MODELING PLACE VULNERABILITY OF HIV/AIDS IN TEXAS

Adam F. Harold

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2011

APPROVED:

David Williamson, Major Professor
Joseph R. Oppong, Minor Professor
Tianji Cai, Committee Member
Ami Moore, Committee Member
Sandra Terrell, Interim Chair of the Department of Sociology
Thomas Evenson, Dean of the College of Public Affairs and Community Service
James D. Meernik, Acting Dean of the Toulouse Graduate School
This study provides a measurable model of the concept of place vulnerability for HIV/AIDS that incorporates both community and structural level effects using data provided at the ZIP code level from the Texas Department of State Health Services. Sociological literature on the effects of place on health has been growing but falls short of providing an operational definition of the effects of place on health. This dissertation looks to the literature in medical/health geography to supplement sociology’s understanding of the effects of place on health, to the end of providing a measurable model. Prior research that has recognized the complexity of the effects of place still have forced data into one scale and emphasized individual-level outcomes. A multilevel model allows for keeping the associated spatial unit data, without aggregating or parsing it out for convenience of model fit. The place vulnerability model proposed examines how exposure, capacity and potentiality variables all influence an area’s HIV/AIDS count. To capture the effects of place vulnerability at multiple levels, this dissertation research uses a multilevel zero-inflated poisson (MLZIP) model to examine how factors measured at the ZIP code and county both affect HIV/AIDS counts per ZIP code as an outcome. Furthermore, empirical Bayes estimates are mapped to display how well the model fits across the state of Texas. Limitations of this research include the need to incorporate time, more specific predictors, and individual level factors. The methodology developed permits a more thorough understanding of place effects on the spatial variation of HIV/AIDS.
Copyright 2011

by

Adam F. Harold
ACKNOWLEDGEMENTS

I would never have been able to finish my dissertation without the guidance of my committee members, the flexibility of my colleagues, and the support from my wife and children. I would like to express my deepest gratitude to my co-chairs, Dr. David Williamson and Dr. Joseph R. Oppong, for their excellent guidance and suggestions during this research. I would like to thank Dr. Ami Moore for her input and expertise and Dr. Tianji Cai for his constant support through the statistical analysis. I would also like to extend thanks to Dr. Chetan Tiwari and the University of North Texas Health & GIS Research Group, as well as the Texas Department of State Health Services, for their critiques as I went through this process. Special thanks goes to my colleagues at UNT-International for their flexibility, patience and willingness to pick up for me while I was working through this degree. Finally I would like to thank my wife, Jill, and my children, Caleb, Toby, and Wesley for standing by me with patience and support.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................... iii

LIST OF TABLES ......................................................................................................................... vi

LIST OF FIGURES ...................................................................................................................... vii

Chapters

I. INTRODUCTION ...................................................................................................1
   Sociology and the Study of Place: Should Sociologists Study Place? ......5
   A Space for Place in Sociology ...................................................................9

II. THE CONCEPTUALIZATION OF PLACE VULNERABILITY .......................13
   The Case for Place in the Sociology of Health: Can Medical/Health
   Sociologists Study Place? .......................................................................13
   Sociology, Health and Place ...................................................................14
   Place Effects and HIV/AIDS: Can Sociologists Study Place Effects on
   HIV/AIDS? ..........................................................................................28
   Towards a Sociological Conceptualization of Place Vulnerability ..........34
   The Complexity of Place .........................................................................41

III. THE OPERATIONALIZATION OF PLACE VULNERABILITY ....................44
   Towards an Operationalization of Place Vulnerability .........................45
   Place and HIV ..........................................................................................52
   Multilevel Models ....................................................................................55
   Expanding Watts and Bohle’s Vulnerability Model ..............................58

IV. METHODS ............................................................................................................65
   Design ...........................................................................................................66
   Specification ...............................................................................................71
   Statistical Method – Model Selection .....................................................90

V. RESULTS AND CONCLUSION ..........................................................................95
   Results .................................................................................................................95
   More than Population Containers: Added Value of Community and
   Structural Effects ..................................................................................98
LIST OF TABLES

1. Selected Theories and their Spatial Implications ............................................................... 35
3. Chi-Square Goodness of Fit for HIV Counts between ZCTAs ............................................ 76
4. Summary of Variables Used in this Analysis ..................................................................... 88
5. Bivariate Correlations of ZCTA Level Predictors and Dependent Variable ....................... 89
6. Bivariate Correlations of County Level Predictors and Dependent Variable Aggregated to County (w/o TDCJ unit cases) ................................................................................. 90
7. Compared Estimates and Model Fit from Poisson, ZIP, ML Poisson and MLZIP .......... 96
8. MLZIP Estimates Adding in Capacity, Exposure and Potentiality Variables ................. 99
9. MLZIP Results .................................................................................................................. 101
10. Five ZCTAs with Highest Standard Errors of Prediction ................................................. 115
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>The space of vulnerability</td>
</tr>
<tr>
<td>2.</td>
<td>Social network approach to HIV transmission (high risk)</td>
</tr>
<tr>
<td>3.</td>
<td>Social network approach to HIV transmission (low risk)</td>
</tr>
<tr>
<td>4.</td>
<td>Geographic approach to HIV transmission</td>
</tr>
<tr>
<td>5.</td>
<td>A typology of choice and area</td>
</tr>
<tr>
<td>6.</td>
<td>The social epidemiology of HIV/AIDS</td>
</tr>
<tr>
<td>7.</td>
<td>A typology of studies for modeling health risk factors</td>
</tr>
<tr>
<td>8.</td>
<td>Expanded Watts and Bohle (1993) vulnerability model to incorporate multiple scales and time</td>
</tr>
<tr>
<td>9.</td>
<td>Multilevel data structure for individuals within areas</td>
</tr>
<tr>
<td>10.</td>
<td>Multilevel data structure for individuals nested within two nested areas</td>
</tr>
<tr>
<td>11.</td>
<td>Multilevel data structure for area nested within another area over time</td>
</tr>
<tr>
<td>12.</td>
<td>Multilevel data structure for area nested within another area</td>
</tr>
<tr>
<td>14.</td>
<td>Layout of data used in this multilevel model</td>
</tr>
<tr>
<td>15.</td>
<td>Standard deviation of HIV/AIDS counts per ZCTA in Texas</td>
</tr>
<tr>
<td>16.</td>
<td>ZCTAs where observed HIV count fall within predicted confidence interval</td>
</tr>
<tr>
<td>17.</td>
<td>ZCTAs where observed HIV count fall within predicted confidence interval within DFW area</td>
</tr>
<tr>
<td>18.</td>
<td>Differences between predicted and observed HIV/AIDS in DFW area</td>
</tr>
<tr>
<td>19.</td>
<td>Differences between predicted and observed HIV/AIDS in Houston area</td>
</tr>
<tr>
<td>20.</td>
<td>ZCTAs where observed HIV count fall within predicted confidence interval</td>
</tr>
<tr>
<td>21.</td>
<td>Standard error of prediction of HIV/AIDS in Texas</td>
</tr>
<tr>
<td>22.</td>
<td>Standard error of prediction of HIV/AIDS in DFW area</td>
</tr>
<tr>
<td></td>
<td>Title</td>
</tr>
<tr>
<td>---</td>
<td>-----------------------------------------------------------------------</td>
</tr>
<tr>
<td>23.</td>
<td>Standard error of prediction of HIV/AIDS in the Houston area</td>
</tr>
</tbody>
</table>
In July 2010, the White House issued a new National HIV/AIDS strategy for the United States (Policy, 2010). One of the overwhelming focuses of this document was that while anyone can be infected with HIV/AIDS, some people, and even some areas, are at greater risks than others. The action plan consists of statements that focus on the disproportionate representation of high risk areas, such as “By focusing our efforts in communities where HIV is concentrated, we can have the biggest impact in lowering all communities’ collective risk of acquiring HIV” (p. viii).

The sociological explanations of health and illness have been extensive from the health belief model, to human ecology, to social capital, to social cognitive theory, and numerous other theories and models (for a review, see (Glanz, Lewis, & Rimer, 1997)). Sociology has also provided major journals (Social Science & Medicine, Sociology of Health & Illness, Journal of Health and Social Behavior) and countless important texts on the subject of health. Moreover, sociology provides valuable explanations for health inequalities, rightly extending stratification and susceptibility models from socioeconomic status to health status. However, until recently, the extension of this research to place effects of health has been underdeveloped, perhaps due to the fear of environmental determinism, or perhaps because it has not been framed properly (Gieryn, 2000). Sampson, Morenoff and Gannon-Rowley (2002) show a dramatic spike in journal articles concerning “neighborhood effects” in the latter 1990s to the tune of over 100 papers per year. The intent of this type of research is not to ignore individual risk factors, but to examine whether or not the risk of contracting a disease varies by area and why. Even still,
sociologists have obviously recognized the need to go beyond research that merely focuses on individual effects of health to theorizing the complexity of area effects.

Of interest for research here is the question of health inequality, or more properly stated, health inequity – the distinction being the former describes measureable differences (e.g. Group X has a higher disease incidence than Group Y) while the latter refers to inequalities that are unfair often as the result of some sort of social injustice. The core of this distinction is based on one’s view of justice and society (Kawachi, Subramanian, & Almeida-Filho, 2002). Sociological explanations can and have been useful in defining which individuals or populations are susceptible to disease. However, a sociological approach may also prove useful in describing what makes places susceptible to a disease. Narrowing this down, one of the areas where both of these perspectives can provide help is in the framework of vulnerability. Kearns and Moon (2002) categorize attempts of general “vulnerability” research into three types: localized studies, landscape studies, and multilevel models. Most of this type of research in vulnerability has been done by geographers or in disaster research (Cutter, Mitchell, & Scott, 2000) with a few notable exceptions (Macintyre, Maciver, & Sooman, 1993; Watts & Bohle, 1993). Even then, most vulnerability research has used either a localized or landscape study, with again very few attempts to combine the two in a multilevel model. Little, if anything, has attempted to examine health place vulnerability using a multilevel model. Still, the direction of health research is moving this way, making this research timely. Since health research tends to be interdisciplinary, one can find studies of the place effects on health in epidemiology, geography, sociology and public health journals. However this research often tends to be either method-heavy with little room given to sociological explanation or filled with overly burdensome conceptualizing leaving little usefulness for operationalization. Therefore a well-grounded
sociological explanation for the methodology chosen in this paper is also timely.

The purpose of this paper is to provide one of the first measureable models using a vulnerability framework specific to infectious disease (specifically HIV/AIDS). The emergence of place effects for understanding health is one of most forefront topics in health research today. However, research in this area lacks in two critical areas. First, little has been done methodologically to combine “composition” measures of a place (i.e., characteristics of people that make up an area) with “context” measures of a place (i.e., characteristics of a place such as political economy and other structural factors). In other words, little has been done to justify “the spatial scales at which human processes occur” (Abler, Adams, & Gould, 1971). Second, very few studies have theoretically justified place beyond a unit of analysis to which all variables are aggregated. Therefore this research examines the use of social theory as a means to develop a model that allows place effects to be measured at the proper scale. The outcomes from this kind of model can then be applied to either an understanding of what makes one place more vulnerable to an infectious disease than another, or whether behavior in one area is more risky than that same behavior in another place.

Sociological explanations of health very often use behavior as a starting point. This is best evidenced by the health belief model which is one of the most widely used conceptual frameworks in the sociology of health (Strecher & Rosenstock, 1997). The model is designed to get at individual risk factors for disease with sociological explanations (such as perceived susceptibility). When analyzing a disease, it is often obvious that behavior plays a vital part, and even the most vital component when dealing with a disease such as HIV/AIDS. While studies may find that behavior that is susceptible to contracting HIV/AIDS does not vary by groups (e.g. adolescent promiscuity produces a similar risk as middle-age promiscuity), these same studies
ignore whether or not the risky behavior or outcome of that behavior will vary dependent upon the area. In other words, a risky behavior in an area with high HIV/AIDS rate is, in fact, more risky than that same behavior in a low HIV/AIDS rate area. The National HIV/AIDS strategy agrees with this statement. It states, “Research has shown that the higher risk for these groups is associated with the sheer number of HIV-positive persons in the communities where they live. As a result, any instance of risk behavior carries a far greater likelihood of infection than other communities with fewer cases of HIV. Thus, unprotected sex even once for individuals in some communities carries a greater risk of HIV infection than for individuals in other communities” (p. 12).

It is a well-studied fact that disease exhibits spatial variation, and that a disease such as HIV/AIDS varies spatially in prevalence and severity. Research has also confirmed that where one lives significantly affects one’s health. While place is crucial to understand health, it should not just be treated as a single variable to be controlled for, as a residual category, or a sum of individual or structural characteristics, but as a dynamic mechanism that can be used to test hypotheses of specific chains of causation that might link place with health (Macintyre, Maciver, & Sooman, 1993).

What occurs in the following is a review of the study of place in sociology followed by the recent calls for a more thorough investigation of place in sociology. Chapter II provides an overview of the literature concerning various sociologies of health and how place has or can be incorporated, to the end that a vulnerability framework yields the best option for a measureable outcome. Chapter III provides an overview of methodologies used in the study of place effects on health with an emphasis on multilevel modeling, to the end that a vulnerability framework combined with an expanded view of place in a multilevel model yields the best option for a
measureable outcome. Chapter IV provides an explanation of variables and the study population while defining and justifying the areas of study and the model selected. Finally, Chapter V discusses the results and analysis of each model and explores the significance of the findings with recommendations for future research.

Sociology and the Study of Place: Should Sociologists Study Place?

Space may not have been the centerpiece of early sociologists, but it was not completely ignored either. Marx understood land ownership and the state with geographic boundaries, even if they did not necessarily play a role in either’s development. Tonnies distinguished between urban and rural, region, national and international, home, village and town, all which clearly can have geographic contexts. Weber formed his Protestant Ethic argument around the dichotomy between Catholic and Protestant Europe before examining the religions of two different countries—India and China. His younger brother Alfred in his *Theory of the Location of Industries* (Weber, 1909) takes into account spatial factors for optimal locations of manufacturing plants. However Simmel is perhaps the first sociologist to attempt to move space into a valuable area of study for sociologists in *The Sociology of Space* and *On the Spatial Projections of Social Forms* (Simmel, Frisby, & Featherstone, 1997). These works have often been used by urban sociologists in developing ideas such as sociation, metropolitan areas, and leisure spaces (Fearon, 2001). Moving through the 20th century finds sociology giving more importance to the concept of space, and some examples of this work are given below.

*Chicago School*

Chicago school sociologists were some of the first to include space in their methodology.
Robert Park (1925) claims that:

Physical geography, natural advantages and disadvantages...determine in advance the general outlines of the urban plan. As the city increases in population, the subtler influences of sympathy, rivalry, and economic necessity tend to control the distribution of population...In the course of time every section and quarter of the city takes on something of the character and qualities of its inhabitants. Each separate part of the city is inevitably stained with the peculiar sentiments of its population. The effect of this is to convert what was at first a mere geographical expression into a neighborhood, that is to say, a locality with sentiments, traditions, and a history of its own. (p. 5-6)

In a following work, Park recognizes that the general outlook on life in different areas is astonishing—noticing the dramatic differences in demographics such as sex and age groups, and regions with excessive divorce, suicide or juvenile delinquency and regions where there is almost none (Park R. E., 1926). Park, Burgess and MacKenzie (1925) developed “concentric zone theory” to model the spatial organization of urban areas.

Drawing from Park, Clifford Shaw (1929) examines one of the first sociological studies of area effects on behaviors in his *Delinquency Areas*. His focus is on juvenile delinquency in the city of Chicago mainly during the 1920s. He states that as a city grows the processes of competition devolves natural segregation patterns along economic, vocational and cultural lines. Over time each group creates a character in the area it occupies that makes it distinct from other communities. Social norms and attitudes create institutions that become part of the community. If the community is disorganized and weak, Shaw says that the institutions will break down and behavior will not be controlled by conventional standards, whereas if the community standards are strong, even if they are contrary to larger society, behaviors within the community will follow accordingly.

It is important to interject here that these were some of the first applications of social disorganization theory which still remains foundational among sociologists studying space today. Using small ecological units such as neighborhoods or census tracts, social disorganization...
theory uses the absence of organization among people to explain crime, illness, delinquency and other social problems. Examples from the early Chicago school include Sutherland (1939) who used social disorganization (and later differential social organization) as a means to explain crime variations between places and cultures and Faris (1955) who used the theory to explain other general social problems such as suicide, violence and mental illness. Many of the concepts in the next chapter (such as social networks, collective efficacy) find their roots in this early Chicago school research.

Shaw goes on to take a sociological approach to the relationship of behavior to the social and cultural setting it appears in. He begins this approach with the caveat that it should be clear that his study does not attempt to show that delinquency is caused by the simple fact of location. Rather, that delinquency tends to occur in a characteristic type of area. Shaw finds marked variations of between area rates of truancy, juvenile delinquents and adult criminals even after considering economic and political factors. Perhaps most interestingly, a sampling of the cases revealed that “delinquency found in areas of low rates revealed the significant fact that many of the delinquents living in these areas had previously become delinquent while living in areas of high delinquency rates” (p. 22). He goes on to conclude that “disintegrative forces” of business and industry invasion creates a situation where a community ceases to function as a means of social control. “Traditional norms and standards of the conventional community weaken and disappear. Resistance on the part of the community to delinquent and criminal behavior is low, and such behavior is tolerated and may even become accepted and approved” (p. 24). Moreover, people who migrate into these deteriorated areas come from cultural and social backgrounds which create conflict between the new and their former cultural and social controls. The result is the breakdown of these controls and little or no forces at work to re-establish a
conventional order which makes for continued social disorganization. Finally, Shaw concludes with a couple important statements regarding delinquency areas.

It has been quite common in discussions of delinquency to attribute causal significance to such conditions as poor housing, overcrowding, low living standards, low educational standards, and so on. But these conditions themselves probably reflect a type of community life. By treating them one treats only symptoms of a more basic process. Even the disorganized family and the delinquent gang, which are often thought of as the main factors in delinquency, probably reflect community situations...In (the) state of social disorganization, community resistance is low. Delinquency and criminal patterns arise and are transmitted socially just as any other cultural and social pattern is transmitted. In time, these delinquent patterns may become dominant and shape the attitudes and behavior of persons living in the area. Thus the section becomes an area of delinquency. (p. 189)

Space was a topic sociologists and social scientists continued to touch on throughout the 20th century. While many examples could be cited to establish that this concept was not neglected, perhaps the most influential work to emphasize space in sociology was by Henri Lefebvre. In his *The Production of Space* (1974), Lefebvre posits that first, social space is a social product and second, that this seemingly obvious statement needs careful investigation. From his Marxist viewpoint, he sees the social production of space as a tool for control, domination and power. However the importance here is that space is not merely abstract nor perceived, but real. Furthermore every society produces its own distinct space. He then goes on to claim that this mode of production is through its containment of two sets of relations—the social relations of reproduction and the relations of production. The product is one that affects spatial practices and perceptions. Though the social relations of reproduction (the “bio-physiological relations between the sexes and between age groups, along with the specific organization of the family”) and the relations of production (“i.e. the division of labor and its organization in the form of hierarchical social functions”) are intertwined, it is important to re-emphasize that each social space is distinctly produced and therefore produces distinct spatial
practices and perceptions. This idea is visited again in Chapter II.

A Space for Place in Sociology

More recently, some sociologists have made it a priority to give space a place in sociology. Gieryn (2000) states that place remains as a fundamental element of social life and historical change and because of this, sociologists measure it even if it is not framed as such. He gives several reasons for the invisibility given to place in sociology including fear of environmental determinism taking away the explanatory power of social and cultural variables, preferring to leave it to geographers to analyze, or worrying that particular place characteristics compromise sociology’s ambition to generalize and abstract. He goes on to carefully define place as not merely “space” or setting but rather a location with material form, meaning and value. It is not just a “context for something else that becomes the focus of sociological attention, nor is it a proxy for demographic, structural, economic or behavior variables” (p. 466). Rather it is “an agentic player in the game—a force with detectable and independent effects on social life.” Perhaps most important for this study, Gieryn recognizes that place sensitive studies must include information about relative location—in other words, not just merely a census tract but the “relative location of that census tract within a metropolitan area” (p. 466).

The important question that Gieryn poses in this article is how does this location with form and meaning intersect with social practices, norms, values, structures, inequality, etc.? His answer delves into a deep discussion with how places are made that must be reviewed here before moving forward.

Places are made, Gieryn states in his review of the primarily sociological literature either by forces of power and wealth, by “professional practices of place experts,” or by perceptions of
ordinary people. Political economists and urban ecologists’ explanations that dominate the first of these often are haunted by structural determinism that leaves out the agency and contingency of place-making. In other words, Gieryn rightly acknowledges that places are not made from faceless forces or natural succession or capitalist accumulation. By “place-professionals” Gieryn means those involved in the design of built places (architects, planners, city officials, etc.). Finally places are personal—derived from ordinary people’s symbolic constructions (sense, perceptions and expectations) of a place.

So if this is what place is, what, Gieryn asks, does place do? In other words does it matter for social life and historical change? Through a variety of literature, places have shown to sustain difference and hierarchy. For example this can occur through land use zoning laws, extending or denying life chances in ethnic enclaves. Places have shown to have “power sui generis, all apart from powerful people or organizations who occupy them: the capacity to dominate and control people or things comes through the geographic location, built form, and symbolic meanings of a place” (p. 475). Perhaps the most persisting sociological contribution to place is the study of proximity, interaction and community. To this Gieryn states,

Whether or not community results from the gathering up of people into proximate face-to-face interactions depends on—sociologists routinely say—on their number, their differentiations along lines of class, race, ethnicity, taste or lifestyle, and the cultural beliefs they share. But is there a ‘place effect’ as well, in which the tight coupling of geography, built-from, and subjective typological understanding mediates the effects of size, demographic patterns, and values on the possibility or achievement of community? Enough studies suggest that the design and serial constructions of places is at the same time the execution of community. (p. 477)

Gieryn goes on to review the plethora of literature that has explored places as normative landscapes. Many studies have looked at deviant behavior, crime, social control, and identity and the way that place shapes behavior. His concluding charge, one that is partly taken up by this study, is to start using a place-sensitive sociology as a way to do sociology in a different key.
A handful of sociologists have taken up and continued this charge laid out by Gieryn. One of the primary areas where space can be included but has until recently been neglected is in the area of inequality. Lobao, Hooks and Tickamyer (2007) summarize the need to include a spatial component into sociology in their volume *The Sociology of Spatial Inequality*. They state that space needs to be included as a fundamental concept in resource distribution, in other words going beyond answering the question “who gets what and why?” to “who gets what and where?”

