

Chapter 19

What Level Macro? Choosing Appropriate Levels to Assess How Place Influences Population Health

Theresa L. Osypuk and Sandro Galea

1. Introduction

I should venture to assert that the most pervasive fallacy of philosophic thinking goes back to neglect of context.

John Dewey

Although it has strong historical roots (Davey Smith, Dorling, & Shaw, 2001; Krieger, 2001), the focus on area or contextual causes of health has only recently resurged in epidemiologic studies (Diez Roux, 2001). Most current epidemiologic inquiry continues to be concerned with studying determinants of health or disease that are proximal to the disease process (compared to causes that are more distal) and causes that are individual-level (compared to those at the population level) (McMichael, 1999). Although higher level causes must be mediated through individual-level and more proximal causes (Diez Roux, 2004b), certain disease causes may not be entirely operationalized at the individual level (Morgenstern, 1985). Increasingly we are recognizing that distal causes manifesting at higher spatial levels (e.g., neighborhoods, states) may present greater potential for health prevention than more proximal causes and, as such, are more fundamental causes of health (Link & Phelan, 1996; Schwartz & Diez Roux, 2001). As discussed throughout this book, studying macro-level causes of health and disease above and beyond individual causes may suggest avenues for disease prevention, intervention, and treatment that would not be evident from inquiry restricted to individual-level determinants.

Despite our growing appreciation that macro- or higher-levels of causes are fundamental for population health and health disparities, it is often unclear at what level to conceptualize and operationalize these macro-level causes. This chapter seeks to address this gap. In this chapter, we are concerned with issues pertaining to “macro-level” factors that exist at levels¹ of spatial aggregation above the individual and that may pertain to social and physical context as related to health.

We begin by discussing historical examples of different spatial levels that have been utilized for examination of health. We next review criteria for choosing what macro level to study, we provide examples of justifications for using various

spatial levels and we present practical issues that must be considered when choosing a level including defining boundaries, data availability, validity, and inference. We focus our discussion on US-based research, informed both by our experience and by the dominant body of work. However, our observations may also be germane to other countries and other levels of analysis. We additionally restrict our focus to examine macro levels pertaining to geography or space, leaving aside institutional contexts (e.g., schools, hospitals, workplaces) which merit separate discussion. We focus principally on a view of space or place that is useful for quantitative statistical analysis, e.g., for multilevel modelingⁱⁱ or ecologic analysis.

2. Historical Operationalizations of Place for Health Inquiry

The idea that place matters for health is not new (Kawachi & Berkman, 2003a, b), and historical examinations of place and health have invoked numerous alternative definitions of “place.” By the mid-19th century, data on geographic health patterns (including mapping of these patterns) had been produced in Britain and other parts of Europe (Macintyre & Ellaway, 2003). We discuss some of these historic examinations with an eye towards what level macro they examined.

In 1826, Louise Rene Villermé documented with census data that variations in annual mortality rates across Parisian neighborhoods (arrondissements) were patterned by poverty and wealth (Krieger, 2001; Macintyre & Ellaway, 2003). In 1837, Farr documented the geographic health variability and social class patterns of mortality in Britain with life-tables by district, observing,

“... the health of all parts of the kingdom is not equally bad. Some districts are infested by epidemics constantly recurring; the people are immersed in an atmosphere that weakens their powers, troubles their functions, and shortens their lives. Other localities are so favourably circumstanced that great numbers attain old age in the enjoyment of all their faculties, and suffer rarely from epidemics.” (Farr, 2001)

Farr thus highlighted healthy and unhealthy districts, noting the excess age and sex adjusted mortality rates in London and other urban towns (Liverpool, Manchester). He discussed several contemporary reports that summarized health variability at different levels of geography, including districts, villages, streets, parishes, towns, cities, and counties (Farr, 2001).

In *The Condition of the Working Class in England* (1845), Engels cited geographic variability in mortality and disease among towns in the mid-19th century, noting the particularly high rates in the largest factory towns (Manchester, Liverpool). Engels reproduced the mortality rate variability by social class of streets and houses originally reported by Holland and discussed variations in epidemic rates, epidemic mortality, and childhood accident mortality by comparing large towns (Manchester, Liverpool, and London) to country districts (Engels, 1845).

In 1902, Charles Booth utilized school board subdivisions, or “blocks”, within districts as his units of inquiry to examine spatial patterns of poverty in London.

Booth implied that districts corresponded to areas of the city (e.g., “the huge district of East London”) and were of a size conducive for school board employees to gain “extensive knowledge of the people” (Booth, 2001):

“The inhabitants of every street, and court, and block of buildings in the whole of London, have been estimated in proportion to the numbers of the children . . . The streets have been grouped together according to the School Board subdivisions or “blocks”, and for each of these blocks full particulars are given . . . The numbers included in each block vary from less than 2,000 to more than 30,000, and to make a more satisfactory unit of comparison I have arranged them in contiguous groups, 2, 3, or 4 together, as to make areas having each about 30,000 inhabitants, these areas adding up into the large divisions of the School Board administration. The population is then classified by Registration districts, which are likewise grouped into School Board divisions, each method finally leading up to the total for all London” (Booth, 2001).

Thus, Booth’s analysis examined place units of several sizes. Booth preserved the school board block unit, but he also created an alternate unit of comparable population size – “a more satisfactory unit of comparison” – by aggregating blocks. Moreover, he superimposed another administrative unit in addition to blocks: the Registration district. It appears that Booth used School Board data as the method for collecting poverty data on the population because the School Board kept extensive records; he noted, “every house in every street is in their books, and details are given of every family with children of school age” and because school board employees (“visitors”) had “very considerable knowledge of the parents of the school children, especially of the poorest among them, and of the conditions under which they live.” Indeed, parents and school board visitors were in daily contact. The results of Booth’s analysis culminated in a map of poverty by blocks (Booth, 2001).

Even among these few historical examples, researchers used many different sizes of “place” to document geographic health patterns, although the justification for the unit of place (even as today) is not often transparent.

3. How Does Spatial Frame Matter?

Central to our concern is the question of what level macro should be examined. Just as the topic of place has not received sufficient attention for health, so too, conceptualizations of what constitutes a place, of what is the appropriate level of analysis, and of how to model space have not occupied much space in epidemiologic writings. Harvey (2006) wonders:

“Are there rules for deciding when and where one spatial frame is preferable to another? Or is the choice arbitrary, subject to the whims of human practice? The decision to use one or the other conception certainly depends on the nature of the phenomena under investigation.” (Harvey, 2006).

As a geographer, Harvey eloquently suggests there is no “one” spatial level that matters (Harvey, 2006); rather, the choice of which specific geographic level to model must be derived from theoretical models for the specific research question at hand (Diez Roux, 2000; Leyland & Groenewegen, 2003). This theoretical

specification of relevant levels must precede data collection and statistical analysis (Diez Roux, 2004b). Generically, this requires defining what it is about a place that may provide adverse or protective exposures that affect health (O'Campo & Kogan, 2005). One must identify the exact hypothesis for examination and the hypothesized pathway by which the exposure translates to the outcome, i.e., via social and biological pathways (O'Campo & Schempf, 2005). Since different processes operate at different scales, issues of spatial scale naturally must be considered when choosing the level or unit of analysis (Macintyre, Ellaway, & Cummins, 2002). For example, the scale at which food availability affects diet may vary depending on the scale of human activity through which people come into contact with food stores. The spatial range of impoverished individuals may be more limited because of access to transportation, as compared with the spatial range of higher income individuals.

Few theories of area influences have been articulated in the health sciences literature, and one of the most vital issues impeding progress in place and health research may be lack of theory articulating the mechanisms of how place affects health (Diez Roux, 2000; O'Campo, 2003; O'Campo & Kogan, 2005). Indeed, since investigators fail to conceptualize, operationalize, and measure place effects with sufficient theory, selection of area variables becomes driven by available data (Macintyre et al., 2002). To advance the field, investigators must rigorously test alternate theories about why place matters for health. It is important not only to frame one's specific hypothesis with appropriate theoretical grounding, but also to test unifying explanations for empirical observations by contrasting one's theory against alternate theories.

Although researchers in epidemiology and public health have largely failed to ground place-health research in relevant theory, researchers in disciplines such as geography, sociology, criminology, and urban planning have articulated theories as to why place matters, and these conceptualizations can be applied to health studies. For instance, geographers have long struggled with what makes a place a place and with the definition of "space". One prominent geographer presents a model of three conceptualizations of space: absolute, relative, and relational (Harvey, 2006). Absolute space is bounded and fixed, symbolized as preexisting and immovable or as the space of Euclid in geometric terms. Through a social lens, absolute space encompasses boundaries of private property or territorial boundaries such as administrative units or states.

Relative space is defined by one thing in relation to another. Relationally, places may be conceived as "nodes in relational settings" or as "articulated moments in networks" (Castree, 2004). In application to health, a relational perspective would examine how people move through space from one destination to another and how the context along these specific routes or nodes may affect health. As an example, a relational perspective of how place affects health may operationalize neighborhood as the path one takes to walk or drive from one point to another – e.g., from home to work. The relational investigator may then model aspects of the neighborhood that one might encounter *on that path* as influencing health. Thus, instead of a spatial area that would be modeled in an absolute framework,

the investigator models individual paths through space. The third conceptualization of relational space is experiential, how people experience space and how a person internalizes information related to a space, including what memories or emotions people bring to or take from a space to construct meaning. This experiential notion of space may be very relevant for mental health outcomes, e.g., where the meaning of place reactivates trauma. Tension exists inherently among these three conceptualizations, and the particular spatial conceptualization to adopt is contingent on the circumstances (Harvey, 1973, 2006). For a relational understanding of place in health research, including a discussion of a relational contrasted with an absolute (Euclidean) viewpoint, we point the reader to other sources (Cummins, Curtis, Diez Roux, & Macintyre, In Press).

From a different tradition, American sociologists concern themselves with understanding how macro-level factors have inhibited and influenced individual behavior and social structure. For instance, many early sociologists argued that neighborhoods were fundamental building blocks of the larger system of stratification in the US and that social relations became manifested in spatial relations (Anderson & Massey, 2001). Illustrating this sociological perspective in his 1987 book *The Truly Disadvantaged*, William Julius Wilson argued that urban poverty was not only perpetuated through individual and family level pathways, but also through structural pathways, including within and between neighborhoods characterized by concentrated poverty (Anderson & Massey, 2001; Wilson, 1987). In a similar vein, some criminologists emphasize the importance of social disorganization theory – a community level explanation for crime that highlights the role of community structures and cultures that cause differential crime rates. This theory is often spatially expressed. For instance, structural features of neighborhoods, including economic status, ethnic homogeneity, and residential mobility, disrupt the social organization of communities, which in turn causes differential crime and delinquency rates across neighborhoods (Sampson & Wilson, 1995). However, the conception of community does not have to be strictly geographic or spatial, since social and organizational networks of residents may transcend smaller levels of geography (Sampson & Wilson 1995).

Transportation, urban design and planning disciplines focus on the importance of neighborhood design and land use development for understanding how and why people make transportation choices, for instance, whether people walk, use public transportation, and/or drive cars. This transportation/urban planning literature discusses how urban form or the built environment may influence physical activity (walking, cycling), for instance, through the layout of the street network or placement of buildings, proximity of residential to retail land uses, and detail of urban design features (Frank & Engelke, 2005, 2001; Saelens, Sallis, & Frank, 2003; Sallis, Frank, Saelens, & Kraft, 2004).

In sum, theory is specific to the phenomena under study and is necessary for choosing the appropriate macro-level unit and variables of interest for one's analysis. Although epidemiology offers few population theories of disease distribution, we add to the writings that have encouraged researchers to draw from the rich theory cultivated in other disciplines, as well as to voice explicitly how these exposures link to the disease outcome of interest through biologic plausibility and prior evidence.

