally, most public agency reports use county or zip code aggregation, so understanding the performance of such aggregation units becomes a priority.

## Which Level of Aggregation Is Best?

Both larger and smaller levels of aggregation hold methodological advantages. Larger geographic units are more likely to be available in the public domain. Many state and federal agencies compile and distribute data at the county or zip code level but are unwilling to do so at lower levels for reasons of confidentiality. An obvious related advantage is the reduction of confidentiality concerns on the part of the researcher as the level of aggregation increases. Another advantage of larger levels of aggregation is the reliability of measures. This is especially true for rare events. Very small areas simply cannot be stable with regard to rates of rare events. For example, if an area includes 100 individuals, is observed for a year, and the median incidence of an event is 1 per 1,000 individuals per year, then the rate in that area will be either 0 or in increments of 10 times the expected incidence—an unstable and unreliable measure. This is, of course, extremely important in the study of child maltreatment. To the degree that data are obtained over long periods, in large areas, and without parsing different categories of abuse, this problem is minimized. However, research on

subsets of maltreatment (for example, substantiated sexual abuse rates) over shorter periods of time and in smaller areas will certainly encounter serious problems with rarity.

Smaller areas may also have advantages. They are presumed to be more homogeneous. Homogeneity is a necessary precondition for a valid contextual variable. To the extent that a given area (for example, a state or a county) contains separate and distinct contexts, any aggregate measure will be suspect. Another advantage of smaller areas in many studies is power. If the unit of analysis is the area, and if the geographic sampling frame remains fixed, then smaller areas will increase the sample size, which may increase the study's power to detect statistical significance (Ouimet, 2000).

## RESEARCH QUESTIONS

This article tests four levels of aggregation with specific regard to size, homogeneity, and rarity of events within geographic areas. The work further extends Lery's (2008, 2009) work by using a state-, rather than a county-level, data set, thus allowing consideration of counties as areal units. This article explores the degree to which counties, zip codes, tracts, and block groups produce materially different results with regard to observed correlations and statistical significance (Question 1). In addition, each of the previously noted characteristics (homogeneity, size, and rarity; Questions 2, 3, and 4, respectively) is parsed and analyzed individually, to better determine the contributions of each areal characteristic. The four research questions are as follows:

1. To what degree do different levels of aggregation produce different results?
2. To what degree does homogeneity within areas affect these results?
3. To what degree do the number of subjects within areas affect these results?
4. To what degree does the rarity of the observed event within the area affect these results?

All of these questions are addressed using weighted and unweighted data. Unweighted data are helpful in generalizing to studies in which a geographic area is the unit of analysis. Weighted data provide better data for generalization to studies done at the individual level of analysis, assuming that the individuals in such a study are representatively drawn from the general population. The data generated will be less generalizable to studies with purposive or otherwise restricted sampling frames.

## METHOD

### Data Sources

**Census Data.** The percentage of children living in poverty was determined for the county, zip code, tract, and block group levels using 2000 U.S. Census Summary File 3 data for Missouri (Field P087, items 3 through 6 and 11 through 14). These same data also provided a total count of the number of children in each geographical unit. All census variables used are available separately at the block group, tract, zip code, and county levels, except PCT038, which is not available at the block or block group levels. Data from Census Field PCT038 (counts of families in specific income strata) were also obtained for use in the MOST3 measure described later, both for Missouri and the United States as a whole. Areas with no children recorded in field P087 were excluded from the analysis. In other words, areas were excluded if they included no children for whom poverty status was determined.

