
REASSESSING RACIAL AND SOCIOECONOMIC DISPARITIES IN ENVIRONMENTAL JUSTICE RESEARCH*

PAUL MOHAI AND ROBIN SAHA

The number of studies examining racial and socioeconomic disparities in the geographic distribution of environmental hazards and locally unwanted land uses has grown considerably over the past decade. Most studies have found statistically significant racial and socioeconomic disparities associated with hazardous sites. However, there is considerable variation in the magnitude of racial and socioeconomic disparities found; indeed, some studies have found none. Uncertainties also exist about the underlying causes of the disparities. Many of these uncertainties can be attributed to the failure of the most widely used method for assessing environmental disparities to adequately account for proximity between the hazard under investigation and nearby residential populations. In this article, we identify the reasons for and consequences of this failure and demonstrate ways of overcoming these shortcomings by using alternate, distance-based methods. Through the application of such methods, we show how assessments about the magnitude and causes of racial and socioeconomic disparities in the distribution of hazardous sites are changed. In addition to research on environmental inequality, we discuss how distance-based methods can be usefully applied to other areas of demographic research that explore the effects of neighborhood context on a range of social outcomes.

Since the mid-1980s, scholarly attention to racial and socioeconomic disparities in the distribution of pollution, environmental hazards, and locally unwanted land uses (LULUs) has been increasing. Many quantitative studies examining environmental inequality have been conducted over the past decade. Although most reviews have found that these inequalities tend to be statistically significant (Lester, Allen, and Hill 2001; Mohai and Bryant 1992; Ringquist 2005), there has been considerable variation in the magnitude of the disparities found. Some studies have found no racial or socioeconomic disparities associated with the distribution of environmentally hazardous sites (Anderton et al. 1994; Bowen et al. 1995; Davidson and Anderton 2000). Uncertainties also exist about the causes of racial and socioeconomic disparities in the distribution of environmental hazards. Indeed, the most fundamental question—which came first, the people or the pollution?—has yet to be satisfactorily answered. That is, are present-day disparities the result of a historical pattern of disproportionately siting polluting facilities in minority and poor communities, or are they the result of demographic changes in communities after siting? The few studies that have been conducted have led to contradictory findings (see, e.g., Been and Gupta 1997; Oakes, Anderton, and Anderson 1996; Pastor, Sadd, and Hipp 2001; Saha and Mohai 2005).

*Paul Mohai, School of Natural Resources and Environment, and Institute for Social Research, University of Michigan, Ann Arbor. Robin Saha, Environmental Studies Program and Programs in Public Health, University of Montana. Address correspondence to Paul Mohai, School of Natural Resources and Environment, University of Michigan, Ann Arbor, MI 48109-1115; E-mail: pmohai@umich.edu. This research was supported by grants from the Sociology Program and Geography and Regional Science Program of the National Science Foundation (#0099123) and the University of Michigan State Policy Research Fund. We thank the following people for their helpful feedback, comments, and suggestions for this article: William Freudenburg, Phil Brown, Steven Brechin, Timmons Roberts, Dorceta Taylor, Jeffrey Morenoff, Sara Cohen, Shannon Brines, and the anonymous reviewers of *Demography*. We are indebted to Professor Vicki Been, of the New York University School of Law, and her coauthor, Francis Gupta, for providing us with the names and addresses of the hazardous waste treatment, storage, and disposal facilities used in their national environmental justice studies.

A principal argument of this article is that much of the source of these uncertainties is related to the failure of the most widely used methodology in environmental inequality research to adequately account for the proximity between environmentally hazardous sites and nearby residential populations. The goal of this article is therefore to identify the reasons for and consequences of the failure of this methodology and to demonstrate ways of overcoming these shortcomings by using alternate, *distance-based* methods. By applying distance-based methods in the reanalysis of a leading national study of the demographic disparities around the nation's hazardous waste treatment, storage, and disposal facilities (TSDFs), we demonstrate how and why findings and conclusions about the magnitude and causes of racial and socioeconomic disparities around such sites are changed. We argue that the application of distance-based methods will help resolve existing uncertainties and improve our understanding of the extent and causes of inequalities in neighborhood environmental conditions. Moreover, such methods, we believe, have applicability to wider areas of demographic research that attempt to understand the effects of neighborhood context on a range of social outcomes (Sampson, Morenoff, and Gannon-Bowley 2002).

THE CLASSIC APPROACH: ANALYZING UNIT-HAZARD COINCIDENCE

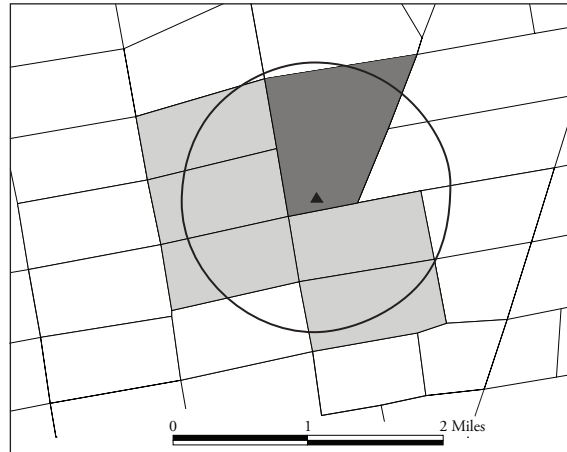
As we suggested earlier, the most widely used approach for assessing demographic disparities in the distribution of environmental hazards and LULUs is also the weakest in its ability to control for the proximity between such hazards and nearby populations. This approach has nevertheless been the typical or "classic" approach to environmental inequality research, including some of the most influential studies (e.g., Anderton et al. 1994; Been 1995; Commission for Racial Justice, CRJ, 1987; Goldman and Fitton 1994). The approach is straightforward. It involves selecting a predefined geographic unit (such as counties, zip code areas, or census tracts), identifying which of the units contain or "host" the hazard, deciding upon an appropriate set of comparison units (essentially, those that do not contain the hazard), and then comparing the demographic characteristics between the two sets. Not taken into account by this method is the exact location of the hazard within the host unit, nor the proximity of the hazard to nearby units. Since this method goes no further than noting whether the general locations of the hazard and host unit coincide, some have referred to this method as "spatial coincidence" (McMaster, Leitner, and Sheppard 1997), although "unit-hazard coincidence" may more precisely describe the approach. Nearly all national-level studies on environmental inequality have used this approach (see Anderton et al. 1994; Anderton, Oakes, and Egan 1997; Been 1995; CRJ 1987; Daniels and Friedman 1999; Davidson and Anderton 2000; Goldman and Fitton 1994; Greenberg 1993; Hamilton 1993, 1995; Hird 1993; Hird and Reese 1998; Lester et al. 2001; Oakes et al. 1996; Perlin et al. 1995; Ringquist 1997; Zimmerman 1993).