4. Different Spatial Levels Relevant to the Study of Population Health

Researchers have considered many different units or levels in exploring why place matters for health, including neighborhoods, counties, states, nations. Decisions about what level of space may be appropriate for analysis should follow from a theoretic appreciation of what characteristic of a particular space may matter for what health indicator. Assumptions and hypotheses that may be valid at one level may not be valid at others. For example, while census tract-level analyses may be appropriate to consider the relation between characteristics of the built environment and mental health, the role of factors such as employment markets, residential segregation, or state policies governing alcohol sales to minors may not be studied meaningfully at the neighborhood level, since those constructs operate at higher levels like metropolitan areas or states. And although an investigator may be able to operationalize variables at the census tract level (e.g. tract-level unemployment rate, tract-level percent black), these variables may capture only part of a process that operates at higher levels (e.g. regional unemployment trends or racial segregation at the metropolitan level).

In the following section, we discuss the historical meaning or definition of different spatial units as originally conceived in the US, why research at such a level is relevant for health, and some limitations involved with considering how characteristics at that level may influence population health. Many of the drawbacks inherent in choosing any one unit of analysis may also be relevant to other levels of analysis. We will address different units working from smallest to largest geographic areas.

4.1. Neighborhoods

4.1.1. Historical Roots and Why the Level May Matter for Health

Some historians posit that the American neighborhood structure emerged in the mid- to late-19th century, as American cities transitioned from small “walking cities” into more segmented, larger urban units. American cities before 1860 were small, compact, and generally integrated. Before technological advances in travel (e.g., the omnibus and the horsecar), settlement sizes were small enough that the city was virtually entirely walkable; the size of cities rarely extended beyond a 2 mile radius from the city center. In the mid-19th century, however, technological innovations allowed the elite to escape the overcrowded conditions of the walking city. The exodus of the elite sorted populations by residence, a sorting which was accelerated by additional transportation innovations, large-scale immigration, and internal migration to American cities in the late 19th century. As the cities grew, their internal structure became characterized by distinct areas, and the residential housing areas came to be known as neighborhoods (Melvin, 1985).

Neighborhoods are a common level for conceptualizing the importance of place (National Research Council, 2002) and are often what is meant by “place”

in the contemporary population health literature (Diez Roux, 2001; Macintyre et al., 2002; National Research Council, 2002). A neighborhood effects literature has developed over the past decade, of which health forms one part; neighborhood effects are defined as outcomes from a causal process of an exposure of living in a particular neighborhood (Altshuler, Morrill, Wolman, Mitchell, & Committee on Improving the Future of U.S. Cities Through Improved Metropolitan Area Governance, 1999).

Several theoretical models discuss why neighborhoods may matter for well-being, highlighting material and social pathways. For example, criminologists have emphasized the role of community social organization and disorganization for producing crime, particularly at the level of neighborhoods. Sociologists have discussed how individuals are embedded within the ecology of neighborhoods, and particularly harmful ecologic contexts for outcomes (e.g., health, education, or employment) include concentrated poverty neighborhoods. The neighborhoods and health literature has been burgeoning lately, so there is no shortage of conceptual frameworks on how neighborhoods may affect health. For an overview on neighborhoods and health, see Kawachi & Berkman 2003a or Ellen, Mijanovich, & Dillman, 2001. As one example of a conceptual framework, from a review of the literature, Ellen et al. (2001) present four mechanisms describing how neighborhoods may matter for health: (1) neighborhood institutions and resources (e.g., food environment, opportunities to promote exercise), (2) stressors in the physical environment (e.g., polluting factories, older housing structures, quality of municipal services), (3) stressors in the social environment (e.g., crime victimization, witnessing crime, noise), and (4) neighborhood-based networks and norms (e.g., for transmitting information, norms, and social support) (Ellen et al., 2001). These pathways may be very important for health, especially if neighborhoods are the dominant social context for an individual or group.

4.1.2. Drawbacks

Chief among the limitations to considering neighborhoods as the key macro-level unit of analysis is the problem of selection in observational studies, whereby individuals with particular characteristics migrate to neighborhoods characterized by the features of analytic interest (Diez Roux, 2002a, 2004a). Selection of individuals into neighborhoods threatens causal inference about neighborhood characteristics and their role in shaping health (Kawachi & Berkman, 2003b). Migration-related selection may be a greater threat to validity in observational studies using neighborhoods than in studies using higher-level spatial units. Residential mobility *between* neighborhoods *within* counties, metropolitan areas (MAs), or states is much more common than mobility *between* counties or MAs or states. The forces operating to sort people into certain locations (e.g., discrimination in housing markets, employment opportunities) are more prominent *within* metropolitan areas *between* neighborhoods than *between* metropolitan areas, states, or counties (Ellen, 2000; Osypuk, 2005). Related to the issue of migration is the relevance of the size and meaning of the neighborhood unit with respect to relevant exposures of interest. Residents often travel beyond the

boundaries of their neighborhood daily, so neighborhood-focused studies will not allow modeling of other aspects of their context (e.g., workplaces, schools, travel routes, social spaces) that may impact their health.

Because of the powerful racial and socioeconomic sorting mechanisms operating in metropolitan housing markets, people of different racial and socioeconomic backgrounds often live in different neighborhoods in the metropolitan area (Acevedo-Garcia, Lochner, Osypuk, & Subramanian, 2003; Massey & Denton, 1993). While this is a complex sociological phenomenon, one implication of this sorting is that it becomes difficult to consider all racial and social class groups together in one neighborhood analysis because such empiric models often violate exchangeability assumptions (Diez Roux, 2004a)(Diez Roux, 2005). A related but distinct concern is non-overlap of populations within neighborhoods with respect to confounders, which can lead to off-support inference (Diez Roux, 2004a; Oakes, 2006). When members of different racial groups live in starkly separate neighborhoods, which is the case in many U.S. metropolitan areas given patterns of racial residential segregation (see (Morenoff, Diez Roux, Osypuk, & Hansen, 2006) for an empiric example with Chicago neighborhoods), stratifying on race may reduce these validity threats (of lack of exchangeability or non-overlap) in neighborhoods analysis. However, residual confounding may remain even after stratification.

Another challenge pertains to ensuring that the sampling scheme that was used to obtain the analytic sample for a particular study appropriately lends itself to the analysis in question. Often neighborhood health studies derive from one of two sampling plans: from national probability samples of individuals or from neighborhood studies within a city or metropolitan area. In the former sampling scheme, the study was typically conceived with individual-level hypotheses in mind, so individuals are typically sampled as the unit of interest and then geocoded to their neighborhood tract. As a result, the application of multilevel modeling to such analysis may be improper; multilevel modeling assumes that the higher-level units are randomly drawn from a larger source population (Duncan, Jones, & Moon, 1998), while in the case of national sampling plans, the assumption is that individuals, not the neighborhoods, are drawn from the larger source population. In addition, this approach plucks neighborhoods from their metropolitan areas and thus models neighborhoods absent their metropolitan context (Diez Roux, 2001); again, this may not be proper if the characteristic of interest is most meaningful relative to its own metropolitan area rather than in an absolute sense compared to the rest of the country. Even if multilevel modeling were used, this design likely precludes complex covariance structures (e.g., random effects models) given the limited analytic power dictated by the small sample size in each neighborhood (Duncan et al., 1998). Therefore, investigators using this design often employ marginal models (or population average models), which allow the modeling of the average effect of a type of neighborhood (characterizing neighborhoods by variables of interest) without modeling the variance attributable to each level (Diez Roux, 2002b).

The second type of neighborhood sampling plan typically involves sampling a sufficient number of neighborhoods within one metropolitan area or city and then sampling persons within these neighborhoods. This type of sample design may be more appropriate for assessing how the neighborhood matters and more generalizable regarding neighborhood effects or associations, at least for that city/MA. Examples of studies that have used such a design include the Project for Human Development in Chicago Neighborhoods (PHDCN) (Raudenbush & Sampson, 1999) and the Los Angeles Family and Neighborhood Study (L.A. FANS, 2006). Since these studies are designed to examine hypotheses about the influence of neighborhood characteristics on population health, they wield sufficient power for analyzing random effects and are appropriate for multilevel statistical analysis. A limitation of this approach is the restriction of the study to only one site or metropolitan area; this hampers generalizability if comparable features of neighborhoods produce different health effects in different places throughout the country.

The spatial autocorrelations of observations within neighborhoods requires special methods to either adjust for or model statistical dependency (e.g. multilevel models). Analyzing neighborhoods may have limited policy relevance unless resources are allocated along these neighborhood boundaries. Adopting administrative definitions that cities use for neighborhoods may be one solution to enhance policy relevance (e.g., the 59 New York City community districts (New York City Department of City Planning, 2006) or the 77 Chicago community areas).

A last important limitation to using neighborhoods regards their definition. Although typically neighborhoods are defined as one's immediate residential environment (Diez Roux, 2001), operationalizing neighborhood boundaries is a challenge. We will discuss these challenges of operationalizing levels in the next section, and in Section 4.3.

4.1.3.1. Historical Roots of Operationalizing Neighborhoods as Census Tracts

The relevant definition of neighborhoods has garnered much attention in neighborhood health research since it is not clear how to bound or define neighborhoods. Neighborhoods have been defined using geographical criteria, historical criteria, administrative boundaries, characteristics of residents, and resident perceptions (Diez Roux, 2001; Diez Roux, 2003). For the purposes of this chapter we discuss the two administrative units most commonly used as proxies for neighborhoods: Census Tracts and ZIP Code postal areas.

Many quantitative researchers have used administrative units defined by the US Census, i.e., census tracts, interchangeably with "neighborhoods", despite the fact that census tracts are only one particular operationalization of neighborhood. The principal reason why census tracts have been adopted enthusiastically as proxies for neighborhoods within the neighborhoods and health literature is that US Census data exists at the census tract level, hence providing readily available administrative data that may be used to characterize these "neighborhoods".

The idea of census tracts originated with Walter Laidlaw, who first divided New York City into tracts for the 1910 Census. Census tracts are artificial units

created only for analyzing the population distribution by different geographies, at a smaller and more consistent scale than political jurisdictions afford. To define census tracts, committees of local data users are asked to create units that follow recognizable boundaries and encompass areas of between 2500 and 8000 population. The Census and community groups draw boundaries based on homogeneity of population in terms of economic status and housing conditions. Once the geographic units of tracts have been established, the Census permits only splits or recombinations of tracts from a previous census. One of the main goals of tracting the population is to provide continuity across time. Therefore, the Census Bureau prioritizes preserving the fixed boundaries over preserving the within-tract homogeneity (Plane, 2004). Hence, the main strength of using tracts relates to the availability of data and the stability of the units' boundaries across time.

4.1.3.2. Drawbacks of Operationalizing Neighborhoods as Census Tracts.

Despite the advantages of census tracts, residents rarely conceive of neighborhoods along census tract boundaries (O'Campo & Kogan, 2005). Local areas like neighborhoods may not be simple places around which one can draw a line (Massey, 1991). Rather, the line to capture a certain exposure perfectly at the neighborhood level may not be fixed for any one person or population but may be contingent on different purposes, processes, or health outcomes (Diez Roux, 2003; Diez Roux & Aiello, 2005; O'Campo & Kogan, 2005). Therefore, the use of census tracts as proxies for neighborhoods is necessarily crude. However, modeling variables at arbitrarily-defined administrative units may still capture some notion of context even if it is not capturing the construct as accurately as possible (Duncan et al., 1998). Since census tracts are artificial units created for administrative purposes, they are probably relevant units for health because their boundaries crudely approximate social and physical phenomena that affect health but that operate at other, unspecified spatial scales. The implications of using census tracts to proxy the neighborhood construct involve measurement error, internal validity, and the Modifiable Areal Unit Problem (MAUP), which we will discuss in Section 5.3. For example, since measurement of groups and of group-level constructs affecting health remains in its infancy, tract-level variables capturing neighborhood-level phenomena likely contain a great deal of measurement error (Diez Roux & Aiello, 2005).