**Child Maltreatment Data.** The number of reported children in each area was calculated in steps. All screened-in maltreatment reports which occurred within a three-year window (January 1, 1999, to December 31, 2001) were extracted from the statewide child abuse and neglect database, a computerized system that includes all screened-in reports in the state of Missouri. The household residences of these cases were also brought forward from that database and were geocoded (first automatically, then interactively) using ArcGIS 9, with 86% of cases being successfully assigned spatial locations. ArcGIS is a comprehensive programming system enabling geocoding, analysis, and visual representation of spatial data. Geocoding is a process whereby addresses can be assigned spatial locations (longitudes and latitudes). Such data can then be used in a variety of ways. For current purposes, counts of events were created within four kinds of geographic boundaries (county, zip code, tract, block group). The final data set included 194,842 reports of child maltreatment. These locations were then used to develop counts for reported children at the block group, tract, zip code, and county levels. These counts were divided into total child population counts from the census (described earlier) to determine a yearly rate of maltreatment reports. Separate databases were constructed at each of the four geographic levels of

analysis and included the total number of children in each geographical unit (block group, tract, zip code, or county), the percentage of children below poverty level in each unit, and the yearly mean rate of unduplicated reports per 1,000 children. A similar, separate database was also constructed to simulate increased rarity (see the next section).

## Data Management

*Homogeneity.* A measure of central tendency was needed within each geographical unit. A measure of standard deviation of income would have been ideal, but it is not available in census data. A measure of central tendency was constructed and labeled "MOST3." MOST3 uses data from the PCT38 fields in the 2000 census, which provide categorical counts of families with children according to their 1999 income in dollars. Values were established representing the percentage of all families who fell within the following seven strata: $0 to $9,9999, $10,000 to $19,999, $20,000 to $29,999, $30,000 to $39,999, $40,000 to $49,999, $50,000 to $59,999, and $60,000 and higher. For each geographic unit, the largest percentage of families who fell within any three adjacent strata was determined. For example, if 30% of all families were in the $20,000 to $29,000 strata, 18% were in the next highest strata, and 12% of the population were in the next highest strata after that, and if these were the three adjacent strata with the highest additive total, then the MOST3 measure would be 60% for that area. This measure is intended as a measure of income homogeneity, with higher values of MOST3 representing more homogeneous areas.

*Rarity.* To simulate increased rarity, a data set was created by recording only those events that happened during a 31-day time frame. This 31-day count was established by counting only those maltreatment events falling on each 36th day during 1999, 2000, and 2001. The final 31-day data set included 5,551 reported children for whom geocoded addresses could be determined. This represents an 87% success rate in geocoding.

*Child Population.* The total number of children in each area was retained to allow analyses of geographic areas with at least 200 or 500 children (see the next section) and for weighting purposes.

## Data Analysis

All analyses were performed using PROC CORR in SAS V.9.1. For research question 1, both unweighted

correlations and weighted (by number of children in the area) Pearson correlations were run on the full sample at all four levels (block group, tract, zip code, county). For research question 2, the full sample was split into two halves, representing the most and least homogeneous areas according to the MOST3 measure. For research question 3, separate analyses were run on the full sample and on a subset of all geographic areas (at each of the four levels) including 200 children or more and on another subset of geographic areas including 500 children or more. For research question 4, the full and 31-day data sets were compared.

## RESULTS

### Correlations at Differing Levels of Aggregation

*Full Sample.* Correlations between poverty and maltreatment report rates for each of the four levels of aggregation are presented in Table 1. Among the unweighted correlations, although all levels of aggregation are statistically significant, there are large differences in the correlations obtained. Tracts showed the strongest correlations (.49), and zip codes showed the weakest (.13). When weights were used, the correlations became much stronger at all levels, with zip code improving dramatically (.70) and all levels of aggregation falling within a smaller (.52 to .70) range.

*Most and Least Homogeneous Areas Compared.* Each level of aggregation was split in half using the MOST3 measure. Weighted and unweighted correlation coefficients are reported for each half and for each level of analysis in Table 2. Detailed income stratification data are not available at the block group level, so no data are reported at that level. In the unweighted data, county- and tract-level data show somewhat stronger correlations in the more homogeneous half of all areas, whereas the correlations for unweighted zip codes remain low regardless of homogeneity.