An implicit assumption in this approach is that people living in the host units are closer to the hazard under investigation than people living in the nonhost units. That this is not always the case, indeed frequently is not the case, becomes apparent when the exact locations of a set of environmental hazards or LULUs are mapped (rarely done in most environmental inequality research) and their proximity to host and surrounding unit boundaries is examined. Figure 1 provides such an illustration. Here, the precise locations of two of the nation's hazardous waste TSDFs are displayed. Although many other TSDFs similarly illustrate the problems of the unit-hazard coincidence method, those presented in Figure 1 provide especially clear examples. A more comprehensive analysis demonstrating the limitations of this method is provided later in the "Results" section.

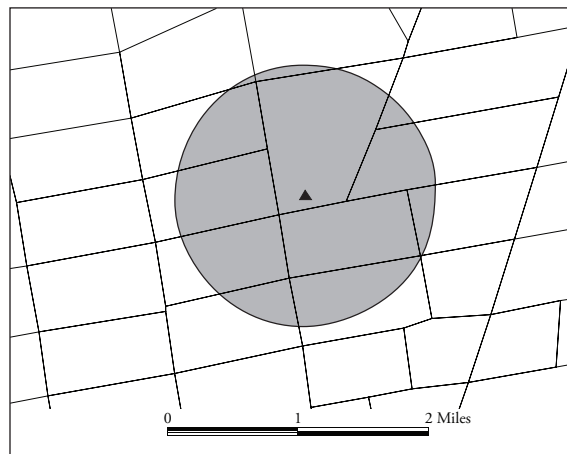
One observation that is apparent from the figure is that, rather than necessarily being located near the host tract's center, the TSDF may be located near a boundary. Indeed, using a national database of 608 TSDFs operating in the early 1990s (see the description below) and 1990 census tracts, we found that 298, or 49%, of them are within 0.25 miles of the boundary of their host tracts, while 433, or 71%, are within 0.5 miles. When the TSDF is

Figure 1. Comparing Neighborhoods Around TSDFs Identified by Unit-Hazard Coincidence and Distance-Based Methods

a. 50% areal containment using a one-mile radius

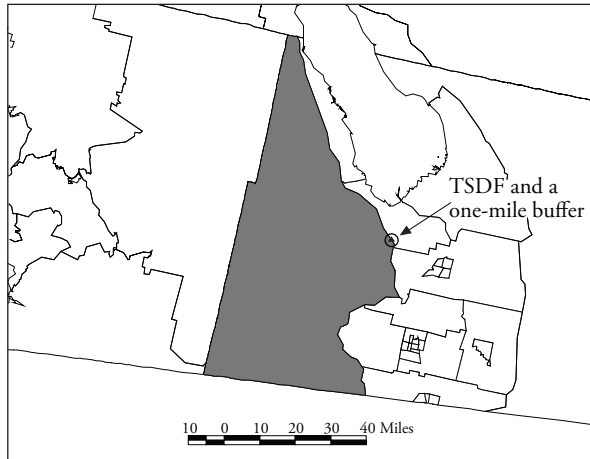


b. Areal apportionment using a one-mile radius



(continued)

near a boundary, much of the adjacent or nearby tracts may be as close to the TSDF as the host tract proper. For example, in Figure 1a, most of the areas of the tracts immediately south and west of the TSDF (shaded light gray) appear to be as near to the TSDF as most of the area of the host tract (shaded dark gray). A one-mile radius captures about as much of the areas of the adjacent and nearby tracts as it does of the host tract. (Consequences of selecting alternate radii are discussed in the “Methods” section.) In spite of their proximity to the TSDF, the unit-hazard coincidence method treats such nearby tracts no differently than nonhost tracts much farther away and places them in the comparison group. However, if

*(Figure 1, continued)***c. Large host tract**

there is a relationship between the location of a TSDF and the demographic characteristics of the neighborhoods surrounding it, then the demographic characteristics of the nearby tracts may be more similar to the host tract proper than to tracts much farther away. Placing such nearby tracts in the comparison group may thus obscure this relationship.

A second observation from Figure 1 is that the sizes of host tracts may vary dramatically. For example, the host tract in Figure 1a is only 0.85 square miles, while the host in Figure 1c is 916.5 square miles. When the host tract is small, such as the tract in Figure 1a, it can be reasonably assumed that almost everyone in the tract lives close to the TSDF. However, when the host tract is large, such as in Figure 1c, little of the tract's population may be close by. Indeed, the TSDF in Figure 1c is 14.9 miles from the centroid of its host. A circle with a one-mile radius captures less than 0.2% of the tract's area. Given that so much of the tract lies far from the TSDF, it is not likely that much of the tract's population lives near it, at least not within the one-mile distance. Because such large host tracts (i.e., those whose areas lie mostly beyond the specified distance) fail to control for proximity between the TSDF and nearby populations, it is uncertain whether such tracts are able to detect a relationship between the presence of the TSDF and the demographic characteristics of the nearby populations. Indeed, the demographic characteristics of such large host tracts may not be much different from the characteristics of nonhost tracts whose areas similarly lie beyond such a distance (a possibility we examine later). If that is the case, then averaging or aggregating population characteristics of such large host units with small host units (where the proximity between TSDF and nearby populations is better assured) may obscure any relationship that might exist between the TSDF and nearby population characteristics.

DISTANCE-BASED METHODS

Standing in contrast to the unit-hazard coincidence studies are a limited number of studies that have employed distance-based methods. These methods overcome a number of limitations of the classic approach. With these methods, the precise locations of

environmental hazards or LULUs are mapped, and their distances to nearby residential populations are specified. The demographics of all units within the specified distances, not just in the host unit proper, are contrasted with the demographics of units farther away. There have been three types of distance-based methods using census data: (1) 50% areal containment, (2) boundary intersection, and (3) areal apportionment. Next we describe each method and its relative advantages over the unit-hazard coincidence approach.