The variable of space in inequality research has been either neglected, controlled away as noise, or have remained the focus of subfields such as rural or urban sociology where inequality itself is given a less central place. These authors then state the question is begged, “Why has space been ignored until recently?” They give three specific reasons. First, classical theorists placed greater emphasis on time and history over space and geography. Second, spatial research conflicted with the liberal social reformist tradition that influenced functionalism. Finally, intellectual priority was given to grand theorizing which assumed generalizations applied everywhere (e.g. globalization) which spatial settings obviously muddled. They go on to state five ways space has been brought back into sociology including a greater emphasis on agency, a move from class stratification to a more general social inequality, a movement towards historical (and spatial) contextualization (as opposed to grand theorizing), employing the cross fertilization of methodologies from other disciplines, and the changes in US society from the 1980s where sociological observations started focusing more on households or areas such as Silicon Valley.

Lobao and Hooks (2007) in the same volume go on to state that because space is mostly dealt with in urban sociology, “Sociology has a poor understanding of people of modest means who reside outside of major metropolitan areas, including their perceived interests and their reasons for voting against the interests sociologists assign to them” (p. 30). Thus there is, they
claim, a great need not only to incorporate space into studies of inequality but also a need to consider human processes occurring at different geographic scales, rather than merely assigning a region or urban/rural tag to an area. The consideration of scale is given more attention later in this paper. For now, it should suffice to say that there has been a call within sociology to at least begin looking at how space is affecting sociological phenomena, and that one of those phenomena includes inequality. Lobao and Hooks state that in order to bridge and develop this type of research one of two avenues can be taken—to begin with spatially oriented traditions and then bring in inequality or to begin with inequality oriented traditions and then bring in space. Regardless of the approach taken, the answer to the question “Should sociologists study space?” must be that “they have and they should continue to do so.” If place can directly or indirectly affect individual or group choices, then sociologists have no choice but to consider it.

Inequality has often been the engine of sociological theory. Race, class and gender have been the traditional foci. As Lobao and Hooks point out, it is time for a fourth focus to be elevated to the level of these three, the spatial one. Inequality varies across space even as race, class and gender proportions remain the same. The mechanism(s) as to why this occurs is difficult to classify and measure, but certainly this is an area needing sociological insight. The hope of the following is to review what has been done and to lay groundwork for future research in this area.
CHAPTER II
THE CONCEPTUALIZATION OF PLACE VULNERABILITY

The Case for Place in the Sociology of Health: Can Medical/Health
Sociologists Study Place?

“Place effects” (operationalized more specifically in many studies to “neighborhood
effects”), serve as an important emphasis in this framework. Even though this has often been the
claimed emphasis of previous research, what really ends up occurring is that individual level data
is obtained and geocoded to an address or small spatial unit to be converted into an area rate.
However this type of research fails to adequately capture the context or composition of a place.
For example, it may show in the study overall that Blacks are more susceptible to a disease than
Whites or that a predominately Black Area is more vulnerable to a disease than a White area.
But this type of analysis does little to advance a concept such as place effects. As MacIntyre et
al. (1993) state, “Unless we try to explore more systematically the ways in which different types
of area differ, we are left without any suggestions for social or public health policies that might
improve the health of those in the worst areas, other than those relating to individual
improvements in lifestyle” (p. 219). Analyzing race while controlling for area ignores the social
and environmental influences on health while analyzing area and controlling for race ignores the
role of political and economic determinants.

Focusing merely on “vulnerable” people in this way often gives way to results that
simply mirror the composition of the population container selected. But this isn’t necessarily the
case. Gatrell (2002) shows that there are significant mortality rate differences between Blacks in
Harlem versus Blacks in Alabama. Fang et al. (1998) show that Whites living in predominantly
White neighborhoods have significantly lower mortality than Whites living in mostly Black
neighborhoods even after controlling for poverty. Cotter, Hermsen, and Vanneman (2007) find
that a family with a mother who is a high school dropout is less likely to be poor in an area with high average earnings than in an area with lower average earnings. Frohlich, Potvin, Gauvin, and Chabot (2002) found that youth tobacco use in poorer neighborhoods with a high percentage of people with a university degree was similar to wealthier neighborhoods. Thus, factors such as race and poverty need to be considered within the area they are located and not merely isolated out as neighborhood effects. In other words, these “fundamental causes” (Link & Phelan, 1995) are necessary but need to be interpreted in light of how they can vary based on the structural effects of the area that contain them.

This is not to say that individual risk factors are not necessary in health research and models. Individual predictors of health such as genetics, psychology, and perceptions have shown great promise in the field. The purpose here rather is to expand on the also promising research that has been completed on place effects of health that is often mentioned or implied in sociological theories of health but not developed. A selection of these theories is reviewed below including examples of studies where the theories are implemented.

Sociology, Health and Place

As mentioned earlier, implications of where one resides can be found in early social theory. Other examples such as Weber’s concept of life chances could include an individual’s access to resources, which today would include health care. Skinner (1953) posited that observed environmental events controlled or influenced behavior. Lewin’s ecological psychology and life space also give room for the study of environmental influences (1936). Even though these early theories served as the basis towards a more micro-theoretical view of health as a sociological conception of health came into being, most recent theories at least give a
passing mention to external and environmental effects of health. These include the health belief model (HBM), human ecology, social capital, structuration, and vulnerability which are briefly described below.

Health Belief Model

The health belief model (HBM) has been one of the most widely used conceptual frameworks in the sociology of health (Strecher & Rosenstock, 1997). While initially developed to explain the lack of program participation in disease detection and prevention with a basis in psychological theory, it was later expanded and extended to include components that Strecher and Rosenstock (1997) summarize into perceived susceptibility, perceived severity, perceived benefits, perceived barriers, cues to action, and self-efficacy.

Still, this model is designed to get at the root of individual risk factors of a disease. Demographic and structural type variables are left as modifying factors that can influence behavior, but not as “fundamental causes.” The components above are individualistic in concept, but one’s perceived susceptibility for example can be applied by defining an at-risk population or one’s perceived barrier measured by proximity to resources. It is important to remember that a place is made up of individuals and therefore understanding the applications of HBM in individual constraints and resources may also be useful when these individuals are aggregated to an area. For example, analysis may find that behavior that is susceptible to HIV/AIDS does not vary by groups such as adolescents, but studies that look at susceptibility in this way ignore whether or not the behavior varies by area. Even if not, again risky behavior in an area with a high HIV/AIDS rate is in greater danger of contracting the virus than that same behavior in an area with a low HIV/AIDS rate.
Sallis and Owen (1997) state that too much energy and resources in health promotion research remains narrowly focused on changing behavior through changes in intrapersonal factors such as knowledge, attitudes, and skills. They go on to promote more ecological models which logically lead to health promotion approaches that supplement behavioral and educational programs with modifications in social climates, in policies, and in physical environments. So while individual risk factors are necessary in the study of vulnerability, the focus of this research is to look beyond the individual to the more ecological effects of vulnerability, though ideally, both levels should be investigated concurrently.

One example of how this framework has been expanded to a spatial understanding of health is in Bettinger et al. (2004). This study is one of many that find a correlation between perceptions of an area and health—in this case, perceived social cohesion and health (Shin, Clark, & Maas, 2006; Cho, Park, & Echvarria-Cruz, 2005; Latkin, German, Hua, & Curry, 2009). While some of these studies have started to attempt to isolate individual factors from more structural or aggregated effects, the focus of most this research remains at the individual level and outcomes.

Human Ecology

Human ecology, which grew out of the early Chicago school sociologists including Park, can capture a broad spectrum of concepts related to health. The purpose here is to briefly understand the ecological models of health as opposed to individual-centric models such as HBM. In other words, human ecology can be viewed as a macro counterpart to the micro HBM. Ecological models consider the environment influences on people. These types of models have not been confined to sociology, but are employed in public health, geography and psychology.
The study of environmental influences of disease has been long recognized. Bronfenbrenner (1979) developed ecological models related to health promotion, recognizing three levels of environmental factors that interact with individual variables: the microsystem, which represents interactions in specific settings such as family members or work colleagues, the mesosystem, which refers to interactions among settings, family and school for example, and the exosystem, which is the larger social system that affects individuals through economic forces, cultural beliefs and values, etc. Sallis and Owen (1997) recognize that a model such as this demonstrates the principle that environments can be conceptualized at various levels of integration. They conclude that among the many research challenges that face ecological theorists, the identification and categorization of environmental factors remains important.

Specific to health, Moos (1979) model recognizes four sets of environmental factors important for health research including physical settings, organizations, the human aggregate, and the social climate. Stokols (1992) also produced assumptions that form the basis for an ecological approach to health promotion which include that health is influenced by physical, social and individual factors, that environments are complex and understanding of this complexity must consider an environment’s multiple dimensions, and that environments can occur at various levels of aggregation. Bronfenbrenner (1979), Moos (1979) and Stokols (1992) all hint at the multileveled complexity of environmental effects and the need to consider this complexity in future models. Place is a concept that is useful in attempting to develop such a complex model.

**Social Capital**

Bourdieu defined social capital as “the aggregate of actual or potential resources linked to
Carpiano (2008) states that this is particularly useful for studies attempting to understand how neighborhood impacts individual health. Some studies, he states have used Bourdieu’s theory to understand social determinants of health (Cockerham, 2007; Frohlich, Corin, & Potvin, 2001; Veenstra, 2007). He goes on to postulate “forms” of social capital, including social support, social leverage, informal social control, and neighborhood organization participation that are consistent with Bourdieu’s theory. Social support captures the social capital that is used to cope with daily problems. Social leverage is the form that gives access to information, opportunities and other possibilities that either minimize socioeconomic hardships or maximize resources related to health. Informal social control is the form that assists in the maintenance of social order while neighborhood organization participation refers to formally organized collective activities. Carpiano emphasizes not only need for theoretical specificity in the area of social capital and health but also the empirical testability.

Research on social capital’s relationship to health has increased exponentially since 1996, mainly using Robert Putnam’s emphasis on social cohesion rather than Bourdieu’s theory (Carpiano, 2008). Kawachi, Subramanian, and Kim (2008) reference Szreter and Woolcock’s (2004) declaration that social capital has become one of the “essentially contested concepts” in social sciences on par with race, class, and gender. Kawachi, Subramanian, and Kim (2008) attempt to whittle down the diverse use of social capital in health research into two distinct conceptualizations. First, social capital conceptualized as resources available to social groups which they refer to the “social cohesion” school of social capital. The important concept in this school is that this is a group attribute (emphasizing contextual influences of the group on the individual) as opposed to the description of individuals who belong to a group. Second, there is
the “network theory” of social capital which emphasizes resources found within an individual’s social networks. The authors state that “Network analysts conceptualize and measure social capital as both an individual attribute as well as a property of the collective (the social network). Most network analysts do not simultaneously assess social capital at both the individual and group levels, but rather they have tended to assess one or the other depending on their method of measurement” (p. 3).

Whether using the cohesion or network schools of thought described above, or distinguishing between bonding and bridging social capital discussed later by these authors, it is again obvious that both schools of thought must include an assumed spatial component. Social networks and group characteristics alike often have a spatial attribute to them. The quantity of research spent on social isolation and differentiation emphasize this point – perhaps most notably in works such as Doreen Massey’s *Spatial Divisions of Labor* (1995). Therefore, one could potentially end up conceptualizing and measuring a subset of social capital in terms of *spatial capital*. The few examples cited above certainly seem to point in this direction. Methodologically speaking then, social capital becomes an issue of scale, as it can be measured at several levels. As the above authors point out, this requires careful theorizing for justifying the scale and interactions at which mechanisms are operating.

*Structuration*

Citing Giddens, Gieryn (2000, p. 467) states that “place stands in a recursive relation to other social and cultural entities: places are made through human practices and institutions even as they help to make those practices and institutions. Place mediates social life; it is something more than just another independent variable.”
In *The Constitution of Society*, Giddens (1984) lays out his theory of structuration that has often been employed by researchers in conceptualization place and health. The core of the theory includes the concepts of structure, system and the duality of structure. Using this theory, researchers have attempted to conceptualize the individual and environmental risk factors of health through Giddens’ understanding of the duality of agency and structure. Giddens is useful for the purpose here based on his view of structure as a constraint and a resource. It can provide an explanation of the relationship between structure and agency in the reproduction of inequalities. “Structure thus refers, in social analysis, to the structuring properties allowing the binding of time–space in social systems, the properties which make it possible for discernibly similar social practices to exist across varying spans of time and space and which lend them systemic form” (Giddens, 1984, p. 17). Thus this structure is not a passive player nor independent of the individual but relational. Many have been resolved to implement this methodologically to cover micro and macro effects, but there is an assumption that has been made that the macro effects are all occurring at the same scale.

Brenner (2001) expands on this, describing the concept of *scalar structuration* that is useful to the present research. He states that previous research has looked at the politico-economic approach to scale “in order to theorize the role of geographical scales as frameworks for a broad range of social activities and struggles, from capital accumulation and state regulation to social reproduction, gender relations, oppositional mobilization (‘jumping scales’) and subjective identity” (p. 599). Brenner then goes on to describe the difference between the singular and plural connotations that politics of scale have been theorized.

The singular meaning can be defined as “the production, reconfiguration or contestation of some aspect of sociospatial organization within a relatively bounded geographical arena –
usually labeled the local, the urban, the regional, the national and so forth. In this singular aspect of the ‘politics of scale’, the word ‘of’ connotes a relatively differentiated and self-enclosed geographical unit.” Scale here is considered as a boundary that separates one geographic unit from another.

The plural meaning of a politics of scale denotes:

The production, reconfiguration or contestation of particular differentiations, orderings and hierarchies among geographical scales. In this plural aspect, the word ‘of’ connotes not only the production of differentiated spatial units as such, but also, more generally, their embeddedness and positionalities in relation to a multitude of smaller or larger spatial units within a multitiered, hierarchically configured geographical scaffolding. The referent here is thus the process of scaling through which multiple spatial units are established, differentiated, hierarchized and, under certain conditions, rejigged, reorganized and recalibrated in relation to one another. (p. 599)

Brenner goes on to state that the main hypothesis that results from these considerations is that geographical scale appears to be most important in those “social processes or transformations which are described through a plural rather than through a singular notion of a politics of scale.”

And finally, quoting Lefebvre, that:

Geographical scale is therefore not to be equated with the totality of sociospatial practices but must be recognized as one crucially important dimension of geographical differentiation, a hierarchically ordered system of provisionally bounded ‘space envelopes’ that are in turn situated within a broader, polymorphic and multifaceted geographical field. (p. 604)

Other researchers, such as Curtis and Jones (1998), provide an important review of the effects of place on health dividing theoretical attempts of contextual health risks into three frameworks. These include the spatial pattern and diffusion of health risk factors, the role of space in social relations important for health, and the more humanist landscapes and senses of place. Most pertinent here is the role of space in social relations which often relies on social theories such as structuration, habitus—adaptable dispositions formed from the structures of an environment (Bourdieu, 1977)—and social control. The variety of hierarchical settings where
people act such as home, work, school are referred by Giddens as locales. An example of using structuration to analyze health inequalities could be accomplished by looking at the constraints on different lives in these different locales, such as how individuals manage their diabetes in different locales, as Curtis and Jones suggest.

Bernard et al. (2007) provide another example of research that uses structuration as the theoretical basis. The authors conceptualize that neighborhoods involve the access to and availability of health resources in a geographically defined area and that these are regulated by four sets of rules including proximity, prices, rights and informal reciprocity. In turn, these rules produce five domains including the physical, economic, institutional, local sociability, and community organization. They claim that specific neighborhood structures influence individual behaviors while resident practice reifies the structure. Even as the authors attempt to conceptualize the neighborhood, they are also quick to point out that many processes occur outside the neighborhood, such as public services, and are spread over larger geographic territories.

**Vulnerability**

Turner (2004) looks at vulnerability as a combination of body vulnerability and the “precariousness of institutions.” While there are good arguments for and against this idea, Turner states that the real argument in the relation between vulnerability and precariousness is found in the connection to social capital. To summarize, he gives examples of vulnerable groups like the elderly who are disconnected from society to emphasize the importance of social capital. “Social capital can be regarded as a collective fund that offsets the vulnerability we experience as individuals and the precariousness of public institutions…The social interconnectedness that
results from human dependency in the life course is in decline as a consequence of the erosion and depletion of social capital” (2004, p. 74).

Building off of Turner’s brief summary on vulnerability, it can be said that there is a need to understand whether or not there is a spatial component to vulnerability. In the example Turner gives, elderly care, it has been shown that proximity is a significant factor that explains whether family extends care to an elderly family member (Joseph & Hallman, 1998). Turner even hints at this underlying concept that serves as a basis for understanding place vulnerability: “Theorists now see disease as having multiple, interactive causes and therefore hold that no simple, single cure is possible or desirable. The AIDS epidemic is a good illustration of this complexity. The AIDS crises in America, Russia and Africa have very different conditions and causes, and hence there can be no single solution to the problem” (2004, p 78). Later on Turner, in discussing the modernization of risk, states that due to globalization, contemporary risks such as HIV/AIDS are not “geographically local niche problems” but “genuinely global epidemics” (2004, p. 246).

But certainly one cannot realistically state that there is equal risk or vulnerability to HIV/AIDS globally or even nationally. When speaking of “the AIDS crises” in America, Russia or Africa in order to develop a measureable model that can produce policy suggestions, a sub-national scale is a more appropriate way to consider the issue. For example, some countries in Africa have less than 1% adult prevalence HIV/AIDS rates others exceed 15% (UNAIDS, 2007). In the US, much of the southeast has HIV/AIDS rates greater than 100 per 100,000 people while the Northern Central area displays rates of less than 16 (per 100,000) (HIV/AIDS Atlas). So while in agreement with Turner’s view that all humans are vulnerable (individually and socially), the distinction must be made that not all humans are equally vulnerable. Place provides a good basis of understanding when speaking of “different conditions and causes” for disease. An
analysis that neglects a spatial component and focuses solely on biological factors or social capital, one would either falsely attribute risk factors obtained in one area to another or be left to develop unique operational models for every new context. More on this and the vulnerability approach is discussed below.

The next step in research appears to be how to model the complexity of spatial or place effects at multiple scales. It seems all previous theories discussed agree on the complexity, but few have provided a way to operationalize it. That is the purpose for the methodological development in Chapter III. For now, the answer to the question “Can medical/health sociologists study place?” seems to be answered with “medical sociologists must study it.” Furthermore, the following shows that a vulnerability framework is best suited for conceptualizing and operationalizing the complexity place effects bring to research.

Examples of Sociological Place and Health Studies

Sociologists have produced volumes that include *The Sociology of Health Inequalities* (Bartley, Blane, & Smith, 1998) and *The Sociology of Spatial Inequalities* (Lobao, Hooks, & Tickamyer, 2007). It should come as no surprise then there have been a few that have attempted to look specifically at the sociology of spatial health inequalities.

MacIntyre et al. (1993) comment that the studies of area’s influence on health can be divided into two categories—studies whose aim is to provide clues to the etiology of disease and focuses on specific properties of the physical environment and, second, studies of the relationship between deprivation and morbidity and mortality which use area level analysis. The critique of the former of these studies is that the work examines a particular cause of death and a particular feature of the environment and that this specificity, while perhaps appropriate for
infectious disease, is less appropriate for disease of multi-factorial origin or chronic in nature, or
even for risk factors that contribute to a range of conditions. These studies use socio-economic
indicators as controls, attempting to explain away the physical environment with socio-economic
factors. The latter of these studies uses aggregate measures of morbidity and mortality as
surrogates for individual measures. The main interest is to see whether deprivation at an area
level can explain observed differences between areas. In sum, the two types of study differ in
that the former studies composition, while the latter studies the context.

The authors go on to claim that few studies directly examine features of places, what
certain places are like or used to be like. This is because researchers make the assumption that
census type classifications actually describe properties of the areas rather than of the residents
that occupy those areas (a compositional approach). “Just as social class is often treated as
though it explains things, rather than providing a starting point for a more detailed examination
of the processes producing health and illness, area may be also treated as though its relationship
with health is obvious” (p. 218). It is worth mentioning again the authors’ statement, “Unless we
try to explore more systematically the ways in which different types of area differ, we are left
without any suggestions for social or public health policies that might improve the health of
those in the worst areas, other than those relating to individual improvements in lifestyle” (p.
219). And again, that analyzing class while controlling for area ignores the role of the social and
physical environment influence on health while analyzing area and controlling for class ignores
the role of political and economic determinants of what a place is like to live in.

Thus MacIntyre et al. (1993, 2002) raise an important distinction in their approach to area
effects on health. While the previous studies discussed address what can be described as a move
from place as a passive place-holding to an active participant (Kearns & Moon, 2002), MacIntyre
et al., move this one step further by going beyond merely the composition of a place to an operationalized context of place. In other words, as stated earlier, place should not just be treated as a single variable to be controlled for, as a residual category, or a sum of individual or structural characteristics, but as a dynamic mechanism that can be used to test hypotheses of specific chains of causation that might link place of residence with health. Few studies have attempted to look at place in this way before MacIntyre.

Since MacIntyre’s seminal sociological work on place and health, there has been a substantial increase in research on the topic that spans the spectrum of the sociological theories discussed above. For the sake of brevity and specificity, only a couple of examples pertinent to the study at hand should suffice. Browning and Cagney (2003) is an example of looking beyond the compositional makeup (specifically socioeconomic status) of a neighborhood and more into structural characteristics, social organization and processes. Browning and Cagney challenge the more traditional social disorganization approach by emphasizing residential stability and its correlation to economic disadvantage using Wilson’s model of neighborhood decline and social isolation. “Stable poverty,” they state, “may encourage the dissemination of health-compromising subcultures” and also may experience “long term declines in the institutional fabric…and informal social control capacity of the community, with consequences for health” (p. 554). Like MacIntyre and others who have stated a need to include both structural and neighborhood (or contextual and compositional) effects, these authors, while acknowledging the need to include structural level indicators of health, also state that structural characteristics alone would leave the question of what mechanisms link these characteristics with health outcomes?

