MEASURING SOCIOECONOMIC AND GEOGRAPHIC DEPRIVATION TO HEALTHCARE:
DEVELOPMENT OF A MISSOURI DEPRIVATION INDEX

A Thesis
presented to
the Faculty of the Graduate School
at the University of Missouri-Columbia

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts

by

JOHN M. THOMAS

Dr. Timothy C. Matisziw, Thesis Supervisor

JULY 2022
The undersigned, appointed by the dean of the Graduate School, have examined the thesis entitled

**MEASURING SOCIOECONOMIC AND GEOGRAPHIC DEPRIVATION TO HEALTHCARE: DEVELOPMENT OF A MISSOURI DEPRIVATION INDEX**

Presented by John M. Thomas,
a candidate for the degree of Master of Geography

and hereby certify that, in their opinion, it is worthy of acceptance.

________________________________________
Professor Timothy C. Matisziw

________________________________________
Professor Mark Palmer

________________________________________
Professor Matthew Foulkes

________________________________________
Professor Eileen Avery
DEDICATION

I would like to dedicate this work to my mother—who cultivated my child curiosity with an endless supply of National Geographic VHS tapes and books from our public library, and to my wife, Heather, and our beautiful daughter, Ruth.

I love you all, thank you.
ACKNOWLEDGEMENTS

I would like to thank the faculty of the Department of Geography at the University of Missouri–Columbia, with special thanks to my advisor, Dr. Timothy Matisziw, for his expertise and guidance throughout, to Dr. Matthew Foulkes for assisting me during the early brainstorming stages, and to Dr. Clayton Blodgett for being such a great mentor to me while working as his teaching assistant. I would also like to extend my sincere appreciation to all of my committee members, including Dr. Mark Palmer and Dr. Eileen Avery. Not only have I truly enjoyed having each of you as professors—you have all contributed to this work in various ways. Thank you.

Long live the Coyotes!
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ......................................................................................................................... ii

TABLE OF CONTENTS .......................................................................................................................... iii

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

LIST OF TABLES ...................................................................................................................................... ix

ABSTRACT ................................................................................................................................................ x

CHAPTER 1: INTRODUCTION .................................................................................................................. 1

1.1 Spatial Analysis and Public Health ......................................................................................................... 1

1.2 The Development of Methods in Health-GIS Research ........................................................................... 2

1.3 Spatial Aggregation & Perceived Patterns of Health ............................................................................. 3

1.4 Fundamental Problems in Spatial Analysis ........................................................................................... 4

1.5 Research Objectives ............................................................................................................................... 6

1.5.1 Overview of the Structure of this Study ............................................................................................ 7

CHAPTER 2: BACKGROUND LITERATURE ............................................................................................... 10

2.1 Socioeconomic Deprivation Indices .................................................................................................... 10

2.1.1 The Townsend Deprivation Index .................................................................................................. 11

2.1.2 The Carstairs Deprivation Index .................................................................................................. 12

2.1.3 Social Determinants of Health .................................................................................................... 13

2.1.4 The Health Opportunity Index .................................................................................................. 13

2.1.5 Area-Based Deprivation Indices .................................................................................................. 14

2.2 Other Variables Used in Constructing Deprivation Indices ................................................................. 16

2.2.1 The Gini Coefficient ..................................................................................................................... 16

2.2.2 Housing Condition Variables ...................................................................................................... 17
2.2.3 Environmental Variables

2.2.4 Migration and Immigration-Related Variables

2.3 Census Block Groups and Neighborhood Definitions

2.4 Concerns with Multivariate Indices, MAUP, and UGCoP

2.5 Other Considerations

CHAPTER 3: MEASURING DEPRIVATION IN MISSOURI

3.1 Motivation

3.2 Selecting and Adapting Established Methodologies

3.3 Areal Units of Analysis

3.4 Socioeconomic Indicators of Health

3.4.1 Median Household Income

3.4.2 Percent of Households Receiving Interest or Rental Income

3.4.3 Percent with High School Diploma & Percent with College Degree

3.5 Missouri Deprivation Index Calculation Methods

3.5.1 Calculating Target Variables

3.5.2 Block Group Exclusions

3.5.3 Standardizing & Normalizing Target Variables

3.5.4 Weighting Variables Prior to Calculating Index Scores

3.5.5 Final Deprivation Score Sums

3.6 Spatial Analysis Using ArcGIS Pro

CHAPTER 4: MEASURING ACCESS IN MISSOURI

4.1 Defining and Measuring Access

4.2 Measuring Access Using Network Analysis

4.3 Data for Measuring Access to Emergency Healthcare Services

4.4 Methods for Determining Healthcare Access Areas

4.5 Calculating Access at the Block Group Level
4.5.1 Intersection of Block Groups and Access Areas ................................................................. 55
4.5.2 Calculating Block Group Access Scores ............................................................................. 56

CHAPTER 5: RESULTS .................................................................................................................. 60

5.1 Deprivation Index Based on Socioeconomic Indicators ......................................................... 60
5.1.1 Hot Spot Cluster Analysis of Deprivation Scores ............................................................. 65
5.2 Access to Hospitals Providing Emergency Healthcare .......................................................... 70
5.2.1 Estimating the Total Population of Access Areas ......................................................... 72
5.3 Deprivation Index Based on Socioeconomic Indicators and Access ...................................... 79
5.4 The Impact of Including Geographic Access in the Deprivation Index ................................. 83
5.4.1 Block Group Z-Score Distribution Change ....................................................................... 83
5.4.2 Analyzing Block Group Index Rank Change ..................................................................... 85
5.4.3 Hot Spot Analysis of Deprivation Score Impact ............................................................... 90

CHAPTER 6: CONCLUSIONS .................................................................................................. 93

6.1 Summary of Findings ............................................................................................................. 93
6.2 Strengths ............................................................................................................................... 94
6.3 Limitations ............................................................................................................................ 95
6.3.1 Transportation Data for Urban Access Measures .......................................................... 95
6.3.2 Rural Clinics and Urgent Care Facilities ......................................................................... 96
6.3.3 Driving Times ................................................................................................................. 96
6.3.4 Qualitative Data ............................................................................................................. 96
6.4 Future Research Opportunities ............................................................................................. 97
6.4.1 Accessibility and Future Infrastructure Development ..................................................... 97
6.4.2 Modifying Index Components for Future Applications ................................................ 98
6.4.3 Measuring COVID-19 Pandemic Effects on Deprivation in Missouri .......................... 99

APPENDIX ............................................................................................................................. 100

REFERENCES ......................................................................................................................... 105
# LIST OF FIGURES

**Figure 1.** Example of the same data in the same place, aggregated two different ways.  
**Figure 2.** Example showing how indicator data can worsen issues related to MAUP. 
**Figure 3.** Locations of excluded block groups labeled with their 12-digit geo-codes. 
**Figure 4.** MDI scores for the greater St. Louis, MO metro region, by block group. 
**Figure 5.** Missouri DHSS designated hospital facilities, 173 in total. 
**Figure 6.** Hospital access areas, semi-transparent to reveal polygon overlap. 
**Figure 7.** Refined hospital access areas. 
**Figure 8.** An illustrated example of aggregating the access area polygons. 
**Figure 9.** Detailed view of access area polygons prior to being aggregated. 
**Figure 10.** Detailed view of access area polygons after being aggregated. 
**Figure 11.** A synopsis of the steps involved in refining access areas. 
**Figure 12.** Final output of hospital access areas shown with counties for reference. 
**Figure 13.** Example of block group access area coverage and percent calculation. 
**Figure 14.** MDI block group scores, classified into equally distributed quantiles. 
**Figure 15.** MDI scores for the Kansas City, MO metro area, by block group. 
**Figure 16.** MDI scores for the greater St. Louis, MO metro area, by block group. 
**Figure 17.** Kansas City deprivation divided east to west with racial dot density map. 
**Figure 18.** St. Louis deprivation divided north to south with racial dot density map. 
**Figure 19.** Hot spot analysis illustrating clusters of high and low deprivation scores. 
**Figure 20.** Hot spot analysis showing high & low MDI score clusters in Kansas City. 
**Figure 21.** Hot spot analysis showing high & low MDI score clusters in St. Louis. 
**Figure 22.** Access scores classified into quintiles shown with highway infrastructure. 
**Figure 23.** Access areas and major roads with darker shades representing lower access. 
**Figure 24.** A detailed view of a very limited access area in South-Central Missouri.
FIGURE 25. A detailed view of a very limited access area in southeast Missouri. ..........................77

FIGURE 26. A detailed view of a very limited access area in northeast Missouri. .........................78

FIGURE 27. A detailed view of very limited access in the central lakes region of Missouri. 78

FIGURE 28. MDI scores when access is included, classified using five EDQs. ................................80

FIGURE 29. MDI scores with access for Kansas City, MO, classified using five EDQs. ...............81

FIGURE 30. MDI scores with access for St. Louis, MO, classified using five EDQs. ........................82

FIGURE 31. Histogram of block group index scores without the access component. ..................84

FIGURE 32. Histogram of block group scores when access is included in the index. .................84

FIGURE 33. Deprivation index ranks not accounting for the access component. .......................86

FIGURE 34. Deprivation index ranks accounting for the access component. ............................87

FIGURE 35. Rates of rank increase or decrease when comparing both indices. .........................88

FIGURE 36. Block group population density & index rank change when including access. ......90

FIGURE 37. Hot spot analysis of change in index rank when access included. ..........................91
LIST OF TABLES

Table 1. Different results when the same data is used but aggregated differently ............ 5

Table 2. Example of MAUP, further complicated with the use of indicator data. .................. 22

Table 3. Variables used in relevant socioeconomic deprivation research......................... 27

Table 4. Neighborhood deprivation variables as outlined by Diez-Roux et al. (2004)........ 29

Table 5. ACS tables & variable information and initial target variable calculations ........... 33

Table 6. Census block groups excluded from calculations............................................... 35

Table 7. Block group access score calculation methods.................................................... 58

Table 8. Population estimates access areas if uniformly distributed within block groups. 74

Table 9. Population estimates for each access area if BG population is not apportioned.... 75

Table 10. The number of Missourians residing in each ranked access category ............... 76
MEASURING SOCIOECONOMIC AND GEOGRAPHIC DEPRIVATION TO HEALTHCARE:
DEVELOPMENT OF A MISSOURI DEPRIVATION INDEX

John M. Thomas

Dr. Timothy C. Matisziw, Thesis Supervisor

ABSTRACT

This study reviews the development and use of socioeconomic deprivation indices in health-related research and the variables and methods used in the construction of some of those indices. Relevant literature is summarized and the most significant methodological contributions to the topic are further described. Those methods are then closely adapted and used to create a block group level composite index of socioeconomic deprivation for the state of Missouri. In addition to the construction of the Missouri Deprivation Index (MDI), spatial analyses of access to emergency healthcare services in Missouri are performed, access scores for each block group in Missouri are generated and estimates of the population residing in those access areas are calculated. The access component is then incorporated into the deprivation index and used to explore the extent to which access may influence or be associated with measures of deprivation. This research will contribute to the literature by providing a background and a framework for the use of socioeconomic deprivation indices as potential explanatory variables in future research.
1.1 Spatial Analysis and Public Health

The use of geographic information systems (GIS) to visualize and analyze population health is long-standing. Development of analysis approaches underlying those implemented in GISs today began in 1830s London, fueled by public health concerns over outbreaks of Cholera. In a novel attempt to reexamine the spread of the disease, Dr. John Snow used spatial analysis to identify the sources of infection—discovering that cases were concentrated near sources of public water. Snow's findings dispelled the theory that Cholera spread through the air and provided the information necessary to affect policy change, slowing infections and saving countless lives (Musa et al. 2013).