50% Areal Containment and Centroid Containment

The 50% areal containment method involves mapping the location of environmental hazards or LULUs and then averaging or aggregating the demographic characteristics of predefined geographic units (such as block groups, census tracts, and zip code areas) captured by circles of a specified distance from the hazards. However, because the units may take up considerable space, often the circle intersects only a portion of the unit, rather than completely encompassing or completely missing it. One rule that researchers have used to decide whether to count a unit as within the specified distance, and hence within the host neighborhood, is to include all units for which at least 50% of the unit's area is captured and exclude all units, including the host, if the captured area is less than 50%. Alternatively, units have been considered within the host neighborhood if the circle captures the geographic center of the unit (the *centroid-containment method*). The resulting area formed around the hazard approximates a circle with "rough edges" (Figure 1a). The demographic characteristics of the units captured by the 50% areal containment (or centroid containment) method are then averaged or aggregated (i.e., the demographics of the units are weighted by the units' population sizes) and compared against the demographics of the averaged or aggregated units not captured. Averaged demographic characteristics (C_{averaged}) of the captured units are computed by

$$C_{\text{averaged}} = \frac{\sum_{i=1}^n c_i}{n},$$

where c_i is the demographic characteristic of unit i , and n is the number of captured units. Aggregated (population-weighted) characteristics ($C_{\text{aggregated}}$) of the captured units are computed by

$$C_{\text{aggregated}} = \frac{\sum_{i=1}^n (p_i)(c_i)}{\sum_{i=1}^n p_i},$$

where p_i is the population in unit i .

Examples of the centroid containment method are provided by Chakraborty and Armstrong (1997). Variations of the centroid containment and 50% areal containment methods are performed by Anderton et al. (1994) and Davidson and Anderton (2000). These studies represent variations of these methods because they do not take into account the actual locations of the hazards under investigation. Instead of centering circles at hazard locations, these studies center their circles at the host tract centroids. Radii of 0.5, 1.0, 2.5, and 3.0 miles have been used in these studies.

Boundary Intersection Method

The boundary intersection method is similar to the 50% areal containment method but without the restriction on including units with captured areas of less than 50%. All units whose boundaries are wholly contained by, partially intersected by, or tangent to a circle of a specified distance centered at the environmental hazard are considered in the host neighborhood. Examples of the application of this method are provided by Boer et al. (1997),

Chakraborty and Armstrong (1997), and Pollock and Vittas (1995). Although the boundary intersection method provides some measure of control for proximity by only including units that have some portion within a certain distance, it captures units that may also have substantial areas that lie well beyond the distance. Because it shares a problem similar to that of the unit-hazard coincidence approach, it is the least effective of the distance-based approaches at controlling for proximity.

Areal Apportionment Method

The areal apportionment method is similar to the version of the boundary intersection method in which the characteristics of all units that are wholly contained by or intersected by a circle of a given radius are aggregated (i.e., weighted by population). However, unlike the boundary intersection method, it does not necessarily give each unit's population full weight in the calculations. Instead, each unit's population is weighted by the proportion of the area of the unit that is captured by the circle. The weighted populations of these units are then used to determine the aggregate demographic characteristics of perfectly circular neighborhoods within a specified distance of the hazard (Figure 1b). The formula for computing the demographic characteristics (C) within the neighborhoods of a given radius is as follows:

$$C = \frac{\sum_{i=1}^n (a_i / A_i)(p_i)(c_i)}{\sum_{i=1}^n (a_i / A_i)p_i},$$

where a_i is the area of unit i captured by a given radius, and A_i is the total area of unit i , and n is the number of units that are wholly or partially contained by a circle of a given radius.

An important assumption of this method is that the proportion of the unit's *area* that is captured by the circle approximates the proportion of the unit's *population* that is captured. This assumes that the population and its characteristics are distributed uniformly within the unit. Of course, this may not necessarily be the case. However, the assumption of uniformity within census units is not unique to this method. It is also implicit in the unit-hazard coincidence method and other distance-based methods and implicit in census data research generally. Furthermore, an important advantage of the areal apportionment method over the other distance-based methods is its avoidance of assigning extremes (i.e., 0% vs. 100%) in weighting partially contained units. Partially intersected units are assigned weights proportional to their intersected areas, reducing the risk that any unit over- or underinfluences the estimated demographic characteristics within a given distance of an environmental hazard. Studies that have employed the areal apportionment method include those by Chakraborty and Armstrong (1997), Glickman (1994), Glickman, Golding, and Hersh (1995), Hamilton and Viscusi (1999), and Sheppard et al. (1999). Radii of one-half, one, two, and three miles have been used with this method.

DISTANCE-BASED VERSUS UNIT-HAZARD COINCIDENCE METHODS

In sum, we have illustrated by examples why the unit-hazard coincidence approach, although it is the principal approach used in the influential national-level environmental inequality studies, is the least likely to adequately control for the proximity between potential environmental hazards and nearby populations. This method fails to control for proximity in two ways. First, it does not take into account a possible association between the potential hazard and the demographic characteristics of nearby nonhost units. It assumes that nonhost units near the hazard are no different demographically from nonhost units much farther away. Second, it assumes that populations in large host units are as near to the potential hazard as populations in small host units, when in reality, populations in the former may be dispersed quite far away. Distance-based methods overcome these limitations by

specifying the locations of environmental hazards and sorting units based on their actual proximity to these sites.

Using a national-level database of the country's hazardous waste TSDFs, we demonstrate, through more comprehensive and systematic means, how the unit-hazard coincidence method fails to control for the proximity between TSDFs and nearby populations and why distance-based methods have a greater ability to do so. In so doing, we provide results of the first national-level analysis of the demographics around hazardous waste TSDFs obtained from such methods. We use these results to demonstrate how using distance-based methods alters assessments of racial and socioeconomic disparities around environmental hazards and LULUs.

DATA AND METHODS

There is no single, definitive source of information about the nation's hazardous waste TSDFs. As a result, previous studies employing the unit-hazard method have relied on different sources and hence have identified different universes of TSDFs (see, e.g., Been 1995 and Been and Gupta 1997, compared with Anderton et al. 1994 and Oakes et al. 1996). To isolate the effects of employing different methodologies, we sought to utilize the same universe of facilities as previous studies. Due to the confidentiality that was promised to companies surveyed by researchers at the Social and Demographic Research Institute of the University of Massachusetts, we were unable to obtain adequate identifying information for facilities used in Anderton et al. (1994) and Oakes et al. (1996). However, we obtained U.S. EPA identifiers and address information for the 608 facilities used for the other leading studies of hazardous waste TSDFs (Been 1995; Been and Gupta 1997). This information, provided by Professor Vicki Been at the New York University School of Law, allowed us to identify precise geographic locations of TSDFs within the host tracts.