Drawing on social disorganization theory and collective efficacy theory, Browning and Cagney cite numerous works and elaborate on previously thought community level mechanisms
that effect health including social support and sociability, normative transmission and health related collective efficacy. First, they review some of the vast literature on informal social supports that has garnered interest for research on social network characteristics and health outcomes. Second, they review the subcultural perspective on health focusing on Wilson’s theory showing that “the combination of poverty and immobility results in the social and spatial isolation of neighborhood residents from mainstream forces of influence” which in turn results in adaptive behavioral strategies (p. 555). Furthermore, within these disadvantaged communities emerge “health-related subcultures.” Browning and Cagney go on to state, “Two aspects of these subcultural orientations may have consequences for health: 1) tolerance for risky lifestyles and 2) anomie or detachment from conventional values” (p. 555).

Third, the collective efficacy perspective emphasizes community goals in action as the critical mechanism. This is distinct from the social support and sociability perspective in that “it emphasizes the sense of attachment to community and the perceived willingness of community residents to intervene on each other’s behalf regardless of pre-existing social ties” (p. 556). Evidence shows that collective efficacy can be a generalizable mechanism influencing many outcomes including physical health. Pathways through which this can occur include the management of physical hazards and decaying infrastructure, access to services, social control of health-risk behavior, capacity to attract health services and recreational space, and the effect of neighborhood trust, fear and self-respect.

The development and expansions of these theories is necessary, and the point here is that there often remains a shortfall when it comes to the operationalization of the theories and the measurement of these mechanisms. As Browning and Cagney state, “While an extensive literature has documented the role of individual level risk behaviors in promoting individual
level health, few studies to date have attempted to assess the health effects of community level
tolerance of risk behavior and anomie” (p. 556). Though previous studies have examined the
impact of these mechanisms, they state that the research has either “focused on a relatively large
units of aggregation…or has used person level operationalizations of processes specified to
operate at the neighborhood level” (p. 553). This present research is an attempt to move this
bulk of theoretical work a small step forward into measurement.

There have been more and more attempts to operationalize the complexities of place on
health. In one well known study (Katz, Kling, & Liebman, 2001), the short term results of
children who moved to a more affluent neighborhood had a significantly reduced prevalence of
injuries, behavioral problems and asthma attacks even after controlling for change in
employment rates, earnings, or welfare usage by for household heads. The same study also
showed significant improvements in the general health status and mental health of the household
heads and the move (Katz, Kling, & Liebman, 2001). Selected other noteworthy examples
include Lillie-Blanton and Laveist (1996), Sampson, Morenoff and Gannon-Rowley (2002),
Ross and Mirowsky (2003), and Browning et al. (2008). Numerous other examples are explored
in Chapter III.

Place Effects and HIV/AIDS: Can Sociologists
Study Place Effects on HIV/AIDS?

The final remaining question is whether place can truly have an impact on a behavior
driven health issue such as HIV/AIDS? Fullilove (2010) has argued that HIV/AIDS is a
collection of social ills with a medical outcome. If place can impact one’s decisions and place
can impact health as cited above, then it should not be out of the question to research whether or
not place impacts the prevalence of HIV/AIDS. Not surprisingly, this very question has been
addressed. For a general example of this, consider Galea, Aherna and Karpati (2005) who find that neighborhood mortality rates whose causes were sensitive social conditions such as HIV/AIDS varied when measured against underlying socio-economic vulnerabilities. Another example, Arnold et al. (2009), found that the disparity in mortality of people living with AIDS (PLWA) between races was negated when introducing the socioeconomic status of a neighborhood. Interestingly, the authors found that the treatment disparity of the entire neighborhood independent of race may have been the reason for the mortality disparity.

Research that emphasizes area effects on HIV/AIDS often begin with a similar format as research that emphasizes area effects on other diseases or health in general. Both look at social and structural factors of a disease and attempt to understand how these shape risk. This often leads to the construction of a “risk environment” as in Rhodes et al. (2005). This particular work defines a risk environment for HIV as “the space, whether social or physical, in which a variety of factors exogenous to the individual interact to increase vulnerability to HIV” (p. 1026). Spaces such as this are constructed through a variety of social or structural factors, such as in this case population movement, neighborhood deprivation and disadvantage, social capital, political and economic transition, social and economic inequities, and laws and policies just to cite a few.

Rhodes et al. provide an important overview for the risk environment specifically to HIV among intravenous drug users (IDUs). They state that HIV infection, while a behaviorally driven disease, is also subject to environmental influence and that the “relativity of risk” and different cultural, social, economic or political environments differentially affect the variation in HIV. Regarding how this has been approached in previous research, they state that:

There have been few attempts to synthesize evidence on how the environment influences the risk of HIV transmission, whether among IDUs or among other populations at risk. There remains a dearth of evidence which in part reflects the historical immediacy of international interest in the role of environmental interventions in HIV prevention, but
also reflects the predominance of ‘individualistic’ models of research and intervention, both in HIV prevention specifically and public health approaches generally. (p. 1027)

When constructing a risk environment, the authors emphasize that understanding the necessity of including both the type (social, political, economic, physical) and the level of environment (micro, meso, macro) as well as potential interaction between types and levels. To summarize they state “We argue that the HIV risk environment is a product of interplay in which social and structural factors intermingle but where political–economic factors may play a predominant role” (p. 1026). Of importance here is again the complexity seen in other sociologies of health in the dynamic between social and structural, composition and context, or micro, meso and macro.

There has been a few ways proposed to approach the sociological significance or context when using HIV/AIDS as an outcome. Any one of the theories reviewed above could be used as a way to analyze HIV/AIDS. An HBM approach to HIV/AIDS could look at availability of knowledge of the disease and the perception of risky behaviors, or the knowledge or lack thereof facilitates. Many studies have used an HBM approach to HIV/AIDS (Brunswick & Banaszak-Holl, 1996; Neff & Crawford, 1998). A Human Ecology approach could look at the changes in human ecology, e.g. hygiene, population density, physiology, etc., and the effects on behavior patterns as it pertains to risk behaviors and HIV/AIDS. Though less research on HIV/AIDS has used this approach then HBM, it still has its following (Holmes, 1994; Wasserheit, 1995). Social capital has been an obvious choice among sociologists who study HIV/AIDS, whether how social capital positively or negatively (or sometimes both) predicts HIV/AIDS (Campbella, B, & Gilgen, 2002; Holtgrave & Crosby, 2003; Gregson, Terceira, Mushati, Nyamukapa, & Campbell, 2004; Rhodes et al., 2005; Pronyk et al., 2008) or how presence of HIV/AIDS predicts lack of social capital (David, 2007). As mentioned earlier, structuration has often been employed among researchers specifically in conceptualizing place and health. It is certainly useful when
attempting to analyze the complexity of factors that have shown to be associated with HIV/AIDS at the micro and macro level. Studies that suggest using this approach include Gatrell (2002, 2005) and Farmer (2006).

The concept of vulnerability is a bit more confusing of an undertaking. The concept itself is used and defined differently among a variety of fields and contexts. Even within sociology its use varies widely (Macintyre, Maciver, & Sooman, 1993; Galea, Ahern, & Karpati, 2005; Hamilton, Broman, Hoffman, & Renner, 1990). Delor and Hubert (2000) review this problem specifically in the context of HIV/AIDS. Their work begins by describing the process of social construction of HIV/AIDS risk through three stages. The first stage singled out specific groups (e.g. homosexuals, Haitians, etc.). The second focused more on the scientific identification of HIV as the cause and singled out specific practices (anal penetration, IV drug use, etc.). Here the authors state there is shift from the group to the individual behavior as the causal mechanism. The final stage emphasized relationship and interaction characteristics where risk occurs. Each stage employs a different intervention strategy. This shifting through stages highlights the difficulty of a concept such as vulnerability because even when dealing with one outcome the definition changes based on the social construction of the time.

The authors go on to define vulnerability in the context of HIV/AIDS through social trajectory (phases in a lifecourse such as pre and post IV drug use), interactions (where two trajectories meet), and social context (the space in which these occur). Synthesized within this is a process of identity construction which is necessary in order to understand the connection between risk, identity and vulnerability. Thus the concept of vulnerability “must focus on examining rigorously the social situations in which vulnerability arises and grows.” The rigorous examination is necessary due to some familiar themes already discussed. In their case,
it is three level—biographical, relational and “macrosocial.” Similarities have already been stated through micro, meso, macro, or individual, compositional, contextual, or any other scales of social processes. The importance here is that vulnerability, they state, can occur at various scales as well.

Delor and Hubert (2000) then analyze how other fields (disaster research, crime, mental health) have conceptualized vulnerability (or have failed to). Disaster research is subject to a reducing vulnerability to a single cause where there is difficulty of real situation applications. They also state that the spatial dimension of vulnerability in this type of research characterizes cities, regions or even whole countries homogenously.

However, Watts and Bohle’s (1993) model abandons this narrow concept and fits better with the “dynamic, relational, and conflictual conception of social relations and ways of coping with risk” and achieves the balance between agency and structure (Figure 1).

![Figure 1. The space of vulnerability (Watts & Bohle, 1993).](image)

This model incorporates what they call the “three basic coordinates of vulnerability,” the
risk of exposure to crises (exposure), the risk of inadequate capacities to cope with crises (capacity), and the risk of severe consequences of, and resiliency from, crises (potentiality).

Many studies, they claim, reveal vulnerability as a multilayered and multidimensional social space. They review three broad approaches, 1) entitlement and capability which views vulnerability from “an economic and juridical vantage point,” 2) empowerment and enfranchisement studies which encompasses the politics and theory of power and 3) class and crisis studies have three basic propositions: a) that “social relations of production are historically specific, b) that historical character of the ways in which surpluses are appropriated and distributed provide a basis to distinguish the broad character of political economies as modes of production,” and c) that political economies “have their own crisis tendencies seen as ‘market failures’ or crises of overproduction.” The strength of this approach is that it links political economy to ecological and spatial processes allowing for incorporating geography and ecology into a vulnerability definition – “a structural historical space which is shaped by the effects of commercialization, proletarianization, and marginalization. This model allows for the graphic representation of both vulnerable groups and vulnerable regions.” Many previous models, they claim, are “flawed by their inability to link structural processes, contingent outcomes, and the space of vulnerability analytically.”

Concerning this approach, Delor and Hubert (2000) conclude that Watts and Bohle’s framework is particularly useful with the HIV/AIDS space of vulnerability, especially when cross cutting it with their social trajectory, interaction, and social context:

This approach definitely has the advantage of looking at various levels of intelligibility and vulnerability and how they are connected. Thus, exposure, capacity, and potentiality call for consideration of elements that come into play at different times and places. This spatio-temporal dimension of vulnerability is important. The relevance of Watts and Bohle's categories is such that we feel that they may be used in the field of HIV/AIDS. Exposure is more upstream from the event, i.e., it concerns the set of factors that
increases the risk of HIV infection. The men and women who have ‘risky behavior’ in contexts where HIV prevalence is already high or becoming so will be deemed particularly exposed to the HIV risk. Capacity concerns the possibility of mobilizing in such contexts the right resources to cope with the situation. Disadvantaged people (migrants, etc.) and young people in search of their sexual identities, for example, may thus be considered particularly vulnerable to the risk of HIV in this regard. Finally, potentiality relates more to what occurs downstream or the consequences of the event, the importance of which will depend on whether the infection occurs in an industrialized country or one in which care is practically non-existent, whether support structures for the PWHA and her/his family and friends exist, etc. (p. 1563)

Gould (2005) is one example of using Watts and Bohle’s framework for understanding HIV/AIDS. Gould noticed the geographic variation of HIV/AIDS in Africa and sought to understand the different “components of vulnerability” using this model. For example he divided factors that influence exposure (biology, location, mobility and sexual behavior), factors that affect capacity (poverty, infection burden, governance and warfare) and factors that influence potentiality (available therapies and care and behavior change) into the framework that Watts and Bohle used for famines. A similar undertaking is considered in Chapter IV.

This section has provided the background and reviewed most of the major frameworks used in the sociology of health and illness. The next section gives a brief justification of why the vulnerability model is best suited in order to understand the sociology of place and health.

Towards a Sociological Conceptualization of Place Vulnerability

To summarize to this point, place has been a factor to be considered in sociology from more classical sociologists (e.g. Weber, Tonnies, Simmel) to the Chicago School where place was presented as a valid research focus. More recently, place has been a focus not to be overlooked (e.g. Gieryn, 2000) and elevated to a priority alongside race, class, and gender (e.g. Lobao, Hooks, & Tickamyer, 2007). Also, the previous section has shown how place has begun to be incorporated into sociologies of health which ever paradigm is being used. As Table 1
shows, each major theory in the sociology of health can be or has been expanded to include place as a factor. Examples of the implications of place effects on each theory have been given which now could lead to a spatial conceptualization that requires an expansion of theoretical development and further research.

Table 1

*Selected Theories and their Spatial Implications*

<table>
<thead>
<tr>
<th>Theory</th>
<th>Implication of Place Effects</th>
<th>Possible Spatial Conceptualization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health Belief Model</td>
<td>Perceived susceptibility and barriers</td>
<td>Perceived susceptibility &amp; barriers in an area</td>
</tr>
<tr>
<td>Human Ecology</td>
<td>Environmental Effects</td>
<td>Spatial Ecology</td>
</tr>
<tr>
<td>Social Capital</td>
<td>Social Networks and Resources</td>
<td>Spatial Capital</td>
</tr>
<tr>
<td>Structuration</td>
<td>Structural Constraints</td>
<td>Scalar Structuration</td>
</tr>
<tr>
<td>Vulnerability</td>
<td>Precariousness of institutions</td>
<td>Place Vulnerability</td>
</tr>
</tbody>
</table>

Furthermore, examples were provided on the major works that have been published by sociologists on place and health (e.g. MacIntyre, 1993; Browning & Cagney, 2003). This led into a review on how previous studies have approached place and HIV/AIDS (e.g. Rhodes) and how HIV/AIDS has been researched using the aforementioned sociologies of health.

At this point it is necessary to look briefly at two potential approaches in the study of HIV/AIDS and place. The first approach falls under the social network approach. Sociologists have used this approach in the study of HIV/AIDS (Latkin, Forman, Knowlton, & Sherman, 2003; Neaigus et al., 1994; Klovdahl, 1985). The social network approach has a longstanding tradition within sociology (as well as other fields) and posits that the direct and indirect links between people should be used to analyze various mechanisms within sociological phenomena. For example, consider two social networks, one with a high rate of HIV/AIDS (Figure 2) and one with a low rate of HIV/AIDS (Figure 3). It is reasonably assumed that the same risky behavior in Figure 2 will in fact be more risky than that same behavior in Figure 3. Sociologists have
furthered this research by looking at an individual’s social network in terms of their social space (Heckathorn, Broadhead, Anthony, & Weakliem, 1999). This interesting work found that the social networks of intravenous drug users most often display a “geographic rootedness” because relationships are based on propinquity. If this is true for injector networks, not surprisingly other research has shown that sexual networks are also geographically driven (Jennings et al., 2010).

\[\text{Figure 2. Social network approach to HIV transmission (high risk).}\]

\[\text{Figure 3. Social network approach to HIV transmission (low risk).}\]

It has previously been suggested that a social network approach is lacking because it does not include spatial arrangement or multidimensional representations (Scott, 1988). Particularly
for HIV/AIDS, it is important to keep in mind how the virus is transmitted. The CDC reports that for all new HIV cases diagnosed in 2007, nearly 85% were transmitted through sexual contact (CDC, 2009). The state of Texas displays nearly the same percentage (84%) for the same year. Furthermore, studies have shown that number of sexual contacts with an HIV positive person increases the chance of contracting the virus (Padian, Shiboski, & Jewell, 1990) and that those in steady relationships are more likely to engage in HIV risk behavior (e.g. unprotected intercourse) than do individuals engaging in just casual sex (Hammer, Fisher, Fitzgerald, & Fisher, 1996; Sherman & Latkin, 2001). Regarding the social network approach, a couple of adjustments must be made in order to identify an individual’s vulnerability to HIV/AIDS through a sexual network. First, family members are generally excluded. Second, while it is evident that social (and potentially sexual) networks can extend over great distances, sexual contact more likely occurs within a reasonable travel distance from an individual’s home. Also, as stated above, higher risk sexual contact more often occurs in a longer term relationship. Furthermore, research has shown that in the cases in which HIV was contracted through sexual contact, that contact was more often than not within a steady, as opposed to a casual or anonymous, relationship (HIV transmission in intimate partner relationships in Asia, 2009; Davidovich, de Wit, Albrecht, Geskus, Stroebe, & Coutinho, 2001; Kegeles, Hays, & Coate, 1996). In one study this type of transmission accounted for 86% of new infections (Xiridou, Geskus, de Wit, Coutinho, & Kretzschmar, 2003). This leaves the occasional out of social network contact (e.g. sexual tourism fling, sexual contact with a commercial sex worker) contributing less to the total of HIV transmission. Even still, a 2009 National Household Travel survey found that in the US, one of the world’s most mobile countries, people still only travel around 17 miles on average for social/recreational purposes (NHTS, 2009).
What may be concluded then is that not only is HIV being transmitted within one’s social space (and network), but within one’s geographic space as well. Thus a second approach in the study of HIV/AIDS and place is a geographic one. Reconfiguring the social space from Figure 2 into geographic space yields the following:

![Geographic approach to HIV transmission.](image)

When comparing the two social network figures above, the statement was made that it can be reasonably assumed that the same risky behavior in Figure 2 will in fact be more risky than that same behavior in Figure 3. Now considering Figure 4, the same type statement can be assumed. Considering two geographic spaces, one with a high rate of HIV/AIDS (Area 3) and one with a low rate of HIV/AIDS (Area 2), it is reasonably assumed that the same risky behavior with someone from Area 3 will in fact be more risky than that same behavior with someone from Area 2. Now, for simplicity, consider combining the two types of choices, risky or safe, in two types of places, high risk or low risk, as in Figure 5.

In order to analyze this type of framework, HBM can be excluded since it does not consider the effects of place outright. However, since there is still an individual choice being made, the human ecology model should be excluded as well.
Figure 5. A typology of choice and area.
If assuming the same behavior in both places, the question must be asked, what factors may be common in areas that have high HIV/AIDS rates? This question is addressed in Chapter IV. However, the crux of this argument is that this question does not go far enough. If it assumed and shown that characteristics of a place are in fact relevant in the transmission of HIV/AIDS, it cannot be assumed that these characteristics are made up of only the context of an area while ignoring its composition; or that micro and macro- factors are important and not the meso-factors. Thus structuration (though not necessarily scalar structuration) is excluded as well.

This leaves social capital and vulnerability as the two options. Social capital does lend itself well especially from a social network approach. It is even more appealing if expanded to include an individual’s spatial capital. However consider again the four interactions in Figure 5.

In the first interaction, risky choices are made in a high risk area. The risk behavior is influencing the place while the place is also influencing the choice. In the second interaction, safe choices are made in a low risk area. The same duality of influence is occurring here. In the third interaction, there is still back and forth influence, but a risky decision is still made. However that decision is less likely to have the consequence of contracting HIV/AIDS.

The fourth interaction is crucial in the justification of choosing a vulnerability approach. This is the case of an individual making safe choices in a high risk area. For example, consider the safe choice as a woman not being promiscuous, and the area as a place with a high HIV rate due to high incarceration of its population. Also due the high incarceration rate (particularly of males), there is a gender imbalance (Adimora, Schoenbach, & Doherty, 2006). Under this circumstance, if a woman desires a man for protection, companionship or provision, she will more likely and eventually see that this area not as a place of bad options but a place of no options. In other words, due to the lack of options this woman finds herself as vulnerable—
willing to give up on her safer lifestyle in exchange for these things. Conversely, the lack of options for individuals facilitates riskier choices and thus makes the place more vulnerable to infection. In this example, the high incarceration rate drives this area to become vulnerable. However, also consider that the state has incorporated a zero tolerance drug conviction policy that dramatically and unfairly targets this area. Thus there can be several levels of vulnerability affecting the transmission of HIV.

Social capital, while useful and worthwhile in exploring, does not necessarily have the flexibility in exploring this multi-scalar type of analysis. The sociology of health must be multileveled to include individuals and behaviors, social networks, environment and structural effects. While social capital can occur and be measured at various levels, it neither captures all risk factors that occur (such as individual risk factors or an area gender imbalance) nor does a social network explain all variables such as political-economic decisions. The vulnerability approach as proposed by Delor and Hubert and Watts and Bohle is best suited for capturing this complexity. Even then it must be expanded to include the multi-scalar complexity which has been called for by a variety of researchers (see Insert 1, italics added).

The Complexity of Place

To summarize, this research has provided a review and given the background for major frameworks used in the sociology of health and illness followed by a justification of why the vulnerability model is best suited in order to understand the sociology of place and health. As Insert 1 captures, a variety of theoretical viewpoints synergistically conclude that risk factors for disease should not be reduced to one level of measurement but expanded to the multiple levels at which the processes are occurring.
Furthermore, expanding to the levels by merely groups of variables is not enough. Within each
group (such as “place effects”) there may be more than one level at which place is affecting
health. Therefore, it is necessary to assign place effects to their proper level of measurement when attempting to model and measure the vulnerability of a place. In order to move this bulk of theoretical work forward into operationalization, a multileveled methodology of place effects on health is pursued—an approach where other fields, including medical geography proves valuable.
CHAPTER III
THE OPERATIONALIZATION OF PLACE VULNERABILITY

Chapter II provided an overview of the literature concerning various sociologies of health and how place has or can be incorporated, to the end that a vulnerability framework yields the best option for a measureable outcome. This chapter provides an overview of methodologies used in the study of place effects on health with an emphasis on multilevel modeling, to the end that a vulnerability framework combined with an expanded view of place in a multilevel model yields the best option for a measureable outcome.

It goes without question that the study of place effects on health is interdisciplinary. This focus has been broadly researched within health geography, epidemiology, public health, and, as shown in the last chapter, sociology. While the last two chapters have shown that sociology provides a proper framework in which to capture place effects, this chapter will borrow from other disciplines in order to provide the best operationalization of place effects. The previous chapter was the first step in developing a sociological explanation for place and health. This step was necessary because of the often method-heavy research of the other disciplines that tend to ignore a well-grounded explanation of why a certain method was chosen. However the purpose of conceptualizing a model for any research is to assist in the categorization of observations that can be measured. While not undermining the necessity to devise a model that allows for place to place comparison, a model by itself cannot be employed on its own. It is the measurements resulting from a model that enables one to advance science and make recommendations for public policy. Thus a model that cannot bring forth operational definitions is not very helpful when attempting to draw place to place comparisons.

Now that it has been argued that the vulnerability approach is most appropriate, this
chapter turns toward determining the best model for operationalizing the complex phenomena of place effects and health. It begins with a review of the methodologies prescribed in epidemiology, sociology and medical geography, looking at some specific to HIV/AIDS, then finishes with the need for using a multilevel modeling methodology.