Nearly two hundred years after Snow's groundbreaking work, spatial epidemiologists and health geographers are up against an ever-growing number of public health concerns. The effects of policy decisions on the health of populations are more apparent than ever—“complicated by long-standing and worsening health inequalities and a rapid spread of misinformation that needs to be countered” (Brownson et al. 2020). Many existing and newly emerging public health concerns are of interest to the public and policymakers alike (e.g., obesity, diabetes, substance abuse and addiction, the opioid epidemic, long-COVID, and an array of mental health issues). Given these growing concerns, it is increasingly important to understand what types of information provide the basis for—or inform public health policy decisions and what health research or socioeconomic knowledge might influence
such decisions. These questions highlight the need for GIS research to capture the patterns and scope of potential public health issues consistently and accurately—with little room for manipulation or misinterpretation.

When spatial data is subjected to different levels of aggregation or different scales, the visual representations of that data consequently change and can sharply impact its interpretation (Nelson and Brewer 2015). Interpretations of public health research can also be affected by uncertainty as to the true geographic context of specific health measures—as they relate to their spatial representation. Fundamental spatial analysis problems such as these (i.e., the Modifiable Areal Unit Problem and the Uncertain Geographic Context Problem) are compounded by the use of multivariate, socioeconomic health indices, and can directly affect the findings of a study and the interpretation thereof (Phillips et al. 2017).

Given that the interpretation of health-GIS research studies can play such a critical role in how public health policy decisions are informed—and affect the health outcomes of countless individuals, it is imperative that health-GIS research is conducted and presented consistently, with a reasonable and responsible level of scrutiny.

1.2 The Development of Methods in Health-GIS Research

In the late 1990s, public health research underwent a spatial turn that led to an expansive field of GIS-informed health research. Further developments in geospatial technologies provided researchers in the academic and private sectors—as well as the general public—access to real-time, complex analysis and visualization of health data (Rushton 2002; McLafferty 2003). Though vital to the success of the
field, advancements such as these have been accompanied by a burgeoning assortment of data sources, increasingly complex socioeconomic health indices, and various spatial analysis techniques (Krieger et al. 2003; Rothenberg et al. 2015).

As a result of the rapidly evolving approaches and research methods, health geographers began to express concern over the need to establish common terminology, develop standards for spatial analysis, increase data interoperability across different fields, and address consistency in spatial analysis techniques (Schuurman 2004; DiBiase et al. 2007; Richardson et al. 2013). Some argue that the failure to address those concerns may exacerbate the effects of well-known issues in spatial analysis, such as the Modifiable Area Unit Problem (Fotheringham and Wong 1990) and the Uncertain Geographic Context Problem (Kwan 2012).

In sum, concerns associated with aggregation, geographic context, and multivariate complexity persist in the literature as they can lead to disparities in results between different health-GIS research studies premised on the same or similar data sources. Given that this type of research can influence policymakers, affect major healthcare decisions and in turn, impact the health of entire populations, it is critical to encourage the investigation of such concerns, rather than hinder them (Rothenberg et al. 2015; Fletcher-Lartey and Caprerllli 2016).

1.3 Spatial Aggregation & Perceived Patterns of Health

Much like the power of visually-stimulating subliminal advertising, data visualization can have comparable influential effects. How spatial data is represented or embodied on a map can affect notions of particular places and influence concepts of—and decisions made about those places. For example, Mark
Monmonier writes “areal aggregation can have a striking effect on the mapped patterns of rates and ratios” (Monmonier 2018, 166). This can become problematic in health-GIS research when discrepancies in the experimental design of different studies using analogous data are significant enough that their results may be conflicting. This is not to suggest that health-GIS researchers intentionally manipulate their methods or analysis techniques to achieve specific, desired spatial patterns, but it is fair to say that much of this research is conducted without specific guidance as to how various types of data and socioeconomic health indices should be aggregated and presented to maximize accuracy and consistency between studies. Unfortunately, these issues are only intensified in research studies conducted with little concern for spatial aggregation and the analysis of multivariate health indices, which may directly impact the legitimacy of such research.

1.4 Fundamental Problems in Spatial Analysis

The Modifiable Areal Unit Problem, or MAUP, is a fundamental and often unavoidable issue in spatial analysis methods (Fotheringham and Wong 1990). MAUP refers to the fact that different representations of a region (and the summary of data therein) can bias visualization, analysis, and analytical outputs (Openshaw 1977). As demonstrated in Figure 1 (accompanied by Table 1), despite no change in the locations or values of the hypothetical illness data, simply modifying the shape of the areal units of analysis within the same region effectively changes how the data is summarized, leading to drastic differences in findings (Grubesic and Matisziw 2006; Matisziw, Grubesic, and Wei 2008). For example, the illness rate in Area x is 100 percent—given the areal units of analysis in Figure 1a—but plummets to 50
percent when subjected to a different representation of the areal units of analysis as shown in Figure 1b. Meanwhile, the overall illness rate in the encompassing area remains 33 percent, and the illness rate in Area y remains at 0 percent (see Figure 1 and Table 1).

![Figure 1](image.png)

**Figure 1.** Example of the same data in the same place, aggregated two different ways.

<table>
<thead>
<tr>
<th>When aggregating the same data by areal representation a (boundary a)</th>
<th>When aggregating the same data by areal representation b (boundary b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall illness rate</td>
<td>33%</td>
</tr>
<tr>
<td>Illness rate in area x</td>
<td>100%</td>
</tr>
<tr>
<td>Illness rate in area y</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table 1. Different results when the same data is used but aggregated differently.

In addition to MAUP, health researchers analyzing the effects of geographic-based variables on the outcomes or behaviors of a population should take into consideration another fundamental problem in spatial analysis: the Uncertain Geographic Context Problem, or UGCoP. According to Kwan (2012), UGCoP occurs...
“because of the spatial uncertainty in the actual areas that exert the contextual influences under study.” In other words, the problem occurs when the values measured in one geographic context are attributed to factors that exist in another geographic context (i.e., the cause of a measured value is situated in a geographic context separate from the one in which that value was observed). Although simplified, the basic concept underlying UGCoP is illustrated in Figure 1 (and Figure 2 in section 2.5) using the dotted Area exerting illness zones. These zones represent the geographic context in which the factors responsible for the illness reside. Although this area is revealed in the illustrations, the factors contributing to health behaviors and outcomes (illness, in this example), not to mention the area in which those factors exist are typically uncertain. The consequence of such uncertainty is, as Kwan puts it, “perhaps a major reason why research findings concerning the effects of social and physical environments on health behaviors and outcomes are often inconsistent” (2012).

1.5 Research Objectives

This study has two main research objectives. First, considering the abundant amount of health-GIS research utilizing multivariate health indices, this study investigates such research to develop a greater understanding of the multivariate indices most often used, their function or purpose, and the methodologies involved in their application. The research studies determined to be distinct, or those representing the most significant contributions to the topic, are cataloged and their methods described further. Of those methods, the most pertinent are adapted to construct a multivariate, socioeconomic deprivation index for the state of Missouri.
The second objective of this study is to measure access to emergency healthcare in Missouri and examine the extent to which access may correspond with—or have an effect on deprivation. To accomplish this, an access score is calculated for each block group based on time-distance measures to the nearest hospitals providing emergency and trauma care. Those scores are then integrated into the initial index and a new index is established. Finally, the indices are compared and analyzed for spatial dependency.

1.5.1 Overview of the Structure of this Study

Chapter two provides background information on the development and use of multivariate, socioeconomic deprivation indices, as well as an overview regarding the definition of “neighborhood” as it is used in the literature as an areal designation, and a review of other considerations and concerns in the use of such indices. The most fundamental deprivation indices—regarding the history of index development, as well as the indices most commonly used in current health-GIS literature—are described in more detail, including a brief overview of their components and construction methods.

Chapter three describes the methods used to develop a deprivation index for Missouri, beginning with the motivation for its construction and the process by which a methodology was chosen for the index calculations. An overview of the socioeconomic indicator data and the areal units of analysis used in the index construction follow. Finally, the selected methods are demonstrated, and additional spatial analyses are conducted and described.
Chapter four introduces the concept of access as it relates to emergency healthcare services and provides a brief background of the methods commonly used to calculate access. The access calculation methods used in this study are then described, including a detailed synopsis of the spatial analyses used to generate access areas (i.e., service areas) and determine block group access scores. Chapter four closes with a description of the process used to integrate access scores into the deprivation index.

Chapter five explores the results of the study in four sections. First, the overall geography of the Missouri Deprivation Index independent of access is described, including the results of hot spot cluster analyses. Next, access measures are described independently of deprivation scores, including a summary of the transportation infrastructure significantly affecting access to emergency healthcare, and an overview of calculating population estimates for the access areas. The results of the index based on socioeconomic indicators and access are then described. The chapter closes with a detailed summary of the impact access has on the index, including observed changes in block group z-scores and their distribution, a cluster analysis comparing index rank change and population density, and the results of a hot spot analysis of index rank change before and after access was included.

Chapter six provides the conclusion to the study. It begins with a summary of the processes involved in the research and the key takeaways regarding deprivation and access in Missouri. Next, research limitations are discussed, including the use of estimates in American Community Survey data, driving time data, hospital location data, and qualitative data. The chapter ends with a discussion on future research.
opportunities, including using this research alongside more recent data to explore changes in deprivation and access before and after the COVID-19 pandemic, as well as modifying index components to address differences in transportation between rural and urban areas.