We identified these locations by geocoding addresses and verifying location information primarily through phone interviews of company employees. For 538 TSDFs, our address and location information was either verified or corrected by the facility personnel. In some cases, we consulted site maps obtained from the companies. For 61 TSDFs that we were not able to contact, state environmental agencies or the U.S. EPA were contacted for this information. For the remaining 9 TSDFs for which insufficient information was available from the above sources, other sources, such as former employees and online commercial mapping services, were consulted.

After TSDF locations were established, we generated one-, two-, and three-mile circular buffers around the TSDFs by using ArcView GIS™ (Version 3.2) and a Lambert's Conformal Conic Projection (for the conterminous United States). We selected these radii because they are within the range used in prior studies and because, as other studies have done, we wanted to examine how demographic characteristics around environmentally hazardous sites vary with varying distances to the site. The radius of influence of larger, more toxic sites is likely to be greater than that of others. Thus, rather than suggesting that there is an ideal radius for all environmental inequality studies, we believe future research will need to explore further how demographic characteristics change with varying distances to hazardous sites of a wide variety, an aim that is currently not possible with the unit-hazard coincidence approach. For hazardous waste TSDFs, however, we found (as we discuss later) that minority and poverty percentages generally decreased with increasing distance from the sites.

To analyze the demographic characteristics within the buffers applying the 50% areal containment and areal apportionment methods, we used 1990 digitized census tracts (Geolytics, Inc. 1999) and 1990 Summary Tape File (STF) 1A and 3A census data (Wessex, Inc. 1994)¹. We examined variables that were used in many prior studies to assess demographic

1. We used 1990, rather than 2000, census data in order to more directly contrast our results that use distance-based methods with the results of prior studies that used the unit-hazard coincidence method.

disparities in the distribution of environmental hazards and LULUs. These included percentage African American, percentage Hispanic, percentage nonwhite, mean household income, mean housing values, percentage living below poverty, percentage without a high school diploma, percentage with a college degree, percentage employed in executive management or professional occupations, and percentage employed in precision production or labor occupations (see the Appendix for details about the construction of these variables).

RESULTS

The first step in our analysis was to determine how well, in comparison with the unit-hazard coincidence method, distance-based methods control for proximity between the TSDFs and nearby residential populations. To accomplish this, we determined the exact areas of the 554 tracts hosting one or more TSDFs. We also determined the precise distances of each of the 608 TSDFs to their respective host tract centroids. The host tract area and the distance of the tract centroid to a TSDF serve as indicators of the degree to which the population in a tract is likely to be close to the TSDF or to be geographically dispersed from it.

Because some host tracts contain more than one TSDF, and hence there is not a one-to-one correspondence between TSDFs and host tracts, we computed the mean, median, and standard deviation of the host tract areas and centroid distances to the TSDFs in two ways. In the first way, values were based on those associated with *each of the 608 TSDFs* (i.e., $N = 608$ in all the calculations). Thus, the distance between each of the 608 TSDFs and its respective host tract centroid was included in calculating the mean, median, and standard deviation of the distances. In the calculation of the corresponding values for host tract areas, each tract was represented as many times in the calculations as it contained TSDFs. In the second way, values were based on those associated with *each of the 554 host tracts*, with any double counting removed (i.e., $N = 554$ in all calculations). Thus, in calculating the mean, median, and standard deviation of host tract areas, we counted each tract area only once, regardless of the number of TSDFs within the tract. In calculating the corresponding values for distances between host tract centroids and TSDFs, we used the distances between the host tract centroid and the nearest TSDF within the tract. Because the results were similar but slightly more conservative in the second set of calculations in which values were based on host tracts, we discuss only the second set of values (both sets of values are provided in Table 1, however; see columns 9 and 10).

Regarding the host tracts proper and confirming what we suggested earlier, we found considerable variation in the size of host tracts and in the location of the TSDFs with respect to the host tract centroids and boundaries. For example, although the smallest tract hosting a TSDF is only 0.07 square miles, the largest is 7,521 square miles (see Table 1). Similarly, while in one case a TSDF is only 0.03 miles from the centroid of its host, in another it is nearly 34 miles from the centroid. The mean and median areas of the host tracts are 58.41 and 4.71 square miles, while the mean and median distances of the TSDFs to their host tract centroids are 1.86 and 0.90 miles.

In contrast, the size of host neighborhoods defined by one-, two-, and three-mile buffers around the TSDFs are generally much smaller than they are for host tracts proper. (Because the results from using a two-mile radius lay in between, only results for one- and three-mile buffers are reported in Table 1 and discussed.) For example, if we apply 50% areal containment, the host neighborhoods formed by aggregating the captured tracts have mean areas of 2.39 and 21.77 square miles, using one- and three-mile radii, respectively (the median areas are similar to the means; see Table 1, columns 6 and 8). Furthermore, neighborhood centroids are closer to the TSDFs within them, with mean distances of only 0.40 miles and 0.71 miles (again, the median distances are similar), respectively, compared with 1.86 miles for host tracts. There is also greater consistency in the size of the host neighborhoods and the location of the TSDFs within them, as evidenced by the standard deviations of the relevant values. At the one- and three-mile radii, the standard deviations of

Table 1. Spatial Comparisons of Neighborhoods Surrounding 608 TSDFs Defined by Areal Apportionment and 50% Areal Containment Methods Versus Raw Host Tracts

Measure	Neighborhoods Defined by a 1-Mile Radius Using Areal Apportionment <i>N</i> = 608		Neighborhoods Defined by a 1-Mile Radius Using 50% Areal Containment <i>N</i> = 608		Neighborhoods Defined by a 3-Mile Radius Using 50% Areal Containment <i>N</i> = 608		Raw Host Tracts <i>N</i> = 608 (<i>N</i> = 554)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Maximum	0.0	3.14	0.0	28.27	1.00	5.60	2.74	45.25	33.95 ^a (33.95) ^b	7,521.43 (7,521.43)
Minimum	0.0	3.14	0.0	28.27	0.03	0.0001	0.00	0.08	0.03 (0.03)	0.07 (0.07)
Mean	0.0	3.14	0.0	28.27	0.40	2.39	0.71	21.77	1.89 (1.86)	73.66 (58.41)
Median	0.0	3.14	0.0	28.27	0.38	2.32	0.54	22.98	0.94 (0.90)	4.90 (4.71)
<i>SD</i>	0.0	0.0	0.0	0.0	0.22	1.06	0.59	7.69	3.47 (3.40)	511.07 (416.14)

^aThe first set of values is based on those associated with each of the 608 TSDFs. Thus, if a host tract contains more than one TSDF, it is included as many times as the number of TSDFs within its boundaries in determining the maximum, minimum, mean, median, and standard deviation of the host tract areas. Similarly, the distance between each TSDF and its host tract centroid is included in determining the maximum, minimum, mean, median, and standard deviation of the distances.