Towards an Operationalization of Place Vulnerability

Epidemiology

Fleischer et al. (2006) state that the global health care system is at a breaking point. One reason for this is because public health research pays “attention to individuals with limited regard for the communities in which they live” (p. 1). McMichael (1995) describes the circle epidemiology has made in light of the development of theories of disease causation discussed earlier. Quoting McMichael:

In the mid-nineteenth century, early epidemiologists had a sturdy interest in accounting for differences in health status among social classes, geographic regions, and occupational groups. Later in the last century, the nature of infectious disease was clarified, and relatively simple agent-environment-host causal models emerged which, with their seductive specificity, eclipsed those earlier ideas. Specific occupational exposures began to be identified as causes of cancers, lung diseases, and so on. During the latter twentieth century, in a similar vein, epidemiologists have come to focus particularly on specific personal behavioral risk factors for non-communicable diseases: smoking, alcohol, diet, physical activity, and sexual behavior. Epidemiology today, in developed countries, thus assigns a primary importance to studying inter-individual variations in risk. By concentrating on these specific and presumed free-range individual behaviors, we thereby pay less attention to the underlying social-historical influences on behavioral choices, patterns, and population health. (p. 633)

He goes on to say that there is an expectation to find a “susceptibility gene” for most diseases leaving epidemiology ignoring the strong undercurrents that may be explaining the differences of risk between populations rather than within them. Krieger (2003) says it this way:

Despite epidemiology's longstanding concern with time, place, and person…place had receded into the background by the mid-20th century, conceptually unmoored from
increasingly influential etiologic frameworks based on characteristics of the individual. Fortunately, GIS has contributed in recent years to a reviving awareness that any epidemiologic explanation worth its salt must encompass geographic—and temporal—variations in population health. (p. 384)

There has been a movement within epidemiology to bring place back in to risk factor studies. Meng et al. (2008), for example, use a logistic regression to show that proximity to traffic density increases the risk of asthma symptoms even after controlling for poverty and other symptom producing variables. Reynolds et al. (2004) use a Cox proportional hazards model (which is a common technique in looking at geographical variation in health outcomes) to show that breast cancer hazard rate ratios varied by region even after controlling for socioeconomic status, urbanization, and personal risk factors. Even regarding HIV/AIDS, epidemiologists are starting to emphasize the importance of the social and economic environment. Adimora and Schoenbach (2002) acknowledge that:

Part of the reason for the failure of current prevention strategies to control STIs among African-Americans may be the exclusive emphasis of most intervention programs on changing individual-level risk behaviors while ignoring the structural features of society that promote transmission of STIs. Individual behavior change is essential for effective STI control, but complementary intervention strategies directed at changing the social environment may be required and should be explored by public health workers. (p. 710)

The authors go on to cite one example of research that found high HIV/AIDS rates in high minority areas of New York City stem from the implementation of specific city policies (Wallace, 1990). Adimora and Schoenbach (2002) then conclude that “Contextual factors are increasingly being taken into account through contextual or multilevel analysis, which seeks to explain how individual-level and group-level variables interact in shaping health outcomes” (p. 710).

Many other examples could be cited in epidemiology’s recent push to include place effects to its individual-centric research focus. The importance here is to emphasize the method
in which place effects are included. Ana Diez-Roux (2007) identifies an individual versus systems dichotomy of approaches by epidemiologists in their attempt to integrate social and biological factors. She provides a worthwhile example to illustrate this dichotomy.

The study of the obesity epidemic provides an illustrative example. The traditional epidemiologic approach begins with a study of the risk factors for obesity in individuals. The genetic, behavioral, and social characteristics of individuals are examined in relation to obesity. The “independent” effects of these variables are isolated in regression models. The assumption is that once we understand these individual-level causes of obesity, we can add them up to understand the causes of the obesity epidemic. In contrast, a system approach would begin by describing the component parts of the system in which the obesity epidemic is embedded. These components would include not only the biologic, behavioral, and social characteristics of individuals but also systems-wide features such as the mass production and marketing of foods, the organization of transportation, and the presence of social norms regarding behaviors and body size. A model would be developed that describes the interrelationship between these components. The model would have to be simple at first but would gain in complexity as the functioning of the system is understood. The model could then be used to predict system changes in response to an intervention. By definition, such an approach necessarily integrates social and biologic factors into the functioning of the system. (p. 573)

Led by Diez-Roux, epidemiologists have used multilevel models to parse out the effects of these systems (and by default in some studies, areas) from individual risk factors. However these “system effects” tend to be population level measurements that do not take into account spatial variation, and researchers do not attribute to these effects the same level of complexity as they do individual risk factors. This statement is unpacked further throughout this chapter.

However before that it is important to briefly mention the potential contribution of sociology to this complexity. Contrary to a more epidemiological approach, sociology should have been at the forefront of devising a relevant methodology in order to measure the complexity of this next level. Sociological research would reasonably view a vulnerable person as not merely biologically driven but an individual embodied in social processes and under institutions that they create. Yet surprisingly, the suggested models from this discipline are still fairly underdeveloped.
As described in detail in Chapter II, there has been no shortage of sociologists of any theoretical bent studying place effects. However the methodological development to accommodate the complexity of place has been relegated to a few outlying journals such as *Sociological Methodology* or *Sociological Methods and Research*. More often than not, these sociologists have relied heavily on other disciplines for methodological input, especially when it comes to place and health.

However there have been a few important works worth mentioning here. Doreian was perhaps the first sociologist to introduce and emphasize the importance of recognizing the flaws in trying to determine linear relationships with spatially distributed data (Doreian, 1981). Loftin and Ward (1981) disputed research that suggested that population density impacts pathology by introducing a technique known as spatial autocorrelation. More recently there has been a greater quantity of literature, specifically on place effects and health, within sociology to the point that there are even now discussions as to what is the best approach for scale and boundary selection (Flowerdew, Manley, & Sabel, 2008; Spielman & Yoo, 2009). Significant other examples have been mentioned in the previous chapter.

The methodology in the sociological literature on HIV/AIDS has largely centered around behavior and social networks. However there are examples of methodologies and models that include larger scale structural effects and the risk environment. Wallace and Wallace (1995) estimate a “spatial threshold” of AIDS diffusion caused by “socio-geographic” factors. They state:

Within New York City in particular, the massive loss of housing to contagious urban decay during the 'planned shrinkage' period of the 1970s accelerated residential outmigration by whites, while seriously disrupting and dispersing previously stable minority communities within the city, and driving urban minority population into long-
standing, previously small suburban minority enclaves. Such disruptions mix previously disparate social networks, both within a between city and suburb, while geographically extending the networks of the displaced populations: expatriate populations may often retain contact with friends and family in the 'old neighborhood', as is certainly the pattern with international migration. (p. 343)

Interestingly they find that “As HIV has intensified in time, it has spread in space so that the fundamental area characteristic of the disease pattern has increased as well” (p.338).

Rhodes et al. (2005) attempt to identify social structural causes of HIV risk through intravenous drug use and develop a definition of an HIV risk environment as “the space, whether social or physical, in which a variety of factors exogenous to the individual interact to increase vulnerability to HIV” (p. 1026). Using a broad range of variables, this risk environment model is comprised of both type and level dimensions. Types of environmental risk include physical, social, economic and policy while including micro, meso and macro levels with interplay between these levels. This type of approach begins to correctly capture the complexities of place, which the authors conclude is a manner in which the complexities of social processes are captured. “HIV risk due to structural factors is often an unintended outcome emerging out of larger social forces which operate on multiple levels and which often have contradictory or synergistic effects on infection rates” (p. 1028).

Along with the model proposed by Watts and Bohle (1993), these examples in the sociology literature suffice to explain the complexities of the social structural processes on health. However, as previously stated, the purpose of conceptualizing a model for any research is to assist in the categorization of observations that can be measured. Operationalizing these models is often where sociology has fallen short, but this the exact area where medical geography has excelled.
Medical Geography

Geography has seen a recent boon in the study of place effects of health and illness though its roots are older than sociology’s. Whereas sociological research is looking towards other disciplines for methodologies that will correctly measure its well conceptualized models of place effects, geographical research generally includes very methodological explanations of the geography of disease and illness and inequalities in health care access. Admittedly however social theory has provided only a late flourish in even human geography research in part due to geography’s pursuit of scientific credentials (Graham, 2005). Though now building theories of health and place through the likes of sociological concepts such as structuration, it is evident that these geographical based empirical models needed sociological input like the one described in Chapter II.

Oppong and Harold (2009) review medical geography’s work in this area.

The earliest work on environmental influences on health has been traced to French geographers Maximilien Sorre (1933) and Jacques May (1950) who wrote of the geographical influences on disease and the pathological aspects of geographical regions. While Sorre is credited with developing the concept of a disease (pathogenic) complex (Learmonth, 1988; Pyle, 1983; Verhasselt, 1983; Barrett, 2000, Brown and Moon, 2004), Jacques May, known as the ‘father’ of medical geography, shaped the concepts of disease ecology into a distinct sub-discipline (Meade and Earickson, 2000). Sorre (1933) argued that the occurrence of diseases may depend on physical, biological and social factors or more specifically, climate, natural biological environment, and the anthropogeographical environment. Thus, environmental conditions, conditions for the life of the pathogen, and characteristics of people influence disease occurrence (Sorre, 1933).

Traditional disease ecology models focus primarily on the transmission and spread of infectious diseases. However, a broader disease ecology perspective can explain the occurrence and prevalence of degenerative diseases as well. Meade and Erickson (2000) formalized such an expanded disease ecology model based on the triangle of human ecology. Habitat, population, and behavior comprise the three vertices of a triangle, with a person’s state of health located somewhere within the enclosed space. These three components are broad, elastic characterizations of the various aspects of human life. For example, habitat encompasses the natural environment -- topography, water, plants, animals, and climate conditions; the built environment -- the urban and residential landscape in which people live and work; and the social environment of people, communities, and societies. Humans continually interact with each of these
habitats, not as passive recipients but active shapers of the natural, built, and social environments. Recent works in disease ecology emphasize this broad definition of environment and see the environment not as static but continually dynamic. (p. 81-82)

However the model is limited in that it does not necessarily account for larger “structural” factors such as political economy nor does it directly incorporate a temporal component. Methodologically speaking, this model lends itself to aggregation of variables to the largest scale when it may be necessary and important to keep them separated. Research has recognized this and has focused on either neighborhood effects (Macintyre, Maciver, & Sooman, 1993) to understand small-area variations in disease or political ecology (Gatrell, 2002) to model disease on a macro scale.

Kearns and Moon (2002) identify that one of the major salient themes in the development of the tradition of health geography is “the emergence of ‘place’ as a framework for understanding health.” Health geography as a discipline has constructed the meaning of place from an “activity container” to an operational living construct which matters. Kearns and Moon (2002) categorize the research done on the socially constructed and complex phenomenon of place into three different attempts, namely, localized studies, landscape studies, and multilevel models. They continue on to critique that the localized versions often enhance perceived place in the world over literal place from which little can be drawn towards neighboring places of study, and that the multilevel models offer little in sociological significance and “tell us more about data collection than realities of place.” The landscape studies on the other hand offer a metaphor for “complex layerings of history, social structure and built environment that converge in particular places,” and a nomenclature that was invented and coined within health geography. Overall, they claim that there is a need for a greater contextualization in health geographic research. This can be understood as a method to combine the localized and landscape studies,
which would also allow for the comparison of places. However, as if this task were not enough, the authors identify one key challenge that faces the future of theorized health geography: the incorporation of time. They conclude by pressing the need to move past “snapshot studies of health in places” and move towards more time incorporating models.

Within each of these approaches (localized, landscape, multilevel) there have been models developed that serve the purpose for each approach. However, for the sake of specificity some of the models developed in HIV/AIDS research are reviewed.

Place and HIV

Aside from the literature that has already been cited on place effects and HIV (e.g. Gould, 2005; Fullilove, 2010; Galea, Aherna and Karpati, 2005; Arnold et al., 2009; Rhodes et al., 2005; Delor & Hubert, 2000), there are a couple other that are worthy of note. Crosby et al. (2003) explore the “state-level association between social capital, poverty, and income inequality and adolescents’ sexual risk and protective behaviors.” Parrado and Flippen (2009) use social disorganization to examine individual risk behaviors and neighborhood contexts influence on HIV/AIDS. They state that, “At the individual level, time in the U.S. alters immigrants’ propensities to engage in risk behaviors. At the neighborhood level, factors such as a highly unbalanced gender composition, large share of recent migrants, concentration of socially isolated individuals, and visible indicators of disorder such as littering are all associated with heightened exposure to risk behaviors” (p. 1). Poundstone et al. (2004) give an example with their heuristic framework how this plays out with HIV/AIDS. As shown in Figure 6, the social epidemiology of HIV/AIDS occurs at multiple levels of analysis. This figure provides a good summary of what all of these studies suggest—that the risk environment for HIV/AIDS is a complex,
multileveled one, just as it was similarly concluded last chapter for any study on place and health.

**Figure 6.** The social epidemiology of HIV/AIDS (from Poundstone et al., 2004).

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Traditional risk factor study</th>
<th>Contextual study</th>
<th>Ecological study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>((y, x))</td>
<td>((y, X))</td>
<td>((Y, x))</td>
</tr>
<tr>
<td></td>
<td><strong>Traditional risk factor study</strong></td>
<td><strong>Contextual study</strong></td>
<td><strong>Ecological study</strong></td>
</tr>
<tr>
<td></td>
<td><strong>((y, x))</strong></td>
<td><strong>((y, X))</strong></td>
<td><strong>((Y, x))</strong></td>
</tr>
</tbody>
</table>

**Figure 7.** A typology of studies for modeling health risk factors (Kim, Subramanian & Kawachi, 2008). \((A)\) This type of study is impossible to specify as it stands. Practically speaking, it will either take the form of \(Y, X\), i.e. ecological study, where \(X\) will now be central tendency of \(x\). Or, if dis-aggregation of \(y\) is possible, so that we can observe \(y\), then it will be equivalent to \(y, x\).
However even though this complexity has been noted throughout the last two chapters in a variety of studies, specifying place in a health risk model has been disparately simplistic. Kim, Subramanian, and Kawachi (2008) give this typology in Figure 7 as the types of studies that have been used for determining risk factors in health. Traditional risk factor studies examine individual risk factors \((x)\) that affect individual health \((y)\), or \((y, x)\)–such as the effects of smoking on lung cancer. Contextual studies are common to what has been reviewed in this paper—looking at how and individual’s context (including place) has an effect on an individual’s health \((y, X)\) (such as proximity to heavy traffic effect on asthma). Ecological studies are interested in aggregated risk factors that affect population health \((X, Y)\), such as areas designated as “food deserts” as a predictor of obesity rate in that area.

Multilevel models have been able to combine these studies so that individual risk factors can be examined simultaneously with contextual risk factors, or, to expand the above typology, to a \((y, x, X)\) study – where individual risk factors \((x)\) are analyzed along with place effects \((X)\). This type of model has been used even in HIV/AIDS research (Rosel, Oliver, Jara, & Caballer, 2000; Ukwuani, Tsui, & Suchindran, 2003; Leaver, Allman, Meyers, & Veugelers, 2004).

However aside from the possibility of including individual and contextual risk factors in the same study, this typology portrays another limitation. The complexity of place has been well documented by now which therefore leads to this observation: *What is missing is the potential for Individual health \((y)\) being affected by contextual factors at multiple levels \((X, X’)\) for a \((y, X, X’)\) study or for population health \((Y)\) being affected by contextual factors at multiple levels \((X, X’)\) for a \((Y, X, X’)\).* For example, individual health \((y)\) can be affected by one’s neighborhood \((X)\) and state level policy decisions \((X’\)). Also, a neighborhood’s aggregate health \((Y)\) can be affected by local effects \((X)\) and state level policy decisions \((X’\)). To the best of this researcher’s
knowledge, neither of these types of studies has been specified in depth. The former is more complex, so the latter is the better place to start.

After a careful review of the literature, only one study has produced a study as specified above in determining multi-scaled effects (Y, X, X’) of contextual factors on HIV/AIDS. Msisha et al. (2008) “examine the extent to which the regional and neighborhood distribution of HIV in Tanzania is caused by the differential distribution of individual correlates and risk factors” and found that the “spatial distribution of individual correlates (and risk factors) of HIV do not explain the neighborhood and regional variation in HIV seroprevalence” (p. 741). In their study, neighborhoods and regions accounted for 14 and 6% of the total variation in HIV which provided “evidence for independent contextual variations in HIV, above and beyond that which can be ascribed to geographical variations in individual-level correlates and risk factors” (p. 741). These authors conclude that because of this that both a group-based and a place-based approach is needed for understanding the epidemiology of HIV rather than the high-risk group approach that is more prevalent in this type of research. This is the type of analysis that is attempted in the next chapter, which can only be performed through the use of a multilevel model. However before specifying this type of model for HIV/AIDS, a brief introduction into multilevel models and their use in public health seems pertinent.

Multilevel Models

A common feature in many studies including ones presented previously is an interest in using both individual-level and community-level independent variables to explain individual-level outcomes while also making an effort to statistically show the effects of place. What multilevel models can accomplish is testing whether, for example, individuals within the same
area are more similar to each other than individuals in different area. If that is true, then special statistical models are needed—called by are variety of names in different disciplines including hierarchical linear models, mixed models, random-effects models, or random-coefficient models. For example, if attempting to determine county level effects on an individual, a technique such as this allows a researcher to determine whether or not counties differ in an individual-level outcome variable. If they do not, then these characteristics may be able to be ignored both in specification and in policy development. However if counties do differ significantly, a model such as this examines what specific county characteristics predict the individual-level outcomes and whether it is the county-level variables, or the individual-level variables, or both, that are affecting the outcome (Mosher, Deang, & Bramlett, 2003).

More technically speaking, data in the social sciences are often “nested” or hierarchical (e.g. individuals are clustered in second-level units—such as students in a school, individuals in a neighborhood). Mosher, Deang and Bramlett (2003) state that “this nested structure may mean that the assumptions of OLS regression—including uncorrelated errors and constant variance—do not hold. Thus a technique is needed to see if there is a clustering effect, and if there is, to use that fact explicitly in the multivariate model” (p. 7). Multilevel analysis uses an “estimation procedure that explicitly models this clustering; correlated errors due to clustering are no longer left in the error term as they would be in a single-level OLS regression model. If a single-level model were used on clustered data, it could produce misleading results, in some cases identifying individual characteristics (for example, race) as a determinant instead of a community-level characteristic (for example, median family income or the unemployment rate)” (p.7).

In other words, a multilevel model is a regression model that has an error term, or
residual, for each level used in the analysis. Thus, in a two level individual within an area example, it can estimate a model in which the value of the outcome variable depends on both the area’s characteristics and the individual’s characteristics, and an error term at both the area level and the individual level.

Obviously multilevel analysis provides a major advantage for public health research. Without this type of analysis data either has to be aggregated to the higher level or an area’s characteristics distributed to all individuals resulting in either the ecological or the atomistic fallacy respectively (Leyland & Groenewegen, 2003).

Leyland and Groenewegen (2003) emphasize that multilevel analysis makes it possible to test three kinds of hypotheses in public health:

First, the hypothesis that only individual characteristics are responsible for health differences between communities. If individual characteristics related to health cluster in some communities, one might mistake this for differences produced by community characteristics or circumstances. For example, some communities may have poorer health outcomes but at the same time have older populations. MLA makes it possible to distinguish these so called compositional effects from real contextual effects. One could pose the question as to why people with certain characteristics cluster together. The identification of compositional effects does not solve the problem of individual choice versus material conditions. Second, if there are contextual effects, MLA enables hypothesis testing about the relationship between contextual characteristics and health, taking individual influences on health into account. This provides better estimates of the relation between context and health. An example would be analyzing the effect of community wealth on population health, taking individual income into account. Finally, MLA makes it possible to study specific combinations of individual and contextual characteristics, so-called cross-level interactions. As an example one could hypothesize that people of low incomes in high income neighborhoods have poor health status (relative to people on low incomes in low-income neighborhoods). The interpretation would be that the absolute level of wealth is not that important for people’s health but relative deprivation is. This would be modeled using an interaction between average and individual income. (p. 270)

They go to state that “it might be argued that in cases where the effect of context on health behavior or outcomes is negligible, MLA is irrelevant. But how are we supposed to know? The relative importance of context and individual is an empirical question and unknown
beforehand. That alone is reason to use MLA, if only to show it to be unnecessary in particular cases” (p. 273). Another critique that has been brought on multilevel modeling is in its mis-estimation of neighborhood effects in favor of “randomized community trials” (Oakes, 2003). However, as Subramanian (2004) points out, this is where the distinction must be made between a methodological strategy and a study design—the former of which is the focus of the current research.

Thus a multilevel approach is necessary and appropriate if doing any study on how place affects health. And even though this explanation (and most research as stated earlier) uses this type of analysis for specifying individuals within areas, it can also be expanded to include areas nested within areas (Y, X, X’) or individuals nested within multiple levels (y, X, X’).

Expanding Watts and Bohle’s Vulnerability Model

Now to specify this type of model for HIV/AIDS, returning to Gould’s (2005) adaptation of Watts and Bohle’s (1993) model proves useful. As mentioned last chapter, Gould divided factors that influence exposure (biology, location, mobility and sexual behavior), factors that affect capacity (poverty, infection burden, governance and warfare) and factors that influence potentiality (available therapies and care and behavior change) into the vulnerability framework that Watts and Bohle used for famines. However this type of framework is limited in two ways. First, there is again the incorrect assumption of one scale of influence which does not lead to the proper measurement of these variables affecting an HIV/AIDS outcome at different or nested scales. For example while governance and warfare are often large scale factors influencing HIV/AIDS, access to care can be a more localized effect. Second, Gould does not incorporate the fact that all of these variables can change over time (though Watts and Bohle do theoretically
address this in their framework, albeit without a suggestion for quantitative analysis). Expanding Watts and Bohle’s framework to account for these two limitations would produce a model such as the one depicted in Figure 8 if data is collected at both Zip code and county levels.

Interestingly, both of these limitations can be specified using a multilevel model approach. As mentioned previously when place and individual effects are considered in a multilevel model, the data is structured as shown in Figure 9 where individual cases are nested within areas (y, x, X).