An appendix section has also been included. This section includes additional map illustrations of the deprivation index with—and without an access component, displayed using alternate classification techniques to highlight the visual differences that can occur when the same data is embodied differently on a map. Details regarding those classification methods can be found in section 3.5.3.
CHAPTER 2: BACKGROUND LITERATURE

2.1 Socioeconomic Deprivation Indices

When investigating spatial patterns of health, researchers often rely on the use of socioeconomic health indicators as proxies for individuals or populations instead of actual health-outcome data (Krieger 1992). Over the last three decades, advancements in data collection, storage and access, as well as geospatial analysis (GIS) software have led to marked growth in the development and use of such socioeconomic health indices in spatial epidemiology. As Diez Roux and Mair (2010) state, “studies attempting to estimate the effects of neighborhood characteristics on health began to appear in the health literature in the late 1980s and early 1990s and grew exponentially over the next 10–15 years.”

Given the well-established associations between socioeconomic status (SES) and health outcomes, socioeconomic deprivation indices can be used in conjunction with additional health outcome data or used independently as prognostic, health indicator data for a given area. These indices are typically constructed from an amalgamation of data sourced from demographic, environmental, or socioeconomic domains and may include hundreds of extrapolative indicators of health (Rothenberg et al. 2015). Factors such as socioeconomic status and social deprivation, material deprivation, and poverty are considered principal indicators of health outcomes. As poverty can be both a cause and a consequence of poor health outcomes, poverty and health are inextricably linked and the most commonly
used indices in health-GIS research include socioeconomic variables (Krieger et al. 2003; Messer et al. 2006; Schuurman et al. 2007).

Overall, the growth, development, and repurposing of health-related indices have been sustained if not amplified in recent years as new socioeconomic health indices, as well as modified versions of existing indices, continue to emerge in health-GIS research (Juarez et al. 2014). The following sub-sections provide an overview of key socioeconomic deprivation and health-related indices that are commonly used or referenced in current health-GIS literature.

2.1.1 The Townsend Deprivation Index

The Townsend Deprivation Index (TDI) was developed to calculate differences in material deprivation across different geographic regions in the United Kingdom. The index uses four variables: unemployment, household overcrowding, non-home ownership, and non-car ownership. The variables are standardized and combined using equal weights to produce an overall deprivation index score for each areal unit. A greater score indicates greater deprivation relative to the surrounding areas (Townsend 1987). Although the TDI does not directly inform measures of health, it has proven to be an effective indicator of material deprivation and the associations between such deprivation and poor health outcomes are well established (Winslow 1951; Haan, Kaplan, and Camacho 1987; Murray 2006). Given the relationship between poverty and health, the TDI has seen widespread use on its own in public health research, and as a component in other socioeconomic deprivation and health-related indices (Jessop 1992; Prendergast, Beal, and
While the TDI was developed for use in the United Kingdom using UK census data and geographic units such as output areas (OAs), enumeration districts (EDs), and wards, equivalent measures can be constructed elsewhere—provided the same census variables are available and aggregated using similar geographic units.

2.1.2 The Carstairs Deprivation Index

Much like the Townsend Deprivation Index, the Carstairs Deprivation Index (CDI) was developed to provide deprivation scores for different areas across the United Kingdom. The CDI uses the same census variables as the TDI except for the non-home ownership variable—which is replaced with the low social class variable. The four UK census variables used are unemployment, household overcrowding, low social class, and non-car ownership. (Carstairs and Morris 1990; Messer et al. 2006). These variables are standardized and summed together using equal weights to produce deprivation index scores for each areal unit.

Although both the Townsend Deprivation Index and the Carstairs Deprivation Index were developed with the United Kingdom and UK census variables in mind, the indices have since been adapted for use in the United States using analogous data from the U.S. Census Bureau. These indices have been used widely by researchers in the United States—both independently, and as components alongside other variables in the construction of other socioeconomic deprivation indices (Kleinschmidt, Hills, and Elliott 1995; Prendergast, Beal, and Williams 1997; Locker 2000; Krieger et al. 2003; Messer et al. 2006; Schuurman et al. 2007). Given
the frequent application of these indices in both past and current literature, they have proven to be fundamental in the establishment of other socioeconomic deprivation indices.

2.1.3 Social Determinants of Health

The Social Determinants of Health (SDH) is a set of environmental factors that have been determined to affect the health and well-being outcomes of people across the United States. It is maintained by the Centers for Disease Control (CDC) and the U. S. Department of Health and Human Services’ Healthy People 2030 initiative but was developed throughout the 1990s in conjunction with the World Health Organization (WHO), and first introduced as a standardized set of metrics in the early 2000s (Marmot 2005). The Healthy People 2030 initiative maintains a dynamic set of 355 national health objectives all of which are meant to measure and improve the health and well-being of U.S. citizens over each decade¹. The SDH is derived from those objectives and grouped into five primary categories: economic stability, education access, health care access, neighborhood or built environment, and social or community environment.

2.1.4 The Health Opportunity Index

The Health Opportunity Index (HOI) was developed by epidemiologist and spatial analyst Rexford Anson-Dwamena at the Virginia Department of Public Health. It was quickly adopted and used by researchers with Ohio State’s Environmental Health Sciences Department and published in July of 2020—making it a relatively recent addition to the pool of such indices (Ogojiaku et al. 2020). The HOI uses a set of thirteen census tract-level input variables derived from the Social
Determinants of Health (SDH) to determine an overall HOI score which is then used to help locate the most vulnerable, at-risk communities (Ogojiaku et al. 2020). The thirteen metrics or indices amalgamated to determine the HOI score include affordability, income inequality, Townsend Deprivation, job participation, employment access, education, population churning, population-weighted density, segregation, food accessibility, walkability, access to care, and the environmental quality index.

2.1.5 Area-Based Deprivation Indices

There are many different area-based deprivation indices (ABDIs) currently used by health-GIS researchers—especially as an increasing number of states, counties and local city governments are developing their own area-based socioeconomic health indices based on community-specific data and issues, e.g., the Virginia Health Opportunity Index (VHOI), the Rhode Island Health Opportunity Index (RIHOI), the Utah Area Deprivation Index (the Utah ADI), the Vancouver Area Neighbourhood Deprivation Index (VANDIX), and the Ohio Opportunity Index (OOI), among others. These index scores are usually constructed with large-scale, small areal unit data in mind, and intended for use in specific areas. It is worth noting that among these types of deprivation indices, the terms deprivation, disadvantage, and disparity are often used interchangeably with respect to neighborhood socioeconomic measures and are not necessarily distinct (Kind et al. 2014; Zuelsdorff et al. 2020; Ursache et al. 2021).

Perhaps the most notable, and actively maintained area-based deprivation index is The Neighborhood Atlas® Area Deprivation Index. The ADI is produced by
the U.S. Health Resources and Services Administration (HRSA) and is based on methods that have been in development for over twenty years\(^2\). A major goal of the Neighborhood Atlas ADI is to “freely share measures of neighborhood disadvantage with the public”, providing national rankings of “census block group neighborhoods” based on socioeconomic deprivation measures\(^2\). Much like other deprivation indices mentioned here, the Neighborhood Atlas utilizes U.S. Census data and American Community Survey (ACS) data that draws from the domains of income, education, employment, and housing. The methods used to construct the Neighborhood Atlas were developed by Singh (2003) and further adapted for HRSA use by Kind et al. (2014). The data comprising the Neighborhood Atlas is maintained and tested by researchers at the University of Wisconsin-Madison School of Medicine and Public Health\(^2\).

It might be worth noting that no significant distinctions between ADIs, ABDIs, and NDIs (neighborhood deprivation indices) were found in the literature beyond their labels. These indices all aim to provide deprivation scores for a specified area, using the smallest feasible areal units, such as “neighborhoods”, census block groups, census tracts, or zip-codes, and most are generated using comparable data and similar calculation techniques (Kind et al. 2014; Zuelsdorff et al. 2020).

Most NDI calculations include anywhere from six to thirteen individual (i.e., non-compounded) variables that are generated using U.S. Census and ACS data. Those variables typically consist of calculated rates of fields such as: employed in management, unemployment, high school graduates, bachelor’s degree or higher,
households without telephone, households without plumbing, family poverty, public assistance, female-headed household, owner-occupied housing units, and housing units receiving interest, dividends, rental income, as well as median values, such as household income and home value. These variables are then standardized, weighted, and summed to provide an index score for each areal unit used (Kleinschmidt, Hills, and Elliott 1995; Roux 2001; Roux et al. 2004; Messer et al. 2006; Lian, Struthers, and Liu 2016). Details regarding standardization and weighting methods are covered in Chapter 3, sections 3.5.3 and 3.5.4, respectively.

2.2 Other Variables Used in Constructing Deprivation Indices

There are several other notable or commonly used variables in the construction of deprivation indices—many of which can be considered indices in and of themselves. These measures span a variety of categorical domains, including the built environment, housing conditions, environmental quality, income inequality, demographics, and immigration—to name a few. A small sample of those variables is briefly described in the following subsections.

2.2.1 The Gini Coefficient

The Gini coefficient is a statistical index commonly used to measure economic inequalities established by Yitzhaki (1979). The Gini coefficient measures the inequality of income in a given area, such as a neighborhood or block group. As stated by Ogojiaku et al. (2020) “income inequality is a critical variable to account for due to its correlation with health outcomes.” A Gini coefficient of 0.00 indicates perfect equality, i.e., all incomes in the population are the same. A Gini coefficient of 1.00—although not technically possible—would be interpreted as complete
inequality. For example, in a population of two individuals with one person having all income and the other having none, the Gini coefficient would be 0.50, in a population of three individuals with one person having all income, the Gini coefficient would be 0.66. This pattern would continue approaching but never ultimately reach a score of 1.0.

2.2.2 Housing Condition Variables

Variables such as incomplete plumbing and phone service available (Singh 2003; Krieger et al. 2003; Andrews et al. 2020), and household crowdedness (Townsend 1987; Messer et al. 2006; Carstairs and Morris 1990) are categorized in the domain of housing conditions. These variables are typically stronger indicators of material deprivation when applied at the census tract and county levels (Krieger et al. 2003; Singh 2003).

Due to confusion regarding the U.S. Census Bureau’s definition of telephone service in survey questions (i.e., landline service vs cellular service), the telephone survey questions were recently reworded and reimplemented\(^3\). However, there remains uncertainty as to the accuracy of this data and how its usage may be impacted in future index construction.

2.2.3 Environmental Variables

The domains of the built environment and environmental quality comprise a growing number of commonly used variables, including walkability, access to green spaces, air quality, water quality, and many more. The Environmental Protection Agency (EPA) provides two indices that significantly contribute to both of these
domains: the Walkability Index and the Environmental Quality Index (Ogojiaku et al. 2020).