^bThe second set of values (in parentheses) is for the 554 host tracts, where each host tract is counted only once, regardless of the number of TSDFs within the tract, in determining the maximum, minimum, means, medians, and standard deviations of areas and distances. When a tract contains more than one TSDF, the distance used in the calculations is from the host tract centroid to the nearest TSDF.

the areas of the resulting neighborhoods are 1.06 and 7.69 square miles, respectively, while the standard deviations of the distances of TSDFs to neighborhood centroids are 0.22 and 0.59 miles. In contrast, the standard deviations of the areas of the host tracts proper and the distance of TSDFs to host tract centroids are 416.14 square miles and 3.40 miles.

With areal apportionment, the neighborhoods that are formed are also smaller than what they are for host tracts proper, forming perfect circles of one- and three-mile radii centered at the TSDFs with areas of 3.14 and 28.27 square miles. Because they are all perfect circles, there is no variation in the size of the neighborhoods or any deviation of the TSDFs from the neighborhood centroids (both with standard deviations of 0.0 square miles).

That the neighborhoods defined by the 50% areal containment and areal apportionment methods are generally smaller than the host tracts proper, and that their centroids are closer to the TSDFs within them, indicates that the populations residing in these neighborhoods are generally closer to the TSDFs than are populations in the host tracts proper. However, does closer proximity to the TSDFs lead to different outcomes, namely, larger proportions of poor people and minorities?

To answer this question, we compared the demographic characteristics of the host tracts proper against those of the neighborhoods surrounding the nation's 608 TSDFs defined by one- and three-mile radii, using the two distance-based methods (see Table 2). We determined demographic characteristics for aggregate populations, rather than computing averages across areas, because averaging skews results toward the less populated areas.² We made the comparisons in two ways. First, we compared the demographic characteristics of the populations residing in the neighborhoods defined by the one- and three-mile radii against the demographic characteristics of the populations within the host tracts proper, taking *all* host tracts into account. Second, we compared the neighborhoods defined by the one- and three-mile radii against only those host tracts *too large to be captured* by these radii. In the first comparison, we wanted to determine whether the proportions of minorities and low-income residents are greater in the neighborhoods defined by the two distance-based methods than those in the host tracts *overall*. In the second comparison, we wanted to see whether these disparities became even greater when the neighborhoods were contrasted against the host tracts that were too large to be captured.

In comparing the neighborhoods defined by a one-mile radius against *all* host tracts proper, we found that the proportion of nonwhites residing in neighborhoods within the circle is over 42.0% (42.8% when applying areal apportionment and 46.2% when applying 50% areal containment), while it is only 25.4% for those living in the host tracts (columns 3, 7, and 1, respectively, in Table 2). Similarly, the proportion of people living in poverty is over 19.0% in the neighborhoods (19.1% when applying areal apportionment; 20.6% when applying 50% areal containment), while it is only 13.6% in the host tracts (see Table 2 for comparisons of other socioeconomic characteristics). Although slightly smaller, the contrasts between neighborhoods defined by the three-mile radius and the host tracts proper remained substantial (compare columns 5, 9, and 1 in Table 2).

As anticipated, in comparing the neighborhoods defined by the one- and three-mile radii against the subset of host tracts too large to be captured by the radii, we found the racial and socioeconomic disparities to be even greater. For example, when the 50% areal containment method is applied, the proportion of nonwhites in host tracts too large to be captured by a

2. We computed the aggregate values by first summing subpopulations in each unit that makes up the host neighborhood, as captured by the distance-based methods, and then using those sums to construct the variables for analysis. For example, in determining the nonwhite percentage by applying the 50% areal containment method, we first determined the total number of nonwhites in the combined neighborhoods captured by the one-mile (or three-mile) radius, the total population (nonwhites plus whites) in the combined neighborhoods, and then divided the two values. This is in contrast to first finding the nonwhite percentage for each neighborhood and then averaging. We similarly determined the demographic characteristics for host tracts, that is, by aggregating their populations rather averaging them.

one-mile radius is only 18.7% (column 8). This is not only substantially less than the 46.2% found in the 50% areal containment neighborhoods but also less than the 25.4% found when all host tracts are taken together. Likewise, the percentage of people living in poverty in the host tracts too large to be captured is smaller than that of either the 50% areal containment neighborhoods or the host tracts taken as a whole (11.4% versus 20.6% or 13.6%, respectively). See Table 2 for comparison of other socioeconomic characteristics. Similar patterns are obtained when we examine tracts captured and not captured by the three-mile radius.

When the areal apportionment method is applied, virtually identical outcomes are produced. However, in keeping with this method, the demographic characteristics of only those portions, or "fragments," of the host tracts lying beyond the one- or three-mile distance were compared against the characteristics of the neighborhoods captured by these radii. In obtaining the demographic characteristics of these fragments, we weighted the contribution of each host tract by the percentage of its area *extending beyond* the radii. As before, when 50% areal containment was applied, the results in Table 2 reveal that the proportions of nonwhites and people of lower socioeconomic status are substantially smaller in the host tract fragments lying beyond the one- and three-mile radii (columns 4 and 6) than they are in the neighborhoods captured by the circle (columns 3 and 5). These proportions are also substantially smaller than in the host tracts taken as a whole (column 1).

The contrasts found when neighborhoods defined by one- and three-mile radii are compared against host tracts or host tract fragments too large to be captured are particularly revealing: this comparison most clearly demonstrates the effects of sorting tracts based on their proximity to TSDFs. Selecting the units based on proximity reveals substantially larger proportions of minorities and poor people near TSDFs than when units are not sorted. Clearly, proximity matters.