When the weighted data were considered, differences by level of homogeneity became stronger. Counties showed the most dramatic differences. In the least homogeneous half of counties, the correlation between poverty and maltreatment was only .19, but for the more homogeneous half of counties, the correlation jumped to .79 and was significant at the $p \leq .0001$ level. Zip codes and tracts also showed the effects of homogeneity, with correlations being .39 and .36, respectively, in the less homogeneous

## Table 1: Correlations between Poverty and Maltreatment Rate: Full Sample

| Areal Unit | Unweighted | Weighted |
|---|---|---|
| County | | |
| Correlation | .32 | .70 |
| n | 115 | 115 |
| p | .0005 | .0001 |
| Zip code | | |
| Correlation | .13 | .70 |
| n | 1,006 | 1,006 |
| p | .0001 | .0001 |
| Tract | | |
| Correlation | .49 | .66 |
| n | 1,303 | 1,303 |
| p | .0001 | .0001 |
| Block group | | |
| Correlation | .37 | .52 |
| n | 4,490 | 4,490 |
| p | .0001 | .0001 |

Note: Weighting is done as a function of child population within the area.

*city*. This independent city functions essentially as a county and has a separate county-level census identifier (510). Essentially, the splitting of St. Louis County resulted in what is today a relatively richer county and a relatively poorer city. If the county had not been split, the resultant single county would have been larger and much more heterogeneous. It is instructive to run the preceding analyses again, to determine how the findings would be different had this historical idiosyncrasy not occurred. The correlation for unweighted data drops modestly from .32 to .27, whereas the correlation for the weighted data drops radically, from .70 to .46. This difference is due to the large populations within the two counties in question.

*Correlations in High- and Low-Population Areas.* Three separate analyses were done with separate thresholds for inclusion on the basis of the number of children in the geographical area (see Table 3). The full sample analysis is presented first (duplicating Table 1), followed by correlation coefficients for only those areas with at least 200 children. Finally, data are presented for only those areas with at least 500 children. The county data do not change in the data presented, as all Missouri counties contain at least 500 children. In all cases, the coefficients for weighted samples do not change dramatically as child population thresholds

areas, and jumping to .81 and .77, respectively, in more homogeneous areas.

*What if?—The Case of St. Louis City.* On July 4, 1876, the city of St. Louis seceded from St. Louis County, forming what is termed an *independent*

## Table 2: Correlations between Poverty and Maltreatment Rate: Areas Compared by Level of Homogeneity

| Areal Unit | | | | |
|---|---|---|---|---|
| County | | | | |
| Correlation | .28 | .19 | .38 | .79 |
| n | 57 | 57 | 58 | 58 |
| p | .0382 | .1598 | .0033 | .0001 |
| Zip code | | | | |
| Correlation | .17 | .39 | .13 | .81 |
| n | 503 | 503 | 503 | 503 |
| p | .0002 | .0001 | .0031 | .0001 |
| Tract | | | | |
| Correlation | .35 | .36 | .52 | .77 |
| n | 651 | 651 | 652 | 652 |
| p | .0001 | .0001 | .0001 | .0001 |
| Block group | | | | |
| Correlation | — | — | — | — |
| n | — | — | — | — |
| p | — | — | — | — |

Notes: Weighting is done as a function of child population within the area. Dashes indicate that data are not available.

**Table 3: Correlations between Poverty and Maltreatment Rate: Areas Compared by Minimum Number of Children**

| County | | | | | | |
|---|---|---|---|---|---|---|
| Correlation | .32 | .70 | .32 | .70 | .32 | .70 |
| n | 115 | 115 | 115 | 115 | 115 | 115 |
| p | .0005 | .0001 | .0001 | .0001 | .0001 | .0001 |
| Zip code | | | | | | |
| Correlation | .13 | .70 | .46 | .73 | .60 | .77 |
| n | 1,006 | 1,006 | 692 | 692 | 445 | 445 |
| p | .0001 | .0001 | .0001 | .0001 | .0001 | .0001 |
| Tract | | | | | | |
| Correlation | .49 | .66 | .63 | .67 | .66 | .68 |
| n | 1,303 | 1,303 | 1,262 | 1,262 | 1,132 | 1,132 |
| p | .0001 | .0001 | .0001 | .0001 | .0001 | .0001 |
| Block Group | | | | | | |
| Correlation | .37 | .52 | .53 | .57 | .59 | .61 |
| n | 4,490 | 4,490 | 3,069 | 3,069 | 565 | 565 |
| p | .0001 | .0001 | .0001 | .0001 | .0001 | .0001 |

Note: Weighting is done as a function of child population within the area.

are increased, as weighting necessarily causes higher population areas to drive the analyses. The largest change in the weighted data was from .52 to .61 for the block group–level data.