However if variables are affecting an individual outcome at more than one scale (y, X, X’), then a multilevel model can be used to measure each of these influences. For example individual cases nested in ZIP codes which are intern nested in Counties and shown in Figure 10. Briefly addressing the limitation of time, data in a multilevel analysis can be structured as in Figure 11. These 3-level models are not only complex in their data collection and specification, but also in their interpretation as well. As mentioned earlier, a thorough review of the literature found sparse models that merely include two nested areas (Y, X, X’) so this is better place to begin (Figure 12).

This chapter provided an overview of methodologies used in the study of place effects on health with an emphasis on multilevel modeling, to the end that a vulnerability framework combined with an expanded view of place in a multilevel model yields the best option for a measureable outcome. The next chapter applies this model using an HIV/AIDS count outcome.
Figure 8. Expanded Watts and Bohle (1993) vulnerability model to incorporate multiple scales and time.
Figure 9. Multilevel data structure for individuals within areas.
Figure 10. Multilevel data structure for individuals nested within two nested areas.
Figure 11. Multilevel data structure for area nested within another area over time.
Figure 12. Multilevel data structure for area nested within another area.
CHAPTER IV

METHODS

Chapter III provided an overview of methodologies used in the study of place effects on health with an emphasis on multilevel modeling, to the end that a vulnerability framework combined with an expanded view of place in a multilevel model yields the best option for a measureable outcome. The purpose of this chapter is to give an example of how such a model can be designed and specified.

Subramanian (2004) states that due to the complexity of multilevel models in place and health research, caution must be used when developing the conceptual justification at the design stage, and that there must be rigorous diagnostics at the analysis stage. To this end he outlines three critical issues that should be addressed before using this type of model. The first issue concerns multilevel design and structures which includes correctly identifying neighborhoods and neighborhood boundaries while also recognizing multiple spatial (and non-spatial) contexts in which neighborhoods operate. He states, “Current applications have failed to recognize this multiplicity of neighborhood contexts and as such do not go beyond the two-level conceptualizations of individuals at lower level hierarchically nested within neighborhoods at a higher level” (p. 65) (Subramanian, 2004). He goes on to state that identifying the appropriate levels of measurement is crucial for interpreting multilevel structures and true neighborhood differences.

The second critical issue for Subramanian (2004) relates to multilevel model specification. This includes an emphasis on “fixed” parameter effects on individual outcomes and problems in specifying “within” and “between” neighborhood models. He suggests conducting sensitivity analysis to determine robust findings across different models.
The final critical issue is found in the interpretation of multilevel model coefficients. Diagnostic tests, he states, are often ignored or not reported in many studies using this type of model. He suggests using and reporting diagnostic testing while also acknowledging power issues. He states, “For instance, how many neighborhoods and individuals within neighborhoods do we need in order to model the average effects of a neighborhood exposure that is hypothesized to have a differential effect on the outcome depending upon individual SES. Or, how many neighborhoods are required to estimate the differential effect of individual SES that is seen to vary across neighborhoods?” (p. 1968).

All three of these critical issues, those of design, specification, and interpretation, are discussed below.

**Design**

Chapter III introduced the 2-level design for nested areas in Figure 12. What a design such as this allows for is the separation of effects of composition from that of context (or defined elsewhere, local versus landscape effects (Kearns & Moon, 2002)) and also neighborhood versus structural effects (MacIntyre, Ellaway, & Cummins, 2002). Many studies focus on one side of this dichotomy or at best resolve this methodologically through the use of aggregation or assigning community level variables to an individual level. Since this type of design allows for keeping community level effects separate from structural level effects, this begs the question: Are there defined areas that can be used that best capture each of these levels?

**Defining Areas of Study**

Research on neighborhoods and health rightly stresses the importance of defining a
geographic area that is relevant to what is being studied. The same data can be aggregated into any number of different boundaries resulting in a statistical bias known as the modifiable area unit problem (MAUP) (Openshaw, 1984). A study such as this one could show how the effects of MAUP are lessened because data analysis is not limited to one geographic scale. A synthesized theoretical basis such as this one can clarify which spatial unit is used for each variable, and if variables happened to be in two or more nested areas, a multilevel model can be employed. Using this approach, spatial misalignment (Cromley, 2003) (or “change of support” (Young & Gotway, 2007)) problems are not solved by aggregating or kriging, but by justifying what spatial scale a human process occurs (Abler, Adams, & Gould, 1971).

Reiterating the fact that capturing the complexity of place effects over multiple geographic scales is a new venture, this research begins by choosing two scales – ZIP codes to capture “neighborhood effects” and counties to capture “structural effects.” A justification for these two selections follows.

Using aggregated area measures in public health research is not only common but often necessary due to privacy laws on data restrictions with many medical records. Extensive studies have tried to determine the advantages and limitations of choosing one scale over another in public health research (Kreiger, Chen, Waterman, Soobader, Subramanian, & Carson, 2002). The use of ZIP codes in US public health research has become quite common (Krieger N., Waterman, Chen, Soobader, Subramanian, & Carson, 2002). More precisely, the US Census Bureau’s zip code tabulation area (ZCTA) has become quite common. Krieger et al. (2002) give the US Census Bureau definition of the ZCTA as such:

ZIP Code Tabulation Areas (ZCTAs) are a new statistical entity developed by the US Census Bureau for tabulating summary statistics from Census 2000. This new entity was developed to overcome the difficulties in precisely defining the land area covered by each ZIP Code. Defining the extent of an area is necessary to accurately tabulate census data.
for that area. ZCTAs are generalized area representations of US Postal Service (USPS) ZIP Code service areas. Simply put, each one is built by aggregating the Census 2000 blocks, whose addresses use a given ZIP Code, into a ZCTA which gets that ZIP Code assigned as its ZCTA code. They represent the majority USPS five-digit ZIP Code found in a given area. For those areas where it is difficult to determine the prevailing five-digit ZIP Code, the higher-level three-digit ZIP Code is used for the ZCTA code. Since we take the ZIP Code used by the majority of addresses in an area for the ZCTA code, some addresses will end up with a ZCTA code different from their ZIP Code. Also, some ZIP Codes represent very few addresses (sometimes only one) and therefore will not appear in the ZCTA universe. (p. 1100)

These authors go to criticize the use of ZCTAs due to spatial and temporal discontinuities. However as Carretta and Mick (2003) correctly point out, spatio-temporal discontinuities occur at all scales of geography. Krieger et al. (2002) also found dissimilarities in measurement between ZCTAs and block groups or census tracts. In spite of this there are some advantages in using ZCTAs. First, because variables computed from ZCTAs encompass larger populations than census tracts or block groups, they are likely to be more stable (Thomas, Eberly, Smith, & Neaton, 2006). Second, if medical records are obtained at the ZIP code level, there is no need to geocode the data to census tracts or any other geographic unit. Third, several studies have shown ZCTAs to be robust in various locales with similarly significant predictors analyzed using census tracts or block groups (Thomas, Eberly, Smith & Neaton, 2006). Fourth, ZCTA boundaries are computer delineated based on addresses as opposed to manually delineated census tracts or gerrymandering as in congressional districts, and thus the use of these boundaries may reduce bias that stems from social or built environments. Fifth, for the purposes of this study, this is the scale which the outcome variable in this study, cumulative new HIV counts, is measured. These boundaries also provide a small scale at which neighborhood effects can be measured and at which HIV/AIDS data are generally not available. Finally, as Thomas et al. (2006) conclude,

It may be that tract-based and ZIP-code-based measures are significant independent
mortality predictors because they capture slightly different effects. Tract-based variables, for instance, could be more relevant predictors of an immediate neighborhood’s “walkability,” a benefit in maintaining cardiovascular health and reducing stress, whereas ZIP code-based variables might better capture the benefits of high-quality health services within easy driving distance. (p. 589)

Based on the variables that were implemented in this study as more “community effects,” ZCTAs serve as a justifying choice. It should be noted however that the primary purpose of the current research is to develop the methodology that will assist in capturing the complexity of place adequately. So if variables are measured at a different scale or can be better justified in a different manner, the following methodology will still prove useful.

Lobao and Hooks (2007) argue that the many significant sociological questions are centered at the subnational scale, and yet this scale is often neglected resulting in missed opportunities specifically in the study of inequalities (they have termed these neglected subnational scales as the “missing middle” (Lobao, Hooks, & Tickamyer, 2008)). These authors mention several important advantages for using counties. First, counties are a good indicator of structural effects on inequality in the U.S. as Lobao and Hooks (2007) indicate, “Counties are more than population containers. They actually shape growth and redistribution, and thus patterns of subnational inequality” (p. 51). Second, their boundaries are relatively stable over time making them ideal for longitudinal studies. Third, there is a wealth of secondary data available at the county level. Fourth, counties cover rural and urban areas and provide a better picture of subnational inequality. Fifth, “Counties are the fastest growing general purpose governments,” increasing employees by 31% from 1980 to 1997 as opposed to 26 and 8% for state and municipal governments respectively. The authors cite in their survey that “Federal economic development, environmental, health and social programs are typically delivered through county based offices,” and “provide an array of services that regulate local economic
development, enhance human capital, and serve social safety net functions” (p. 51). Finally counties allow for a structural level analysis that does not restrict the research to a conflict theoretical framework that is prominent in country level research (Gould, 2005). Based on the variables that are implemented in this study as “structural effects,” counties thus serve as a justifying choice.

Vulnerability Framework

Using the multilevel design to capture the complexity of place effects leads to how predictor variables should be specified with respect to an HIV/AIDS outcome. The vulnerability framework that has already been emphasized serves as a convenient way to categorize these variables. As previously mentioned, Gould (2005) is one example of research that used Watts and Bohle’s vulnerability framework for understanding HIV/AIDS. He sought to understand the different “components of vulnerability” using the model shown in Figure 1 and divided factors that influence exposure (biology, location, mobility and sexual behavior), factors that affect capacity (poverty, infection burden, governance and warfare) and factors that influence potentiality (available therapies and care and behavior change).

Expanding this to include the fact that these variables may be occurring over different levels of geography provides a model such as the one proposed in Figure 8. For example, capacity (what influences the ability to prevent or cope with the disease) may be affected by community level effects such as poverty as well as structural level effects such as the infection burden on a health department. Exposure (factors that influence the risk of contracting the disease) may be affected by available sexual partners at the community level and the sexual mores of the culture at a structural level. Potentiality (factors that affect the ability to recover
from a disease or epidemic) may be influenced by transportation availability at a community level and public assistance offered at a structural level. This strays from Gould’s (2005) model which views exposure merely at a biological level, capacity only at a structural level, and potentiality as interactions between the two (alleviating/exacerbating the problems from exposure and resiliency of systems and structures to cope with the problem). The specification of this vulnerability framework with an HIV/AIDS outcome is provided in the following.

Specification

HIV/AIDS exhibits spatial variation. In 2008, the latest data available, the CDC reports that the rate of new HIV diagnoses in the U.S. was 19.4 per 100,000 people with the state of Texas only slightly higher at 19.5 per 100,000 (Prevention, Diagnoses of HIV infection and AIDS in the United States and Dependent Areas, 2008, 2008). However, there is not an even geographic distribution of HIV/AIDS throughout the U.S. In 2007, 46% of all new AIDS diagnoses in the U.S., 40% of all people living with HIV/AIDS (PLWHA), and 50% of all deaths resulting from AIDS were in the South, which includes Texas (Prevention, 2009). Within Texas, HIV/AIDS cases are not evenly distributed either. In 2008, PLWHA were concentrated in metropolitan areas, such as Houston and Dallas. Furthermore South and West Texas have lower rates than East Texas (Branch, 2010). The following is an attempt to explain this spatial variation in Texas, by analyzing community and structural level effects at the two nested areas (ZCTAs and counties) within a vulnerability framework. This yields the following study questions and hypotheses.

Study Questions and Hypotheses

1) Is it necessary to separate the relative roles of community level effects and structural
level effects in explaining the geographic variation in the occurrence of HIV/AIDS in Texas?

*Hypothesis:* Adding structural level effects to a community effects only model will improve the prediction power of HIV/AIDS counts per ZCTA.

2) Are exposure, potentiality and capacity variables each necessary to include in determining HIV counts per ZCTA in Texas even after controlling for population characteristics of a ZCTA?

*Hypothesis:* All categories should be included with exposure having a greater impact on HIV/AIDS counts than both potentiality and capacity.

3) Does this model (community + structural level effects) accurately predict HIV/AIDS counts without spatial variation in Texas?

*Hypothesis:* The model will predict with greater accuracy urban area HIV/AIDS counts and will not provide as good a fit in rural areas.

It needs to be reiterated that while research is rife with individual-level risk factors, those are outside the scope of the current research. But as Poku (2002) points out, “Sexual behavior is undoubtedly an important factor in the transmission of any sexually based disease. Alone, however, it appears totally inadequate in explaining HIV prevalence as high as 30 percent of the adult population in some African countries and less than 1 per cent anywhere in the Western world” (p. 533). The focus here, rather, is to build on the research that has already shown that community level and structural level factors play significant roles as risk factors for HIV/AIDS and to construct a measurable model that includes the magnitude and difference for each.

The 2000 Census file contains 1,939 ZCTAs for the state of Texas. Out of these:

- 43 were special ZCTAs assigned to hydrographic features
- 29 were special ZCTAs assigned to large land areas with no 5 digit ZIP postal area such as forests, parks, etc.
- 2 others had population = 0
- 1 is primarily located in another state

Therefore the above 75 ZCTAs are excluded in the following analysis. This leaves 1,864 Texas ZCTAs (mean population = 11,172; min = 12; max = 114,124) that contained a population
and is to be included in the following analysis. There are 254 Texas counties that are analyzed as boundaries for structural data obtained from the Texas Department of State Health Services (DSHS). Combining this level of analysis with the ZCTA level data results in a data structure as depicted previously in Figure 12.

**Dependent Variable – HIV Counts per ZCTA**

Texas DSHS has provided ZCTA level HIV/AIDS counts for the state of Texas representing every reported HIV/AIDS case since 1980 totaling 109,284 cases. Out of these:

- 935 were diagnosed with HIV via perinatal transmission, blood transfusion or hemophilia
- 63076 were diagnosed with HIV before 1999 and after 2008
- 580 were missing ZIP codes for HIV and AIDS diagnoses
- 304 had their HIV diagnosed outside the state of Texas
- 7 were ZIP codes erroneously listed outside the state of Texas
- 15 were missing HIV diagnosis ZIP, and the corresponding AIDS diagnosis ZIP did not fall within the stated HIV diagnosis county

All of these cases above are excluded in the following analysis. For the purpose of research using a vulnerability framework, it is appropriate to only include “horizontal” transmission of HIV while excluding those cases contracted via medical procedures. Before 1999, Texas did not have name based reporting of HIV or AIDS diagnoses so there is a good chance that the cases contained in this data set are underreported before that year and are not underreported in a manner that is representative of the entire population. Data for 2009 and 2010 are incomplete due to delays in reporting. Also, the 1999 to 2008 reporting mechanisms and laws remained relatively stable so there is reasonable confidence that year-to-year reporting has not been influenced by major legislative shifts or practices in reporting. HIV cases that were diagnosed outside the state of Texas were excluded since the analysis is only interested in vulnerability.
predictors within the state of Texas. The remaining cases in the above list are considered missing or erroneous as noted.

A total of 2,587 cases had a ZIP code at AIDS diagnosis but were missing the diagnosis ZIP code for HIV. In these cases an assumption was made that the ZIP code of AIDS diagnosis was also the ZIP code of HIV diagnosis (this was true 86.8% of the time when both years are recorded) so these were copied over. Also, 2,407 did not have the HIV diagnosis year so their HIV report year was used instead (HIV diagnosis year equaled HIV report year 71.4% of the time when both were recorded). Finally 3,184 of these remaining cases are diagnosed as TDCJ cases and are removed since prison presence is one of the predictor variables. This leaves 41,183 horizontal transmission HIV cases with a ZIP code in the state of Texas and a range from 1999 – 2008. Table 2 presents some characteristics of this data.

Table 2

<table>
<thead>
<tr>
<th>Gender</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
<th>Total</th>
<th>Total</th>
<th>Mode of Exposure</th>
<th>Total</th>
<th>Year</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>31341</td>
<td>9842</td>
<td>16126</td>
<td>2770</td>
<td>8380</td>
<td>Male-Sex-Male (MSM)</td>
<td>18178</td>
<td>1999</td>
<td>4540</td>
</tr>
<tr>
<td>Hispanic</td>
<td>11673</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Intravenous Drug Use (IDU)</td>
<td>3506</td>
<td>2000</td>
<td>4571</td>
</tr>
<tr>
<td>Other</td>
<td>481</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MSM + IDU</td>
<td>11119</td>
<td>2003</td>
<td>3964</td>
</tr>
<tr>
<td>Unknown</td>
<td>133</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Heterosexual</td>
<td>3995</td>
<td>2007</td>
<td>3830</td>
</tr>
</tbody>
</table>

The above characteristics portray that HIV in Texas is primarily a Male (76.1% of all cases), homosexual (52.6%), Black (39.2%) problem that has been declining since 1999 until a recent upturn in 2007. However as mentioned previously there is also a spatial problem as well.
and this analysis will attempt to determine whether or not the spatial variation is determined by factors other than the male, Black and homosexual population characteristics of an area.

A total of 312 of the HIV cases did not fall into any of the 1,864 ZCTAs from the US Census file. This leaves a total of 40,871 HIV cases to work with. This is the number of cumulative new HIV diagnoses per ZCTA from 1999 – 2008 and serves as the dependent variable for this analysis (µ = 21.93, σ² = 54.239, N = 1,864). Using cumulative new HIV diagnosis is better justified than using a total people living with HIV/AIDS (PLWHA) count for that fact that it is posited that the PLWHA count is significantly underestimated. Also, using new diagnoses is better predicted in a model using census and other static data measured around the same time period due to the fact that this model is concerned with the environment of the person at the time of HIV exposure.

Diagnostics on HIV Counts per ZCTA

Out of the 1,864 Texas ZCTAs with a population, 557 reported no new HIV diagnoses while 58.5% (1,091) of the ZCTAs had five or less new diagnoses. On the other end, 25 ZCTAs (1.3% of total) account for greater than 20% (21.3%) of all the new HIV diagnoses from 1999 – 2008. This leaves a highly negatively skewed and leptokurtic (skewness = 7.285, kurtosis = 95.427) distribution as shown in Figure 13.

![Figure 13. Frequency distribution of new HIV diagnoses (1999 – 2008) by ZCTA.](image)
With such a highly skewed distribution, techniques such as OLS are not appropriate. As is common with count data, a Poisson analysis can be used. A chi-square goodness of fit test shown in Table 3 supports the non-similarity of HIV counts between ZCTAs. New HIV counts is a relatively rare event and the population at risk is large. Since the distribution of new HIV Counts is a highly skewed distribution with an excess of zero counts, even a standard Poisson analysis is not suitable. In such situations, a zero inflated Poisson analysis is a natural choice. Since this analysis includes county-specific effects, and there are multiple zip-codes within each county, I used a mixed-effects zero-inflated Poisson (ZIP) regression model (Hedeker and Gibbons (2006). A mixed-effects ZIP regression model is discussed shortly.

Table 3

<table>
<thead>
<tr>
<th>Chi-Square Goodness of Fit for HIV Counts between ZCTAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIVCount</td>
</tr>
<tr>
<td>Chi-Square</td>
</tr>
<tr>
<td>df</td>
</tr>
<tr>
<td>Asymp. Sig.</td>
</tr>
</tbody>
</table>

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 10.4.

**Independent Variables**

As stated previously, the model conceptualized by Watts and Bohle (1993) and implemented by Gould (2005) provides a valuable way of organizing independent predictors of an area’s HIV burden. Gould divided factors that influence exposure (biology, location, mobility and sexual behavior), factors that affect capacity (poverty, infection burden, governance and warfare) and factors that influence potentiality (available therapies and care and behavior change). However influences on an area can occur at multiple scales. Because the statistical model being proposed is complex in its specification and interpretation and stands at the frontier
of place and health research, one predictor at each level (ZCTA and county) is selected for each of the points of Watts and Bohle’s vulnerability framework, namely, exposure, capacity and potentiality. The model can be then easily expanded in future research. Watts and Bohle (1993) state that all three spaces of vulnerability simultaneously exist but the weight of each is an empirical question. This is one of the questions that is answered in the following analysis.

Exposure

Exposure can simply be defined as factors that positively or negatively influence an area’s introduction to a disease. From a public health standpoint it is frivolous to determine when and where HIV was first introduced into a Texas county or ZCTA. The more pressing need for analysis is the “cumulative” exposure which is being measured by new HIV diagnoses per ZCTA (precisely the dependent variable in this analysis). Again Watts and Bohle (1993) examine the risk of exposure to crises as “those groups of society who are most exposed to market failures,” while Delor and Hubert (2000) expound on this in the HIV context by adding that “exposure is more upstream from the event, i.e., it concerns the set of factors that increases the risk of HIV infection. The men and women who have “risky behavior” in contexts where HIV prevalence is already high or becoming so will be deemed particularly exposed to the HIV risk.” Finally, Gould (2005) describes this as physical exposure to the virus evident in biology (age, gender), location (urban vs. rural), mobility and sexual behavior (and translating knowledge into behavior).

Late Testers

The Texas Department of State Health Services (DSHS) in their Statewide Coordinated
Statement of Need 2008 – 2010 (2008) state that late testers are a cross-cutting issue. Late testers – defined as those who are diagnosed with AIDS within one year of HIV diagnosis – can unknowingly and therefore easily increase exposure to HIV. As DSHS states, “early testing is critical in preventing the further spread of HIV/AIDS. Those unaware of their status are more likely to transmit the disease to others, resulting in missed opportunities for the prevention of new HIV infections” (p. 11) (Texas Department of State Health Services, 2010). This same report found that over one-third of all new diagnosis in Texas were late testers from 2003 to 2007 with Hispanics having larger proportion (43%) than Whites (32%) and Blacks (33%). This dataset shows a similar trend in that from 1999 – 2008 over one third (37%) of HIV diagnoses had AIDS diagnosed less than one year later. The late tester rate of new HIV diagnoses widely varying from 0 to 275 per 100,000. This variable will calculate late tester rate (from 1999 – 2008) per 100 population (US Census 2000 SF1). Late Testers Rate (ltrate) will serve as the exposure predictor at the ZCTA level (\(\mu = .0055, \sigma^2 = .014, N = 1,864; \text{min} = 0, \text{max} = .276\)).