The Walkability Index is available at the county level and is calculated using four variables: design (the built environment that promotes or impedes walking), diversity (the variety of land use and activities available in a walkable area), distance (the distance to pedestrian-accessible public transit), and density (the residential and employment concentration in an area) (Ogojiaku et al. 2020).

With respect to air quality, the Environmental Quality Index (EQI) is a measure of air pollutants calculated using data from the EPA’s National Air Toxins Assessment (NATA) and is available at the census tract level, higher pollutant levels result in a “higher possibility that the exposure to environmental conditions will result in a negative health outcome” (Ogojiaku et al. 2020).

2.2.4 Migration and Immigration-Related Variables

Variables from the domains of population, migration, and immigration are used in some socioeconomic deprivation indices—though far less often than variables from the domains previously mentioned. One reason for the lack of incorporation of such variables into deprivation indices may be due to difficulty in obtaining accurate migration and immigration estimates, especially when U.S. Census and American Community Survey data are used as primary data sources (Massey 2010).

Although accurate immigration data can be difficult to obtain, Cebrecos et al. (2018) argue that immigration is “a statistically significant component for socioeconomic deprivation.” The area-based deprivation index constructed by
Cebrecos et al. (2018) use two indicators directly related to immigration, including "recent foreigners born in a low-income country" and "foreigners whose father or mother were born in low-income country", although it is not clear how those variables are derived.

The neighborhood socioeconomic deprivation index constructed and tested in Lian, Struthers, and Liu (2016) uses "the percentage of the population who is foreign-born" as a socioeconomic indicator variable for census tract and block group geographies, however, it is used in the index as a racial component proxy and is by no means a direct measure of immigration.

To leverage the effects that net migration can have on neighborhoods and communities, the deprivation index developed by Ogojiaku et al. (2020) includes a variable to represent the Churning Index (i.e., the population churning rate). Population churning rates represent the overall rate of migration for a given area and are calculated by dividing the sum of in-migration and out-migration by the total population (Ogojiaku et al. 2020).

2.3 Census Block Groups and Neighborhood Definitions

Although the geographic units used for generating deprivation indices vary, most often they use Census tract or Census block group geographies. County-level data and geographies are also used, but less often (see Andrews et al. 2020). Compared to county-level geographies, block group level data can provide greater insight into specific neighborhood deprivation for a number of factors—simply due to their typical differences in scale.
The same may be said when comparing census tracts and block groups; differences in their geographic scale can translate into differences in data granularity. Census tracts—although also used as neighborhood proxies, are typically larger than residents' subjectively defined neighborhood boundaries (Campbell et al. 2009). As such, census block groups are frequently used as proxies for neighborhoods in geospatial research as they often represent the largest scale for which many socio-economic variables are reported.

Debates over what constitutes a neighborhood, both spatially-quantitatively, and socially or qualitatively are long-standing and continue to appear in the literature (Weiss et al. 2007; Campbell et al. 2009; Root 2012). Nevertheless, in much of the research that involves socioeconomic deprivation indices, census block group units are considered to be “the closest approximation to a neighborhood”1. Regarding area deprivation indices specifically, Kind et al. (2014) state, "‘Neighborhood’ is defined as a Census Block Group.” In some cases, block group geometries approximately align with local neighborhood designations, however, what constitutes a neighborhood can be dynamic and subjective, and there is often disagreement between residents as to the spatial boundaries of the neighborhoods in which they share (Campbell et al. 2009).

2.4 Concerns with Multivariate Indices, MAUP, and UGCoP

Regarding potential problems with multivariate indices, Schuurman et al. (2007) state that “despite their widespread use, comparably less attention has been focused on their geographic variability and practical concerns surrounding the Modifiable Area Unit Problem (MAUP) than on the individual attributes that make
up the indices”. This suggests that despite known problems associated with the increased complexity that accompanies amalgamated socioeconomic health indices, fundamental questions of geographic context and aggregation might still be overlooked.

As the variables that comprise these indices can differ greatly, the maximum effective scale of spatial analyses that use different health indices may also differ. For example, some applications of socioeconomic health indices use data that is reported both at the county level, as well as at the Census block group level (Andrews et al. 2020). Some indices may include data that is reported at larger scales with finer resolution, e.g., metropolitan statistical areas, communities, neighborhoods, block groups, or even at the individual level (Campbell et al. 2009). Moreover, in addressing the status of health indicator usage—specifically urban health indicators, Rothenberg et al. (2015) state:

“Problems of summarization and interpretation persist, and scale is one of the critical factors in developing indicators and indices. The type and number of indicators, how they are presented, transformed, and combined, the size of the targeted area, the relative placement of geographic units—all are scalable factors in the construction of an overall assessment. Indeed, the audience for the assessment is also scalable—from neighborhood groups to global agencies. Issues of scale, and the tension between multiple indicators and single statistics, suggest the need for a variety of alternative approaches.”
Accordingly, the importance of identifying an effective scale of indicator data, as well as respecting the limiting factors of data availability should be acknowledged. The usage and efficacy of the variables most commonly used in deprivation score calculations are discussed in great detail in detail in Messer (2006) and to a lesser degree in Lian, Struthers, and Liu (2016).

A hypothetical socioeconomic health indicator for a particular illness is depicted in Figure 2 to illustrate discrepancies in results that can occur between different studies using the same data but applying different areal units of analysis.

![Figure 2](image.png)

**Figure 2.** Example showing how indicator data can worsen issues related to MAUP.

<table>
<thead>
<tr>
<th></th>
<th>When aggregating the same data by areal representation a (boundary a)</th>
<th>When aggregating the same data by areal representation b (boundary b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall estimated illness rate</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>Estimated illness rate in area x</td>
<td>66%</td>
<td>75%</td>
</tr>
<tr>
<td>Estimated illness rate in area y</td>
<td>33%</td>
<td>0%</td>
</tr>
</tbody>
</table>

**Table 2.** Example of MAUP, further complicated with the use of indicator data.
Much like the example shown previously in Figure 1, despite no change in the locations of the hypothetical illness data or health indicator data, simply modifying how the data is geographically summarized leads to vastly different findings. For example, the estimated illness rate for Area x is 66 percent when given the representation of the areal units of analysis as shown in Figure 2a but increases to 75 percent when represented by the areal units of analysis as shown in Figure 2b. At the same time, the indicated illness rate for Area y is 33 percent when represented by the areal units of analysis in Figure 2a, but drops to 0 percent when it is represented by the areal units of analysis in Figure 2b.

As demonstrated in Figure 2, the findings and interpretations of spatial analysis may largely be contingent on the interactions between multiple methodological components and data selection, and can easily produce conflicting results. Moreover, both Figure 1 and Figure 2 depict scenarios in which none of the illness or indicator data represent the areal units from which they originated, i.e., the area exerting the illness. Although overly simplified, this represents the problem surrounding uncertainty in the geographic context of the data measured, i.e., UGCoP.

2.5 Other Considerations

There seems to be little consensus on when to use variables that indicate positive associations to socioeconomic deprivation and when to use variables that indicate negative associations to socioeconomic deprivation, i.e., reverse-coded variables (Diez-Roux et al. 2001; Roux et al. 2004; Andrews et al. 2020; Lian, Struthers, and Liu 2016). Many indices include a variable such as the percent of households below the poverty line—a measure of poverty, while also including a
reverse-coded variable such as the percent of households receiving interest or payment from net rental income, dividends, or investments—a measure of wealth. Moreover, some indices measure socioeconomic deprivation by only analyzing the inverse scores of negatively associated variables, e.g., only analyzing the inverse of wealth (Roux et al. 2004). The latter practice—measuring the inverse of wealth indicators—is common in the literature (Roux et al. 2004; Henderson et al. 2005; Nordstrom et al. 2007; Rudolph et al. 2014) and is also the practice applied in this study. Although there is ample opportunity for further investigation into this observation, it was only included here as a discussion point so as not to be overlooked.
3.1 Motivation

As mentioned in the previous sections, neighborhood, or area-based deprivation indices provide insight into socioeconomic disparities that often run parallel with disparities in health access and health outcomes. One of the objectives of this study is to review different ways in which socioeconomic deprivation is measured, identify which index calculation methods have been frequently recognized and applied in health-GIS research, and address the extent to which those indicators and their calculation methods may influence the assessment of deprivation.

There are reasons why analyzing neighborhood level (block group) data within the context of an entire state is less common than analyzing census tract or county-level data (e.g., data collection & availability issues, smaller sample sizes, greater margins for error, etc.). Also, neighborhood characteristics may vary dramatically over a relatively large area. However, there are many common threads connecting neighborhoods located within the same state, regardless of their proximity to one another or their distinctness (e.g., political redistricting, public health policy, social affiliations, assistance programs, historical context, cultural connectedness, and even sports rivalries). As such, there may be even more to gain from establishing such an index. In recent years, researchers have constructed deprivation indices for a number of states, including Rhode Island, Virginia, Utah, and Ohio, to name a few (Kim 2018; Ogojiaku et al. 2020; Knighton et al. 2016;
Fareed et al. 2020). To date, no such socioeconomic deprivation scores have been produced for the state of Missouri at the neighborhood level (i.e., block group level), and including a healthcare access component in the deprivation index could provide an innovative technique to shed light on the different challenges that rural and urban populations face regarding healthcare access and socioeconomic deprivation.

3.2 Selecting and Adapting Established Methodologies

While reviewing the literature relevant research articles on socioeconomic deprivation indices were documented and summarized. Of the roughly fifty articles relating to the spatial analysis of SES and health that utilized multivariate indices, twelve articles applied methods using deprivation indices deemed plausible for creating a neighborhood deprivation index at the block group level for Missouri. Variables used in those twelve articles are summarized in Table 3.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>Percentage of persons age 25+ with high school diploma or higher</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of persons age 25+ with bachelor's degree or higher</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of population without high school education</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Income,</td>
<td>Median household income in the past 12 months</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Wealth</td>
<td>Percentage of household with low income</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Poverty</td>
<td>Percentage of households below the federal poverty line</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of households receiving interest, dividend, rental income</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of households earning &lt;$30,000 per year</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of families receiving public assistance income, past 12mos</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of population below federal poverty line</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of households with no vehicle</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Median household value</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of female-headed households with children under 18</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Demographics</td>
<td>Percentage of population in working class</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of persons age 16+ employed in mgmt, science, or arts</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of males in management</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of females in management</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Employment &amp; Occupation</td>
<td>Percentage of males in professional occupations</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of females in professional occupations</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of persons age 16 and older that are unemployed</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of males no longer in work force</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Housing &amp; Residential Stability</td>
<td>Percentage of households that are owner-occupied</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of households that are renter-occupied</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of vacant housing units</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of renter or owner costs in excess of 50% of income</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of households with no phone</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of households with no kitchen</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of households with incomplete plumbing</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of households with more than 1 person per room</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of boarded up housing units</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of persons living in the same housing for at least 5yrs</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of non-Hispanic blacks</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of the population that was less than 18 years of age</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>Percentage of residents aged 65 or older</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
The most commonly used and well-established methods were determined based on a.) the frequency of other peer-reviewed research articles applying those methods, b.) the overall usage and most recent applications of those methods, c.) the number of citations, and d.) the fewest variables spanning multiple domains.

