The Modifiable Areal Unit Problem (MAUP) and Health Disparities

Peter Martin Hayward

B.A., Indiana University of Pennsylvania, 2002

M.A., University of California—Santa Barbara, 2004

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut 2009

The Modifiable Areal Unit Problem (MAUP) and Health Disparities

Peter Martin Hayward, PhD

University of Connecticut, 2009

The modifiable areal unit problem (MAUP) is the idea that the interpretation of a geographical phenomenon within analysis depends on the scale and partitioning of the areal units that are imposed on the analysis. The problem is inherently linked to minority health disparities as differences in health outcomes by race or ethnicity are often reported using these areal units, which may lead to inaccurate statistical measures and spatial patterns of disparities. The purpose of this research was to investigate the MAUP and its potential impact on the interpretation of minority health disparities. Several models were developed to determine if the MAUP influenced the interpretation of minority health disparities in Connecticut, and how the effects were manifested statistically and spatially. The models highlighted that the MAUP was an important problem to be accounted for when investigating minority health disparities. The research also showed that preferred mortality area designs could be used to model minority health disparities more consistently across scales, especially when compared to actual district designs. By doing so, this research simultaneously contributes to geographic and public health studies. This research adds to the discipline of geography by outlining innovative models to examine the MAUP. Through these models, it is shown that minority health disparity studies need to take into account the confounding effects of the MAUP.

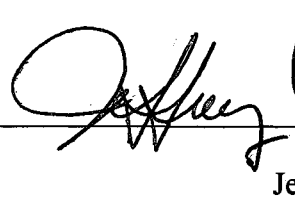
APPROVAL PAGE

Doctor of Philosophy Dissertation

The Modifiable Areal Unit Problem (MAUP) and Health Disparities

Presented by

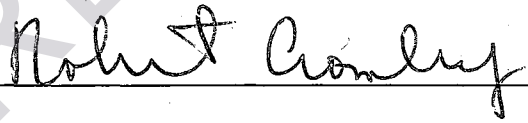
Peter Martin Hayward, B.A., M.A.

Major Advisor 

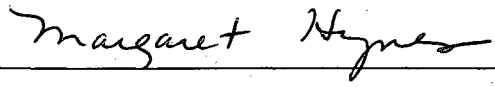
Jeffrey Osleeb

Associate Advisor 

Alexander Vias

Associate Advisor 

Robert Cromley

Associate Advisor 

Margaret Hynes

University of Connecticut

2009

Acknowledgements

I would like to thank my family, especially my parents Malcolm and Nancy Hayward. I truly appreciate their continued support, advice, and love, while I only hope that one day, I can be as good of a parent as they each have been to me.

In helping me prepare this dissertation as well as leading me through a rigorous doctoral program, I'd like to thank all members of my advisory committee: Professor Robert Cromley, Professor Alexander Vias, and Dr. Margaret Hynes. I have gained a wealth of knowledge and experience through time spent with each member. For being a great friend, I would also like to acknowledge Professor William Berentsen. Most importantly, I want to acknowledge and thank my advisor, Professor Jeffrey Osleeb. Jeff has offered steady guidance throughout the last four years and especially during this project, while providing me unwavering support even in periods of disagreement. We may soon be considered colleagues, but I will always view Jeff as a mentor.

Table Of Contents

List Of Tables

List Of Figures

1 Chapter One

Introduction And Statement Of Problem

1.1 Introduction

The modifiable areal unit problem (MAUP) is the idea that the interpretation of a geographical phenomenon within analysis depends on the scale and partitioning of the areal units that are imposed on the analysis. The problem is inherently linked to minority health disparities as differences in health outcomes by race or ethnicity are often reported using these areal units, which may lead to inaccurate statistical measures and spatial patterns of disparities. The purpose of this research was to investigate the MAUP and its potential impact on the interpretation of minority health disparities.