With regard to unweighted data, imposing population thresholds did cause the correlations obtained from zip codes to change radically. Of the 1,006 zip codes, only 692 had at least 200 children, and only 445 had at least 500 children. Restricting the analysis of zip codes from the full sample to a 200-child minimum threshold raised the observed unweighted coefficient from .13 to .46. Moving to the 500-child minimum, the zip code coefficient raises to .60. Changes at the tract and block group levels of analysis were more moderate. Coefficients for tracts ranged from .49 (full sample) to .66 (500-child minimum). Corresponding coefficients for block groups were .37 and .59.

*Rarity.* Rarity was assessed by restricting the number of days on which reports were counted to 31 (see Table 4). Effects at the county level were minimal, as might be expected given their large populations. Unweighted zip code data showed weak correlation coefficients across both conditions, and weighted coefficients dropped modestly, from .70 in the full sample to .58 in the 31-day window. Weighted and unweighted tract data correlations

both dropped moderately, from .49 (unweighted, full sample) and .66 (weighted, full sample) to .27 (unweighted, 31-day sample) and .49 (weighted, 31-day sample). Block group correlations were lower than tract correlations in all cases, with full sample correlations being .37 (unweighted) and .52 (weighted) and 31-day samples being .24 (unweighted) and .32 (weighted).

*Characteristics of Levels of Analysis—Missouri and the United States.* Numbers, means, and standard deviations for child populations and of the MOST3 measure described earlier are presented in Table 5. With regard to child population, counties and zip codes show great variation both in Missouri and in the United States as a whole, with their population standard deviations exceeding their population means by two or three times. Missouri has somewhat smaller mean populations at the county and zip code levels than the United States as a whole. At the tract and block group levels, Missouri and the United States look quite similar with regard to mean and standard deviations of child populations, and the population distributions are more normal, with standard deviations being roughly two-thirds of the population means. Although tracts contain relatively large numbers of children (about 1,000), block groups contain a

| Area Unit | Full (Three Year) Sample Unweighted | Full (Three Year) Sample Weighted | 31-Day Sample Unweighted | 31-Day Sample Weighted |
|---|---|---|---|---|
| County | | | | |
| Correlation | .32 | .70 | .28 | .67 |
| n | 115 | 115 | 115 | 115 |
| p | .0005 | .0001 | .0021 | .0001 |
| Zip code | | | | |
| Correlation | .13 | .70 | .12 | .58 |
| n | 1,006 | 1,006 | 1,006 | 1,006 |
| p | .0001 | .0001 | .0001 | .0001 |
| Tract | | | | |
| Correlation | .49 | .66 | .27 | .49 |
| n | 1,303 | 1,303 | 1,303 | 1,303 |
| p | .0001 | .0001 | .0001 | .0001 |
| Block Group | | | | |
| Correlation | .37 | .52 | .24 | .32 |
| n | 4,490 | 4,490 | 4,490 | 4,490 |
| p | .0001 | .0001 | .0001 | .0001 |

Note: Weighting is done as a function of child population within the area.

mean of 311 children (Missouri) or 341 children (United States). In short, counties and zip codes vary radically in population, with many zip codes having very few child residents (see Table 3). Block groups are noteworthy because they average only about 300 children, far fewer than any other level.

The MOST3 variable in both Missouri and the United States as a whole shows that counties are less homogeneous than either zip codes or tracts. In counties, the MOST3 value is 50.55% for the United States and 50.81% for Missouri. MOST3 values for tracts and zip codes are between 59% and 61% for both Missouri and the United States.