An additional drawback to using census tracts for operationalizing neighborhoods relates to one of the strengths: use of available data. The most common variables investigated in neighborhood health research to date have been indicators of aggregate socioeconomic position derived from administrative (e.g., Census) data (Diez Roux, 2001). Although census tracts allow one to access and model Census data, the variables one is able to model with Census data (e.g., tract-level, census defined economic and demographic variables) are not themselves the literal cause of health or disease patterns. Census variables are likely crude proxies for other more direct social or physical characteristics, resources, services, or hazards of a place which affect health (Diez Roux, 2007). Choosing tracts and tract-level census data might be important for initial

descriptive analysis, for replicating prior findings, or for monitoring place-based disease patterns; however, an important direction for neighborhood health studies at this stage includes utilizing data sources and variables that are not derived from the census. Non-census based variables may capture better the specific social and physical features of the neighborhood that cause health and illness, thereby informing the mechanisms by which neighborhood is relevant for health and informing interventions to improve health (Diez Roux, In press). We note that although we discuss this issue with respect to census tracts, the issue is equally relevant to other units (e.g., ZIP codes, discussed below) and other census-based units, such as block groups, that have also been used in some studies as proxies for neighborhoods.

4.1.4.1. Historical Roots of Operationalizing Neighborhood as ZIP Code Postal Areas

Between 1863 and 1963, individuals in the U.S. addressed letters by the street address of the recipient, including the city and state, or by denoting the rural delivery route. When thousands of postal employees left the U.S. Postal Service (USPS) to serve with the military during World War II, the USPS created a zoning address system in 124 of the largest Post Offices to facilitate mail sorting. For this system, 1 or 2 numbers were placed between the city and state in the addressing of a letter – e.g., Birmingham 7 Alabama. In 1963, the Post Office improved upon this rudimentary coding system by implementing the Zoning Improvement Plan (ZIP) code. Beginning on July 1, 1963, every address in the US was assigned a 5-digit ZIP code:

“The first digit designated a broad geographical area of the United States, ranging from zero for the Northeast to nine for the far West. This number was followed by two digits that more closely pinpointed population concentrations and those sectional centers accessible to common transportation networks. The final two digits designated small Post Offices or postal zones in larger zoned cities.” (U.S. Postal Service, 2003)

In 1983, the USPS introduced the ZIP + 4 code, with a hyphen and four digits added to the existing 5-digit ZIP code:

“The first five numbers continued to identify an area of the country and delivery office to which mail is directed. The sixth and seventh numbers denoted a delivery sector, which may be several blocks, a group of streets, a group of Post Office boxes, several office buildings, a single high-rise office building, a large apartment building, or a small geographic area. The last two numbers denoted a delivery segment, which might be one floor of an office building, one side of a street between intersecting streets, specific departments in a firm, or specific Post Office boxes. This ZIP + 4 code again improved the efficiency of mail delivery by reducing the number of times a piece of mail is handled, and by reducing the amount of time letter carriers spent organizing the mail for delivery.” (U.S. Postal Service, 2003)

A good deal of data is collected and available at the ZIP code level, also making this level a convenient candidate for considering the influence of space on health. As the USPS notes, “Today’s use of a ZIP Code extends far beyond the mailing industry. ZIP Code numbers are embedded into the way that businesses

work and have become an integral element of the 911 emergency system that uses ZIP code mailing codes as an aid in saving lives” (U.S. Postal Service, 2006a). For instance, businesses collect customer ZIP codes for establishing new store locations and for direct mail marketing (where businesses target products and services to certain addresses). The Census released population data based on ZIP codes in 1990 (Summary Tape File 3) (U.S. Census Bureau, 2002, 2006b), and the US Economic Census releases data on American businesses by ZIP code (U.S. Census Bureau, 2006a). ZIP codes may be a rough proxy of a person’s location of residence for research purposes, and with the inclusion of “+ 4”, the precision increased substantially. Therefore, ZIP codes may approximate neighborhood of residence.

Since the public has requested statistics at the ZIP code level, the Census Bureau has created a new statistical area called the ZIP Code Tabulation Area (ZCTA) for Census 2000 data. This statistical geographic entity approximates the area defined by a 3 or 5-digit postal service ZIP code (U.S. Census Bureau, no year).

“ZCTAs were designed to overcome the operational difficulties of creating a well-defined ZIP Code area by using Census blocks (and the addresses found in them) as the basis for the ZCTAs.” (U.S. Census Bureau, 2005b).

For the most part, ZCTAs coincide with ZIP codes (U.S. Census Bureau, 2001a). We refer the reader to the US Census (U.S. Census Bureau, 2000) for further detail about the ZCTA methodology.

One potential benefit of the ZIP code relates to how it is defined: ZIP code boundaries are conceptually related to how people (postal employees) move through space. For a given scale at which ZIP codes operate, ZIP code delineations may be more relevant than Census geography because of people’s routes through them. Since ZIP code boundaries were defined with a mail carrier’s route in mind (and this is often by foot in densely settled areas), the ZIP code’s rear property boundaries indicate that people living across a street from each other share the same “neighborhood”. With Census geography, streets are often boundaries, thus dividing people across the street from one another into separate blocks or tracts.

The USPS has several sources of data that might be relevant for historical study. Postal routes, post office locations, or postal employees that existed before ZIP codes may inform historical health questions involving population settlement or geography. The USPS has cultivated sources for archival research on aspects of historical postal service, post offices, postal routes, mail contracts/contractors, and personnel (e.g., postmasters, salaries) (U.S. Postal Service, 2006b). For instance, historical post route maps, available for the 1830s to the 1940s at the National Archives and the Library of Congress (U.S. Postal Service, 2006b), may indicate settlement patterns related to population density and shipping patterns (U.S. Postal Service, 2006b).

4.1.4.2. Drawbacks of Operationalizing Neighborhoods as ZIP Code Postal Areas

Unfortunately, the key drawback to the use of ZIP codes for the purposes of defining characteristics of space that may influence population health is that ZIP

codes are very heterogeneous in terms of area and population size and population composition. ZIP code addresses correspond to historical post office locations, which may not correspond to current political nomenclature (e.g., names of places for the purposes of mailing addresses may differ from the name of the political unit). Historically, “Post Office names were typically suggested by prospective patrons; there are no postal records that explain their origin”(U.S. Postal Service, 2006b). Moreover, given that some ZIP codes refer to points or routes of delivery, it may be difficult to model spatial effects of ZIP codes. Although in some areas ZIP codes conform neatly to polygons (urban areas especially), in other areas (e.g., rural areas) polygons cannot be created accurately (U.S. Census Bureau, 2006b). Therefore, they are primarily a collection of linear or point features (e.g., carrier routes).

Although the system was geographically derived, the US Postal Service had created ZIP codes to facilitate efficient mail delivery, as groupings of mailing addresses or delivery nodes. According to the Census, “ZIP codes do not respect political or census statistical area boundaries. ZIP codes usually do not have clearly identifiable boundaries, often serve a continually changing area, are changed periodically to meet postal requirements, and do not cover all the land area of the United States” (US Census Bureau). ZIP code “boundaries do not necessarily follow clearly identifiable visible or invisible map features; also, the carrier routes for one ZIP code may intertwine with those of one or more other ZIP codes, and therefore this area is more conceptual than geographic” (US Census Bureau, 2001b). ZIP code boundaries often cross state, place, county, census tract, block group, and census block boundaries (US Census Bureau, 2006a). Although the geographic units by which the Census Bureau tabulates data are relatively stable across time, ZIP codes are altered periodically to meet day-to-day operational needs of the US Postal Service; thus, continuity across time may be a problem (U.S. Census Bureau, 2006a).

ZIP codes may approximate neighborhoods or cities, although that representation varies considerably. ZIP codes cover large geographic areas, including upwards of 30,000 people (Krieger, Williams, & Moss, 1997), though ZIP + 4 codes certainly improve precision. We recently documented that the median ZCTA in US metropolitan areas in year 2000 contained about 15,000 people. Ordering ZCTAs by population, the 1st quartile was at 6,723 people and 3rd quartile at 28,146. However, these sample sizes per ZCTA varied considerably (Osypuk, 2006).

4.2. *Cities*

4.2.1. Historical Roots and Why the Level May Matter for Health

Cities or townships are Minor Civil Divisions in the US and are considered tertiary political subdivisions (with states as primary political units below the federal government and counties as secondary political units). As such, this municipal level comprises one of the lower political units in the United States. Reporting data for the municipal level dates back to the first Census in 1790, with data

provided for local governmental units (Plane, 2004). In 2002, there were 35,933 municipal or town/township general-purpose governments in the United States (US Census Bureau, 2004a). The largest cities in America house a good proportion of the US population. For instance, in the year 2000, 8.5 percent of the US population lived in the largest 10 cities, and 16 percent of the population lived in the largest 50 cities (U.S. Census Bureau, 2001c). Twenty-seven percent of the U.S. population lived in incorporated places (e.g., cities) of over 100,000 population (Katz & Lang, 2003).

The city or municipality is a politically relevant unit for education and service distribution, local public health regulation, and other domains of local control (e.g., zoning). Cities matter for health because of the immediate physical and social environments and the range of available services that they provide for residents, as well as economic and political factors (Galea, Freudenberg, & Vlahov, 2006). For instance, Galea and colleagues posit that three broad municipal-level determinants influence health, including government (policies and practices implemented in cities), markets (food, housing, labor), and civil society (community organizations, community capacity, social movements) (Freudenberg, Galea, & Vlahov, 2006; Galea et al., 2006).

Government influences population health via provision of municipal services, regulation of health activities, and dictation of parameters for urban development. Municipal services include education, social services, policing, courts/jails, fire services, housing, parks and recreation, sanitation, transit, environment protection/water supply, economic development, and zoning and urban planning (Freudenberg et al., 2006). As depicted in Figure 19.1, governmental expenditures at the city level accounted for 25 percent of all expenditures at the city, county, and state levels. Cities spend more on average than counties (comprising 16 percent of total expenditures) and spend less than half as much as states (comprising 59 percent of total expenditures). Three of the main expenditure categories for cities are for public utilities (water, electric, gas, transit), public safety (police protection, fire protection, corrections, protective inspection and regulation), and environment/housing (natural resources, parks and recreation, housing and community development, sewerage, solid waste management). As Figure 19.2 shows, cities spend more in these three categories than counties or even states (U.S. Census Bureau, 2004b). Therefore, the city level might be relevant for studying determinants of health in these or other domains where cities dominate governmental spending relative to other governmental levels. If one's research hypothesis relates to governmental provision of services or resources, then comparing governmental expenditures at different levels of government might be one criterion for deciding on an appropriate unit of analysis.

The network of cities comprising the core of every current major US region except Las Vegas was established in the 19th century (Katz & Lang, 2003), so historical factors certainly played a role in the evolution of US urban form. Freudenberg et al. (2006) emphasize four social trends that contributed to historical geographic explanations for health patterns: migration, suburbanization,

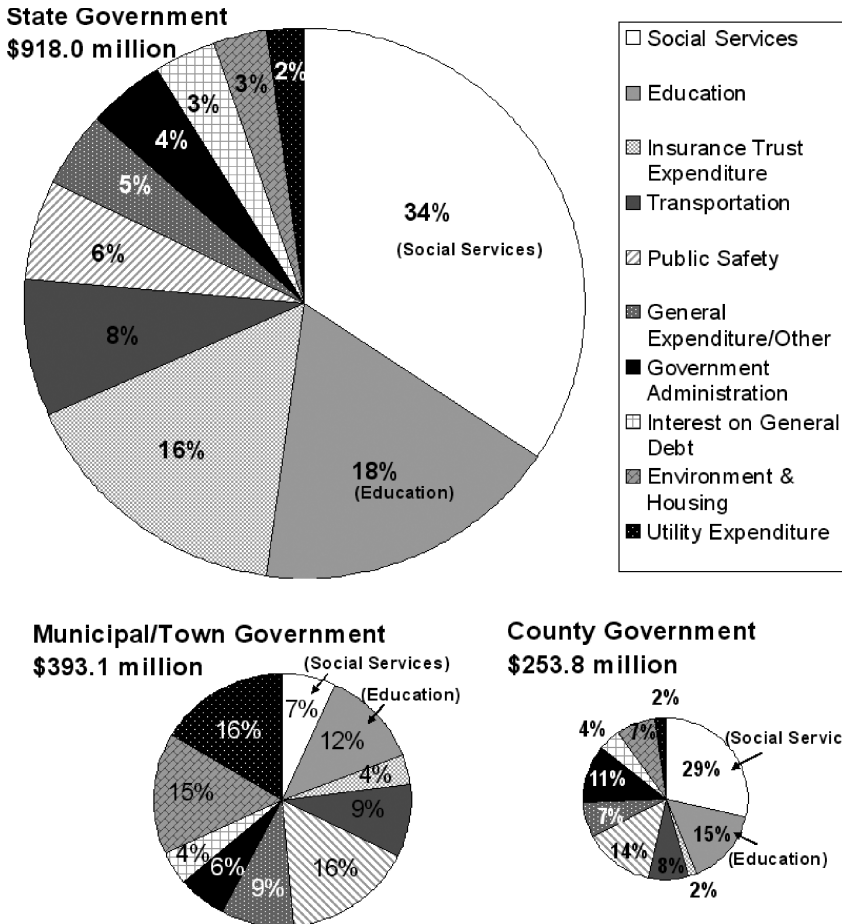


FIGURE 19.1. Expenditure categories for state, county, and local governments, proportional to size of total expenditures, 2002 Census of Governments Social services (public welfare including cash assistance payments, vendor payments and other public welfare; hospitals; health; social insurance administration; veteran’s services); education (higher, elementary, and secondary education; other education; libraries); insurance trust expenditure (unemployment compensation, employee retirement, workers’ compensation, other); public safety (police protection, fire protection, correction, protective inspection and regulation); transportation (highways, airports, parking facilities, sea/inland port facilities, transit subsidies,); environment and housing (natural resources, parks and recreation, housing and community development, sewerage, solid waste management); general expenditure (miscellaneous commercial activities, other and unallocable); governmental administration (financial administration, judicial and legal, general public buildings, other); utility expenditure (water supply, electric supply, gas supply, transit); interest on general debt.