Prisons

While linking preventative and coping strategies to a structural level indicator may be obvious, introducing political-economic measures of exposure is a difficult and sensitive endeavor. Political and economic decisions certainly influence where and when a prevention program or public health disease burden coping strategy is implemented. Even still, reviewing factors that affect exposure at the structural level is necessary especially when investigating environmental risk factors and diseases affected by proximity (e.g. air pollution and asthma). When examining the risk areas of a behavior driven disease such as HIV/AIDS, extreme caution must precede any variable selection.
Prisons are well known to have a high concentration of HIV/AIDS (Baillargeon et al., 2007). Furthermore in the last thirty years there has been a prison building boom in Texas which has constructed 100 new prisons since 1980 (Engel, 2007). Before 1980, no state prison was located west of Dallas, but now the state is beset with them. Furthermore the size of these prisons has grown from an average population for prisons built in the 1980s 770 inmates to 1350 inmates for those built in the 1990s. Since 2000, Texas has added 61 facilities (sixteen of which are state-run jails), which account for 84,000 new inmates (Engel, 2007). More so, the trend of new prison construction has shifted to nonmetropolitan areas. With this shift came a major community concern over inmate visitors and the fear of associated impacts (Martin & Myers, 2005). But was this perception a reality and did this fear come to fruition? This is a topic that may warrant future study.

With respect to a community’s exposure to HIV/AIDS due to prison location, there are only two ways that this could directly occur. First, prison visitors have sexual relations with inmates, become infected and then spread the virus to the surrounding community. This is unlikely since the Texas Department of Criminal Justice (TDCJ) does not support conjugal visits (Jones, 2010). Second, inmates are released into the community where they are incarcerated and have sexual relations with someone in that community. This is not a possibility for private facilities since TDCJ limits private units to pre-release centers (Maruono, 2000). However with state run jails inmates are released from the place of incarceration (Justice, 2010). This alone is nowhere near needed evidence to justify prison location as a source of exposure. However, there are several factors that turn this evidence from plausible to possible. First, according to TDCJ site selection for prisons must be based upon, among other factors, “close to an area that would facilitate the release of inmates to area of residence” (Maruono, 2000), meaning access to major
highways and other infrastructure. Second, one four state study documented that 50% of inmates surveyed had sex within the 1st 12 hours after release from prison (Gomez, Marin, & Morales, 1996). A follow up study found that more than 51% had sex the first day of release showed an immediacy of risk taking following release from prison and those have sex the first day had riskier sex than the rest of the released inmates in the study (Grinstead et al., 2005). These findings are important regarding the relationship between prison location and HIV especially if inmates are released to the communities where which they were incarcerated.

Furthermore, a final study reported that only 30% of released Texas prisoners infected with HIV filled prescriptions for highly active anti-retroviral (ARV) drugs within 60 days of their release which is a major public health concern due to higher viral loads and infectivity for those not on ARVs (Baillargeon et al., 2009).

Even, if all of these factors led to an increased presence of HIV in a prison community, it would be difficult to separate this from findings that suggest “prisons are more likely to be built in densely populated towns with prior proximate prison with a higher than average percentages of poverty, Blacks, and Hispanics” (Eason, 2010). However if this latter explanation is the reason why in Texas prison communities seem to have a higher rate of HIV/AIDS as some suggest (Kutch, 2010), then this variable simply needs to be moved from a factor that influences exposure, to one that influences capacity or potentiality. This may end up being appropriate since prison communities from the recent rural prison building boom have been found to be “rural concentrations of disadvantage” (Eason, 2010).

There are 113 TDCJ units in 63 of the 254 Texas counties. These units include prisons, private prisons, private state jails, state jails, transfer facilities, pre-parole transfer facilities, and pre-release facilities (Unit Directory). Thus the presence of a prison (prisonDV) in a county will
serve as the exposure predictor at the county level (where 1 = prison located in the county, 0 = not).

Capacity

Watts and Bohle (1993) understand capacity by “the risk of inadequate capacities to cope with crises.” They state that the vulnerable are those groups whose “coping capacity with respect to unfavorable terms of exchange is low.” In the HIV context, Delor and Hubert (2000) understand capacity as it “concerns the possibility of mobilizing in such contexts the right resources to cope with the situation.” They say “disadvantaged people (migrants, etc.) and young people in search of their sexual identities, for example, may thus be considered particularly vulnerable to the risk of HIV in this regard;” this can include socio-cultural capital, knowledge of the risk, prior experiences, etc. Gould then defines capacity as “Structural conditions that affect the rate and direction of spread” of HIV which includes poverty reduction, access to health care and development context and policy (p. 480). Also are included “structural variables associated with the general conditions of high mortality which set the context of national and local capacity to cope and manage the epidemic.” Capacity, Gould says, is “differentiated through ‘entitlements’ and that some regions and groups have more and better access to the benefits of development than others. Furthermore the allocation of these entitlements will emerge through the workings of political economy and the empowerment (or lack of it) of disadvantaged groups” (p. 480), or as in this current research, disadvantaged areas. Gould (2005) uses poverty, infection burden, government and warfare in his analysis.

Poverty

The association between poverty and HIV has been well documented (Gillespie, Greener,
Whiteside, & Whitworth, 2007; Krueger, Wood, Diehr, & Maxwell, 1990). Poverty not only restricts an area’s ability to cope with the burden of a disease such as HIV/AIDS, it also is a crucial factor in creating a spatial concentration of disadvantage (Buck & Gordon, 2004). A neighborhood’s poverty level has been shown to have a negative impact even on the health of those who are relatively wealthier but living in a low income area. The same trend has been shown that racial characteristics of an individual or area do not affect HIV as much as poverty (Krueger, Wood, Diehr, & Maxwell, 1990). While poverty may be a mechanism for poor health in a variety of ways, one undoubted impact it has on an area is how it affects that area’s ability to cope with a disease burden.

Thus poverty is used as a measure of capacity at the ZCTA level using 2000 US Census Data SF1 which includes the number of people in poverty as well as the total population between 18 and 64 years of age. This proportion constitutes the poverty variable (perpov1864) in the following analysis ($\mu = .078, \sigma^2 = .052, N = 1,864; \text{min} = 0, \text{max} = .357$).

Public Health Assistance Funding

Relying again on Gould (2005), government plays an important part in the health service delivery system, and as mentioned earlier, Lobao and Hooks state that counties typically deliver health and social programs (2007). In the state of Texas, around 35 of the 254 counties rely on the county health department as the sole location for HIV testing and educational outreach, yet public health departments as a whole in Texas only received 9% of the Ryan White Funding in 2007 (Administration). More generally though, the Texas Department of State Health Services allocates a portion of funding for public assistance to each county that is used for categories such as family planning services, hospitals, and community service programs. Unfortunately data on
these categories are not available at this specificity. However a county’s per capita allocation can be used as an indicator to measure the effect that a county’s health resources have on the health of its general population. The dollar amount of public assistance funding in the year 2000 obtained from the Texas Comptroller website (Texas Comptroller of Public Accounts, 2000) was divided by the county population obtained from the US Census 2000 Summary File 1 to get County Health Public Assistance Funding per Capita (Fundper) ($\mu = $360.78, $\sigma^2 = 218.36, N = 254; \text{min} = $26.0, \text{max} = $2078.21$).

Potentiality

Watts and Bohle (1993) define potentiality as the risk of severe consequences of, and resiliency from, crises (recovery potential and damage control). The vulnerable are those groups who are insufficiently integrated into social security arrangements. Delor and Hubert (2000) expand this to state that “potentiality relates more to what occurs downstream or the consequences of the event, the importance of which will depend on whether the infection occurs in an industrialized country or one in which care is practically non-existent, whether support structures for the PWHA and her/his family and friends exist, etc.” Gould (2005) furthers this by noting that potentiality is “the ability to do something to alleviate the problems raised by both exposure to the disease and also capacity to be resilient in the face of these structures of poverty, weak governance and high disease burden.” The focus is on “control and prevention through willingness to cope with conditions HIV creates” (p. 481). Gould uses therapies and care (e.g. provision of ARVs), service delivery and domestic care (e.g. public health messages and information distribution) and behavior change (e.g. whether the state creates a context of political will for behavioral and cultural change) in his analysis. Watts and Bohle’s (1993)
description of potentiality seems to lead to duplicity since “coping,” “control,” and “resiliency” can obviously be operationalized similarly. It appears that Gould’s (2005) interpretation focusing on prevention is a better fit and the missing piece when it comes to a complete vulnerability approach.

Service Provider Outreach

The availability of services remains a crucial component in the prevention of HIV. The Texas Department of State Health Services (DSHS) in their Statewide Coordinated Statement of Need 2008 – 2010 (2008) state that a significant barrier to care is that “clients often do not know where to go to get the services they need, or what services are needed. Assessment data consistently rate access to transportation as a primary barrier to care. Publicly funded HIV health care services in Texas are concentrated in larger cities and individuals living outside these communities must travel long distances to access needed care and services.” Barriers exist within cities due to large geographic areas and lack of public transportation. Research has seldom looked at the availability of educational outreach as a means to reduce the infection burden of HIV/AIDS within a given area.

DSHS provides a list of self report HIV/STD Services in the state of Texas. This list provides 325 service providers throughout the state who provide educational programs and HIV testing to the general public. Thus the presence of a service provider who provides these programs (SrvPrDV) in a ZCTA will serve as the potentiality predictor at the ZCTA level (where 1 = service provider located in the ZCTA, 0 = not). This is consistent with Thomas et al. (2006) assessment above, that “ZIP code-based variables might better capture the benefits of high-
quality health services within easy driving distance” (p. 589). In Texas, 208 ZCTAs have at least one of these service providers located within its boundaries.

**Social Norms**

County level HIV prevention programs may seem an intuitive place to look in order to complete the list of independent predictors. However federal funding programs such as Ryan White do not provide any prevention measures and HIV prevention is still surprisingly very much limited to ad hoc community programs or school health education. Unfortunately the former is difficult to capture over a large area like Texas. The latter proves difficult in the state of Texas as well since almost all of the school sexual education programs are abstinence only thus not providing enough variability to provide any true measurement.

Aside from outreach programs and prevention funding, another place to turn to operationalize prevention is using a proxy measure related to the social norms of an area. If social norms that promote risky behavior as normal and acceptable are ingrained in an area, this could lead to less focus on preventative measures. The formation and perpetuation of social norms can facilitate social interaction which in turn is a mechanism through which location and health can be tied. Culture and social norms are argued to be rooted in place (Mills et al., 2001; Gesler & Kearns, 2002). These can be by way of descriptive norms which are perceptions of an individual about their peers engaging in a specific behavior or injunctive norms which are perceptions about approval of a behavior (Cialdini, Reno, & Kallgren, 1990). Numerous studies with over a wide range of populations and ethnicities have reported positive associations between both descriptive and injunctive norms and HIV-risk behaviors, including condom use (Albarracin, Johnson, Fishbein, & Muellerleile, 2001; Hart & Peterson, 2004), use of shooting
galleries (Tobin, Davey-Rothwell, & Latkin, 2010), sharing injection equipment (Andia, Deren, Robles, Kang, & Colon, 2008), and sex exchange (Davey-Rothwell & Latkin, 2008). Understanding the spatial context of norms, of which there has been little research, could inform placement of interventions designed to change norms.

Thus to operationalize this county level prevention variable, a measure of social norms that affects the increase or decrease of HIV presence is appropriate. Intuitively something like the number of gay bars in an area would be useful. But again there is a limitation due to the size and scope of this study. More on limitations are discussed in chapter V. One measure that is available is the number of gay and gay friendly churches. This measure is perhaps even a better measure of risky behavior becoming normative due to the fact that whereas bars often offer an escape from the norms of an area in exclusion and exclusivity, the presence of a gay or gay friendly church suggests that social norms accepting this group and/or behavior has been mainstreamed.

A public list (Welcoming Gay Friendly Churches in Texas) of gay “affirming” churches provides a list of 215 gay or gay friendly churches in 39 of the 254 Texas counties. This website defines a gay affirming church as a church that “does not view homosexuality in and of itself as a sin and therefore they would welcome and treat a homosexual person no differently than any other person who walked through their church doors seeking Christ.” To create this variable (gaychurch) the number of gay affirming churches in a county was divided by the 2000 US Census population for that county and multiplied by 10,000. The result will serve as the potentiality measure for the county level (μ = .034, σ² = .10, N = 254; min = 0, max = .62).
Controls

Earlier it was mentioned that HIV in Texas looks primarily like a Black homosexual male problem. Many public health policies and HIV studies target and focus on these groups. One of the main purposes of this study is to show that while HIV is more likely in this population, the vulnerability of an area also plays a part outside these characteristics. Thus race, gender and homosexual populations will serve as controls in the following analysis. Race is operationalized by the Black population percentage of each ZCTA (perblk) ($\mu = 8.66$, $\sigma^2 = 13.77$, $N = 1864$; min = 0, max = 97.83). Gender (permale) is measured by the proportion of males in total ZCTA population ($\mu = .50$, $\sigma^2 = .04$, $N = 1864$; min = .32, max = .90). The homosexual measure is obtained by the percentage of same-sex households (perSSHH) per total households for each ZCTA ($\mu = .51$, $\sigma^2 = .47$, $N = 1864$; min = 0, max = 6.67). Although using this as a measure for homosexual population is not without problems, it serves an adequate available proxy measure for the time being (O’Connell & Lofquist, 2009). Two other control variables – a rural/urban dummy variable ($URDV$) and the proportion of unmarried households (perUMHH) – will also be included since sexual behavior in an area with a high unmarried population is assumed to be less monogamous while the urban/rural distinction is included to attempt to show that place vulnerability cannot merely be drawn along an urban/rural divide. Each of these variables has been obtained from the US 2000 Census SF1 and SF3.

Table 4 displays some descriptive statistics and the source for each variable used in the following analysis while Tables 5 and 6 provide the bivariate correlations. Table 5 displays that all the variables at the ZCTA level are significantly correlated to the new $HIVCount$ at the $p < .01$ level, with the exception of permale and perUMHH which were significant at the $p < .05$ level.
### Table 4

**Summary of Variables Used in this Analysis**

<table>
<thead>
<tr>
<th>Name</th>
<th>Measured by</th>
<th>N</th>
<th>μ</th>
<th>σ²</th>
<th>Range</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variable</strong></td>
<td>HIVCount # of newly diagnosed HIV cases in a ZCTA from 1999 - 2008</td>
<td>1864</td>
<td>21.93</td>
<td>54.24</td>
<td>Min = 0 Max = 1079</td>
<td>Texas DSHS (special request)</td>
</tr>
<tr>
<td><strong>ZCTA Level Independent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exposure</td>
<td>Ltrate Late tester rate per 100</td>
<td>1864</td>
<td>.006</td>
<td>.014</td>
<td>Min = 0 Max = .276</td>
<td>Texas DSHS (special request)</td>
</tr>
<tr>
<td>Capacity</td>
<td>Per1864pov Proportion of population age 18 – 64 in poverty</td>
<td>1864</td>
<td>.078</td>
<td>.052</td>
<td>Min = 0 Max=.357</td>
<td>2000 Census SF1</td>
</tr>
<tr>
<td>Potentiality</td>
<td>SrvPrDV Presence of HIV/STD Service provider (yes = 1)</td>
<td>1864</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Texas DSHS website</td>
</tr>
<tr>
<td><strong>County Level Independent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exposure</td>
<td>prisonDV Presence of a prison (yes = 1)</td>
<td>254</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>TDCJ website</td>
</tr>
<tr>
<td>Capacity</td>
<td>Fundper $ amount of public health assistance funding in the year 2000 per total county population</td>
<td>254</td>
<td>360.78</td>
<td>218.36</td>
<td>Min = 26.0 Max = 2078.21</td>
<td>Texas comptroller website</td>
</tr>
<tr>
<td>Potentiality</td>
<td>gaychurch # of gay churches in a county per 10,000 population</td>
<td>254</td>
<td>.03</td>
<td>.10</td>
<td>Min = 0 Max = .62</td>
<td>gaychurch.org</td>
</tr>
<tr>
<td><strong>Controls</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td>Perblk % of black population in ZCTA</td>
<td>1864</td>
<td>8.66</td>
<td>13.77</td>
<td>Min = 0 Max = 97.83</td>
<td>2000 Census SF1</td>
</tr>
<tr>
<td>Gender</td>
<td>Permale % of male population in ZCTA</td>
<td>1864</td>
<td>.50</td>
<td>.04</td>
<td>Min = .32 Max = .90</td>
<td>2000 Census SF1</td>
</tr>
<tr>
<td>Homosexual</td>
<td>perSSHH % of same-sex households per total households in ZCTA</td>
<td>1864</td>
<td>.51</td>
<td>.47</td>
<td>Min = 0 Max = 6.67</td>
<td>2000 Census SF3</td>
</tr>
<tr>
<td>Unmarried</td>
<td>perUMHH % of unmarried households per total households in ZCTA</td>
<td>1864</td>
<td>.04</td>
<td>.02</td>
<td>Min = 0 Max = .31</td>
<td>2000 Census SF1</td>
</tr>
<tr>
<td>Urban/Rural</td>
<td>URDV &gt;50% of total ZCTA pop is rural = 1</td>
<td>1864</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2000 Census SF1</td>
</tr>
</tbody>
</table>

The strongest correlations are found with the ltrate (exposure) variable and the perblk (control) variable. All of the correlations are in the expected direction with the exception of SrvPrDV.

Although significantly correlated, this predictor shows that when a service provider is present in a ZCTA, that ZCTA has a larger number of HIVCount. This may be due to the fact that service
providers select locations based on known locations of high HIV/AIDS presence and also because urban areas with high HIV/AIDS counts also have a larger variety of services as a result of a sufficiently large enough population to support that service.

Table 5

*Bivariate Correlations of ZCTA Level Predictors and Dependent Variable*

<table>
<thead>
<tr>
<th></th>
<th>HIVcNP</th>
<th>Lr100</th>
<th>PerPov164</th>
<th>SnPPOV</th>
<th>PerBk</th>
<th>Permale</th>
<th>PerSSHH</th>
<th>PerUMHH</th>
<th>URDV</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIVcNP</td>
<td>1</td>
<td>.397*</td>
<td>.168*</td>
<td>.330*</td>
<td>.366*</td>
<td>.053</td>
<td>.325</td>
<td>.049</td>
<td>-0.079*</td>
</tr>
<tr>
<td>Stg. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.021</td>
<td>0.00</td>
<td>0.034</td>
<td>0.00</td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
<tr>
<td>Lr100</td>
<td>.397*</td>
<td>1</td>
<td>.161*</td>
<td>.159*</td>
<td>.256*</td>
<td>.210*</td>
<td>.170*</td>
<td>-0.036</td>
<td>.172*</td>
</tr>
<tr>
<td>Stg. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
<tr>
<td>PerPov164</td>
<td>.168*</td>
<td>.161*</td>
<td>1</td>
<td>.202*</td>
<td>.170*</td>
<td>.004</td>
<td>.093*</td>
<td>-0.12*</td>
<td>.020</td>
</tr>
<tr>
<td>Stg. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
<tr>
<td>SnPPOV</td>
<td>.330*</td>
<td>.159*</td>
<td>.202*</td>
<td>1</td>
<td>.172*</td>
<td>.004</td>
<td>.093*</td>
<td>-0.12*</td>
<td>.020</td>
</tr>
<tr>
<td>Stg. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
<tr>
<td>PerBk</td>
<td>.366*</td>
<td>.256*</td>
<td>.170*</td>
<td>.172*</td>
<td>1</td>
<td>.000</td>
<td>.079*</td>
<td>.028</td>
<td>-0.063*</td>
</tr>
<tr>
<td>Stg. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
<tr>
<td>Permale</td>
<td>.053*</td>
<td>.210*</td>
<td>.036</td>
<td>.004*</td>
<td>1</td>
<td>.043</td>
<td>.056</td>
<td>.021</td>
<td>.039</td>
</tr>
<tr>
<td>Stg. (2-tailed)</td>
<td>0.021</td>
<td>0.00</td>
<td>0.118</td>
<td>0.876</td>
<td>0.987</td>
<td>0.001</td>
<td>0.066</td>
<td>0.521</td>
<td>0.003</td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
<tr>
<td>PerSSHH</td>
<td>.325*</td>
<td>.172*</td>
<td>.139*</td>
<td>.093*</td>
<td>.079*</td>
<td>0.043</td>
<td>1</td>
<td>0.004</td>
<td>0.030</td>
</tr>
<tr>
<td>Stg. (2-tailed)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.066</td>
<td>0.063</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
<tr>
<td>PerUMHH</td>
<td>.049*</td>
<td>.202*</td>
<td>.001</td>
<td>.012</td>
<td>.028</td>
<td>.005</td>
<td>.004</td>
<td>1</td>
<td>-0.307*</td>
</tr>
<tr>
<td>Stg. (2-tailed)</td>
<td>0.034</td>
<td>0.941</td>
<td>0.500</td>
<td>0.604</td>
<td>0.229</td>
<td>0.821</td>
<td>0.883</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
<tr>
<td>URDV</td>
<td>-0.079*</td>
<td>-0.026</td>
<td>0.008</td>
<td>0.020</td>
<td>0.063*</td>
<td>0.014</td>
<td>0.030</td>
<td>-0.307*</td>
<td>1</td>
</tr>
<tr>
<td>Stg. (2-tailed)</td>
<td>0.001</td>
<td>0.290</td>
<td>0.741</td>
<td>0.388</td>
<td>0.007</td>
<td>0.539</td>
<td>0.196</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed).**

* Correlation is significant at the 0.05 level (2-tailed).

Table 6 displays all the variables at the county level with HIVCount aggregated to the county level, with the exception of the TDCJ cases (this is due to the fact that prison presence is one of the predictor variables). As expected, the county-level exposure variable prisonDV is significantly correlated to the number of HIV Counts at the county level. The capacity measure public assistance health funding per person in a county (fundper) shows a significant ($p < .01$) negative correlation with the number of HIV counts as expected. Also, the presence of gay
churches in a county (gaychurch) has a weak, significant positive correlation with the number of HIV counts.