The methods for calculating socioeconomic deprivation developed by Diez-Roux et al. (2001; 2004) were also applied in Henderson et al. (2005), Nordstrom et al. (2007), Rudolph (2014), and applied to an extent in Andrews et al. (2020). Aside from their common use, these methods also use six variables that are obtainable at the block group level—which may be beneficial given that there are known issues that accompany complex indices containing many compound variables (Rothenberg et al. 2015). Ultimately, the methods used by Diez-Roux et al. (2004) were chosen as the basis for constructing the current Missouri Deprivation Index.

In Diez-Roux et al. (2004), deprivation scores are calculated using a refined pool of six socioeconomic variables acquired from the American Community Survey (ACS) 5-year estimates: median household income; median value of housing units; percentage of households receiving interest, dividends or net rental income; percentage adults that finished high school and higher; percentage adults that completed a 4-year college degree and higher, and percentage of persons in management, professional, and specialty occupations. A list of the variables used in Diez-Roux et al. (2004) can be found in Table 4. As most of these variables must be computed from multiple fields in the ACS data tables, this study refers to these six variables as the target variables.
Table 4. Neighborhood deprivation variables as outlined by Diez-Roux et al. (2004).

3.3 Areal Units of Analysis

The smallest areal unit for which all the specified variables are available from the ACS data is the Census block group unit. Block groups are frequently used as proxies for neighborhoods in geospatial research as they often represent the largest scale for which many socioeconomic variables are reported (Kind et al. 2014; Zuelsdorff et al. 2020). In some cases, block group geometries do align approximately with local neighborhood designations (Campbell et al. 2009).

Compared to county-level data, block group data can provide greater insight into specific neighborhood deprivation for several variables. The geographic unit used for generating deprivation indices can vary, but most often these indices use county, Census tract, or block group geographies. Nothing was found in the literature that indicates these variables are less effective indicators of deprivation when aggregated at the census tract or block group level. Thus, unlike the Neighborhood Deprivation Index of Andrews et al. (2020) which uses ACS data aggregated at the county level, a Missouri Deprivation Index (MDI) was generated.
using block group level data, and aggregated at that same level, as done by Diez-Roux et al. (2001; 2004).

Block group boundaries for the state of Missouri were downloaded from the U.S. Census Bureau’s TIGER/Line® Shapefiles website. Since this research is focused on conditions of deprivation in Missouri prior to the potential effects of the ongoing COVID-19 pandemic, this study used 2019 boundary data for the state, county, and Census tract and block groups geographies to coincide with the ACS 2019 5-year estimate data.

3.4 Socioeconomic Indicators of Health

As described earlier, a wide range of socioeconomic indicators of population health have been considered. Based on the methods used by Diez-Roux et al. (2004), six socioeconomic variables from the ACS 2019 5-year estimates were used in the MDI calculation. These variables have been determined to be ideal indicators of deprivation at the neighborhood level (Diez-Roux et al. 2001; Roux et al. 2004) and some are described here in more detail in the following subsections.

3.4.1 Median Household Income

Income measures are key indicators of overall economic conditions; the income information obtained by the ACS is used to determine poverty status, assess economic well-being, and evaluate the need for financial assistance programs. Household income data originated from questions 43 and 44 of the 2019 ACS. This variable represents the sum of reported amounts for the following: salaries or wage incomes; net self-employment incomes; interest, dividends, net rental or royalty incomes; incomes from estates and trusts; Social Security or Railroad Retirement
incomes; Supplemental Security Incomes (SSI); public assistance or welfare payments; retirement, survivor, or disability pensions; and all other incomes (ACS Subject Definition List, 2019). The ACS provides these household income figures as the total for all individuals reported in a household that are 15-years old or older, regardless of their relation to the householder (ACS Subject Definition List, 2019).

3.4.2 Percent of Households Receiving Interest or Rental Income

Of the eight different income categories reported by the ACS, Interest, dividends, or net rental income includes “interest on savings or bonds, dividends from stockholdings or membership in associations, net income from rental of property to others and receipts from boarders or lodgers, net royalties, and periodic payments from an estate or trust fund” (ACS Subject Definition List, 2019). This ACS measure is a key indicator for use in the domain of income, wealth, and poverty, as described by Diez-Roux et al. (2001).

3.4.3 Percent with High School Diploma & Percent with College Degree

Educational attainment is frequently used as a key indicator of overall socioeconomic status, deprivation and opportunity (Krieger et al. 2003). Both variables used here—the percent of adults that completed high school or higher, and the percent of adults that completed a 4-year college degree or higher, were gathered from answers to question 11 on the 2019 ACS and were tabulated for all individuals over the age of 18 years.

The total number of those who completed high school or higher includes all individuals who received a high school diploma or equivalent, such as a G.E.D. (General Educational Development), as well as those who pursued post-secondary
tracks but have not received a degree and all others who have received an associate’s, bachelor’s, master’s, professional, or doctoral degree.

The total number of individuals completing a bachelor’s degree or higher includes all of those who completed a four-year college degree (i.e., bachelor’s degree), as well as those who have completed a master’s, a professional, or doctoral degree, but does not include those who have only completed an associate degree.

3.5 Missouri Deprivation Index Calculation Methods

3.5.1 Calculating Target Variables

Tables containing the relevant ACS 2019 5-year estimate data for all Missouri block groups were downloaded from the U.S. Census Bureau data portal as CSV (comma separated values) files and imported into Tableau Desktop Pro v2021 statistical software. Using Tableau, tables were generated from the CSV files and values from the required block group fields from each table were compiled. Table 5 contains information on the ACS tables used, as well as the initial basic calculations applied to acquire the target variables.
<table>
<thead>
<tr>
<th>Target Variable</th>
<th>Data source details &amp; calculations</th>
</tr>
</thead>
</table>
| Percent employed in management professions | **Table 5. ACS tables & variable information and initial target variable calculations.**

**Table 5. ACS tables & variable information and initial target variable calculations.**
3.5.2 Block Group Exclusions

As defined by the U.S. Census Bureau, block groups (BGs) typically contain between 600 to 3,000 people, still, this definition is not a requirement, and many block groups report zero residents. Although block groups can be added, removed, or altered over time they don’t always retain a population. For example, a block group area that previously reported residential features and a population may now be the site of an airport, a nature reserve, or an industrial zone. In 2019, the U.S. Census reported 4,506 block groups for the state of Missouri. Of those, ten reported zero residents, households, or families (less than 2 percent of all Missouri block groups). As those block groups contain zero residents—let alone neighborhoods—they were omitted prior to any calculations or analysis, leaving 4,496 block groups remaining for analysis. If included, the zero-population block groups could significantly skew calculations and the resulting index scores. Table 6 provides the list of excluded block groups and Figure 3 shows where these block groups are located. Some examples of what is located in the excluded block groups are the St. Louis Lambert International Airport, Kansas City International Airport, the professional sports complex near Kansas City that includes both Kauffman Stadium and Arrowhead Stadium, Swope Park in Kansas City, Forest Park in St. Louis, as well as several state-managed conservation areas.
Table 6. Census block groups excluded from calculations.

<table>
<thead>
<tr>
<th>State FIPS</th>
<th>County FIPS</th>
<th>Census Tract</th>
<th>Block Group</th>
<th>Block Group Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>027</td>
<td>070300</td>
<td>6</td>
<td>Block Group 6, Census Tract 070300, Callaway County, MO</td>
</tr>
<tr>
<td>29</td>
<td>095</td>
<td>980101</td>
<td>1</td>
<td>Block Group 1, Census Tract 980101, Jackson County, MO</td>
</tr>
<tr>
<td>29</td>
<td>095</td>
<td>980802</td>
<td>1</td>
<td>Block Group 1, Census Tract 980802, Jackson County, MO</td>
</tr>
<tr>
<td>29</td>
<td>095</td>
<td>989100</td>
<td>1</td>
<td>Block Group 1, Census Tract 989100, Jackson County, MO</td>
</tr>
<tr>
<td>29</td>
<td>095</td>
<td>989200</td>
<td>1</td>
<td>Block Group 1, Census Tract 989200, Jackson County, MO</td>
</tr>
<tr>
<td>29</td>
<td>165</td>
<td>030307</td>
<td>1</td>
<td>Block Group 1, Census Tract 030307, Platte County, MO</td>
</tr>
<tr>
<td>29</td>
<td>183</td>
<td>980000</td>
<td>1</td>
<td>Block Group 1, Census Tract 980000, St. Charles County, MO</td>
</tr>
<tr>
<td>29</td>
<td>189</td>
<td>210803</td>
<td>1</td>
<td>Block Group 1, Census Tract 210803, St. Louis County, MO</td>
</tr>
<tr>
<td>29</td>
<td>189</td>
<td>213102</td>
<td>2</td>
<td>Block Group 2, Census Tract 213102, St. Louis County, MO</td>
</tr>
<tr>
<td>29</td>
<td>510</td>
<td>112100</td>
<td>3</td>
<td>Block Group 3, Census Tract 112100, St. Louis City, MO</td>
</tr>
</tbody>
</table>

Figure 3. Locations of excluded block groups labeled with their 12-digit geo-codes.
3.5.3 Standardizing & Normalizing Target Variables

Z-score standardization is used for converting a group of values to the same scale by dividing the difference between each variable (x) and the mean (μ) by the standard deviation (σ). The resulting normalized values represent the departure from the mean—which will always be zero. A resulting z-score value of zero indicates that the value is equal to the mean of all values, a positive z-score signifies the value is greater than the mean of all values, and a negative z-score signifies the value is lower than the mean of all values. In other words, z-score standardization scales a group of values such that we can more easily gain insight into how each individual value relates to the others in a group and compare them, regardless of their initial range or differences in their magnitude. The formula for the z-score calculation is illustrated in Equation (1).

\[ Z_i = \frac{(x_i - \mu)}{\sigma} \]  

Some of the research reviewed uses quintile normalization to further standardize their index values (e.g., Andrews et al. 2020). A quintile—not to be confused with a quantile—is itself a type of quantile. Each quintile (Q) represents one-fifth of the total set of values which are divided into five equal interval subranges with each representing 20 percent of all values (x_i). This classification method can facilitate continuity in the presentation of data in the form of maps, charts, or graphs, but is more commonly used for recognizable data such as
precipitation or temperature. That said, when normalizing in this way, there is a loss of numeric resolution—detail in the values themselves. In this study, both the quintiled values and the non-quintiled values have been analyzed and mapped to illustrate the differences between the two. The process of normalizing the values into quintiles is performed using the formula shown in Equation (2).

\[ Q_i = \frac{Z - \min(Z_i)}{\max(Z_i) - \min(Z_i)} \times 5 \]  

Maps of the MDI before and after access was included in the index and normalized using the equal interval quintile (EIQ) method discussed above can be found in the Appendix section of this study, illustrated in Appendix 1 through Appendix 4.