Source: U.S. Census Bureau, 2004b.

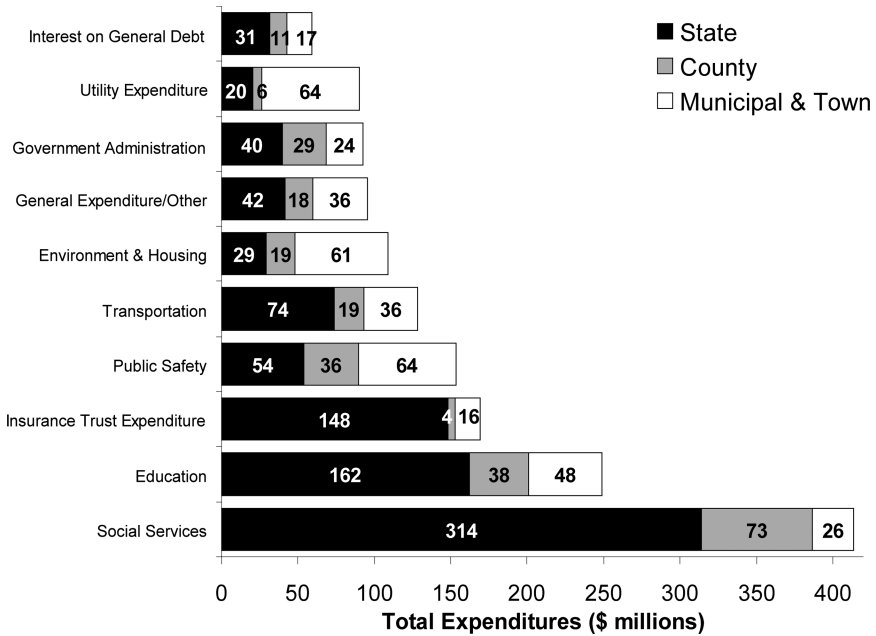


FIGURE 19.2. Expenditure by function and government level for state, county, and local governments, Census of Government Expenditures 2001–2002. Social services (public welfare including cash assistance payments, vendor payments and other public welfare; hospitals; health; social insurance administration; veteran’s services); education (higher, elementary, and secondary education; other education; libraries); insurance trust expenditure (unemployment compensation, employee retirement, workers’ compensation, other); public safety (police protection, fire protection, correction, protective inspection and regulation); transportation (highways, airports, parking facilities, sea/inland port facilities, transit subsidies,); environment and housing (natural resources, parks and recreation, housing and community development, sewerage, solid waste management); general expenditure (miscellaneous commercial activities, other and unallocable); governmental administration (financial administration, judicial and legal, general public buildings, other); utility expenditure (water supply, electric supply, gas supply, transit); interest on general debt.

Source: U.S. Census Bureau, 2004b.

changes in the role of government, and changes in the global economy. Finally, Cities hold significant historical and non-historical meaning, which might be integral for the identity or pride of residents.

4.2.2. Drawbacks

Some residents in the US are not served by a municipal government because they live outside of incorporated areas. Therefore, examination of cities will exclude some portions of the population. A related drawback of choosing cities as a unit

of analysis relates to truncation of the population distribution, leading to reduced variance or limited generalizability. For instance, choosing the city as the unit of analysis for a nationwide study of the largest US cities omits suburban populations. If one is interested in assessing the macrosocial determinants of racial or socioeconomic inequality, excluding suburban populations means omitting the most privileged group from the analysis, since development patterns and policies promoting suburban development have facilitated migration of higher social class, white populations to the suburbs.

Another central limitation to the use of cities as an analytic spatial unit of interest is the tremendous heterogeneity among cities and in their definitions; this is a drawback for other spatial units of analysis as well (counties, states, nations). Cities are complex communities of heterogeneous individuals, and although multiple factors are important determinants of population health, some of these factors may be unique to a specific city. Assessing how urban living may affect health raises issues often not easily addressed through the application of simple analytic methods. Empiric inquiry in health presupposes that there are identifiable factors that influence average levels of health. Although certain aspects of city living may be universal (e.g., population density) and relevant for the health of populations in different cities, other urban characteristics that are important in one city (e.g., local transportation patterns and air pollution) may not be important in others, limiting the generalizations that can be drawn about how urban living influences health.

Further complicating this task is the fact that the populations of cities change over time due to migration, which has implications for the relative contribution of different factors in affecting population health. For example, municipal taxation of alcohol and cigarettes may be an important determinant of alcohol and cigarette consumption in a particular city at one point in time (Grossman, 1989). However, changing social norms around smoking and alcohol use may either obviate or reinforce the influence of taxation. Therefore, when considering urban characteristics that affect health, it is important to note both the prevailing context within which such characteristics operate and that the role of these characteristics may change over time.

Finally, definitions of cities may vary. Particularly with regard to cross-national analysis, investigators must ensure consistent definition of cities. As discussed elsewhere in this book, the US Bureau of the Census (2006c) defines an urbanized area as “an area consisting of a central place(s) and adjacent territory with a general population density of at least 1,000 people per square mile of land area that together have a minimum residential population of at least 50,000 people.” (US Census Bureau, 2006c) However, this definition, as all other possible definitions of “city”, has inherent limitations. Most cities are actually far-reaching densely populated areas, containing peri-urban and suburban areas, which continue relatively uninterrupted for miles beyond the municipal city boundaries and the city center. Therefore, any empiric analysis that considers how characteristics of cities may influence population health must consider an appropriate definition of city that is relevant for all the urban units under consideration or cross-nationally.

4.3. *Counties*

4.3.1. Historical Roots and Why the Level May Matter for Health.

Counties are the primary divisions of US states. The United States has 3,066 counties, and 48 of the 50 US states maintain operational county governmentsⁱⁱⁱ (US Census Bureau, 2004a). Historically, the county is rooted in the function of English shires, which served a dual function as both an administrative body of the national government and the residents' local government. The framers of the US Constitution relegated local government to the power of states and thus did not call for local governments. State constitutions created the counties as an extension of state government. (National Association of Counties, 2006b)

Historically, the state mandated that counties perform tasks such as keeping property and vital statistics records, assessing property, administering elections and judicial functions, maintaining rural roads, and directing poverty programs. For instance, counties continue to handle responsibility for elections at all levels, including voter registration, education, and ensuring integrity of the voting process. Currently, counties also run programs including economic development, child welfare, consumer protection, zoning and planning, water quality, and training/employment. According to the 2002 US Census of Governments (depicted in Figure 19.1), county expenditures are allocated mostly to social services and health services – 29 percent for public welfare (e.g., cash assistance, hospitals, health, social insurance administration, and veteran's services), 15 percent for education services, 14 percent for public safety (e.g., police protection, fire protection, corrections), and 11 percent for governmental administration (financial administration, judicial and legal, general public buildings) (U.S. Census Bureau, 2004b). Relevant to analyses that consider the role of county-level characteristics that may influence population health, the majority of county employees have jobs in education, hospitals, and police protection, but other large job categories include streets/highways, corrections, public welfare, fire protection, health, justice/legal system, and financial administration (National Association of Counties, 2003, 2006b). Relative to state and city expenditures, however, county governmental expenditures budgets are small.

4.3.2. Drawbacks

As we discussed with regard to cities, counties vary greatly in their functions and service delivery, in terms of both infrastructure (utilities, water, roads) and social service responsibility. For instance, although most counties build and maintain roads as a core function, North Carolina counties have no responsibility to do so. In terms of service delivery, certain states entirely administer Medicaid (e.g., Alabama), alleviating counties of fiscal and administrative responsibility. However in Iowa, counties play a much larger fiscal role with Medicaid, e.g., funding 100 percent for certain waiver services not covered by the federal government and half of the case management funding not covered by the federal government. (National Association of Counties, 2003, 2006b)

The number and size of counties varies substantially throughout the states. For instance, Texas has the most counties (254), while Delaware and Hawaii have the least (3). County areal size ranges from 26 square miles in Arlington, VA, to 87,860 square miles in North Slope Borough, AL. County population size ranges from 67 residents in Loving County, TX, to 9.5 million residents in Los Angeles County, CA. (National Association of Counties, 2006a).

There are many differences among the states and counties in the form of their political areas. For instance, not all geographic areas called counties have county governments. In some places, like New York City and Philadelphia, municipal and county governments have merged or have been consolidated and are therefore counted as municipal governments for the purposes of government statistics. Some cities are outside areas of adjacent counties, and these independent cities are considered county equivalents – e.g., Baltimore, MD, Carson City, NV, St. Louis, MO, and 40 cities in Virginia (Plane, 2004). Therefore, almost 10 percent of the US population is not served by a county government, although in these cases they are served by a large municipal government (U.S. Census Bureau, 2004a). The majority of counties are sparsely populated; nearly three quarters of all county governments in 2000 were comprised of counties with populations under 50,000 (National Association of Counties, 2006b). Thus, an unweighted county-level analysis will overemphasize rural areas.

As noted in Figures 19.1 and 19.2, among sub-national government levels, counties expend fewer dollars than either cities or states, at least in the aggregate. Counties comprise only 28 percent of the size of state expenditures, but this may vary substantially in certain places or by expenditure category.

4.4. Metropolitan areas

4.4.1. Historical Roots and Why the Level May Matter for Health.

A metropolitan area is defined by the United States Office of Management and Budget (OMB) as a core area with a large population nucleus (an area containing at least one urbanized area with a population of at least 50,000) combined with adjacent communities that share a high degree of economic and social integration with that core (Office of Management and Budget & Spotila, 2000). The 1910 Census first officially recognized the metropolitan concept and designated metropolitan districts for cities with populations over 100,000. By 1930, the government expanded metropolitan district definitions to cities with populations of more than 50,000, creating 140 recognized metropolitan districts in 1940. Population density was the basis of metropolitan district boundaries from 1910 to 1940, and minor civil divisions (municipalities/cities/townships) were used as the building blocks. However, since 1950, counties have been the building blocks for delineating metropolitan units, except in New England, where towns are more powerful governmental units. Therefore, although they are not political units themselves, Metropolitan areas are defined using standard political areas (counties or towns) as their building blocks (Plane, 2004). Metropolitan areas are therefore much larger than cities and counties but smaller than states.

The Census defines three different types of metropolitan areas: Metropolitan Statistical Areas (MSAs), Consolidated Metropolitan Statistical Areas (CMSAs), and Primary Metropolitan Statistical Areas (PMSAs). CMSAs are large regions with a population size over 1 million, comprised of several smaller component parts (called PMSAs). MSAs are stand-alone metro areas that are not part of CMSAs. Since CMSAs are so much larger than typical employment markets, often metropolitan-level analyses include PMSAs and MSAs, not CMSAs (Jargowsky, 2003; Osypuk, 2005; Subramanian, Acevedo-Garcia, & Osypuk, 2005).