## DISCUSSION

### Research Questions

The first research question dealt with the degree to which different levels of aggregation provided different results in the model used. Substantial differences, highlighting the weakness of unweighted zip code data, are presented in Table 1. Tract data performed well at both weighted and unweighted levels. With regard to the second research question, the effect of homogeneity, it appears that this is a very important determinant of the utility of aggregate data. In the weighted data, correlations between poverty and child maltreatment were at least *twice as high* in the more homogeneous areas as compared with the less homogeneous areas (see Table 2). The third research question addressed increased reliability, resulting from higher population levels. Although county-level data are not affected by this concern, other levels of aggregation showed increased correlations at higher population levels, particularly in the unweighted data. The final question dealt

### Table 5: Population Means, Standard Deviations, and Mean MOST3 Measure: Missouri and the Full United States

| Area and Areal Unit | N | Child Population M | Child Population SD | MOST3 Measure M (%) | MOST3 Measure SD |
|---|---|---|---|---|---|
| Missouri | | | | | |
| County | 115 | 12,181 | 30,057 | 50.81 | 5.38 |
| Zip code | 1,006 | 1,392 | 2,394 | 59.75 | 13.12 |
| Tract | 1,303 | 1,075 | 591 | 59.05 | 12.69 |
| Block group | 4,490 | 311 | 229 | — | — |
| United States | | | | | |
| County | 3,218 | 22,376 | 75,791 | 50.55 | 8.35 |
| Zip code | 31,757 | 2,267 | 3,518 | 59.27 | 13.99 |
| Tract | 65,506 | 1099 | 652 | 60.97 | 13.93 |
| Block group | 211,267 | 341 | 273 | — | — |

Note: Dashes indicate that data are not available.

with the degree that rarity of events might affect observed correlations. Increased rarity degraded correlation levels in all cases but was most notable at the block group level, as the low populations in these areal units made them more vulnerable to this threat to reliability.

### County-level Aggregation in Detail

The lack of homogeneity in county data (see Table 5) makes counties an unattractive choice for researchers. Furthermore, because some counties include very large populations, single counties or locations of individual county boundaries can radically affect weighted models. In the Missouri data, had St. Louis not seceded from the county, the data would have looked very different, and the utility of county-level data would have appeared even more suspect. States with large, heterogeneous counties should not be analyzed at the county level. However, counties are largely invulnerable to reliability threats associated with small populations. Said simply, county data are highly reliable but of very suspect validity.

### Zip Code-level Aggregation in Detail

Weighted zip code data performed as well as or better than weighted tract-level data in all cases. Many researchers will undoubtedly be surprised by this finding. Unweighted zip code data performed poorly in all cases, except when minimum population cutoffs (see Table 3) were used (which might itself be seen as a very crude form of weighting). It is interesting that, as measured by the MOST3 variable, zip codes are as homogeneous as tracts. On the basis of these data, the long-standing and widespread belief that zip codes are too heterogeneous to provide acceptable areal delineation appears to be false. The threat to the utility of zip codes lies in the large number of zip codes with small populations and the attendant threats to reliability, not on heterogeneity. On the basis of these data, we see no reason to avoid use of zip code data if the zip codes used are adequately large, if minimum population cutoffs are used, or if the data are weighted.

### Tract-level Aggregation in Detail

Tracts appear large enough to be reliable and small enough (or sufficiently well conceptualized) to be acceptably homogeneous. This makes them an attractive choice to researchers. Tract data perform adequately in all cases, and they appear to be the safest choice overall. This is especially true for re-searchers who may not spend a lot of time coming to know the intricacies and foibles of the level of analysis they select. In other words, tracts are a good "default" choice.

### Block Group-level Aggregation in Detail

A major weakness of this study is the inability to derive a homogeneity measure at this level. Despite this limitation, the observed correlations using block group aggregation were always smaller than those observed at the tract level. In addition, as the rarity of observed events increases (see Table 4), the relative weakness of block group data becomes more evident. The problem with block group data appears to be one of reliability, at least with regard to rare events. Clearly, the issues of rarity and areal size are inextricably linked mathematically, so the ability to use small areal units is largely dependent on the rarity of the event studied. Some work-arounds may exist, such as increasing the sampling frame over time to allow for more occurrences of rare events or allowing other kinds of events to proxy for the rare event. However, the question arises, are there *empirical* grounds to favor these small areas? This study finds none.