3.5.4 Weighting Variables Prior to Calculating Index Scores

Deprivation indices are constructed using a variety of methods to weight standardized indicators (i.e., variables) before calculating final index scores. One technique uses a normative approach to weighting variables—assigning equal weights to each standardized indicator value before summing them to generate index scores (Allik et al. 2020; Ursache et al. 2021). This weighting approach considers each variable to be equally significant to the overall index score regardless of geographic variability, and it is the method used in both the Townsend deprivation index and the Carstairs deprivation index, among others (Allik et al. 2020).

Other indices use data-driven weighting approaches whereby a factor analysis or a principal component analysis (PCA) is conducted across all variables to determine weights to assign to the variables before summing them (Lian, Struthers,
and Liu 2016). To date, the use of factor analysis and principal component analysis techniques to weight variables is still debated (Ursache et al. 2021). And, despite the common practice of using principal component or factor analysis to assign weights to indicator variables, it may not always be the best-suited method. Allik et al. (2020) argue:

“A major shortcoming of factor analysis is poor replicability across time and space. Correlations between indicators vary across time, space and geographic scale, meaning that the different indicators will have different weights at different time points and for different levels of area aggregation. This makes factor analysis less suited for longitudinal research, or for work that aims to develop a deprivation measure for different countries or levels of analysis.”

Given the potential issues that may arise in using factor analysis or principal component analysis to weight variables (as mentioned above), the Missouri deprivation index constructed here uses a normative weighting approach in which all indicator variables are equally weighted.

3.5.5 Final Deprivation Score Sums

For the primary six-variable deprivation index scores, each of the target variables’ z-scores are given equal weights and summed together, i.e., each z-score value is multiplied by $\frac{1}{6}$ and then summed (see Equation (3)). It may be worth noting that like the index calculation methods used by Diez-Roux et al. (2004), a natural log transformation was applied to the median household income variable

38
and the median home value variable to offset their skewness before their z-scores were calculated.

\[ \text{Index score} = \sum [Z_i \times (1 / n)] \]  

The completed table contained each target variable’s z-scores, summed z-scores, quintile-normalized values, summed quintile-normalized values, and an overall index score rank for each block group ranging from 1 to 4,496 in order from least deprived to most deprived. The data table was then imported as a CSV (comma-separated values) file into ArcGIS Pro version 2.9 for spatial analysis.

3.6 Spatial Analysis Using ArcGIS Pro

The block group geometries mentioned in the previous section were joined with the table containing the target variable calculations using ArcGIS Pro. The initial Missouri Deprivation Index values were computed by summing the equal-weighted z-score values for each block group. As an example, Figure 4 illustrates the resulting scores for the greater St. Louis metro area categorized into one of five equally distributed quantile classes—each containing roughly 900 block groups. The color classes illustrate the block groups in Missouri with the highest levels of deprivation (dark brown) to those with the least deprivation. The MDI outputs for the entire state are discussed in greater detail in Chapter 5.
Figure 4. MDI scores for the greater St. Louis, MO metro region, by block group.
CHAPTER 4: MEASURING ACCESS IN MISSOURI

Once Missouri block group index scores were calculated, an additional access to health services component was created and used to construct a new index and explore the extent to which geographic access to healthcare services may influence – or be associated with deprivation measures. This chapter describes access and accessibility and outlines the process of calculating block group access scores for the newly created Missouri Deprivation Index.

4.1 Defining and Measuring Access

Access plays a major role in the delivery of practically all services of all forms, whether measured by time, cost, distance, cultural or societal acceptability, or by differences in communication (McLafferty 2003; Sable et al. 2009; Griffiths et al. 2012; Douthit et al. 2015). Healthcare access is typically considered across five categories: accessibility, availability, accommodation, acceptability, and affordability (Penchansky and Thomas 1981). As this study is interested in measuring access to emergency healthcare services (i.e., services for which a prognosis requires immediate care), accessibility represents the most pertinent access category.

There are many ways to measure the geography of service access and accessibility using a GIS (Schuurman et al. 2006). Accessibility to healthcare is typically measured based on area or distance. Area measures may focus on the characteristics within an area, such as population, demand, and the number of available doctors or medical facilities, whereas distance-based measures may use Euclidian distances (as the crow flies, straight-line distances, and radii), travel times,
such as isochronal driving distances, or the travel costs incurred by a population to receive healthcare services (McLafferty 2003).

Although deprivation indices can account for socioeconomic barriers that may also affect access to healthcare (e.g., poverty, vehicle ownership, health insurance coverage) most do not specifically address geographic access to healthcare. In other words, these indices do not typically incorporate a variable that accounts for the cost (e.g., time, distance, etc.) required for an individual to travel to receive healthcare services. Moreover, no such indices were found in the literature that incorporates cross-area travel access measures (i.e., between neighborhoods, block groups, or census tracts). As McLafferty (2003) states, “most area measures do not take into account cross-area travel—an important factor when the area units are small; nor do they assess differences in access within areas—an important factor when the units are large.” To the contrary, distance-based measures can “avoid many of the problems that arise from area-based access measures” (McLafferty 2003).

Some access measurement methods, such as two-step floating catchment area methods (2SFCA) incorporate both area-based and distance-based measures (Luo and Wang 2003; Bagheri, Benwell, and Holt 2005). Two-step floating catchment area methods are commonly used to measure spatial accessibility to healthcare and most utilize a gravity model comprising two components: 1.) a ratio of physicians to population, 2.) a demand location based on travel thresholds of distance or time (Luo and Wang 2003; Wang and Luo 2005; Amiri et al. 2021; Chen and Jia 2019). Although 2SFCA methods do account for cross-area travel, they also
rely on the notion that access to healthcare is contingent on the number or availability of physicians for a given demand area, and the fact that medical care may still be unattainable, regardless of whether an individual arrives at a healthcare service facility.

According to the Emergency Medical Treatment and Active Labor Act of 1986 hospitals must provide immediate medical attention to individuals whose prognosis is deemed an emergency or in active labor, regardless of the patient’s insurance status or ability to pay. As this study is interested in access to hospitals for providing urgent, emergency, or trauma care (where capacity and physician availability are not normally a limiting factor outside of catastrophic, mass-casualty events) access measures should not necessarily be based on a ratio of physicians to population, but rather on the cost in distance or travel time required to access those health services. Given the objectives of this study and the reasons for incorporating area-based access measures mentioned above, distance-based methods utilizing travel times were chosen to measure access in this study.

4.2 Measuring Access Using Network Analysis

The distance-based methods used to measure access in this study are based on those used by several authors (see Carr et al. 2009; Luo and Wang 2003; Bejleri et al. 2017; Xi, Miller, and Saxe 2018) and were adapted for further use as a component to add to the deprivation index. To provide a practical range of driving times in which emergency care intervention might still be successfully administered, Carr et al. (2009) use 30-minute, 45-minute, and 60-minute driving time intervals to calculate block group access to emergency care facilities. In that
study, all block groups in the United States were included and classified as rural, urban, or sub-urban based on block group population densities. It is unclear how Carr et al. (2009) generated driving time catchment areas stating that “drive times in our study were derived from a meta-analysis of prehospital times for trauma” (pp5), but their access results were determined by summing the population of block groups that were within each driving time interval of an emergency care facility. As Carr et al. (2009) use a 30-minute transport time to represent the minimum (i.e., the fastest travel time), access scores for most urban and suburban areas were equally characterized as having the greatest (best) access, whereas rural areas were typically found to have the most limited access.

During a preliminary network analysis using the MO DHSS hospital data, several drive time catchment areas were generated to test the consistency of the processing algorithms. Analysis indicated that a 60-minute drive time catchment area covered over 99 percent of Missouri’s total area, excluding bodies of water and those block groups reporting zero population (as discussed in section 3.5.2). For this reason, the primary network analyses used to determine access in this study did not include the 60-minute drive time threshold.

When measuring cross-area access over a relatively large area—such as Missouri, rural expanse and urban clustering should both be accounted for. To better represent urban and suburban driving times—and differentiate them from rural driving times—5-minute isochrone increments were also calculated based on access measurement methods used in Bejleri et al. (2017) Luo and Wang (2003), Escribano (2021), and Xi, Miller, and Saxe (2018). The final driving time intervals
used in the network analysis of hospitals were as follows: 0- to 5-minutes, 5- to 10-minutes, 10- to 15-minutes, 30- to 45-minutes, and 45-minutes or more.

4.3 Data for Measuring Access to Emergency Healthcare Services

To aid in the measuring of access to health services, specifically urgent, emergency and trauma services, hospital location data for the state of Missouri was downloaded from the Missouri Department of Health and Senior Services (DHSS) GIS data portal\textsuperscript{10} and cross-references with records from the MSDIS data portal\textsuperscript{11} for the year 2019. A total of 173 hospitals were included, with their corresponding point geometries, i.e., their latitudes and longitudes (see Figure 5). It should be noted that this dataset does not include all health care facilities and clinics in the state, rather, it only reports facilities that are formally designated as hospitals by the Missouri DHSS. This study is interested only in the facilities designated as hospitals, as these are the facilities in which emergency and trauma care can be administered.
4.4 Methods for Determining Healthcare Access Areas

To measure access to health care services in this application, network-based access area polygons were generated for each hospital based on driving-time intervals. In particular, hospital arrival-time isochrones (or catchments) were generated via ArcGIS Pro using the TravelTime ArcGIS add-in\textsuperscript{12}. These polygons represent the areas from which individuals could arrive at a designated hospital within a stipulated drive time in average peak traffic. Much like using turn-by-turn navigation that yields an arrival time from a point using a multitude of dynamic

Figure 5. Missouri DHSS designated hospital facilities, 173 in total.
variables and conditions, this platform uses similar algorithms but processes every possible departure point in the given timeframe, and accounts for variables like speed limits, stop signs and traffic lights, average and peak traffic (depending on the time of day chosen for travel in the analysis), and is frequently updated with actual drive-time data for routes. Five different drive-time catchment areas were generated for each of the 173 hospitals at 5-minute, 10-minute, 15-minute, 30-minute, and 45-minute intervals. Using 15-minute and 30-minute driving times is an established practice when measuring driving modes of access (e.g., Bosanac, Parkinson, and Hall 1976; Metz 2013; Bejleri et al. 2017; Rothman et al. 2020; Escribano 2021). To enhance the precision of cross-area access measures, 5 and 10-minute catchments were included for more urban areas and 45-minute catchments were included for more rural, remote areas.