Over 80 percent of the US population resides in one of the 331 metropolitan areas defined in 1999 for the 2000 Census (Orfield, 2002). As such, the metropolitan area level captures how the majority of US residents interact with the housing market, job market, and educational systems, which operate beyond one city, county, and sometimes state boundary but usually within a regional area proxied by the metropolitan area spatial unit. Metropolitan areas may therefore be an appropriate level of analysis for racial inequality related to the housing market (including racial residential segregation, housing, lending) or to employment markets that may operate outside of any local jurisdiction.

The sociology literature suggests that the metropolitan area is a conceptually relevant unit for studying racial inequality (and links between racial inequality and health) because the significance of race may be powerful in terms of local race relations, residential housing and neighborhood processes, and labor market outcomes – processes that operate beyond individual-level attitudes or acts of discrimination. Metropolitan areas are the context within which geographic inequalities in the housing and education markets developed during the 20th century along a central city-suburban spatial divide (Frey, 2001; Jargowsky, 1997). Residential suburbanization and deconcentration of industry has transformed the geography of urban settlement, shifting new housing development and new jobs toward suburban areas – changes that have disproportionately harmed minorities (O'Connor, 2001). The concentration of minorities, especially African Americans, in impoverished neighborhoods of metropolitan areas embodies the spatial distribution of metropolitan racial inequality (Altshuler et al., 1999).

Metropolitan areas are frequently referred to as geographic regions; since much advertising is purchased locally, the metropolitan area may be a meaningful unit for analyzing advertising, marketing campaigns, and other media related health effects. Metropolitan areas may also be relevant for analyzing the built environment, given regional development patterns across the 20th century, e.g., suburban sprawl. Transportation infrastructure is often something that transcends one municipality to affect neighboring regional areas, as is sewage infrastructure or pollution.

4.4.2. Drawbacks

Unfortunately, metropolitan areas are often redefined at every Census, which complicates temporal comparisons. For instance, after the 2000 Census, the Census Bureau defined a new system for 2003, the Core Based Statistical Area (CBSA)

classification, and this system is not entirely comparable to prior definitions of metropolitan areas. MA definition changes have chiefly occurred due to (1) recognition of new areas reaching the minimum required city area population and (2) addition of counties (or cities and towns in the case of New England) to existing areas as a result of new decennial Census data (U.S. Census Bureau, 2005a).

Americans often view the two components of metropolitan areas, suburbs and central city areas, as separate types of places; however, the interests and future of both places are intertwined, financially and otherwise (Orfield, 2002). Some problems require regional solutions, for instance transportation, water infrastructure, real estate prices. However, metropolitan areas are not very politically relevant units. A metropolitan area is not a legal or political entity but a statistical entity. There is rarely a political body corresponding to their boundaries. Indeed, state government deference to local governments with respect to land use, incorporation of municipalities, and fiscal systems has resulted in fragmented, decentralized structures that have exacerbated socioeconomic and racial inequality (Altshuler et al., 1999). Despite the current political limitations to metropolitan governance, some advocate political strategies to pursue a common regional agenda, e.g., “Metropolitics” and regionalism for community growth (Orfield, 2002). Some places have developed regional governance bodies to share the cost burdens of transportation or to contain growth patterns (e.g., Minneapolis, MN, Portland, OR), but they remain the exception. For researchers positing political hypotheses for health, this suggests that MAs are unlikely to be the most relevant unit of analysis.

Additional drawbacks of the metropolitan unit include exclusion of the rural population, and as with other units discussed here, there is vast heterogeneity in the physical size and population composition of MAs across the US.

4.5. States

4.5.1. Historical Roots and Why the Level May Matter for Health

States are the primary political units in the US. There are 50 states in the United States proper, although the District of Columbia is treated equivalently to states by the Census Bureau (Plane, 2004). Since states are responsible for service provision to their populations, they are a relevant level for many public health programs. The majority of public health statutes in the US are found at the state level, including the missions, functions, structures, and powers of public health agencies (Gostin, Koplan, & Grad, 2003). Indeed, many responsibilities not set forth explicitly for the (relatively weak) federal government by the US Constitution fall to states. As a result, judicial case law often reinforces the realm of states to exert power in areas of health and public health.

Government at any level, although especially the state level, has several tools to promote public health, including tax and spend policy; the power to shape the informational environment; direct regulation of individuals (e.g., behavior such as seat-belt use), professionals (licensing), or businesses (restaurant health codes); indirect regulation through tort litigation (to incite businesses to engage in less risky activities, e.g., the tobacco Master Settlement Agreement, litigation to restrict tobacco

advertising and promotion); and deregulation (Gostin et al., 2003). For instance, states are relevant for health because elected officials as part of state legislatures or regulatory bodies pass numerous laws and regulations (Osypuk, Kawachi, Subramanian, & Acevedo-Garcia, 2006). States are therefore particularly relevant units for statistical analyses because of the policies they enact.

States also directly implement and evaluate health programs to ensure the health of their populations. To that end, states routinely collect information for monitoring, planning, and evaluation of programs. States also use health data to inform program eligibility (e.g., Medicaid, SCHIP). Identifying geographic patterns in health at the state level allows the prioritizing of programs and allocating of resources for high-risk populations, as well as planning for future disease treatment programs (Osypuk et al., 2006). As Figure 19.1 and Figure 19.2 demonstrate, the state expenditures far outweigh those of county and local governments, at over twice the size of local governments and 4 times the size of county government expenditures (US Census Bureau, 2004b). States spend the majority of their budget on social service and insurance trust provision, including for welfare cash assistance payments, hospitals, health, social insurance administration, veteran's services, unemployment compensation, employee retirement, and workers' compensation (US Census Bureau, 2004b).

4.5.2. Drawbacks

As with other political levels we have discussed here, the size, form, and function of states and their governments differ considerably across the United States. Some states are almost entirely rural, while others are dominated by very large urban areas. States differ along many dimensions aside from size, including the composition of the population, economic activity, urban/rural residence, and climate – all of which likely affect health. For instance, some states are destination points for large numbers of immigrants (New York, California, Texas), while others contain practically no immigrants. Some states house large numbers of so-called “snow-birds” – retired persons migrating from northern states to warmer states, e.g., Florida or Arizona. As such, Florida has the highest average population age. Therefore, without standardizing the population by age, Florida and other snowbird states would appear to contain an excess of disease, since age is a strong individual-level predictor of morbidity and mortality.

Some phenomena under study are less relevant for state level analysis because the state has failed to step in with policy or action. For instance, land use policies are typically localized, and although states theoretically could regulate this area, they typically do not. Therefore, in this instance, state would be the less appropriate level for analysis, and localities or municipalities would be the more appropriate level. Further, although states may be attractive and appropriate for some types of policy analysis, if the responsibility for funding programs or enforcing the laws falls to lower governmental units (e.g., counties), then modeling these lower units is more fitting. Lastly, many state laws or regulations (e.g., state health codes) are quite fragmented because they evolved independently, resulting

in profound differences among states in the structure, content, and procedures of regulations for disease monitoring, control, and prevention (Gostin et al., 2003).

4.6. Nations

4.6.1. Historical Roots and Why the Level May Matter for Health

Countries are clearly the relevant spatial unit of interest when we are concerned with the influence of macro-level social and economic factors that relate to national governmental decisions and actions. Nations are appropriate units because of the power of federal governments to create policies, to tax and spend, and to regulate activities in all sectors of life. For instance, some countries have passed comprehensive social policies to aid their populations, and these policies likely affect health through their influence on socioeconomic position or provision of concrete resources to the population (e.g., subsidies to buy housing or health care). Since they are financed by federal government dollars (raised by taxation), these social policies are redistributive in domains such as education, child care/maternity policy, unemployment, and health care. Taxation is largely controlled at the national level, although variation within any nation may exist (e.g., in the US there is variation in state and local income tax rates), and taxation is one main income redistribution mechanism. Because the nation is such a relevant level for income distribution (e.g., minimum wage policy) and redistribution (e.g., taxation, payroll deductions for social programs), some investigators have found the unit appropriate for examining how income inequality between countries may affect population health (Wilkinson, 1992b, 1992a; Wilkinson & Pickett, 2006). Health economists have studied national factors such as economic development, political economy, and employment structure in relation to health. For instance, capitalism (Cereseto & Waitzkin, 1986a, 1986b) or the degree of political freedom (e.g., democracy) (Franco, Alvarez-Dardet, & Ruiz, 2004; Martyn, 2004) may be associated with population health.

The federal government plays a key role in ensuring quality of consumer products and of the food supply in the US by regulating corporate practices. For example, some impacts on food supply relate to the nutritional content of nationally distributed packaged goods. Therefore, the inclusion of harmful or less nutritionally beneficial ingredients in this food supply (e.g., trans fats) may only be regulated at the federal level. Federal governments can also add nutrient additives to the food supply to prevent disease, for instance, fortification of bread with folic acid to prevent neural tube defects or fortification of water with fluoride to prevent tooth decay (Grosse, Waitzman, Romano, & Mulinare, 2005). Government at this level also can regulate the safety of products, for example, standardizing automobile design, mandating lower vehicle emissions, mandating flame-retardant children's clothing, or regulating firearms (Hahn, et al., 2005). Economic development is beneficial for population health and economic cycles may operate at the national level (or above – e.g., the Great Depression). Some have researched how between-country differences in diet, family structure, social custom, or distribution

of income within a society may influence health (Kawachi, Wilkinson, & Kennedy, 1999; Rose, 1992). The national level may also be meaningful for examining between-country corporate practices that may affect health but are not adequately regulated by government – e.g., tobacco marketing, product design, or advertising (Ballin, 1994; Freudenberg, 2005; Gilmore & McKee, 2004). It will be interesting to observe whether national boundaries may decline in importance in places where international migration is easier (e.g., the European Union).

4.6.2. Drawbacks

One of the primary drawbacks of cross-national studies is the availability of data. Although data is more often available for ecologic studies, multilevel studies require individual-level data to examine contributions to individual health. Cross-national comparisons are difficult given the variations in the fundamental geographic units for which data is gathered in different countries (variations in name, function, power of primary and secondary political units) (Plane, 2004). International comparisons are additionally complicated by the many other factors by which nations differ from each other; issues such as the meaning and measurement of constructs may be complicated by multiplicities of language and cultural differences. As discussed in relation to other levels, countries are incredibly heterogeneous and impart different historical meaning and identities to their residents.

4.7. *Other Administrative Units*

4.7.1. Historical Roots and Why the Level May Matter for Health

Spatial areas delineated for administrative purposes may be appropriate for studying health if the causal processes are related to that unit. Examples include Congressional districts (the districts with an elected representative in the U.S. House of Representatives), legislative districts (those represented by lawmakers in the state legislature), school districts, traffic zones, and neighborhood planning districts (Plane, 2004). For instance, Congressional districts may be relevant for examining policy making, voting districts may be relevant for measuring political empowerment or disenfranchisement, and police precincts are relevant if police monitor these areas and provide crime statistics at that level.

There were 34,683 specialized-function governments (aside from school districts) according to the 1997 Census of Governments (Plane, 2004). Special district governments provide a single service or one group of related specific services not supplied by general purpose governments, for instance, hospitals, fire protection, mosquito abatement, or cemetery maintenance (U.S. Census Bureau, 2004a). Analyzing special district governmental units would therefore be relevant when examining such purposes as they relate to public health (e.g., mosquito abatement related to vector-borne infectious disease). Once again, use of an administrative unit makes sense if the exposure or the outcome is conceptually patterned at this level.

4.7.2. Drawbacks

Since specialized governmental entities are very specific, the causal pathways must be conceptually relevant. Further, depending on the unit and its usage by administrative bodies, these units may or may not be stable across time; for instance, in some western states, election precincts may be altered after each election based on the number of votes cast (Plane, 2004).