Table 6

Bivariate Correlations of County level predictors and Dependent Variable Aggregated to County (w/o TDCJ unit cases)

<table>
<thead>
<tr>
<th></th>
<th>HIV (no TDCJ - 3571)</th>
<th>PrisonsDV</th>
<th>Fundper</th>
<th>gaychurch</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV (no TDCJ - 3571)</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>.474**</td>
<td>-.083**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
<tr>
<td>PrisonsDV</td>
<td>Pearson Correlation</td>
<td>.474**</td>
<td>1</td>
<td>-.032</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.164</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
<tr>
<td>Fundper</td>
<td>Pearson Correlation</td>
<td>-.083**</td>
<td>-.032</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.164</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
<tr>
<td>gaychurch</td>
<td>Pearson Correlation</td>
<td>.253**</td>
<td>.288**</td>
<td>-.158**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
<td>1864</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

Statistical Method – Model Selection

Due to the nature of the HIVCount variable being overdispersed, estimates of a Poisson model although perhaps unbiased, may be inefficient with standard errors biased downward (Cameron & Trivedi, 1998; Long, 1997). There is a likelihood ratio test for overdispersion that examines the null hypothesis of alpha=0. This test statistic follows the chi-squared distribution with one degree of freedom. If the null hypothesis is rejected, a negative binomial model is preferred to a Poisson model. In general, a negative binomial regression does better with overdispersed data or where the data has a variance which is much larger than the mean.
As mentioned earlier 557 of the 1,864 ZCTAs had no new HIV counts from 1999 – 2008. Thus a zero inflated model must also be considered. “Zero-inflated models attempt to account for excess zeros. In other words, two kinds of zeros are thought to exist in the data, "true zeros" and "excess zeros". Zero-inflated models estimate two equations simultaneously, one for the count model and one for the excess zeros. This might be good if there are more zeros than would be expected by either a Poisson or negative binomial model” (UCLA Academic Technology Services).

Zero-inflated models handle overdispersion by changing the mean structure to explicitly model the production of zero counts. These models assume two latent groups. One is the always-zero group and the other is not-always-zero or sometime-zero group. Thus, zero counts come from the former group and some of the latter group with a certain probability. (Park, 2005)

Multilevel Zero Inflated Poisson

Raudenbush and Bryk (2002) state “in a sense, one of the basic problems of sociology is to relate the properties of individuals and properties of groups and structures in which the individuals function” (p xix). Multilevel modeling (MLM) is used when data is organized in a hierarchical structure. There are many advantages to using MLMs over other methods. One advantage is that error independence is not required. This is useful because often when analyzing areas, spatial autocorrelation occurs which violates this assumption. Another advantage is that treating data organized into hierarchical levels as if they are on the same level can result in statistical and interpretational errors. MLMs take these issues into account by allowing for means and IV-DV relationships to vary between higher level units (Tabachnick & Fidell, 2006). For example, measuring a county’s health budget as a neighborhood effect is not
A multilevel model however can provide a way to incorporate county level measures with neighborhood level effects.

Multilevel modeling is used due to the organization of data into two nested spatial scales – ZCTA and county. This technique improves on traditional OLS regression approaches to place effects on health due to MLMs ability to analyze higher level influences on lower level differences and due to OLS regression requiring independence of errors, which is violated by the hierarchical nature of the data. The use of a multilevel model allows for the development of place effects research in that it does not treat place as merely a container in which all aggregated variables are held. It is a statistical technique that can allow for leaving data at the justifiable scale at which human processes occur (Abler, Adams, & Gould, 1971). A multilevel model can provide a way to incorporate county level measures with neighborhood level effects as shown in Figure 14.

\[
\begin{align*}
\text{County}_1 & \quad \text{County}_2 & \quad \ldots & \quad \text{County}_k \\
\text{ZIP}_1 & \quad \text{ZIP}_2 & \quad \ldots & \quad \text{ZIP}_n
\end{align*}
\]

Figure 14. Layout of data used in this multilevel model.

Due to the nested structure of the data and the frequency distribution of the dependent variable, a multilevel zero inflated Poisson (MLZIP) model is appropriate with the current data (the statistical justification for this is provided in Chapter V). Fitting this into the current multilevel data structure produces the following.

Let \( y_{ij} \) denote the response variable and \( a_{ij} \) and \( x_{ij} \) be the vectors of covariates of sizes \((m+1) \times 1\) and \((q+1) \times 1\) for the logistic and the Poisson components respectively, from the \(j^{th}\) zip-
code in the $i^{th}$ county. Let $\alpha$ and $\beta$ be the vectors of parameters associated with the logistic and Poisson components. Assuming normally distributed random effects $w_i \sim N(0, \sigma_1^2)$ and $v_i \sim N(0, \sigma_2^2)$, we can define our mixed-effects ZIP model as:

$$P(y_{ij} = 0) = \varphi_{ij} + (1 - \varphi_{ij}) \exp(-\lambda_{ij})$$

$$P(y_{ij} = k) = (1 - \varphi_{ij}) \frac{\exp(-\lambda_{ij})\gamma_{ij}^{y_{ij}}}{\gamma_{ij}!}, k = 1, 2, \ldots$$

$$\log \left[ \frac{\varphi_{ij}}{1 - \varphi_{ij}} \right] = \alpha' a_{ij} + \sigma_1 \zeta_{1i}$$

$$\log(\lambda_{ij}) = \eta_{ij} = \beta' x_{ij} + \sigma_2 \zeta_{2i}$$

Note that the covariates in the logistic and Poisson components can be the same but do not have to be as is the case in our model. The two random-effects in this model $w_i$ and $v_i$ for the logistic and Poisson components are standardized as $\zeta_{1i} = w_i / \sigma_1$ and $\zeta_{2i} = v_i / \sigma_2$, allowing the variation in random-effects to be different corresponding to the logistic and Poisson components of the model. In this model $\alpha$ measures the change in conditional logit of the response with the observed covariates for zip-codes in each county described by the random-effect. Similarly, $\beta$ measures the change in the conditional log mean of the response with the observed covariates for zip codes in each county described by the random-effect. The above model, which is used to divide the place effects over two levels, can be implemented in SAS. To justify the MLZIP selection, a single level Poisson and zero inflated Poisson as well as a multilevel Poisson model is investigated to determine which model provides the best fit and is robust. The results and discussion of each of these models are presented and discussed in the next chapter.

This chapter provided an explanation of variables selected, the justification for the selection of ZCTAs and Counties as units of analysis, and the validation for choosing a
multilevel zero inflated Poisson model. Chapter V presents and discusses results for various model estimations to determine which provides the best fit and proves robust.
CHAPTER V

RESULTS AND CONCLUSION

Chapter IV gave the design and specification of the multilevel zero-inflated Poisson model and provided an explanation of the variables selected and the justification for selecting these measurements at the ZCTA and county levels. This chapter provides the results, analysis and limitations of the model that has been selected.

Returning to the study questions from chapter IV, the interest of this research is to provide support for separating exposure, potentiality and capacity into community (ZCTA) and structural (county) spatial units, show that population characteristics do not alone explain the variation in HIV counts per ZCTA, determine whether exposure, potentiality and capacity are all necessary in determining HIV counts per ZCTA, and determine how well this type of model fits throughout the state of Texas.

Results

Two-Level Vulnerability: A Multilevel Zero Inflated Poisson (MLZIP)

As mentioned in Chapter IV, a multilevel zero inflated Poisson model appears to be the obvious choice with the data structure and theoretical approach that this research has undertaken. This model can be implemented using proc nlmixed in SAS®. The syntax for the final model can be found in the appendix. To justify this choice a series of models have been selected. One of the main emphases thus far is the need to separate out place effects over multiple scales. Most research up to this point has been content with leaving these place effects all contributing at the same spatial scale. This type of regular Poisson model would attempt to measure count data using contributing factors measured or aggregated to the same unit of analysis while also not accounting for an inflated number of zeros in the data. A zero inflated Poisson (ZIP) model
would be analyzed if the data suggests that zero counts may be overdispersed as mentioned in the previous chapter. A multilevel Poisson model would be used in the case of count data that features nested predictors. Finally, a multilevel zero inflated Poisson (MLZIP) would combine both of these aspects (zero inflation and nested data) into the model that has been proposed here.

The results of all four models are found in Table 7.

Table 7

*Compared Estimates and Model Fit from Poisson, ZIP, ML Poisson and MLZIP*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model</th>
<th>Poisson</th>
<th>ZIP</th>
<th>ML Poisson</th>
<th>MLZIP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intercept</strong></td>
<td>1.7086***</td>
<td>1.8787***</td>
<td>.0306***</td>
<td>.3028</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0695)</td>
<td>(.0707)</td>
<td>(.2019)</td>
<td>(.2051)</td>
<td></td>
</tr>
<tr>
<td><strong>ZCTA Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lrtrate</td>
<td>6.9069***</td>
<td>4.0820***</td>
<td>1.7096***</td>
<td>.7132**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.1901)</td>
<td>(.2205)</td>
<td>(.2241)</td>
<td>(.2571)</td>
<td></td>
</tr>
<tr>
<td>Per1864pov</td>
<td>.0152***</td>
<td>.0257***</td>
<td>.0305***</td>
<td>.0332***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0010)</td>
<td>(.0011)</td>
<td>(.0011)</td>
<td>(.0012)</td>
<td></td>
</tr>
<tr>
<td>SrvPrDV</td>
<td>.8649***</td>
<td>.6255***</td>
<td>.7233***</td>
<td>.6471***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0115)</td>
<td>(.0117)</td>
<td>(.0122)</td>
<td>(.0123)</td>
<td></td>
</tr>
<tr>
<td><strong>County Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>prisonDV</td>
<td>1.4473***</td>
<td>1.1940***</td>
<td>1.0135***</td>
<td>.9741***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0153)</td>
<td>(.0153)</td>
<td>(.1885)</td>
<td>(.1840)</td>
<td></td>
</tr>
<tr>
<td>fundper</td>
<td>-.0003***</td>
<td>-.0003***</td>
<td>-.0005</td>
<td>-.0006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0000)</td>
<td>(.0000)</td>
<td>(.0004)</td>
<td>(.0004)</td>
<td></td>
</tr>
<tr>
<td>gaychurch</td>
<td>1.7952***</td>
<td>1.1354***</td>
<td>4.3772***</td>
<td>3.6617***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0391)</td>
<td>(.0407)</td>
<td>(.8095)</td>
<td>(.5535)</td>
<td></td>
</tr>
<tr>
<td><strong>Controls (ZCTA)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>perblk</td>
<td>.0169***</td>
<td>.0160***</td>
<td>.0089***</td>
<td>.0094***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0002)</td>
<td>(.0002)</td>
<td>(.0003)</td>
<td>(.0002)</td>
<td></td>
</tr>
<tr>
<td>permale</td>
<td>-1.5362***</td>
<td>-1.0303***</td>
<td>.7138***</td>
<td>.4911**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.1309)</td>
<td>(.1331)</td>
<td>(.1386)</td>
<td>(.1391)</td>
<td></td>
</tr>
<tr>
<td>perSSHH</td>
<td>.4752***</td>
<td>.5792***</td>
<td>.3710***</td>
<td>.4068***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0060)</td>
<td>(.0071)</td>
<td>(.0077)</td>
<td>(.0080)</td>
<td></td>
</tr>
<tr>
<td>perUMHH</td>
<td>6.2396***</td>
<td>5.5782***</td>
<td>3.1009***</td>
<td>2.9313***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.2645)</td>
<td>(.2608)</td>
<td>(.3115)</td>
<td>(.3112)</td>
<td></td>
</tr>
<tr>
<td>URDV</td>
<td>-2.024***</td>
<td>-.1498***</td>
<td>-.0795***</td>
<td>-.0725***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.0106)</td>
<td>(.0107)</td>
<td>(.0117)</td>
<td>(.0118)</td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>51672</td>
<td>42510</td>
<td>30565</td>
<td>28111</td>
<td></td>
</tr>
<tr>
<td>AICC</td>
<td>51672</td>
<td>a</td>
<td>30565</td>
<td>28111</td>
<td></td>
</tr>
<tr>
<td>BIC</td>
<td>51738</td>
<td>a</td>
<td>30611</td>
<td>28160</td>
<td></td>
</tr>
</tbody>
</table>

***<.0001; **<.005; *<.05; Note: Standard Error reported in parentheses; aAICC and BIC not reported in SAS® for ZIP model
Table 7 provides parameter estimates and significance on four different models. The parameter estimates for the Poisson model and the ZIP model are similar with each independent variable significant at the \( p < 0.001 \) value. Comparing these two model estimations with the ML Poisson displays several differences. First, the parameter estimate value for gaychurch is dissimilar than the values from the first two models. Second, the variable fundper is no longer significant. Third, the direction of permale changes from negative to positive. Finally the intercept changes in magnitude. The MLZIP model results also display some similarities with the ML Poisson model. The main difference is that the intercept is no longer significant.

While comparing parameter estimates, direction and significance over these four types of models can be useful, the real importance in comparing these four models is to gain insight into which of the four has the best fit, or, in other words, best predicts HIVCount. This is the purpose of Akaike’s Information Criterion (AIC) which is an index used to determine which competing models has the best goodness-of-fit (Akaike, 1974). According to the Cambridge Dictionary of Statistics (Everitt, 2006), it is defined as 

\[ -2L_m + 2m \]

Where \( L_m \) is the maximized log-likelihood and \( m \) is the number of parameters in the model. The index takes into account both the statistical goodness of fit and the number of parameters that have to be estimated to achieve this particular degree of fit, by imposing a penalty for increasing the number of parameters. Lower values of the index indicate the preferred model, that is, the one with the fewest parameters that still provides an adequate fit to the data. (p. 13)

Thus a smaller AIC indicates the preferred model. AIC will alone be used for model comparison here and below since it virtually equals the corrected AIC (AICC) in every model and is often preferred theoretically over the Bayesian Information Criterion (BIC) (Burnham & Anderson, 2004). Looking again at Table 7, two conclusions can be drawn. First, accounting for zero inflation is preferable (AICs for the non-zero inflated models are 51762 and 30565 versus 42510 and 28111 respectively for the zero inflated models). Second, a multilevel model is preferable to
a single level model (AICs for the single level models are 51672 and 42510 versus 30565 and 28111 respectively for the multilevel models).

The first study question and hypothesis from Chapter IV was as follows:

1) Is it necessary to separate the relative roles of community level effects and structural level effects in explaining the geographic variation in the occurrence of HIV/AIDS in Texas?

*Hypothesis:* Adding structural level effects to a community effects only model will improve the prediction power of HIV/AIDS counts per ZCTA.

Thus, the fact that the lowest AIC is found in the MLZIP model confirms that separating out the community (ZCTA) level effects and structural (county) level effects is necessary.

More than Population Containers: Added Value of Community and Structural Effects

As mentioned in Chapter IV, one of the main purposes of the current research is to determine whether or not the spatial variation of HIV/AIDS in Texas is determined by factors other than the male, Black and homosexual population characteristics of an area. In order to pursue this objective a model is first run using these predictors with perUMHH and URDV. Chapters III and IV went into to detail as to how place vulnerability characteristics measured by exposure, capacity and potentiality can be a useful way in determining how place affects health beyond merely population characteristics of an area. Again using the AIC, adding exposure measures to a controls only model is useful in ascertaining whether or not this set of measures adds value to the model—and so forth when again adding capacity and then potentiality measures. Output from all four of these models is displayed in Table 8.
Table 8

MLZIP Estimates Adding in Capacity, Exposure and Potentiality Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Controls Only</th>
<th>Controls + Exposure</th>
<th>Controls + Exposure + Capacity</th>
<th>Full Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>.1539 (.1155)</td>
<td>.2295 (.1294)</td>
<td>.4586* (.2170)</td>
<td>.3028 (.2051)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perblk</td>
<td>.0169*** (.0002)</td>
<td>.0159*** (.0002)</td>
<td>.0109*** (.0003)</td>
<td>.0094*** (.0002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Permale</td>
<td>.6955*** (.1121)</td>
<td>.0213 (.1287)</td>
<td>.4755** (.1456)</td>
<td>.4911** (.1391)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>perSSHH</td>
<td>.6098*** (.0067)</td>
<td>.5730*** (.0073)</td>
<td>.5377*** (.0075)</td>
<td>.4068*** (.0080)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>perUMHH</td>
<td>1.1983** (3259)</td>
<td>1.0745* (.3275)</td>
<td>1.7055*** (.3196)</td>
<td>2.9313*** (.3112)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>URDV</td>
<td>-.1457*** (.0118)</td>
<td>-.1512*** (.0118)</td>
<td>-.0901*** (.0118)</td>
<td>-.0725*** (.0118)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ltrate</td>
<td>2.6791*** (.2271)</td>
<td>.5959* (.2480)</td>
<td>.7132** (.2571)</td>
<td>.9741*** (.1840)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PrisonsDV</td>
<td>1.0612*** (.1985)</td>
<td>1.1045*** (.1985)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>per1864pov</td>
<td>.0477*** (.0011)</td>
<td></td>
<td>.0332*** (.0032)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fundper</td>
<td>.0009* (.0004)</td>
<td></td>
<td>-.0006 (.0004)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SrvPrDV</td>
<td></td>
<td></td>
<td>.6471*** (.0123)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gaychurch</td>
<td></td>
<td></td>
<td>3.6617*** (.5535)</td>
<td></td>
</tr>
</tbody>
</table>

**Note: Standard Error reported in parentheses**

The controls only model displays that all five control variables are significant predictors of HIVCount while the model’s AIC equals 32660. Adding exposure measures Ltrate and PrisonsDV to the model lowers the AIC to 32510 while keeping all predictors in the model significant except for permale. The third model adds in the capacity measures per1864pov and
fundper and again lowers the AIC to 30794 with all predictors again significant in the model. Finally, the potentiality measures SrvPrDV and gaychurch are added in further lowering the AIC to 28111 with fundper the lone remaining non-significant predictor. This full model is the MLZIP model used in the previous section.

Referring back to the second study question and hypothesis from Chapter IV:

2) Are exposure, potentiality and capacity variables each necessary to include in determining HIV counts per ZCTA in Texas even after controlling for population characteristics of a ZCTA?

Hypothesis: All categories should be included with exposure having a greater impact on HIV/AIDS counts than both potentiality and capacity.

This model confirms that adding exposure, capacity and potentiality measures as defined in this model to the population characteristics greatly increases the model fit. It also confirms that the exposure variables resulted in the greatest impact due to the largest change in AIC when added.

Since the full model has been confirmed as appropriate and necessary in determining the parameter estimates for HIV/AIDS as defined Chapter IV, a closer examination of this output is displayed in Table 9.

The estimates in Table 9 are the estimated Poisson regression coefficients for the model. The HIV/AIDS dependent variable is a count variable, and Poisson regression models the log of the expected count as a function of the predictor variables. To interpret the Poisson regression coefficient one can say that for a one unit change in the predictor variable, the difference in the logs of expected counts is expected to change by the respective regression coefficient, given the other predictor variables in the model are held constant. However these coefficients can also be interpreted as for a one unit change in the predictor variable, the difference in the expected counts is expected to change by $e^x$, where $x$ is the coefficient, given the other predictor variables.
in the model are held constant. For clarity purposes the latter interpretation is used below for the interpretation of Table 9.

Table 9

**MLZIP Results**

| Variable   | Estimate | Standard Error | t Value | Pr>|t| |
|------------|----------|----------------|---------|-----|
| Intercept  | .3028    | .2051          | 1.48    | .1411 |
| perblk     | .0094    | .0002          | 33.16   | <.0001 |
| permale    | .4911    | .1391          | 3.53    | .0005 |
| perSSHH    | .4068    | .0080          | 50.97   | <.0001 |
| perUMHH    | 2.9313   | .3112          | 9.42    | <.0001 |
| URDV       | -.0725   | .0118          | 6.16    | <.0001 |
| ltrate     | .7132    | .2571          | 2.77    | .0059 |
| PrisonsDV  | .9741    | .1840          | 5.29    | <.0001 |
| per1864pov | .0332    | .0012          | 28.06   | <.0001 |
| fundper    | -.0006   | .0004          | -1.47   | .1433 |
| SrvPrDV    | .6471    | .0123          | 52.73   | <.0001 |
| gaychurch  | 3.6617   | .5535          | 6.62    | <.0001 |
| a0         | -1.4714  | .0752          | -19.56  | <.0001 |
| s2u        | 1.3742   | .1668          | 8.24    | <.0001 |
| Inflation probability | .1867 | .0114 | 16.35 | <.0001 |

The intercept is the Poisson regression estimate when all variables in the model are evaluated at zero. In other words, for ZCTAs that do not contain an HIV/AIDS service provider and do not fall within a county that contains a prison, assuming all other variables in the model are zero, the expected count of newly diagnosed HIV/AIDS cases from 1999 to 2008 is $e^{.3028}$ or
1.35. Table 9 also shows the significant result \((p < .0001)\) for the variance of the random effect \((s2u)\) which confirms the need for including the county level variables in the model. Finally, the zero inflation control \((a0)\) and inflation probability measures are also significant \((p < .0001)\) which confirm the same results that Table 7 displays in that a zero inflation model is necessary.

Population Characteristics Results

The four ZCTA population characteristics are represented in Table 9 above which displays the Poisson regression estimates for the percentage of black \((\text{perblk})\), male \((\text{permale})\), same sex households \((\text{perSSHH})\), unmarried household \((\text{perUMHH})\) within a Texas ZCTA as well as the place characteristic as to whether that ZCTA is rural or urban \((URDV)\). All four population characteristics were positive predictors of HIVCount and significant. Looking at the characteristics individually, first, if a ZCTA increases perblk by one percentage point, the difference in HIVCount would be expected to increase by \(e^{.0094}\) or 1.009 counts, while holding the other variables in the model constant. Second, if a ZCTA increases permale by one percentage point, the difference in HIVCount would be expected to increase by \(e^{.4911}\) or 1.634 counts, while holding the other variables in the model constant. Also, if a ZCTA increases perSSHH by one percentage point, the difference in HIVCount would be expected to increase by \(e^{.3082}\) or 1.502 counts, while holding the other variables in the model constant. These results confirm what was shown in the descriptive data from Chapter IV – that HIV/AIDS in Texas tends to be more common in black, male, homosexual people and therefore as these populations increase within an area the new HIV/AIDS count from 1999 – 2008 increases in that area as well. Furthermore, if a ZCTA increase perUMHH by one percentage point, the difference in HIVCount would be expected to increase by \(e^{2.9313}\) or 18.752 counts, while holding the other variables in the model.
constant. Finally, a rural ZCTA without a prison or HIV/AIDS service provider is expected to have fewer HIV cases than those in urban areas, holding everything else in model constant. But again the argument of this research has been that there are additional factors outside of these population characteristics and a rural/urban dichotomy that also have an effect on HIV/AIDS within an area.