Individual level spatial data are integral to geographic research. Data collected at distinct locations are reflective of processes operating in space and time. Despite this relative importance, individual level spatial data in geographic research are often unsuitable for analysis and display. There are several possible reasons for this. For

example, there may be privacy concerns and confidentiality issues associated with using and displaying individual level spatial data. Also, these data may contain extreme outliers that can render the data inefficient. Outliers can influence the results of analyses and overemphasize trends that are not reflective of all observations. Finally, there may be an increased level of computational complexity in dealing with individual level spatial datasets, which often contain thousands or even millions of observations.

To combat these problems associated with using individual level spatial data, researchers aggregate the data to areas (also known as areal units or spatial units). Data aggregation can mask the confidential locations associated with individual level spatial data, thus alleviating the privacy concerns. The noise created by extreme outliers can be alleviated by combining those observations with other data points. This combination subsequently permits the analyses to use datasets with more observations, thereby creating more meaningful results. The aggregation of data also lessens the computational complexity of using large spatial datasets in that the problem size will be reduced. Guo and Bhat (2004) note that, "we collapse and aggregate observations in order to make the data more workable to the problem at hand, to gain understanding of the phenomenon in question, and to uncover patterns confounded by the noise typically found in observations."

Despite its usefulness, aggregating individual level spatial data comes at a cost in that the scale of analysis is changed. This results in a loss of information. A more serious problem is the idea that in many cases, the point information is combined and reported using arbitrarily defined areal units resulting in the generation of artificial spatial patterns (Heywood, 1998). This problem has been identified as the MAUP (Openshaw,

1984). The imposition of artificial areal units on a map may not only affect the map's appearance, but it may also influence the interpretation of the statistical measures and spatial patterns of a geographical phenomenon. The MAUP is the idea that these interpretations may differ due to the scale or to the partitioning of space. These are the two major components of the MAUP, and they are referred to as the scale effect and the zoning effect (Openshaw, 1984). The former refers to how changing the number of areal units imposed on the map can affect the interpretation of a phenomenon. The latter refers to how the partitioning of space within a map, while maintaining the same number of areal units, can affect the interpretation of a phenomenon.

In terms of resolving the MAUP, studies have found that the scale and zoning effects are inherently insoluble unless individual level spatial data are used for analysis and display. Despite this, geographers and others have derived unique methods using Geographic Information Systems (GIS) by which the scale and zoning effects can be tested. These include spatial autocorrelation procedures and exploratory spatial data analysis (ESDA), automated zoning procedures (AZPs), and computer-intensive resampling (a form of automated redistricting). Together, these methods have been used to help alleviate some of the problems associated with the MAUP (Martin, 2003; Openshaw, 1984; Openshaw and Rao, 1994; Openshaw, Alvanides, and Whalley, 1998).

The MAUP is especially problematic to health geography studies as epidemiologists, health practitioners, and geographers have all recognized the importance of the MAUP to health data, which are often spatially represented using areal units instead of individual points. However, most of this research has tried to identify the preferred areal unit to use when studying a particular health process or outcome without

first statistically or spatially determining if the MAUP is a problem (Gregorio *et al.*, 2005; Odoi *et al.*, 2003). Another dilemma is that several themes in health geography still remain unattended.

One such theme is the problem of minority health disparities, which is the study that defines differences in health outcomes according to race and ethnicity. Research on minority health disparities has recently gained significant interest as there now exists numerous governmental initiatives and programs geared towards reducing minority health disparities. These include the Healthy People 2000 and Healthy People 2010 initiatives, the Office of Minority Health and Health Disparities, and the National Center on Minority Health and Health Disparities. Most of these programs not only outline the issues and causes surrounding minority health disparities, but they also provide specific health disparity statistics. These statistics are reported using areal units such as towns, counties, or states. Such reporting makes the study of minority health disparities susceptible to the MAUP. However, to date, no research has specifically examined if and how the MAUP may influence the interpretation of minority health disparities. The purpose of this research was to investigate this relationship.

1.2 Problem Statement And Research Questions

This research addressed the MAUP and its potential influence on the interpretation of minority health disparities within the state of Connecticut. The main research question asked was the following:

- Is the MAUP an important problem to be taken into account when trying to understand minority health disparities?