A total of 865 hospital access zone multipart polygons were generated using the TravelTime ArcGIS Pro add-in: five for each of the 173 hospitals. The features were then added to the geodatabase in ArcGIS Pro and mapped. Those zones are shown in Figure 6 with a semi-transparent symbology applied to show the overlap of polygons.
Figure 6. Hospital access areas, semi-transparent to reveal polygon overlap.

As this study concerns block groups located in Missouri and includes only those hospitals located within Missouri, the portions of access area polygons existing outside of the state boundary were removed. This was done not merely for aesthetic reasons, but rather to simplify the computational burden (see Figure 7).
Figure 7. Refined hospital access areas.

Given that there are many instances where two or more hospitals may be within just a few minutes of one another, overlapping access area polygons may exist. These overlapping areas become more frequent closer to urban areas due to the clustering of hospitals and transportation infrastructure. As it is not possible to travel to two destinations simultaneously—and hospital capacity is not a component measured in this study, redundancy in hospital access area coverage for a particular area (i.e., a block group) is not considered beneficial to the overall access of that area. Accordingly, overlapping areas of coverage were not given
additional weight in the index calculations used in this study. Overlapping polygon features were subsequently aggregated (i.e., merged together) based on their respective drive-time classes by using the Dissolve function\textsuperscript{13} in ArcGIS Pro. The Dissolve function can aggregate multiple polygons feature classes, their features, and their attributes into single, multipart features. An example of the process performed in this study that shows access area polygons before and after the dissolve is illustrated in Figure 8.

![Figure 8. An illustrated example of aggregating the access area polygons.](image)

Not allowing for multipart features can produce an excessive number of individual features for each polygon part—potentially overwhelming the computer’s physical memory. In this implementation, the Create multipart features
parameter was selected so that all polygons for each drive-time class were aggregated into five separate polygon features, each consisting of multiple parts. This resulted in a total of five multipart polygon feature classes, each representing the portions of the state covered by each drive-time class. Figure 9 shows the access area polygons prior to aggregation in much more detail. Note the white areas on the map revealing areas of very limited access, i.e., areas where driving time to a hospital is 45-minutes or more. Figure 10 shows the same area after the access area polygons have been dissolved—or aggregated.

Figure 9. Detailed view of access area polygons prior to being aggregated.
Figure 10. Detailed view of access area polygons after being aggregated.

To summarize this procedure, Figure 11 provides a synopsis of the steps involved: 1) Plot hospital coordinate features to the map; 2) Create 5-minute, 10-minute, 15-minute, 30-minute, and 45-minute driving-time access areas (i.e., generate polygons for each hospital feature using network analysis tools); 3) Limit the polygon features to the Missouri boundary; 4) Use dissolve to aggregate overlapping access areas. The final access area output is shown in Figure 12.
Figure 11. A synopsis of the steps involved in refining access areas.
4.5 Calculating Access at the Block Group Level

After generating the hospital access area polygons, further spatial analysis was needed to calculate access scores for each block group. The methods used for those calculations include measuring the rates coverage (i.e., overlap) for each of the five access areas for each block group, then summing those rates in a way that is meaningful with respect to rural, suburban, and urban population drive times, and finally, normalizing those scores using a z-score standardization. After analysis of access independent of deprivation is completed, each block group’s z-standardized access score is then incorporated into the initial deprivation index as an additional,
equally-weighted variable. The following sections describe the methods used in these calculations in more detail.

4.5.1 Intersection of Block Groups and Access Areas

To analyze the access area coverage of each block group, ArcGIS Pro’s Intersect analysis tool\textsuperscript{14} was used to combine the access area polygon features with the block group features. The percentage overlap for each block group and each access area polygon was automatically generated during this process and those rates were then used to assess the extent to which each block group was within each drive-time threshold. For example, block groups having no access area coverage were given values of 0.0, whereas block groups for which there was complete coverage of a 5-minute access area would be given a value of 100 percent, regardless of additional access area overlap. To demonstrate how the intersect analysis was used to compute access area coverage percentages, hypothetical block groups units, along with access area polygons that mirror the actual analysis of this study are illustrated in Figure 13, as well as the percent access area coverage for each block group that was computed via the intersect.

Sometimes, access measures using point and polygon geometries are calculated based on where the centroid of that polygon falls, however, given the methods used in generating the access area polygons (i.e., the buffer distances from roads), this could be problematic if a block group centroid falls within a “dead zone”, or a pocket where no transportation infrastructure exists, i.e., the middle of a lake, the top of a mountain, inside a national park or land reserve, etc.). For this reason measuring access this way was not considered.
4.5.2 Calculating Block Group Access Scores

After the percent of each block group’s access area coverage was computed, final access scores were calculated by summing the weighted values of block group coverage rates for each access area threshold. The methods used to calculate access scores for block groups is illustrated using the example shown in Figure 13 whereby the block group coverage for each access is depicted. The calculations used are outlined in detail in

<table>
<thead>
<tr>
<th>Block Group #</th>
<th>Percent of access area coverage &amp; weights applied</th>
<th>Access score calculation</th>
<th>Final Access Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 to 5-min</td>
<td>5 to 10-min</td>
<td>10 to 15-min</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>100</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>65</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 7.
Figure 13. Example of block group access area coverage and percent calculation.

<table>
<thead>
<tr>
<th>Block Group #</th>
<th>0 to 5-minutes</th>
<th>5 to 10-minutes</th>
<th>10 to 15-minutes</th>
<th>15 to 30-minutes</th>
<th>30 to 45-minutes</th>
<th>&gt; 45-minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>50</td>
<td>50</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>92</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>65</td>
<td>25</td>
<td>10</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 7. Block group access score calculation methods.

This method assumes the population within each block group is uniformly distributed, and although the population densities of block groups are not figured into these index calculations, access is ultimately computed based on rates of access area coverage for each block group. When a block group is only partially covered by an access area, or covered by more than one access area, the population within that block group is considered as a single areal unit. For instance, imagine block group #3 from the example in Figure 13 and Table 7 has a population of 3,000 and has 100 percent access area coverage: 50 percent of its areal coverage from the 5- to 10-minute access area, and another 50 percent of areal coverage from 10 to 15-minute access area. In this example, the block group’s total population of 3,000 is assigned the same access score rather than be divided proportionally such that 1,500 are considered to be in a 5- to 10-minute access area and another 1,500 in the 10- to 15-minute access area. Additionally, a block group of larger area with low population density can yield the same access score as a block group of smaller area with higher
population density. The inverse of this also applies, as access scores are not calculated relative to other block groups. It should be noted that there are limitations when assuming the population within an areal unit is spatially distributed uniformly—this notion is addressed in more detail in the limitations section of Chapter 6.

Final access scores ranged from 0.0 to 100.0, with 0.0 representing the least access and 100.0 representing the most access. These scores were standardized using the z-score method (see Section 3.5.3) and incorporated into the MDI as an additional, equal-weight variable. Spatial patterns observed when reviewing access scores independent of deprivation (i.e., prior to adding them to the MDI) are illustrated and discussed in Chapter 5, Section 2, and the impact the access scores have on the MDI is discussed in Chapter 5, Section 4.
5.1 Deprivation Index Based On Socioeconomic Indicators

After completing calculations for the Missouri Deprivation Index (MDI) based on the six socioeconomic indicators discussed in Section 3.5, scores were classified into five equally distributed quantiles (EDQs)—five groups of roughly nine-hundred block groups. MDI scores for the entire state, as well as the Kansas City and St. Louis areas are illustrated in Figure 14, Figure 15, and Figure 16, respectively. The lowest scores (those with the highest deprivation) are represented in dark brown.

Figure 14. MDI block group scores, classified into equally distributed quantiles.
Figure 15. MDI scores for the Kansas City, MO metro area, by block group.
Figure 16. MDI scores for the greater St. Louis, MO metro area, by block group.
The results depict clear and consistent spatial patterns of high and low socioeconomic deprivation between block groups across Missouri. Many of the very lowest deprivation scores are densely concentrated in urban areas that are often delineated by geographic or infrastructure boundaries—such as rivers, rail lines, or major roadways, e.g., Troost Avenue in Kansas City and Olive and Delmar Boulevards in St. Louis (see Figure 17 and Figure 18, respectively). These spatial patterns are consistent with—and embody complex histories of racial discrimination, redlining, deindustrialization, economic decline or transition, and infrastructure expansion (Hillier 2003; Trivers and Rosenthal 2015; Massey and Tannen 2017).

Figure 17. Kansas City deprivation divided east to west with racial dot density map.
Figure 18. St. Louis deprivation divided north to south with racial dot density map.
5.1.1 **Hot Spot Cluster Analysis of Deprivation Scores**

Using ArcGIS Pro, an optimized hot spot analysis was performed to test the spatial dependency of the index, i.e., whether—and to what degree did block groups that are close to one another share similar index values. The hot spot analysis uses the Getis-Ord Gi* statistical framework to detect significant spatial clustering of high standard deviation and low p values, i.e., hot spots, as well as low standard deviation and high p values, i.e., cold spots (Getis and Ord 1992; Ord and Getis 1995). Compared to the standard hot spot analysis, the optimized hot spot analysis tests the source data to determine an optimal distance band to use for the cluster analysis, as opposed to manually specifying a fixed distance band. For this analysis, an optimal fixed distance band of 8.96 miles (14.42 km) was determined. The fixed distance band is basically the spatial extent to which features are analyzed in the context of their neighboring features at any given point in time. As an analogy, consider the optimal scale or viewable extent required for a microbiologist to visually compare particular cells under a microscope.

Output from the optimized hot spot analysis is illustrated in Figure 19, Figure 20, and Figure 21. The red features represent hot spots—areas where high block group scores (low deprivation) are clustered, and the blue features represent cold spots—where low block group scores (high deprivation) are clustered.