To determine whether the application of distance-based methods also leads to different assessments about the relative importance of racial and socioeconomic factors in the distribution of the nation's TSDFs, we performed logistic regression analyses using both unit-hazard coincidence and 50% areal containment methods and compared the results. We used individual tracts as the units of analysis.³ In applying the unit-hazard coincidence method, we assigned the dependent variable in the logistic regression a value of 1 if a census tract hosted a TSDF and a value of 0 if it did not. For the 50% areal containment method, the dependent variable took a value of 1 if the tract lay within one (or three) mile(s) of a TSDF and a value of 0 if the tract lay beyond that range. The independent variables included the race and socioeconomic variables described in the "Methods" section, excluding some of the variables (e.g., percentage nonwhite and percentage without a high school diploma) to avoid multicollinearity problems (see Table 3).⁴

3. Tracts, rather than neighborhoods, were used as the units of analysis in the logistic regression because of the difficulty of otherwise defining meaningful units to represent the nonhost neighborhoods in the case of the 50% areal containment method. Because of this difficulty, we follow the precedent of prior environmental inequality studies employing the 50% areal containment method (e.g., Anderton et al. 1994; Davidson and Anderton 2000; Pastor, Sadd, and Morello-Frosch 2004).

4. Because of possible spatial autocorrelation among the census tracts, there is some risk that the statistical significance levels of the independent variables may be inflated due to an underestimation of the standard errors. As Pastor et al. (2004) pointed out, prior environmental inequality research has tended not to employ spatial regression models. One of the difficulties is that spatial regression methods are currently performed using linear regression models that assume a continuous dependent variable. However, logistic regression models are the correct specification for environmental inequality analyses because the unit-hazard and 50% areal containment methods involve dichotomous dependent variables. To conduct a spatial regression, one must assume that the dichotomous dependent variable employed in the logistic regression is continuous and that the model is linear. Despite this limitation, Pastor et al. proceeded to assume a continuous dependent variable and to conduct a spatial regression in a version of their analyses. They found that the outcomes were not appreciably different from those obtained from logistic regression. We similarly replicated our logistic regression analyses using spatial regression methods and found the pattern of results to be similar. Because the results are not appreciably different and because logistic regression is

Table 3. Logistic Regression Results Comparing Unit-Hazard Coincidence and 50% Areal Containment Methods

Variable	Unit-Hazard Coincidence		50% Areal Containment (1-Mile Radius)		50% Areal Containment (3-Mile Radius)	
	Coefficient (1)	Significance (2)	Coefficient (3)	Significance (4)	Coefficient (5)	Significance (6)
% African American	-.003	.986	.698	.000	1.522	.000
% Hispanic	.431	.066	1.482	.000	1.960	.000
Mean household income (\$1,000s)	.012	.000	-.025	.000	-.015	.000
Mean property value (\$1,000s)	-.002	.058	.005	.000	.004	.000
% With a college degree	.338	.673	-1.704	.012	-.409	.046
% Employed in executive, managerial, and professional occupations	-3.215	.002	-.872	.282	.010	.970
% Employed in precision production or labor occupations	2.323	.000	1.787	.000	.073	.684
Constant	-5.052	.000	-4.197	.000	-2.220	.000
-2 Log-Likelihood	6,010.2		8,077.3		40,995.556	
Model chi-square	153.743	.000	548.233	.000	2786.536	.000
Sample size	59,050		59,050		59,050	

An examination of the results in Table 3 reveals important differences obtained from applying the two methods. For example, when we apply the unit-hazard coincidence method, the race and ethnicity variables are not at all significant predictors (at the .05 level) of the location of TSDFs (see column 2), while the occupation variables are (other socioeconomic variables are either not significant or predict TSDF location in the unexpected direction). Specifically, percentage employed in executive management or professional occupations is negatively associated with TSDF location, while percentage employed in precision production/labor is positively associated. Such results might suggest that a disproportionate presence of hazardous waste TSDFs near where minorities live may be a function of the tendency of TSDFs to be concentrated near where industrial labor pools exist (see, e.g., Anderton et al. 1994).

In contrast, when 50% areal containment is applied (using a one-mile radius), the African American and Hispanic percentages of the tract become highly statistically significant predictors of TSDF location (column 4), suggesting that racial disparities in the distribution of TSDFs are not solely a function of the labor force or other socioeconomic characteristics of nearby neighborhoods. Other factors associated with race (e.g., racial targeting or housing discrimination) may also be linked to TSDF locations (possibilities currently being examined in our research). Nevertheless, a number of socioeconomic variables (e.g., percentage employed in precision production/labor, mean household income, and percentage with a college degree) remain or increase in significance. Moreover, the

the correct specification for models with a dichotomous dependent variable, we display and discuss the results for the latter in this article. The results for the spatial regression, however, are available upon request.

model chi-square increases from 153.7 to 548.2 (compare columns 1 and 3), indicating that the model's overall ability to predict TSDF location is improved when 50% areal containment is applied. When a three-mile radius is used, similar results are obtained, except that percentage employed in precision production/labor is no longer a statistically significant predictor of TSDF location (column 6).

SUMMARY AND CONCLUSIONS

Many studies of racial and socioeconomic disparities in the distribution of environmental hazards and LULUs have been conducted. Although the majority of these have found that racial and socioeconomic disparities in these distributions exist (Ringquist 2005), there has been considerable variation in the magnitude of disparities found, with some studies finding no disparities. We have argued that a principal reason for the variation in findings is the wide reliance on unit-hazard coincidence methodology. We have demonstrated in this article that this method fails to control adequately for the proximity between environmental hazards and LULUs and nearby populations.

The unit-hazard coincidence method fails to control for proximity in two ways. First, it does not take into account the proximity of the hazard to adjacent or nearby units. Nonhost units that are nevertheless close to such sites are treated in the analyses in the same way as units much farther away. Second, it does not take into account the great variation in the size of the units of analysis typically used in such studies, such as tracts and zip code areas, and implicitly assumes that people living in large host units necessarily live as close to the hazard under investigation as people living in small host units. Although it is reasonable to expect that people living in small host units live close to the hazards within them, the same expectation cannot be made about large host units.

Distance-based methods overcome these limitations by including in the defined neighborhoods nearby units that are within a specified distance of a hazard while excluding units, including host units or unit fragments, whose areas lie mostly or entirely outside those distances. In contrast to "raw" units, such as tracts, the neighborhoods defined by distance-based methods are generally smaller and have greater consistency in their size and shape and greater consistency in the location of the hazards within them. We also note the considerable robustness of the results in estimating the demographic characteristics of the defined neighborhoods. The results are very similar regardless of which of the two distance-based methods is employed. Furthermore, although not shown in this article, the results are also very similar regardless of which predefined geographic units are used as the building blocks: block groups, census tracts, or zip code areas.