### Limitations

These data were based on one state, Missouri. Counties and zip codes in Missouri are substantially smaller than national averages, but the proportion of standard deviations to means is similar to nationwide figures. Tract and block group data from Missouri are close to national data with regard to population means and standard deviations. Homogeneity of counties, zip codes, and tracts are similar between Missouri and the United States. These data are therefore most generalizable at the tract and block group levels. County and zip code level results should be applied to other states with some caution.

Statistical analyses were intentionally restricted to simple correlations, and this is a study strength, not a limitation. Correlations are well understood by most researchers, and the use of simple statistics provide an elegant set of tests with maximum generalizability to other models. Other statistical tools are applied to geographic modeling (for example, controls for spatial autocorrelation; see Lery, 2009), but it is important to first approach core issues—such as homogeneity, population, and sample size—using simple models. Even simple correlations are more sophisticated than the simple cross-tabulations most

commonly found in nonacademic products such as state-level reports, which are commonly reported at zip code or county levels.

## CONCLUSION

Current findings are consistent with findings by Lery (2008, 2009), both in that neighborhood risk strongly predicts child welfare involvement, and that different areal units can be useful. This work extends Lery's work by moving beyond a single urban context, including the county level, and teasing out different threats (rarity, low population, homogeneity) specifically.

Results support the use of tract-level data as the safest general choice. However, the most unexpected and, arguably, important finding is that in many instances, zip codes work very well. For example, the zip codes often outperformed tracts when weighting or minimum population cutoffs were used. It would appear that zip code–level aggregation may have gotten a "bad rap," at least if strong protections (cutoffs, weighting) can be used to protect against unreliable low-population zip codes introducing error. Recognizing the utility of zip code data may have very positive practical implications. Zip code data are vastly more available than tract data, and increased use of zip code data may well enable more, and more timely, research on key issues. As more public databases go online, including many at the zip code level (for example, crime, child maltreatment, foreclosure rates), more contextual variables can be easily acquired by researchers at the zip code level, allowing for richer models. The timeliness issue arises from a key difference between tracts and zip codes: Zip codes generally exist in address data without the need for geocoding. This allows for "off-the-shelf" use of such data. No geocoding also means no loss of observations at this step and no geocoding errors. As we move further toward a future in which near-real-time analyses are expected (Drake & Jonson-Reid, 1999) and even automated, this becomes a useful feature, especially for nonacademic researchers such as state agencies or the federal government. As such reports become more sophisticated, increased use of contextual controls or grouping might be contemplated.

Although this study was unable to assess block group homogeneity, block group aggregation *might* be a good choice when the events being studied are very common, leading to increased reliability. County-level aggregation is almost always a subop-

timal approach, and use of county, regional, or state data should occur only following a very thorough analysis of the homogeneity of key constructs within those areas. The current study remains limited by consideration of only the constructs of child welfare and poverty, but there appear to be no necessary reasons why the key principles explored (rarity, homogeneity, size) would manifest differently in another substantive area. **SWR**

## REFERENCES

Ards, S. D., Myers, S. L., Chung, C., Malkis, A., & Hagerty, B. (2003). Decomposing black–white differences in child maltreatment. *Child Maltreatment, 8,* 112–121.

Coulton, C., Korbin, J., & Su, M. (1999). Neighborhoods and child maltreatment: A multi-level study. *Child Abuse & Neglect, 23,* 1019–1040.

Coulton, C., Korbin, J., Su, M., & Chow, J. (1995). Community level factors and child maltreatment rates. *Child Development, 66,* 1262–1276.

Drake, B., & Jonson-Reid, M. (1999). Some thoughts on the increasing use of administrative data in child maltreatment research. *Child Maltreatment, 4,* 308–315.

Drake, B., & Jonson-Reid, M. (2008). *Social work research methods: From conceptualization to dissemination.* Boston: Pearson.

Drake, B., Jonson-Reid, M., Way, I., & Chung, S. (2003). Substantiation and recidivism. *Child Maltreatment, 8,* 248–260.

Drake, B., Lee, S., & Jonson-Reid, M. (2009). Race and child maltreatment reporting: Are blacks overrepresented? *Children and Youth Services Review, 31,* 309–316.