5. Practical Issues: Operationalizing Constructs and Units

Choice of unit or level for analysis affects all aspects of one's study. Many of these issues overlap, so although we organize them here in certain categories, they should not be considered mutually exclusive. We recommend that a number of factors be taken under consideration when an investigator chooses the levels and units of analysis in his/her study. The spatial level at which an exposure of interest is measured may influence the construct validity of the exposure measure, the internal validity and external validity of the study, and the potential threats to statistical conclusion validity due to data availability and power. We will discuss each of these topics in turn below. Ultimately, as discussed, one must refer to theory to explicate constructs of interest and to specify the causal model.

5.1. *Validity*

First, choice of level affects construct validity, which is defined as whether the measured variable accurately represents the higher order construct (Shadish, Cook, & Campbell, 2002). The two issues involved with construct validity include initially conceptualizing a construct and then operationalizing or measuring it (Shadish et al., 2002). The first is theoretical, the second empirical. If the construct is operationalized or estimated with error, a biased estimation results. An important, and perhaps the primary, threat to construct validity here is measurement error; this is discussed below as it relates to threats to validity arising from how one operationalizes levels and boundaries.

Second, choice of level affects internal validity. Internal validity refers to inferences about whether the covariation between two variables reflects an unbiased causal relationship (Shadish et al., 2002). Serious threats to internal validity include selection and confounding, which we will discuss below. Another issue with regard to internal validity is the validity of inferences that derive from the data and study design, since the study design (including level of data) bears greatly here.

Third, external validity will be affected by one's choice of unit, defined as the inferences of whether a presumed causal effect is maintained across variations in persons, settings, treatment variables, and measurement variables (Shadish et al., 2002). Fourth, statistical conclusion validity concerns the use of statistics for inferring covariation of two variables, including whether the observed covariation of

two variables is likely due to chance (Shadish et al., 2002). Statistical conclusion validity concerns power, sample size, and effect sizes.

5.2. Data Availability and Power

Even if one identifies the ideal level, definition of units at that level, or spatial range of exposure, the availability of data at that level may be limited (Macintyre et al., 2002). Else, if primary data collection is desired, cost may be limiting. Secondary data may be helpful. Some secondary data is not released as individual-level units or at units small enough to be manipulated in the way researchers may want; often data is aggregated into larger units, e.g., for confidentiality protections, so the researcher may have to manage with the unit available to him/her.

Data used by researchers is often collected for other purposes (e.g., administrative purposes) (Diez Roux, 2005). Therefore, it is important to examine whether the data is etiologically relevant for the health outcome or exposure process under study. If not, construct validity is threatened. If cost required for primary data collection limits data availability, one solution may be to collaborate efforts across studies in the same area for shared data collection efforts (O'Campo, 2003).

Power may be limited in studies where data was intended for only individual-level analysis, and this threatens statistical conclusion validity. In multilevel analysis, power is determined by the number of groups and by the number of observations per group (Diez Roux, 2005). The former determines the power to assess between-context associations; the latter determines power for within-context associations (Duncan et al., 1998). Sparse data means analyses are possible only at higher levels (e.g., cities, counties) rather than at lower levels (census tracts). Power is also influenced by the variability or homogeneity of the sample. The same techniques that increase power in an individual-level study will also generally improve power in a multilevel study, including larger heterogeneity and range of units on the independent variable; matching, stratifying or blocking; multivariate analysis; using larger sample sizes; using equal cell sample sizes; improving measurement; reducing random setting irrelevancies; and ensuring that assumptions are met for powerful statistical tests (Shadish et al., 2002). For instance, one will have more power to detect area socioeconomic effects on health with homogeneous socioeconomic areas, and areas tend to be more internally homogeneous at smaller levels (Greenland, 2002).

5.3. Operationalizing Levels and Boundaries

Although operationalizing or defining the relevant unit of analysis is straightforward in some cases, such as institutional settings (e.g. schools) or political boundaries, it can be more challenging in other cases – especially with regard to neighborhoods. The bounding or definition of neighborhoods in neighborhood health research is a topic of contention (O'Campo, 2003). As with any exposure,

bias can result if the neighborhood construct of interest does not map onto the units chosen for operationalizing neighborhoods and their constructs. Defining neighborhoods (or any relevant unit) mainly threatens construct validity, although it threatens internal validity as well (e.g., non-differential exposure misclassification tends to bias the effect estimate towards the null).

One important methodologic problem for operationalizing neighborhood boundaries and interpreting the conclusions in spatial studies is the Modifiable Areal Unit Problem, commonly known as MAUP. MAUP is defined as “a problem resulting from the imposition of artificial units of spatial reporting on continuous geographical phenomenon resulting in the generation of artificial spatial patterns” (Heywood, 1998). MAUP is a concern in spatial analyses when the definitions of boundaries may be unclear, and as a result, differences in how one draws spatial boundaries may lead to artificial statistical spatial patterns – an issue of exposure misclassification. Two issues are at play with MAUP – scale and aggregation. First, the scale problem results in different statistical answers when the same data is aggregated at different scales (corresponding to different sizes of areal units). If one does not have a clear idea about the level at which a certain phenomenon is operating, one might choose a certain unit of analysis because it minimizes aggregation bias. For example, Bailey and Peterson (1995) chose cities instead of metro areas or states in an analysis of how gender socioeconomic inequality affects female homicide victimization because cities minimized aggregation bias; as crime statistics are collected at the city level, modeling crime statistics at higher levels (metro areas or states) would have required aggregating data collected at the city level.

The second MAUP issue is the aggregation problem or zoning effect, which occurs when one receives a different answer when drawing the line differently at a particular scale (Yang, 2005). For instance, a neighborhood social phenomenon that may affect health (for instance, social capital) is not likely bounded by the edges defined by tract boundaries. In other words, social capital may not be sharply higher or lower as one crosses the arbitrary census tract boundaries. *Some* census tract boundaries may coincide with sharp breaks in the true construct of social capital, e.g., breaks that occur across a busy street or highway that is difficult to traverse or across a river or ravine, because this geographic boundary meaningfully separates different people from these places, and there may be little social mingling across the barrier. However, social capital in other places may gradually decrease or increase across space, despite the fact that our methods model the construct as abruptly decreasing or increasing across these boundaries.

The issue of how to define neighborhoods may also be a conceptual one. For instance, how an outsider defines neighborhood boundaries might be different from how a resident him/herself does. There is also likely heterogeneity within a certain neighborhood as to how residents define their neighborhood. It is difficult to claim that these differences are simply a result of measurement error; it is more likely that the underlying construct of neighborhood varies for different people. If one is interested in social exposures that affect health (e.g., social interactions), then resident perceptions or definitions of neighborhood might be more relevant

for that examination than for, say, resource distribution, which may occur according to administrative boundaries (Diez Roux, 2001). It is also important to keep in mind that some processes occurring in neighborhoods that an investigator hypothesizes to affect health (e.g., social networks) are not necessarily contained within any given spatial boundary (O'Campo, 2003).

Although the challenges for operationalizing relevant neighborhoods are important to consider as one weighs level of analyses, they should not paralyze one's empiric investigation (Diez Roux, 2001). Since different phenomena may operate at different scales to affect health, multiple appropriate neighborhood units may be defined to accommodate inclusion in one's study of multiple processes that operate in the neighborhood environment, including institutional, political, economic, or cultural phenomena (O'Campo, 2003). Indeed, a multitude of different groups/units may be relevant for a specific research question (Diez Roux, 2004b).

One satisfactory solution for bounding neighborhoods is to define them according to how the city planners define them, which may also relate to resource or service distribution (e.g., public utilities, public safety services like policing). For instance, some neighborhood health studies have identified neighborhoods by how cities identify them for administrative purposes – e.g., the 59 New York City community districts (New York City Department of City Planning, 2006) or the 77 Chicago community areas.

Finally, although neighborhoods have dominated the discussion in the literature with regard to proper bounding of areas, it is certainly not the only level featuring challenges to operationalization. For example, another level that might be difficult to operationalize is media markets, which may be relevant for health-related communications messages – e.g., the reach of television stations in a metropolitan area. There may be difficulties in defining a market based on topography that dictates signal strength or on individual subscription to cable versus satellite providers.

5.4. Measurement and Measurement Error

Although epidemiologists have become sophisticated at measuring individual-level phenomena, measuring group-level phenomena related to health remains in its infancy (Diez Roux, 2004b; Diez Roux et al., 2005; Oakes & Kaufman, 2006). As a result, existing studies have measured group-level effects with much measurement error, and these exposures may be grossly misspecified, either because of error in defining the group or in operationalizing the group-level variable (Diez Roux et al., 2005). Many population-level causes of health are irreducible to individual level analogs and are thus constitutive properties of only populations. Examples of such causes are income inequality and herd immunity, two determinants of population health that have no individual analog. Although some group-level variables have analogs at the individual level or at multiple higher-levels, their meaning and theoretical connection to health likely differs at each level (Diez Roux, 2000). We refer the reader to other sources for further discussion of group-level factors in epidemiologic study (Diez Roux, 2001, 2004b).

Measurement reliability and study replicability are affected by choice of macro-level units, since both refer to the reproducibility of measurement or results across different variables or studies. For instance, one reason to employ a certain level of analysis may be because the rest of the literature does it that way. Replicating analytic approaches from prior studies improves the comparability of results across studies and communicates transparency of methods. Thus, one might choose cities as the unit of analysis because all other studies in the extant literature do so as well (Bailey & Peterson, 1995). Types of reliability include interrater (how different people define the same construct in the same way) and test-retest reliability (the stability of the construct and its measure across time). These dimensions of reliability relate to measurement of the construct of interest (construct validity); while the issues involved with macro-level constructs may be extended from the individual-level psychometric literature, there are other reliability issues specific to the econometrics of context. Consider an example where we assess the reliability of a measure of school context derived from a scale reported by individual teachers. Psychometric principles suggest that the reliability of a measure can be increased by adding more items to a scale; more important for improving the reliability of the school level measure is increasing the degree of rater agreement within a school and increasing the number of raters per school (Raudenbush & Sampson, 1999). While some investigators have proposed statistical methods to assess the validity and reliability of group-level measures (e.g., Raudenbush et al., 1999), creation of valid and reliable measures of context as they relate to health remains a need in the public health literature (Diez Roux, 2001, 2004b; Raudenbush et al., 1999).

5.5. Inference

The level of data aggregation for the variable affects the inference and meaning of the variable. For instance, the unemployment variable at census tract-level seems to indicate the relative deprivation or affluence of the tract population; however, tract-level unemployment would not represent employment markets since people travel well beyond their own tract of residence for work. An unemployment rate at the city, metro, or state level is more likely representative of structurally-patterned employment opportunities since employment markets operate at these levels; moreover, unemployment rates at these levels seem orthogonal to other aspects of relative affluence or deprivation levels at these levels (Land, Cantor, & Russell, 1995). At the national level, unemployment rates fluctuate with business cycles (Land et al., 1995) and would not be appropriate measures of relative affluence or deprivation.

The level of variable operationalization holds implications for the level at which one may derive inferences about study effects. False inferences arise when there is a mismatch between the level at which one has data and the level at which one wants to infer effects; these errors have been classified as the ecologic fallacy, atomistic fallacy, psychologistic fallacy, or sociologistic fallacy, depending on the mismatch. Ecologic fallacy is committed when one uses group level data to draw

inferences about relationships at the individual level. Conversely, the atomistic fallacy may be present when one makes inferences about groups or group-level variables based on individual-level data (or more generally, drawing inferences about units at a higher level based on data collected for units at a lower level).

Fallacies arising from mismatch between data and inferences have received considerable attention in epidemiology; however, other more substantive problems regarding ignoring variables at different levels have not been as much discussed (Diez Roux, 2000). For instance, the psychologistic fallacy derives from failing to consider group characteristics when drawing inferences about causes of variability among individuals. In other words, this fallacy arises when one assumes that exclusively individual-level characteristics explain individual-level outcomes. Conversely, the sociologistic fallacy arises when one fails to consider individual level characteristics when making inferences about the causes of variability among groups (Diez Roux, 2002b). Ultimately, to avoid incorrect inferences, the researcher needs to use the right level of data as well as appropriate methods to ask the right question (Schwartz & Carpenter, 1999). Further, although this discussion has focused on choosing one level of analysis for study, disease processes related to social structure invariably concern multiple levels of constructs that may operate independently, successively, or simultaneously to affect health. As discussed in other chapters of this book, a more comprehensive understanding of disease processes will be gained by considering these multiple levels of causes and understanding their causal associations with health.