We demonstrated that when racial and socioeconomic disparities around the nation's TSDFs are analyzed by applying distance-based methods, such disparities are found to be greater than when the unit-hazard method is applied. Furthermore, distance-based methods lead to different assessments about the relative importance of racial and nonracial factors in the distribution of TSDFs. Such outcomes demonstrate the importance of applying distance-based methods in future efforts to determine the extent and causes of racial and socioeconomic disparities in the distribution of not only hazardous waste TSDFs but of a wide variety of environmental hazards and LULUs. Furthermore, because distance-based methods produce consistently sized and shaped geographic areas around hazardous sites, such methods should prove very useful in dealing with the irregularities of census boundary changes in longitudinal studies that seek to track demographic changes around environmentally hazardous sites over time, such as recently done by Pastor et al. (2001) and Saha and Mohai (2005).⁵

5. Although the distance-based methods that we describe in this article have clear applicability and usefulness to environmental inequality research, we anticipate that other distance-based approaches may also prove useful and warrant investigation. One such method that is not currently used in environmental inequality

In addition to environmental inequality research, distance-based methods are applicable to other areas of demographic research, particularly studies of the effects of neighborhood context on various social outcomes. Such studies have been increasing in number and involve examination of the relationship of neighborhood physical and social characteristics to such outcomes as the incidence of crime, adolescent sexual activity, neighborhood attachment, mobility, chronic stress, disease, and others (see, e.g., Barrett, Oropesa, and Kanan 1994; Ford and Beveridge 2004; and Scribner, Cohen, and Farley 1998; for a comprehensive review of such studies, see Sampson et al. 2002). The physical characteristics of neighborhoods in such studies are typically assessed by noting the presence or absence of such features as housing projects, liquor stores, vacant lots, and malls. Proximity to such features is assumed if they are coincident with some geographic unit, such as census tracts. In using the unit-hazard coincidence method, these studies face the same difficulties in discerning neighborhood effects that we examined in this article. In the one exception we are aware of, McNulty and Holloway (2000), taking actual distance into account, found that distance to public housing projects was a more important predictor of the incidence of crime than the racial and socioeconomic characteristics of neighborhoods. Given current widespread availability of GIS technology, we believe that distance-based methods should prove feasible and useful in these, as well as environmental inequality, studies.

APPENDIX: DEFINITION AND CONSTRUCTION OF KEY VARIABLES

Data are from STF 3A data files unless otherwise indicated:

1. Percentage African American: The total number of African Americans (Table P-8, Category 2) divided by the total number of persons (Table P-1).
2. Percentage Hispanic: The total number of persons of Spanish/Hispanic origin (Table P-10) divided by the total number of persons (Table P-1).
3. Percentage nonwhite: The difference between the total number of persons (Table P-1) and number of non-Hispanic whites (Table P-12, Category 1) divided by total number of persons (Table P-1). Includes the four nonwhite racial categories (black, American Indian/Alaskan Native, Asian/Pacific Islander, other nonwhite) and white Hispanic.
4. Percentage living in poverty: The number of persons below the poverty line in 1989 (Table P-119, Categories 36-70) divided by the number of persons for whom poverty status was determined (Table P-119, Categories 1-70). The poverty line is prescribed by Directive 14 of the Office of Management and Budget.
5. Mean household income: The total aggregate household income in 1989 (Table P-81, Categories 1 and 2) divided by the total number of households (Table P-5).
6. Mean housing values: The total aggregate value of specified owner-occupied housing in 1989 (STF 1A; Table H-24) divided by the total number of specified owner-occupied housing units (Table H-25, Categories 1-5).
7. Percentage without a high school diploma: Derived from Table P-57, Categories 1 and 2; represents persons aged 25 years and older without a high school diploma or its equivalent.
8. Percentage with a college degree: Derived from Table P-57, Categories 6 and 7; represents persons aged 25 years and older with at least a four-year college degree.

research is to measure the distance of census tract centroids to the nearest hazard (we thank one of the anonymous reviewers for pointing this out). Although such an approach would involve solving some logistical difficulties (e.g., is it necessary or practical to measure the distance of every one of over 60,000 tracts in the nation to each of the 608 TSDFs?), it would provide the advantage of producing a continuous dependent variable. Other alternate distance-based approaches may also become apparent in the future. Their common elements will be that they take the precise location of the hazards into account and control for proximity between the hazards and nearby residential populations.

9. Percentage employed in executive management or professional occupations: Derived from Table P-78, Categories 1 and 2; represents employed persons aged 16 years and older belonging to either of the following two occupational categories: (1) executive, administrative, and managerial, or (2) professional specialty.

10. Percentage employed in precision production or labor occupations: Derived from Table P-78, Categories 10–13; represents employed persons aged 16 years and older belonging to any of the following four occupational categories: (1) precision, production, craft, and repair, (2) machine operators, assemblers, or inspectors, (3) transportation and material moving operators, or (4) handlers, equipment cleaners, helpers, and operators.