**Exposure Variable Results**

Two variables that represent exposure to HIV/AIDS are represented in Table 9 above which displays the Poisson regression estimates for the rate of late testers within the ZCTA population (ltrate) and whether the ZCTA is located within a county that has a prison (PrisonsDV). Both of these exposure variables were positive predictors of HIVCount and significant (p < .0001) as mentioned earlier is this chapter, these exposure variables together provide a better explanation of HIV counts per ZCTA than just population characteristics alone. Looking at the characteristics individually, first, if a ZCTA increases ltrate by one per 100 population, the difference in HIVCount would be expected to increase by \( e^{.7132} \) or 2.041 counts, while holding the other variables in the model constant. This finding may be the result of late testers – defined as those who are diagnosed with AIDS within one year of HIV diagnosis – unknowingly increasing an area’s exposure to HIV. As mentioned in the previous chapter, “Those unaware of their status are more likely to transmit the disease to others, resulting in missed opportunities for the prevention of new HIV infections” (p. 11) (Texas Department of State Health Services, 2010).

Second, a ZCTA within a county that contains a prison has an expected HIVCount that is \( e^{.9741} \) or 2.649 counts higher than those ZCTAs that lie within a county that does not contain a
prison, holding the other variables in the model constant. This finding is expected based on the research discussed in Chapter IV, however as mentioned before caution must be taken to conclude that the presence of prisons have an effect on the STD and HIV rates of a surrounding area rather than prisons being located into areas that area show a spatial concentration of disadvantage.

**Capacity Variable Results**

Two variables that represent a ZCTAs capacity to cope with an HIV/AIDS infection burden are represented in Table 9 above which displays the Poisson regression estimates for the percentage of the ZCTA population aged 18-64 that are in poverty (per1864pov) and the amount in dollars a county gives towards public health funding per capita for the ZCTAs located within this county (fundper). Only the poverty measure was a significant predictor ($p < .0001$) of HIVCount and as mentioned earlier in this chapter, these capacity variables together provide a better explanation of HIV counts per ZCTA than a model that contains both population characteristics and exposure variables. Looking at the characteristics individually, first, if a ZCTA increases per1864pov by one percentage point, the difference in HIVCount would be expected to increase by $e^{0.0332}$ or 1.034 counts, while holding the other variables in the model constant. As mentioned last chapter, the association between poverty and HIV/AIDS has been well documented because it not only restricts an area’s ability to cope with the burden of a disease such as HIV/AIDS, but because it also is a crucial factor in creating a spatial concentration of disadvantage (Buck & Gordon, 2004). The importance of this finding here is that this shows that area poverty is a significant contributor (albeit most likely endogenous) to an area’s HIV/AIDS counts even after controlling for the population characteristics of that area.
Second, although the fundper estimate was in the expected direction, the variable was not found to be a significant predictor of HIVCount in this model. This is potentially due to the fact that a more specific funding measure was unable to be obtained but further investigation with a more specific measure of funding would be needed to confirm this.

**Potentiality Variable Results**

Two variables that represent an area’s potential to prevent HIV/AIDS are represented in Table 9 above which displays the Poisson regression estimates for the presence of an HIV/AIDS Service Provider within a ZCTA (SrvPrDV) and the number of gay friendly churches per 10,000 within a county (gaychurch). Both of these potentiality variables were positive predictors of HIVCount and significant \( p < .0001 \), and as mentioned earlier in this chapter, these potentiality variables together provide a better explanation of HIV counts per ZCTA than a model that contains population characteristics plus exposure and capacity variables. Looking at the characteristics individually, first, a ZCTA that contains an HIV/AIDS Service Provider has an expected HIVCount that is \( e^{.6471} \) or 1.910 counts higher than those ZCTAs that do not contain one, holding the other variables in the model constant. This finding does not support that proximity to an HIV/AIDS Service provider serves as measure of prevention to HIV and may be the result of service providers selecting locations based on known locations of high HIV/AIDS presence and also because urban areas with high HIV/AIDS counts also have a larger variety of services as a result of a sufficiently large enough population to support that service.

Second, if a ZCTA resides within a county and that county increases the number of gay churches by 1 per 10,000, that ZCTA has an expected HIVCount increase of \( e^{3.6617} \) or 38.927, holding the other variables in the model constant. While the justification for the use of this
variable came by way of an attempt to understand the spatial context of norms, it also possible
that population density and proximity to urban centers would play a mediating factor on this
result. As mentioned in Chapter IV, future research is necessary to look not only into spatial
contexts of norms but also into other spatial variables that could affect the prevention of
HIV/AIDS.

Mapping Empirical Bayes Estimates to Display Model Fit across Texas

So far this study has shown that it is justifiable to separate the relative roles of
community level and structural level effects and to include exposure, capacity and potentiality
variables in a model that attempts to estimate the geographic variation of HIV/AIDS in Texas.
The final question of interest in this study is how well this model does in predicting areas of
Texas in particular, rather than just in Texas overall as Tables 7 – 9 show.

In order to show the prediction power of this model for each ZCTA in Texas, SAS®
version 9.2 allows for the construction of predictions for every observation in our dataset using
empirical Bayes estimation (Wolfinger). The syntax for this is included in the appendix. The
output from this gives several useful measures including the predicted HIV Counts for every
ZCTA, the standard error of each prediction, and the significance of each predictor based off the
t-value of each predictor and standard error. These values allow for several useful visualizations
when mapped using ArcGIS ® version 10.

Figure 15 displays the HIV Counts per ZCTA to give a general picture of the geographic
distribution of HIV/AIDS in Texas. This map shows that HIV/AIDS is more evident in the
eastern and urban areas of Texas and this is expected based on the population characteristics of
these areas as mentioned in Chapter IV.
Figure 15. Standard deviation of HIV/AIDS counts per ZCTA in Texas.

The actual (observed) values of new HIV/AIDS counts from 1999 – 2008 can be subtracted from the predicted values obtained from the model which allows for a visualization of which areas in Texas our model over predicts, which areas the model under predicts and which areas it does fairly well (Figures 18 and 19 use this approach for the Dallas and Houston areas). However since there are so many zero and one count ZCTAs, a better model fit visual representation is given by mapping whether the observed HIVCount value of the ZCTA fell within a predicted 95% confidence interval. Figure 16 shows that our MLZIP model does fairly
well in its prediction power for much of the rural areas in West Texas, though with many of these areas having 0 counts this is not much of a significant finding. Also, since \textit{HIVCount} is a count variable the upper and lower bounds of the confidence interval were rounded to the nearest whole number. Taking a closer look into the Dallas (Figures 17 and 18) and Houston (Figures 19 and 20) areas gives a clearer picture of how the model performs in these two urban areas.

\textit{Figure 16.} ZCTAs where observed HIV count fall within predicted confidence interval.
Figure 17. ZCTAs where observed HIV count fall within predicted confidence interval within DFW area.

Figure 17 above shows that much of the DFW area is not predicted within a 95% confidence interval. So while the overall results of the model shows that it does fairly well throughout the state of Texas, there are only a few ZCTAs in the DFW area where it does very well. Using a visualization such as this is beneficial since this shows that more research needs to be conducted in this area in order to examine if HIV exposure, capacity, or potentiality is affected differently here than in other areas of Texas. Still, the model is not useless for the DFW area as mapping the over and under predictions as in Figure 18 displays.
Figure 18. Differences between predicted and observed HIV/AIDS in DFW area.

It is interesting that in the Dallas area the model tends to over predict in the suburban areas of Denton, Collin and Tarrant counties but under predict in the downtown areas of these three counties whereas in Dallas county, the model grossly under predicts and grossly over predicts in ZCTAs that are right adjacent to each other. An opposite trend is also observed in that the further from the downtown Dallas area one goes the model tends to under predict.
Figure 19. Differences between predicted and observed HIV/AIDS in Houston area.

Turning attention to the Houston area displays large over estimations in pockets of Harris County and most of Walker County (where Huntsville prison, the largest prison in Texas – and the highest HIV Count in Texas – are located). There is a similar trend in Harris County as there is in Dallas County with a ring of under estimation around the downtown area and
overestimation beyond that. There also appears to be an even distribution of over-, under-, and well estimated ZCTAs throughout the rest of the counties.

Figure 20. ZCTAs where observed HIV count fall within predicted confidence interval.

Taking a look at the 95% confidence interval for the Houston area shows that the model does much better in the Houston area than in Dallas but still does not do extremely well. Again further research into the potential uniqueness of this area should be undertaken in order to determine how exposure, capacity, and potentiality should be operationalized.
Overall what Figures 16-20 display is that underestimated areas possibly portray areas in which better variable specification may be needed or that more investigation is required that would lead towards a greater understanding into the uniqueness of that area.

![Map of Texas with standard error of prediction colors](image)

*Figure 21. Standard error of prediction of HIV/AIDS in Texas.*

Overestimated areas may show that all the variables in the model are not necessary or that too much weight is given to a particular variable in that area (for example a ZCTA with a relatively small population lies within a county that has a gay friendly church). Since there is such a wide range of \( HIVCount \), another useful visualization is to use the standard error of prediction for each ZCTA. Therefore, Figures 21 portrays just this for the state of Texas. The standard error of
prediction per ZCTA from this model is the standard deviation of the sampling distribution which tells us the variability of the estimated parameters. In other words, Figure 21 displays that our model estimates HIV/AIDS rather well throughout the state of Texas with a pocket of larger standard error in west Texas (near Lubbock) and along the I-45 corridor between Dallas and Houston (and specifically nearer to Huntsville).

Taking a closer look into the Dallas (Figure 22) and Houston (Figure 23) areas gives a clearer picture of how the model performs in these two urban areas. The model does fairly well throughout the Dallas-Fort Worth area with the exception of mid and south Dallas. The model also shows to be efficient in its estimation of HIV/AIDS in the Houston area with the exception of a small area around downtown Houston as well as with the exception of again the Huntsville area in Walker County. It is interesting to note that the model does better in the areas other than the most urban areas. Maps like these may be useful to demonstrate the distinctiveness of areas when dealing with a disease, even a disease such as HIV/AIDS. In this case further research may be necessary to investigate what other factors are causing the spread of HIV/AIDS in the most urban areas than the ones that have been included in this model.

One way to use this data is to take a closer look into the data for each ZCTA with the highest standard error of prediction as in Table 10. For example, ZCTA 75247 is potentially over-predicted due to the abnormally high proportion of males combined with relatively high late tester and poverty rate. This warrants future research as to why this area has maintained an unpredictably low HIV/AIDS count. ZCTA 75219 is an outlier with the highest non-prison HIV counts in Texas while ZCTA 76469 needs further investigation as to the cause of the flagrant over-prediction. Maps such as Figures 22 and 23 are useful to find distinctiveness between areas
that can guide further investigation. Of interest for future research is why six of the ten highest standard errors all fall within Dallas County.

Table 10

Five ZCTAs with Highest Standard Errors of Prediction

<table>
<thead>
<tr>
<th>ZCTA</th>
<th>Predicted</th>
<th>Observed</th>
<th>Perblk</th>
<th>Permale</th>
<th>perSSHH</th>
<th>Per1864pov</th>
<th>Ltrate</th>
<th>Stderr</th>
</tr>
</thead>
<tbody>
<tr>
<td>75215</td>
<td>472</td>
<td>313</td>
<td>85.26</td>
<td>47.21</td>
<td>0.75</td>
<td>23.30</td>
<td>.048</td>
<td>9.56</td>
</tr>
<tr>
<td>75247</td>
<td>171</td>
<td>25</td>
<td>56.69</td>
<td>77.16</td>
<td>0.00</td>
<td>25.98</td>
<td>.276</td>
<td>9.86</td>
</tr>
<tr>
<td>77006</td>
<td>585</td>
<td>496</td>
<td>4.30</td>
<td>57.50</td>
<td>4.46</td>
<td>12.33</td>
<td>.082</td>
<td>15.29</td>
</tr>
<tr>
<td>75219</td>
<td>771</td>
<td>1079</td>
<td>8.48</td>
<td>58.54</td>
<td>4.43</td>
<td>12.25</td>
<td>.118</td>
<td>18.66</td>
</tr>
<tr>
<td>76469</td>
<td>43</td>
<td>0</td>
<td>0.00</td>
<td>48.48</td>
<td>6.67</td>
<td>30.30</td>
<td>0.00</td>
<td>22.05</td>
</tr>
</tbody>
</table>

Figure 22. Standard error of prediction of HIV/AIDS in DFW area.
Figure 23. Standard error of prediction of HIV/AIDS in the Houston area.

Referring back to the third study question and hypothesis from Chapter IV:

3) Does this model (community + structural level effects) accurately predict HIV/AIDS counts without spatial variation in Texas?

Hypothesis: The model will predict with greater accuracy urban area HIV/AIDS counts and will not provide as good a fit in rural areas.
Based on the maps above (specifically Figures 16 and 21), this hypothesis is not confirmed. This model actually does a better job of predicting with greater accuracy HIV/AIDS in rural and suburban areas than in urban areas. Further research is necessary in order to investigate what factors need to be taken into consideration in these urban areas. Using maps such as these are of great benefit to determine how a model can be improved as a researcher may use them as a launching off point into further research into areas that do not necessarily fit the mold of the model. Without them, a “table-only” approach leaves research to other perhaps less concentrated foci on how to improve a model.

Limitations and Considerations for Future Research

Two articles by leading researchers in this area have thoroughly discussed the limitations that the current research must consider (Cummins, Curtis, Diez-Roux, & MacIntyre, 2007; Diez-Roux, 2004). Thus it is necessary that each of these limitations be discussed and addressed here. First, Cummins and others suggest that in order to empirically address how place affects health, three things should be considered.

First, there is a need to “collapse the false dualism of context and composition by recognizing that there is a mutually reinforcing and reciprocal relationship between people and place” (p. 1835). This is in fact what is undertaken in this current research. By understanding the complexity of place as more than an “activity container,” the model proposed allows for the measurement of how place can affect individuals at multiple levels, in multiple ways. While this model allows for a more proper measurement of activity and population containers to measure composition or contextual effects, it also goes beyond this by allowing for a method to analyze the interaction between people and place. The limitation here is that only a “bare bones” model
was explored. Future research would have to include three items for a proper analysis. First, individual effects should be considered. As mentioned previously, a three-level model of the type proposed in Figure 10 has yet to be considered and thus the current simpler two-level approach was a necessary first step. Only by adding in individual effects can the interaction between place and people be fully understood. Second, more specific variables need to be included. Again due to the pioneering nature of this model only one readily available variable was selected for each node of vulnerability. Future models should consider the best variable or variables that can affect exposure, capacity and potentiality. Finally, interaction effects could be explored further to examine how the nested spatial structure interacts with each other. This would assist in not only determining how place is affecting an individual but how places are affecting each other.

Cummins and others go on that, second, there is a need to “recognize that ‘context’ and ‘place’ varies in time and space” (p. 1835). Although not analyzed in this particular model, the suggested approach in Figure 11 demonstrates precisely a structure that recognizes the need to incorporate time in this type of research. Diez-Roux (2004) suggests that incorporating time into place effect studies may also shed light on whether or not individual level variables are acting as confounders or mediators. Future research would add time to the first level of this hierarchical data structure.

Third, these authors conclude with the importance of “incorporat(ing) scale into the analysis of ‘contexts’ relevant for health. Understanding the appropriate level, from the local to the global, at which ‘contextual’ processes and actors operate as well as the spatial scale at which their impacts are expressed, is important in order to deliver effective ‘contextual’ policy interventions” (p. 1835). This is precisely the reason why a section is dedicated towards defining
the areas of study in chapter IV. Previous research has often aggregated place data to the most convenient scale. This current research justifies the reason for the scales selected. The limitation here however is the attempt to provide a landscape view of HIV/AIDS in Texas. Future research on this subject may need to scale down this approach to smaller areas of study which will provide a more reasonable opportunity to collect data and gain insight into the idiosyncrasies of selected areas (e.g. Dallas-Fort Worth vs. Houston).

Diez Roux’s (2004) responses to limitations must be considered next. In this research, Diez Roux readily admits that “estimation procedures for multilevel logistic, Poisson, and survival models are an area of ongoing research” (p. 1955). The main limitation that needs to be addressed is whether neighborhood or area effects can be analogized to an individual-level construct. However it is crucial to revisit two important themes throughout this research in answering this critique that Diez Roux (2004) and many others suggest. First, this present research is interested foremost in what place effects have on a particular area, not an individual. The outcomes of such a study are useful in public health understanding and preparation of programs and funding. At no point has any attempt been made to extrapolate these findings down to individual level outcomes. That being the case, future research could use such a model as the one suggested in Figure 10 to further the research that has already begun on how place can effect an individual outcomes. The model proposed here is useful in advancing this type of research by not limiting these place effects to one aggregated scale of analysis.

Diez Roux (2004) first states that it is important not to infer causality among the correlations found in neighborhood effect research. Much care has been taken to not make that leap here. She further points out that even small correlations do not mean something is not worth studying. It just means that better measures need to be undertaken instead of crude census proxy
measures in most research. This is a limitation in this research, but necessary in order to get a broad picture of HIV/AIDS in Texas. As Figure 16 displays, the model does not predict each area in Texas equally and further research can and should be undertaken to determine why these areas differ from one another with more specific and individualized measures. Finally Diez Roux (2004) identifies four “methodological obstacles” when performing place effect research using multilevel modeling.

The first of these how social stratification confounds comparisons of individuals between neighborhoods. Since an individual outcome is not measured here, some of this obstacle is removed. However there is also the concern that neighborhood socioeconomic status is confounding the differences in neighborhood effects of HIV/AIDS. Special care was taken to make sure that there was not high multicollinearity between the ZCTA poverty measure and the other variables in the model. Also, it was shown that other ZCTA potentiality and capacity measures still improve model fit even after considering poverty. Still, better and more specific measures should be used in future research to determine the true effects of poverty on HIV/AIDS at individual, neighborhood and structural levels.

The second obstacle Diez Roux (2004) mentions is “that neighborhood effects are by definition endogenous to the compositional characteristics of neighborhoods. Another way to state this is to say that neighborhood characteristics are determined by the individual characteristics of the residents” (p. 1956). Again, since individual characteristics are not considered in the current model, this obstacle is not as serious as it otherwise would be. However even if individual characteristics or outcomes were considered it would be hard to argue that the presence of prisons, service providers, or county level funding are endogenous measures of individuals within a specified area.
The third obstacle she responds to is that of extrapolation. She responds to the critique of whether people from two different neighborhoods are “exchangeable,” even after controlling for the individual-level covariates usually included in contextual or multilevel models. In the absence of the exchangeability assumption, one cannot conclude that the associations observed between neighborhood characteristics and health reflect a causal effect of neighborhoods on health… if there is little overlap between individual level SES characteristics across neighborhoods, adjustment for these variables may not lead to a valid estimate of the neighborhood effect because it will be based purely on extrapolation. (p. 1957)

This is a limitation of any research (not just multilevel models) that compares categories of any kind including this research and is a well-known limitation (Diez Roux, 2004). Future research should be cautious especially when adding individual level effects.

The final obstacle that needs to be considered is the stable unit treatment value assumption (SUTVA). “SUTVA implies that the response for a given unit depends on the treatment assigned to that unit but not on the treatment assigned to other units” (p. 1958). Part of this assumption can be alleviated by adjusting for spatial autocorrelation and other random effects as this model does. Still, Diez Roux (2004) mentions that these assumptions are more likely to be violated when dealing with contagion processes as the current research does and can be minimized after running sensitivity analysis. Future research should consider this as well.

One of the purposes of Chapter II was to explain how the area in which one lives can affect the spread of a contagious disease, even when that disease is primarily behavior driven. The main purpose of the current research was to develop a more robust measure for place effects for any disease. If a model such as this can be shown useful with behavior driven diseases the possibility of using this type of model for other infectious and even non-communicable diseases should be promising. Therefore even though there are many limitations to this research, the model has been developed that allows for more specific measures to be easily plugged in for any type of place effect research on any disease.
Conclusion

Even if area effects can be reduced to the impact of composition as some suggest (King, 1997; Sloggett & Joshi, 1994), justifying at which scale composition should be measured is still enhanced by this model. MacIntyre et al.’s (1993) answer to “Should we be focusing on places or people?” is often answered by including both. Moon et al. (2005) warn that multilevel models need to have large sample sizes at each level, theoretical justification for contextual factors, and use complex variance-covariance structures in order to fully represent the nature of place effects.

As shown in the previous pages, this dissertation has the potential to do all of this. This research endeavor is a unique opportunity to go beyond traditional methods of measuring place effects by expanding place to more than one measure and more than one scale. This is something that simply has not been done in previous research. While previous research using multilevel models has been used to inform public health policy, it has either focused on community level influence on individuals (Mosher, Deang, & Bramlett, 2003) or country differences (Or, Wang, & Jamison, 2005). This research overcomes the limitations of these approaches and advances knowledge by addressing the “missing middle” (Lobao & Hooks, 2007). Thus it makes suggestions to better inform public health policies by addressing neighborhood and structural effects at their proper scale. Only by using the methodology described in this dissertation can the relative roles of neighborhood and structural effects be measured while identifying the spatial variation of a disease such as HIV/AIDS.
SAS Syntax for MLZIP with EB estimates

```sas
proc nlmixed data=work.adam TECH=NEWRAP;
parameters b0=0 b1=0 b2=0 b3=0 b4=0 b5=0 b6=0 b7=0 b8=0 b9=0 b10=0 b11=0
a0=0 s2u=1;
/* linear predictor for the inflation probability */
linpinfl = a0;
/* infprob = inflation probability for zeros */
/* = logistic transform of the linear predictor*/
infprob = 1/(1+exp(-linpinfl));
/* Poisson mean */
lambda = exp(b0 + b1*ltrate+ b2*Per1864pov + b3*SrvPrDV + b4*permale + b5*perSSHH +
b6*PerBlk + b7*PrisonsDV + b8*fundper + b9*gaychurch + b10*perUMHH + b11*URDV + u);
/* Build the ZIP log likelihood */
if HIVcount=0 then
  ll = log(infprob + (1-infprob)*exp(-lambda));
else ll = log((1-infprob)) - lambda + HIVcount*log(lambda) - lgamma(HIVcount + 1);
model HIVcount ~ general(ll);
random u ~ normal (0, s2u) subject=cntyfip;
predict lambda out = ebpred;
estimate "inflation probability" infprob;
run;
```
REFERENCE LIST


Carpiano, R. M. (2008). Actual or potential neighborhood resources and access to them: Testing hypotheses of social capital for the health of female caregivers. Social Science and Medicine, 67(4), 568-582.


Delor, F., & Hubert, M. (2000). Revisiting the concept of 'vulnerability.' *Social Science and Medicine, 50*, 1557-1570.