Generalizability, or external validity, is another relevant inferential issue to consider when choosing one's macro-level data. External validity is determined by the structure and execution of the sampling frame/plan with respect to the theoretical study population. For example, if data are based on a survey, then that data might not be representative at smaller levels than the specified sampling frame. In the same vein, the sampling frame must be examined closely to guide methodologic issues regarding how certain levels should be represented in a multilevel model. For example, in a multistage sample design it is unclear whether a level that was not part of the sampling frame should be used as a level in the analyses. If it is used, certain variance correction techniques may need to be undertaken to adjust properly for clustering (e.g., in the outcome); investigators of multistage samples often view clustering as a necessary nuisance for more efficient survey execution instead of as an inherent piece of information that may relate to common causes. However, the capacity is currently limited for common multilevel statistical programs to adjust for multistage survey designs and simultaneously model variance at different levels, particularly if the primary sampling units are masked for confidentiality (Osypuk et al., 2006).

5.6. Policy or Intervention Relevance

Other criteria that might guide decisions about level of analysis may relate to the purpose of the analysis in terms of the salience of one level for action versus another (O'Campo & Kogan, 2005). From this perspective, policy actionability is

constrained by political or administrative boundaries, since political boundaries are relevant for distribution or exclusion of rights, services, and resources. For instance, one drawback to examining the city level concerns the ability of policy at the city level to address regional issues – issues that transcend one city – e.g., sprawl or transportation infrastructure. Even if cities represent a conceptually-relevant unit, cities have fewer resources to address public health problems and their antecedents, including a declining local tax base, declining federal financial support, and limited national leadership to advance an urban agenda (Freudenberg et al., 2006). Alternately, as we have discussed, examining neighborhoods has limited policy relevance, although it may be related to administrative distribution of resources (e.g., targeting of programs for the elderly based on the population distribution of elders by neighborhood).

6. Conclusion

The goal of this chapter was to provide conceptual and practical guidance for choosing a macro-level unit that is relevant for examining how place affects population health. We aim to stimulate investigation of multiple systems, including various structural, political, physical, social, or organizational attributes of place, as potential upstream determinants of population health (Diez Roux, 2001). By articulating theoretical and empirical justifications for using alternate definitions of place, including neighborhoods, cities, counties, metropolitan areas, states, and nations, we strive to broaden the conceptualization of how place affects health in a way that may be useful for future research. We hope that research that considers the potential influence of factors at multiple and different levels of space can contribute to a more complete understanding of the macrosocial determination of the health of populations.

Endnotes

- i. For the purposes of this chapter, levels are composed of spatial units that can be observed, sampled, and analyzed (Leyland & Groenewegen, 2003).
- ii. Multilevel modeling, or hierarchical linear modeling, refers to a method whereby variables are statistically modeled at 2 or more levels, the outcome variance is partitioned into 2 or more levels, and units are often nested (e.g. individuals within neighborhoods). Multilevel modeling allows examination of group level and individual level variables on individual-level outcomes and accounts for the dependence of observations within groups (Diez Roux, 2002b).
- iii. Although Rhode Island and Connecticut are divided into counties, the counties do not employ functioning governments as defined by the Census Bureau. Certain other areas in the US may also be called counties but lack county governments. Alaska calls its county government a borough, and Louisiana a parish, and both are classified as counties for the purposes of the census statistics of US governments (U.S. Census Bureau, 2004a).

Acknowledgments. The authors would like to thank M. Maria Glymour, Steve Cummins, and the anonymous reviewer for their helpful comments. We also thank the Robert Wood Johnson Foundation for their financial support of this work via the Health and Society Scholars Program.

References

- Acevedo-Garcia, D., Lochner, K. A., Osypuk, T. L., & Subramanian, S. V. (2003). Future directions in residential segregation and health research: A multilevel approach. *American Journal of Public Health, 93*(2), 215–221.
- Altshuler, A., Morrill, W., Wolman, H., Mitchell, F., & Committee on Improving the Future of U.S. Cities Through Improved Metropolitan Area Governance (1999). *Governance and opportunity in metropolitan America*. Washington, DC: National Academy Press.
- Anderson, E., & Massey, D. S. (2001). The sociology of race in the United States. In E. Anderson & D. S. Massey (Eds.), *Problem of the century: Racial stratification in the United States*. New York: Russell Sage Foundation.
- Bailey, W. C., & Peterson, R. D. (1995). Gender inequality and violence against women. In J. Hagan & R. D. Peterson (Eds.), *Crime and inequality*. Stanford, CA: Stanford University Press.
- Ballin, S. D. (1994). Thirty years of tobacco industry domination of tobacco control efforts in the federal government. *Circulation, 89*(2), 543–544.
- Booth, C. (2001). On the city: Physical pattern and social structure. Introduction. In G. Davey Smith, D. Dorling, & M. Shaw (Eds.), *Poverty, inequality, and health in Britain 1800–2000: A reader*. Bristol, UK: The Policy Press.
- Castree, N. (2004). Differential geographies: Place, indigenous rights, and “local” resources. *Political Geography, 23*, 133–167.
- Cereseto, S., & Waitzkin, H. (1986a). Capitalism, socialism, and the physical quality of life. *International Journal of Health Services, 16*(4), 643–658.
- Cereseto, S., & Waitzkin, H. (1986b). Economic development, political-economic system, and the physical quality of life. *American Journal of Public Health, 76*(6), 661–666.
- Cummins, S., Curtis, S., Diez Roux, A. V., & Macintyre, S. (In press). Understanding and representing “place” in health research: a relational approach. *Social Science & Medicine*.
- Davey Smith, G., Dorling, D., & Shaw, M. (2001). *Poverty, inequality, and health in Britain 1800–2000: A reader*. Bristol, UK: The Policy Press
- Diez Roux, A. V. (2007). Neighborhoods and health: where are we and where do we go from here? *Revue d'Epidemiologie et de Sante Publique, 55*(1), 13–21.
- Diez Roux, A. V. (2000). Multilevel analysis in public health research. *Annual Review of Public Health, 21*, 171–192.
- Diez Roux, A. V. (2001). Investigating neighborhood and area effects on health. *American Journal of Public Health, 91*(11), 1783–1789.
- Diez Roux, A. V. (2002a). Invited commentary: Places, people, and health. *American Journal of Epidemiology, 155*(6), 516–519.
- Diez Roux, A. V. (2002b). A glossary for multilevel analysis. *Journal of Epidemiology and Community Health, 56*(8), 588–594.
- Diez Roux, A. V. (2003). Issues related to multiple levels of organization. In I. Kawachi & L. F. Berkman (Eds.), *Neighborhoods and health*. New York: Oxford University Press.
- Diez Roux, A. V. (2004a). Estimating neighborhood health effects: The challenges of causal inference in a complex world. *Social Science & Medicine, 58*(10), 1953–1960.

- Diez Roux, A. V. (2004b). The study of group-level factors in epidemiology: Rethinking variables, study designs, and analytical approaches. *Epidemiologic Reviews*, 26(1), 104–111.
- Diez Roux, A. V. (2005). Commentary: Estimating and understanding area health effects. *International Journal of Epidemiology*, 34(2), 284–285.
- Diez Roux, A. V., & Aiello, A. E. (2005). Multilevel analysis of infectious diseases. *Journal of Infectious Disease*, 191(Supplement 1), S25–S33.
- Duncan, C., Jones, K., & Moon, G. (1998). Context, composition, and heterogeneity: Using multilevel models in health research. *Social Science & Medicine*, 46(1), 97–117.
- Ellen, I. G. (2000). Is segregation bad for your health? The case of low birth weight. In W. G. Gale & J. R. Pack (Eds.), *Brookings-Wharton papers on urban affairs 2000*. Washington, DC: Brookings Institution Press.
- Ellen, I. G., Mijanovich, T., & Dillman, K. N. (2001). Neighborhood effects on health: Exploring the links and assessing the evidence. *Journal of Urban Affairs*, 23(3–4), 391–408.
- Engels, F. (1845). The condition of the working class in England. The great towns. In G. Davey Smith, D. Dorling, & M. Shaw (Eds.), *Poverty, inequality, and health in Britain 1800–2000: A reader*. Bristol, UK: The Policy Press.
- FANS, L. A. (2006). The Los Angeles Family and Neighborhood Survey: Overview of L.A. FANS Survey Design. RAND. (August 2, 2006); <http://www.lasurvey.rand.org/thesurvey.htm>.
- Farr, W. (2001). Vital statistics: A memorial volume of selections from the reports and writings of William Farr. In G. Davey Smith, D. Dorling, & M. Shaw (Eds.), *Poverty, inequality, and health in Britain 1800–2000: A reader*. Bristol, UK: The Policy Press.
- Franco, A., Alvarez-Dardet, C., & Ruiz, M. (2004). Effect of democracy on health: Ecological study. *British Medical Journal*, 329(7480), 1421–1423.
- Frank, L. D., & Engelke, P. O. (2001). The built environment and human activity patterns: Exploring the impacts of urban form on public health. *Journal of Planning Literature*, 16(2), 202–218.
- Frank, L. D., & Engelke, P. (2005). Multiple impacts of the built environment on public health: Walkable places and the exposure to air pollution. *International Regional Science Review*, 28(2), 193–216.
- Freudenberg, N. (2005). Public health advocacy to change corporate practices: Implications for health education practice and research. *Health Education and Behavior*, 32(3), 298–319.
- Freudenberg, N., Galea, S., & Vlahov, D. (2006). Changing living conditions; changing health: US cities since World War II. In N. Freudenberg, S. Galea, & D. Vlahov (Eds.), *Cities and the health of the public*. Nashville, TN: Vanderbilt.
- Frey, W. H. (2001). *Melting pot suburbs: A census 2000 study of suburban diversity*. *Census 2000 Series*. Washington, DC: The Brookings Institution Center on Urban & Metropolitan Policy.
- Galea, S., Freudenberg, N., & Vlahov, D. (2006). A framework for the study of urban health. In N. Freudenberg, S. Galea, & D. Vlahov (Eds.), *Cities and the health of the public*. Nashville, TN: Vanderbilt.
- Gilmore, A., & McKee, M. (2004). Tobacco and transition: An overview of industry investments, impact and influence in the former Soviet Union. *Tobacco Control*, 13, 136–142.
- Gostin, L. O., Koplan, J. P., & Grad, F. P. (2003). The law and the public's health: The foundation. In R. A. Goodman, M. A. Rothstein, R. E. Hoffman, W. Lopez, & G. W. Matthews (Eds.), *Law in public health practice*. New York: Oxford University Press.