REFERENCES

- Anderton, D.L., A.B. Anderson, J.M. Oakes, M.R. Fraser. 1994. "Environmental Equity: The Demographics of Dumping." *Demography* 31:229–48.
- Anderton, D.L., J.M. Oakes, and K.L. Egan. 1997. "Environmental Equity in Superfund: Demographics of the Discovery and Prioritization of Abandoned Toxic Sites." *Evaluation Review* 21(1):3–26.
- Barrett, A.L., R.S. Oropesa, and J.W. Kanan. 1994. "Neighborhood Context and Residential Mobility." *Demography* 31:249–70.
- Been, V. 1995. "Analyzing Evidence of Environmental Justice." *Journal of Land Use and Law* 11(1):1–36.
- Been, V. and F. Gupta. 1997. "Coming to the Nuisance or Going to the Barrios? A Longitudinal Analysis of Environmental Justice Claims." *Ecology Law Quarterly* 24:1–56.
- Boer, J.T., M. Pastor, J.L. Sadd, and L.D. Snyder. 1997. "Is There Environmental Racism? The Demographics of Hazardous Waste in Los Angeles." *Social Science Quarterly* 78:793–810.
- Bowen, W.M., M.J. Salling, K.E. Haynes, and E.J. Cyran. 1995. "Toward Environmental Justice: Spatial Equity in Ohio and Cleveland." *Annals of the Association of American Geographers* 85:641–63.
- Chakraborty, J. and M.P. Armstrong. 1997. "Exploring the Use of Buffer Analysis for the Identification of Impacted Areas in Environmental Equity Assessment." *Cartography and Geographic Information Systems* 24:145–57.
- Commission for Racial Justice (CRJ). 1987. *Toxic Waste and Race in the United States: A National Report on the Racial and Socioeconomic Characteristics of Communities With Hazardous Waste Sites*. New York: United Church of Christ.
- Daniels, G. and S. Friedman. 1999. "Spatial Inequality and the Distribution of Industrial Toxic Releases: Evidence From the 1990 TRI." *Social Science Quarterly* 80:244–62.
- Davidson, P. and D.L. Anderton. 2000. "The Demographics of Dumping II: Survey of the Distribution of Hazardous Materials Handlers." *Demography* 37:461–66.
- Ford, J.M. and A.A. Beveridge. 2004. "'Bad' Neighborhoods, Fast Food, 'Sleazy' Businesses, and Drug Dealers: Relations Between the Location of Licit and Illicit Businesses in the Urban Environment." *Journal of Drug Issues* 34(1):51–76.
- GeoLytics, Inc. 1999. StreetCD 98 [Digital data files on CD-ROM] (1990 TIGER Files). New Brunswick, NJ: Author [producer and distributor].
- Glickman, T. 1994. "Measuring Environmental Equity With GIS." *Renewable Resources Journal* 12(3):17–21.
- Glickman, T.S., D. Golding, and R. Hersh. 1995. "GIS-Based Environmental Equity Analysis—A Case Study of TRI Facilities in the Pittsburgh Area." Pp. 95–114 in *Computer Supported Risk Management*, edited by G.E.G. Beroggi and W.A. Wallace. Dordrecht, Netherlands: Kluwer Academic.
- Goldman, B.A. and L. Fitton. 1994. *Toxic Wastes and Race Revisited: An Update of the 1987 Report on the Racial and Socioeconomic Characteristics of Communities With Hazardous Waste Sites*. Washington, DC: Center for Policy Alternatives.

- Greenberg, M. 1993. "Proving Environmental Inequity in Siting Locally Unwanted Land Uses." *Risk: Issues in Health and Safety* (Summer):235–52.
- Hamilton, J.T. 1993. "Politics and Social Costs: Estimating the Impact of Collective Action on Hazardous Waste Sites." *Rand Journal of Economics* 24:101–25.
- . 1995. "Testing for Environmental Racism: Prejudice, Profits, Political Power?" *Journal of Policy Analysis and Management* 14(1):107–32.
- Hamilton, J.T. and W.K. Viscusi. 1999. *Calculating Risks? The Spatial and Political Dimensions of Hazardous Waste Policy*. Cambridge, MA: MIT Press.
- Hird, J.A. 1993. "Environmental Policy and Equity: The Case of Superfund." *Journal of Policy Analysis and Management* 12:323–43.
- Hird, J.A. and M. Reese. 1998. "The Distribution of Environmental Quality: An Empirical Analysis." *Social Science Quarterly* 79:694–716.
- Lester, J.P., D.W. Allen, and K.M. Hill. 2001. *Environmental Injustice in the United States: Myths and Realities*. Boulder, CO: Westview Press.
- McMaster, R.B., H. Leitner, and E. Sheppard. 1997. "GIS-Based Environmental Equity and Risk Assessment: Methodological Problems and Prospects." *Cartography and Geographic Information Systems* 24:172–89.
- McNulty, T.L. and S.R. Holloway. 2000. "Race, Crime, and Public Housing in Atlanta: Testing a Conditional Effect Hypothesis." *Social Forces* 79:707–29.
- Mohai, P. and B. Bryant. 1992. "Environmental Racism: Reviewing the Evidence." Pp. 163–76 in *Race and the Incidence of Environmental Hazards: A Time for Discourse*, edited by B. Bryant and P. Mohai. Boulder, CO: Westview Press.
- Oakes, J.M., D.L. Anderton, and A.B. Anderson. 1996. "A Longitudinal Analysis of Environmental Equity in Communities With Hazardous Waste Facilities." *Social Science Research* 25:125–48.
- Pastor, M., J. Sadd, and J. Hipp. 2001. "Which Came First? Toxic Facilities, Minority Move-In, and Environmental Justice." *Journal of Urban Affairs* 23:1–21.
- Pastor, M., J. Sadd, and R. Morello-Frosch. 2004. "Waiting to Inhale: The Demographics of Toxic Air Release Facilities in 21st-Century California." *Social Science Quarterly* 85:420–40.
- Perlin, S.R., W. Setzer, J. Creason, and K. Sexton. 1995. "Distribution of Industrial Air Emissions by Income and Race in the U.S." *Environmental Science & Technology* 29:69–80.
- Pollock, P.H. and M.E. Vittas. 1995. "Who Bears the Burdens of Environmental Pollution? Race, Ethnicity, and Environmental Equity in Florida." *Social Science Quarterly* 76:294–310.
- Ringquist, E.J. 1997. "Equity and the Distribution of Environmental Risk: The Case of TRI Facilities." *Social Science Quarterly* 78:811–29.
- . 2005. "Assessing Evidence of Environmental Inequities: A Meta-Analysis." *Journal of Policy Analysis and Management* 24:223–47.
- Saha, R. and P. Mohai. 2005. "Historical Context and Hazardous Waste Facility Siting: Understanding Temporal Patterns in Michigan." *Social Problems* 52:618–48.
- Sampson, R.J., J.D. Morenoff, and T. Gannon-Bowley. 2002. "Assessing 'Neighborhood Effects': Social Processes and New Directions in Research." *Annual Review of Sociology* 28:443–78.
- Scribner, R.A., D.A. Cohen, and T.A. Farley. 1998. "A Geographic Relation Between Alcohol Availability, and Gonorrhea Rates." *Sexually Transmitted Diseases* 25:544–48.
- Sheppard, E., H. Leitner, R.B. McMaster, and H. Tian. 1999. "GIS-Based Measures of Environmental Equity: Exploring Their Sensitivity and Significance." *Journal of Exposure Analysis and Environmental Epidemiology* 9:18–28.
- Wessex, Inc. 1994. 1990 Census of Population and Housing. Summary Tape Files 1A and 3A [Computer file on CD-ROM]. Washington, DC: U.S. Bureau of the Census [producer]. Winnetka, IL: Author [distributor].
- Zimmerman, R. 1993. "Social Equity and Environmental Risk." *Risk Analysis* 13:649–66.