Within this framework, several other questions were asked that were broken down into two categories: the influence of the scale effect on minority health disparities and the influence of the zoning effect on minority health disparities. The research questions related to the scale effect focused on how changing scale, that is the number of areal units imposed on a map, influenced the interpretation of minority health disparities:

- Are the minority health disparities recognized using the larger scales such as census block groups or census tracts still recognized using the smaller scales such as towns and counties?
- Do the numbers and locations of minority health disparity clusters change according to scale of analysis?

The research questions related to the zoning effect focused on how changing the partitioning of space while maintaining the same number of areal units imposed on a map influenced the interpretation of minority health disparities:

- Are the minority health disparities recognized using actual district designs such as census tracts or towns still recognized when using different district designs (referred as "mortality area designs" herein) that contain the same number of areal units?
- Do the numbers and locations of minority health disparity clusters change according to the partitioning of space?

While these questions addressed *if* the MAUP was a problem for minority health disparities and *how* it may have influenced the patterns of disparities, they did not approach the question of *what* could be done to overcome some of the MAUP's effects. With this in mind, the research asked:

- Can preferred mortality area designs be derived in order to display minority health disparities more effectively?

Finally, it must be noted that across the analysis, the term "district design" is used often. This refers to the scale and partitioning of areal units in a particular map.

1.3 Structure Of The Dissertation

This dissertation is comprised of eight chapters. Chapter 2 offers a definition of the MAUP. The scale and zoning effects are described, with the discussion focusing on methodologies that have been used in studies to resolve the two effects. In doing so, the MAUP's applicability to a range of topics is revealed. One such topic is health geography, and a large portion of Chapter 2 is devoted to reviewing the health geography studies that have attempted to overcome problems associated with the MAUP. Despite this work, it is shown that no study has explored the inherent link between the MAUP and minority health disparities. Given this relationship, this dissertation provides a review of minority health disparities. Particular attention is given to those studies that have focused on disparities at the local level, and the problems these studies have encountered. This review of the literature provides a foundation for the experiments performed in this dissertation.

Chapters 3 and 4 describe the methods and data that were used in this dissertation respectively. Chapter 3 outlines the research model that was utilized. Following a description of how minority health disparities were calculated, Chapter 3 highlights the unique statistical and spatial methods that were used to test if the scale and zoning effects influenced the interpretation of minority health disparities. The chapter concludes by

describing how the research model was slightly changed in order to provide examples of preferred mortality area designs to view minority health disparities. It is shown that these preferred mortality area designs can be used to display minority health disparities across scales with more consistency. Chapter 4 then describes the specific data that were used within the model. These data included Connecticut mortality data, state population data, census data, and census boundary files.

Chapter 5 summarizes and discusses the results from the scale effect analysis, while Chapter 6 outlines the results from the zoning effect analysis. Visual elements such as maps, charts, and tables are presented. Together, the results for both effects are outlined to show how the MAUP as a comprehensive problem can influence the interpretation of minority health disparities. Chapter 7 expands upon this discussion by showing the preferred mortality area designs that can be used represent minority health disparities in Connecticut. It is shown that despite their artificial nature, these designs can still be related to actual neighborhood data. Chapter 8 offers a conclusion to the dissertation with a discussion of implications of the study, the limitations experienced, and recommendations for possible future research directions.

2 Chapter Two

Literature Review

2.1 Introduction

In defining the literature surrounding the relationship between the MAUP and minority health disparities, several themes deserve recognition. First, a large number of geography studies have attempted to derive GIS-based methods to overcome the problems associated with the scale effect. These studies have offered ideas ranging from simplistic methods like displaying data at multiple scales to more complex methods such as automated zoning procedures (AZPs) or exploratory spatial data analysis (ESDA). Next, there exists a related base of geographic research which has focused on deriving methods to overcome the zoning effect. While AZPs offer one such method by which the zoning effect can be approached, more robust statistical analyses such as computer-intensive resampling (redistricting) procedures have also been used.

Recently, health geography studies have come to recognize the influence of the MAUP on a variety of health phenomena. Because of this, health geography research has