- Greenland, S. (2002). A review of multilevel theory for ecologic analyses. *Statistics in Medicine*, *21*, 389–395.
- Grosse, S. D., Waitzman, N. J., Romano, P. S., & Mulinare, J. (2005). Reevaluating the benefits of folic acid fortification in the United States: Economic analysis, regulation, and public health. *American Journal of Public Health*, *95*(11), 1917–1922.
- Grossman, M. (1989). Health benefits of increases in alcohol and cigarette taxes. *British Journal of Addiction*, *84*, 1193–1204.
- Hahn, R., Bilukha, O., Crosby, A., Fullilove, M., Liberman, A., Moscicki, E., et al. (2005). Firearms laws and the reduction of violence: A systematic review. *American Journal of Preventive Medicine*, *28*(2 Supplement 1), 40–71.
- Harvey, D. (1973). *Social justice and the city*. London: Edward Arnold Publishers Ltd.
- Harvey, D. (2006). Space as a keyword. In N. Castree & D. Gregory (Eds.), *David Harvey: A critical reader*. Malden, MA: Blackwell Publishing.
- Heywood, D. I. (1998). *Introduction to geographical information systems*. New York: Addison Wesley Longman
- Jargowsky, P. A. (1997). *Poverty and place: Ghettos, barrios, and the American city*. New York: Russell Sage Foundation
- Jargowsky, P. A. (2003). Stunning progress, hidden problems: The dramatic decline of concentrated poverty in the 1990s. In Center on Urban and Metropolitan Policy (Ed.), Washington, DC: Brookings Institution.
- Katz, B., & Lang, R. E. (2003). Introduction. In B. Katz & R. E. Lang (Eds.), *Redefining urban & suburban America: Evidence from census 2000*. Washington, DC: Brookings Institution
- Kawachi, I., Wilkinson, R. G., & Kennedy, B. P. (1999). Introduction. In I. Kawachi, B. P. Kennedy, & R. G. Wilkinson (Eds.), *The society and population health reader: Income inequality and health*. New York: The New Press.
- Kawachi, I., & Berkman, L. F. (2003a). *Neighborhoods and health*. New York, NY: Oxford University Press
- Kawachi, I., & Berkman, L. F. (2003b). Introduction. In I. Kawachi & L. F. Berkman (Eds.), *Neighborhoods and health*. New York: Oxford University Press.
- Krieger, N. (2001). Historical roots of social epidemiology: Socioeconomic gradients in health and contextual analysis (letter to the editor). *International Journal of Epidemiology*, *30*, 899–900.
- Krieger, N., Williams, D. R., & Moss, N. E. (1997). Measuring social class in US public health research: Concepts, methodologies, and guidelines. *Annual Review of Public Health*, *18*(1), 341–378.
- Land, K. C., Cantor, D., & Russell, S. T. (1995). Unemployment and crime rate fluctuations. In J. Hagan & R.D. Peterson (Eds.), *Crime and inequality*. Stanford, CA: Stanford University Press.
- Leyland, A. H., & Groenewegen, P. P. (2003). Multilevel modeling and public health policy. *Scandinavian Journal of Public Health*, *31*, 267–274.
- Link, B. G., & Phelan, J. C. (1996). Editorial: Understanding sociodemographic differences in health — The role of fundamental social causes. *American Journal of Public Health*, *86*(4), 471–473.
- Macintyre, S., & Ellaway, A. (2003). Neighborhoods and health: An overview. In I. Kawachi & L.F. Berkman (Eds.), *Neighborhoods and health*. New York: Oxford University Press.
- Macintyre, S., Ellaway, A., & Cummins, S. (2002). Place effects on health: How can we conceptualise, operationalise, and measure them? *Social Science & Medicine*, *55*, 125–139.

- Martyn, C. (2004). Politics as a determinant of health. *British Medical Journal*, 329(7480), 1423–1424.
- Massey, D. (1991). The political place of locality studies. *Environment and Planning Series A*, 23, 267–281.
- Massey, D. S., & Denton, N. A. (1993). *American apartheid: Segregation and the making of the underclass*. Cambridge, MA: Harvard University Press
- McMichael, A. J. (1999). Prisoners of the proximate: Loosening the constraints on epidemiology in an age of change. *American Journal of Epidemiology*, 149(10), 887–897.
- Melvin, P. M. (1985). Changing contexts: Neighborhood definition and urban organization. *American Quarterly*, 37(3), 357–367.
- Morenoff, J. D., Diez Roux, A. V., Osypuk, T., & Hansen, B. (2006). *Residential environments and obesity: What can we learn about policy interventions from observational studies? National Poverty Center Working Paper*. Ann Arbor, MI: University of Michigan.
- Morgenstern, H. (1985). Socioeconomic factors: Concepts, measurement, and health effects. In A. M. Ostfeld & E. D. Eaker (Eds.), *Measuring psychosocial variables in epidemiologic studies of cardiovascular disease: Proceedings of a workshop*. Washington, DC: National Institutes of Health.
- National Association of Counties (2003). A brief overview of county government. Washington DC: NACO.
- National Association of Counties (2006a). About counties. July 8, 2006; http://www.naco.org/Template.cfm?Section = About_Counties.
- National Association of Counties (2006b). An overview of county government. (July 8, 2006); http://www.naco.org/Content/NavigationMenu/About_Counties/County_Government/Default271.htm.
- National Research Council (2002). *Equality of opportunity and the importance of place: Summary of a workshop*. Washington DC: National Academy Press
- New York City Department of City Planning (2006). New York: A city of neighborhoods. New York (2006); <http://www.nyc.gov/html/dcp/html/neighbor/neighbor.shtml>.
- Oakes, J. M. (2006). Commentary: Advancing neighbourhood-effects research—selection, inferential support, and structural confounding. *International Journal of Epidemiology*, 35, 643–647.
- Oakes, J. M., & Kaufman, J. S. (2006). Introduction: Advancing methods in social epidemiology. In J. M. Oakes & J. S. Kaufman (Eds.), *Methods in social epidemiology*. San Francisco: Jossey Bass.
- O’Campo, P. (2003). Invited commentary: Advancing theory and methods for multilevel models of residential neighborhoods and health. *American Journal of Epidemiology*, 157(1), 9–13.
- O’Campo, P., & Kogan, M. D. (2005). Multilevel modeling in perinatal research. (February 2, 2005); <http://www.uic.edu/sph/cade/mchept/meetings/feb2005/index.htm>.
- O’Campo, P., & Schempf, A. (2005). Racial inequalities in preterm delivery: Issues in the measurement of psychosocial constructs. *American Journal of Obstetrics and Gynecology*, 192, S56–63.
- O’Connor, A. (2001). Understanding inequality in the late twentieth-century metropolis: New perspectives on the enduring racial divide. In A. O’Connor, C. Tilly, & L. D. Bobo (Eds.), *Urban inequality: Evidence from four cities*. New York: Russell Sage Foundation.

- Office of Management and Budget, & Spotila, J. T. (2000). Part IX, Office of Management and Budget, standards for defining metropolitan and micropolitan statistical areas; notice. *Federal Register*, 65(249), 82228–82238.
- Orfield, M. (2002). *American metropolitics: The new suburban reality*. Washington, DC: Brookings Institution Press
- Osypuk, T. L. (2005). *Demographic and place dimensions of racial/ethnic health disparities. Unpublished dissertation*. Boston, MA: Harvard University School of Public Health, Department of Society, Human Development, and Health.
- Osypuk, T. L. (2006). Inequality in exposure to fast food outlets by race/ethnicity & social class across 325 US metropolitan areas. *American Journal of Epidemiology*, 163(11 Supplement), S144.
- Osypuk, T. L., Kawachi, I., Subramanian, S. V., & Acevedo-Garcia, D. (2006). Are state patterns of smoking different for different racial/ethnic groups? An application of multi-level analysis. *Public Health Reports*, 121(5), 563–577.
- Plane, D. A. (2004). Population distribution: Geographic areas. In J. S. Siegel & D. A. Swanson (Eds.), *The methods and materials of demography*. 2nd edition. San Diego, CA: Elsevier Academic Press.
- Raudenbush, S. W., & Sampson, R. J. (1999). Ecometrics: Toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. *Sociological Methodology*, 29, 1–41.
- Rose, G. (1992). *The strategy of preventive medicine*. New York, NY: Oxford University Press
- Saelens, B. E., Sallis, J. F., & Frank, L. D. (2003). Environmental correlates of walking and cycling: Findings from the transportation, urban design, and planning literatures. *Annals of Behavioral Medicine*, 25(2), 80–91.
- Sallis, J. F., Frank, L. D., Saelens, B. E., & Kraft, M. K. (2004). Active transportation and physical activity: Opportunities for collaboration on transportation and public health research. *Transportation Research Part A: Policy and Practice*, 38(4), 249–268.
- Sampson, R. J., & Wilson, W. J. (1995). Race, crime, and urban inequality. In J. Hagan & R.D. Peterson (Eds.), *Crime and inequality*. Stanford, CA: Stanford University Press.
- Schwartz, S., & Carpenter, K. M. (1999). The right answer for the wrong question: Consequences of type III error for public health research. *American Journal of Public Health*, 89(8), 1175–1180.
- Schwartz, S., & Diez Roux, A. V. (2001). Commentary: Causes of incidence and causes of cases — a Durkheimian perspective on Rose. *International Journal of Epidemiology*, 30(3), 435–439.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. New York: Houghton Mifflin.
- Subramanian, S. V., Acevedo-Garcia, D., & Osypuk, T. L. (2005). Racial residential segregation and geographic heterogeneity in black/white disparity in poor self-rated health in the U.S.: A multilevel statistical analysis. *Social Science & Medicine*, 60(8), 1667–1679.
- U.S. Census Bureau. (2000). Census 2000 ZCTAs ZIP code tabulation areas technical documentation. Washington, DC: U.S. Census Bureau.
- U.S. Census Bureau. (2001a). ZIP code tabulation areas (ZCTAs). U.S. Census Bureau. Washington, DC (April 13, 2001); <http://www.census.gov/geo/ZCTA/zcta.html>.
- U.S. Census Bureau. (2001b). Geographic changes for Census 2000 + glossary. Washington, DC (September 20, 2001); <http://www.census.gov/geo/www/tiger/glossary.html>.

- U.S. Census Bureau. (2001c). Ranking tables for incorporated places of 100,000 or more: Population in 2000 and population change from 1990 to 2000 (PHC-T-5). Table 2: incorporated places of 100,000 or more ranked by population: 2000. Washington, DC (July 31, 2002); <http://www.census.gov/population/www/cen2000/phc-t5.html>.
- U.S. Census Bureau. (2002). U.S. Gazetteer: 2000 and 1990: 1990 Census gazetteer files. Washington, DC (May 3, 2005); <http://www.census.gov/geo/www/gazetteer/gazette.html>.
- U.S. Census Bureau. (2004a). 2002 Census of Governments, Volume 3, Number 1, employment of major local governments GC02(3)-1. Washington, DC: U.S. Government Printing Office.
- U.S. Census Bureau (2004b). Federal, state, and local governments, state and local government finances, 2002 census of governments: Viewable data: State and local summary tables by level of government and by type of government. U.S. Census Bureau. Washington, DC (October 14, 2004); <http://www.census.gov/govs/www/estimate02.html>.
- U.S. Census Bureau (2005a). About metropolitan and micropolitan statistical areas. U.S. Census Bureau. Washington, DC (June 7, 2005); <http://www.census.gov/population/www/estimates/aboutmetro.html>.
- U.S. Census Bureau (2005b). Answers to frequently asked questions about census bureau geography, maps and mapping engines: ZIP code information. Washington, DC (August 19, 2005); <http://www.census.gov/geo/www/tiger/tigermap.html#ZIP>.
- U.S. Census Bureau (2006a). ZIP Code Statistics. July 10, 2006. <http://www.census.gov/epcd/www/zipstats.html>.
- U.S. Census Bureau (2006b). TIGER Frequently Asked Questions. July 10, 2006. <http://www.census.gov/geo/www/tiger/faq-index.html>.
- U.S. Census Bureau (2006c). American factfinder. Urbanized area. U.S. Census Bureau. October 8, 2006.
- U.S. Census Bureau (no year). Appendix A: Census 2000 geographic terms and concepts.
- U.S. Postal Service (2003). The United States Postal Service: An American history 1775 – 2002. Publication 100. U.S. Postal Service. (September, 2003); <http://www.usps.com/cpim/ftp/pubs/pub100/pub100.htm#thepostalrole>.
- U.S. Postal Service (2006a). Mr ZIP: The nation's original "digital" icon. July 10, 2006. <http://www.usps.com/postalhistory/mrzip.htm>.
- U.S. Postal Service (2006b). Sources of historical information on post offices, postal employees, mail routes, and mail contractors: Publication 119. U.S. Postal Service. (October, 2006); <http://www.usps.com/cpim/ftp/pubs/pub119/welcome.htm>.
- Wilkinson, R. G. (1992a). National mortality rates: The impact of inequality? *American Journal of Public Health*, 82(8), 1082–1084.
- Wilkinson, R. G. (1992b). Income distribution and life expectancy. *British Medical Journal*, 304(6820), 165–168.
- Wilkinson, R. G., & Pickett, K. E. (2006). Income inequality and population health: A review and explanation of the evidence. *Social Science & Medicine*, 62, 1768–1784.
- Wilson, W. J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago: University of Chicago Press
- Yang, T. C. (2005). Modifiable areal unit problem. GIS resource document 05–65 (GIS_RD_05–65). (February, 2005); [http://www.pop.psu.edu/gia-core/pdfs/gis_rd_05–65.pdf](http://www.pop.psu.edu/gia-core/pdfs/gis_rd_05-65.pdf).