Drake, B., & Pandey, S. (1996). Understanding the relationship between neighborhood poverty and child maltreatment. *Child Abuse & Neglect, 20,* 1003–1018.

Drake, B., & Zuravin, S. (1998). Revisiting the myth of classlessness. *American Journal of Orthopsychiatry, 68,* 295–304.

Freisthler, B. (2004). A spatial analysis of social disorganization, alcohol access, and rates of child maltreatment in neighborhoods. *Children and Youth Services Review, 25,* 803–819.

Freisthler, B., Needell, B., & Gruenewald, P. (2005). Is the physical availability of alcohol and illicit drugs related to neighborhood rates of child maltreatment? *Child Abuse & Neglect, 29,* 1049–1060.

Garbarino, J. (1976). A preliminary study of some ecological correlates of child abuse: The impact of socioeconomic stress on mothers. *Child Development, 47,* 178–185.

Garbarino, J., & Sherman, D. (1980). High-risk neighborhoods and high-risk families: The human ecology of child maltreatment. *Child Development, 51,* 188–198.

King, G. (1997). *A solution to the ecological inference problem: Reconstructing individual behavior from aggregate data.* Princeton, NJ: Princeton University Press.

Krieger, N. (1992). Overcoming the absence of socioeconomic data in medical records: Validation and application of a census-based methodology. *American Journal of Public Health, 82,* 703–710.

Lery, B. (2008). A comparison of foster care entry risk at three spatial scales. *Substance Use and Misuse, 43,* 223–237.

Lery, B. (2009). Neighborhood structure and foster care entry risk: The role of spatial scale in defining

neighborhoods. *Children and Youth Services Review, 31,*
331–337.

Nakaya, T. (2000). An information statistical approach to
the modifiable areal unit problem in incidence rate
maps. *Environment & Planning, 32*(1), 91–109.

Ouimet, M. (2000). Aggregation bias in ecological research:
How social disorganization and criminal opportuni-
ties shape the spatial distribution of juvenile delin-
quency in Montreal. *Canadian Journal of Criminology
and Criminal Justice, 42,* 135–156.

Sampson, R., Morenoff, J., & Gannon-Rowley, T. (2002).
Assessing neighborhood effects: Social processes and
new directions in research. *Annual Review of Sociology,
28,* 443–478.

Sedlak, A., & Broadhurst, D. (1996). *The third national
incidence study of child abuse and neglect: NIS 3.*
Washington, DC: U.S. Department of Health and
Human Services, Government Printing Office.

Subramanian, S., Lochner, K., & Kawachi, I. (2002). Neigh-
borhood differences in social capital: A compositional
artifact or contextual construct? *Health and Place, 9*(1),
33–34.

Veenstra, G. (2005). Location, location, location: Contextu-
al and compositional health effects of social capital in
British Columbia, Canada. *Social Science and Medicine,
60,* 2059–2071.

Wooldredge, J. (2002). Examining the (ir)relevance of ag-
gregation bias for multilevel studies of neighborhoods
and crime with an example comparing census tracts
to official neighborhoods in Cincinnati. *Criminology,
40,* 681–709.

---

**Sarah B. Aron, MSW,** is *teen services coordinator, Children's
Aid Society Family Wellness Program, New York. Jean Mc-
Crowell, MA, MSG, lives in Lexington, VA.* **Alyson Moon,
MSW,** *lives in Seattle, WA.* **Ryoichi Yamano, MSW,** *is a
social worker, Kanagawa Child Guidance Center, Yokohama,
Japan.* **Duston A. Roark, MSW,** *is a medical case manager,
Evergreen AIDS Foundation, Everett, WA.* **Monica Simmons,
MSW,** *lives at Minot Air Force Base, ND.* **Zurab Tatanashvili,
MD, MSW,** *is an expert in social work and organizational
development, Georgian Association of Social Workers, Tbilisi,
GA.* **Brett Drake, PhD, MSW, LCSW,** *is professor, George
Warren Brown School of Social Work, Washington University
in St. Louis, One Brookings Drive, St. Louis, MO, 63130;
e-mail: brettd@wustl.edu.*