Clustering of low deprivation scores (1,410 BGs) tends to be situated within or adjacent to major metropolitan areas, especially in their immediate suburbs. This is easily observed in the Kansas City and St. Louis areas—and is shown in much greater detail in Figure 20 and Figure 21, respectively. Outside of the major
metropolitan areas, there is also notable low deprivation clustering in and to the south of Columbia and Jefferson City, as well as in the southern outskirts of Springfield. Although clustering of high deprivation (1,135 BGs) can also be observed in major metropolitan areas, they tend to be situated much closer to the greater, downtown areas around city centers, with the minor exception of the northern outskirts of Springfield. Other, smaller clusters of high deprivation can be found throughout rural Missouri—especially in the south-eastern most “Bootheel” region of the state, as well as in and around smaller, non-metropolitan cities such as Farmington, Sedalia, Poplar Bluff, Joplin, and Moberly, to name a few. The remaining 1,951 features indicated no significant clustering of high or low deprivation scores relative to neighboring block groups when analyzed using the distance band of 8.96 miles.
Figure 19. Hot spot analysis illustrating clusters of high and low deprivation scores.
Figure 20. Hot spot analysis showing high & low MDI score clusters in Kansas City.
Figure 21. Hot spot analysis showing high & low MDI score clusters in St. Louis.
5.2 Access To Hospitals Providing Emergency Healthcare

Block group access scores pertaining to driving times to emergency healthcare services were calculated using the distance-based methods described in Chapter 4. Raw access scores—ranging from 0.0 (least access) to 100.0 (most access) were standardized using the z-score method (see Section 3.5.3). The resulting z-scores ranged from -5.32 to 0.95. In this section, notable spatial patterns produced by those access measures are highlighted, and the results of analyzing access independent of the deprivation index are illustrated and described in more detail.

As demonstrated in Figure 22 and Figure 23, when access to emergency healthcare is displayed based on each block group’s access score, spatial patterns emerge, revealing underlying transportation infrastructure and more populated areas quite clearly. As to be expected, block groups with greater access to hospitals providing emergency care are situated within, or closer to more densely populated areas, and thus, are inherently closer to major interstates, U.S. highways, and transportation infrastructure.

As shown in Figure 22, the I-44 corridor running between St. Louis and Springfield effectively increases access scores to the areas within or near its path and divides a large portion of southern Missouri where areas of very limited access remain both northwest of I-44 and southeast of I-44. In contrast, access scores in the north-central part of the state appear to be less dependent on major interstate transportation. Given the similarities regarding the rurality of both areas (the southeast and the north-central regions), the differences in access scores observed here
are likely due to two factors: better network connectivity through the presence of more—and more uniformly distributed hospitals and U.S. highways in the north-central region, as well as less mountainous terrain than that found in the southeast region—terrain which can make roadway infrastructure more challenging to construct, and impede travel where it does exist.

Figure 22. Access scores classified into quintiles shown with highway infrastructure.
5.2.1 Estimating the Total Population of Access Areas

The initial network analysis that generated the access areas has provided a great deal of valuable information, and estimating the number of people residing within those access areas is equally—if not more important. Although block group access scores alone are useful in identifying areas, or clusters of very limited access, gaining insight into the population that may be represented by those block groups can provide us with a greater understanding of the number, and the proportion of people that may be affected by limited emergency healthcare access. Such knowledge could also be used to identify areas of abundance in emergency
services—where resources could be reimplemented or redeployed to have greater accessibility to relatively close pockets of limited access. Several approaches were considered when exploring ways to calculate these estimates. The methods used to do so, as well as the resulting figures are outlined in this section.

If the population of each block group is assumed to be uniformly distributed, the access area coverage rates of each block group could be used to calculate the rate of each block group's population that falls within each access area. In other words, the number of individuals in each block group that reside in each access area would be proportional to the rate of areal block group coverage by each access area. Although there are obvious flaws in assuming a uniformly distributed population, if this notion is applied, the total number of Missourians within each access coverage area can be calculated simply by multiplying the rates of access coverage for a block group (as described in Section 4.5.1) by the block group's total population.

Table 8 provides the population estimates when using this method of calculation. In using the approach described above, an estimated 449,276 Missourians live at least 30-minutes away from a hospital capable of providing emergency or trauma care, and of those, an estimated of 99,331 individuals live beyond the 45-minute threshold.
Population estimations of those residing in each access area may also be calculated in a way that does not assume the population in each block group is uniformly distributed. If the population of each block group is only recognized as an attribute value, the block groups and their population values are indissoluble and can be classified using the access areas that cover the majority of each block group. For example, if 75 percent of a block group is within (i.e., covered by) the 5- to 10-minute access area and 25 percent of the block group area is within the 0- to 5-minute access area, the entire block group is classified as being within the 5- to 10-minute access area. Table 9 provides the results for calculating the population affected by each access area using this method.

Table 8. Population estimates access areas if uniformly distributed within block groups.

<table>
<thead>
<tr>
<th>Accessibility rating</th>
<th>Access coverage area</th>
<th>Population of Missourians within access coverage area</th>
<th>Percent of Missouri population within access coverage area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>0 to 5-minutes</td>
<td>917,064</td>
<td>15.03%</td>
</tr>
<tr>
<td></td>
<td>5 to 10-minutes</td>
<td>2,163,490</td>
<td>35.46%</td>
</tr>
<tr>
<td></td>
<td>10 to 15-minutes</td>
<td>1,154,076</td>
<td>18.92%</td>
</tr>
<tr>
<td></td>
<td>15 to 30-minutes</td>
<td>1,417,118</td>
<td>23.23%</td>
</tr>
<tr>
<td></td>
<td>30 to 45-minutes</td>
<td>349,945</td>
<td>5.74%</td>
</tr>
<tr>
<td>Poor</td>
<td>Over 45-minutes</td>
<td>99,331</td>
<td>1.63%</td>
</tr>
</tbody>
</table>
Table 9. Population estimates for each access area if BG population is not apportioned.

In applying this method, an estimated 367,932 Missourians live at least 30-minutes away from a hospital that can provide emergency or trauma care, and of those, an estimated 46,532 individuals live beyond the 45-minute driving time threshold.

The above examples are meant to illustrate that there are several ways in which the population of block groups may be considered with respect to the access areas. When the access scoring method outlined in Section 4.5.2 is used to calculate the population within each access category, a sum of the population for all block groups in each category provides us with yet another way to calculate access populations. The access scores are categorized across five equal intervals of standardized scores (illustrated in Figure 22) and the results of population sums for each category are shown in Table 8.
The total population of Missourians with limited access is 87,528 and comprises 86 block groups, and of those, 42,396 reside in block groups with very limited access—comprising 41 block groups. It may be worth noting that nearly half of the population with very limited access is situated in one of four counties: Douglas County, Ozark County, Lewis County, and Reynolds County.

To illustrate the population density in and around the areas where access to emergency healthcare services is most limited, Figure 24, Figure 25, Figure 26, and Figure 27 provide detailed views of those more rural areas. In these figures, population is represented visually using a randomly generated dot distribution at a rate of one dot per twenty-five people. The boundaries of notable counties where access is extremely limited are in bold.

Table 10. The number of Missourians residing in each ranked access category.
Figure 24. A detailed view of a very limited access area in south-central Missouri.

Figure 25. A detailed view of a very limited access area in southeast Missouri.
Figure 26. A detailed view of a very limited access area in northeast Missouri.

Figure 27. A detailed view of very limited access in the central lakes region of Missouri.
5.3 Deprivation Index Based on Socioeconomic Indicators and Access

To incorporate geographic access into the deprivation index, MDI scores for each block group were recalculated such that each of the initial six variables, as well as the seventh, z-standardized access score variable, were multiplied by \( \frac{1}{7} \) (one-seventh) to give them equal weights and summed together to produce new index scores that include a measure of geographic access. The new deprivation index scores—based on socioeconomic indicators and access—were classified into five EDQs each containing roughly nine hundred block groups. These scores are illustrated in Figure 28, Figure 29, and Figure 30 with the lowest scoring block groups (those with the most deprivation) represented in dark brown.

When access is included in the index, large metropolitan areas such as Kansas City and St. Louis exhibit spatial patterns that are nearly indistinguishable from those described in section 5.1: high concentrations of low-scoring block groups exist closer to the downtown city centers, while higher scores tend to be concentrated in the suburban areas immediately outside of the greater metro areas (see Figure 29 and Figure 30). The same can be said regarding smaller metropolitan areas, such as Springfield and Columbia; low scoring block groups tend to be concentrated closer to city centers, while higher scores are located along the periphery of the cities and typically remain high or moderately high as they extend out towards their respective county boundaries.

As for more rural areas when access is included in the index, patterns of deprivation are similar to those described in sections 5.1 and 5.2: the I-44 corridor
running between St. Louis and Springfield can still be observed as a band of higher scoring block groups (lower deprivation) dividing a large area of lower-scoring block groups (higher deprivation) to northwest and southeast of I-44. However, the lower-scoring rural areas appear to be far less isolated and more densely concentrated than in either of the previous independent measures, i.e., MDI scores or access scores. As expected, this indicates access measures had a much greater impact on the index scores of more rural, remote areas than urban and suburban areas. The next section describes that impact in more detail.

Figure 28. MDI scores when access is included, classified using five EDQs.
Figure 29. MDI scores with access for Kansas City, MO, classified using five EDQs
Figure 30. MDI scores with access for St. Louis, MO, classified using five EDQs
5.4 The Impact of Including Geographic Access in the Deprivation Index

After the standardized access scores were incorporated into the initial deprivation index, spatial analyses were conducted to determine if, and to what extent, accounting for access to healthcare may impact deprivation scores. These analyses also provided insight into the strength of the access component as an equal-weighted variable in the deprivation index. This section describes the statewide patterns observed when comparing these two indices and considers block group characteristics such as population, population density, rural and urban contexts, as well as the physical size of block groups.

5.4.1 Block Group Z-Score Distribution Change

The following histograms show the distribution of block group z-scores for both indices—the MDI based only on socioeconomic indicators (Figure 31), and the MDI after access scores were incorporated (Figure 32). When access scores were included, the absolute range of z-score values decreased from 5.846028 to 5.010881, indicating that—overall, the access variable slightly lowered the differences between all block groups. In other words, including access in the deprivation index decreased the maximum achieved scores from 3.482517 standard deviations (SDs) above the mean (0.0) to 3.069237 SDs above the mean, while increasing the minimum scores from -2.363511 SDs below the mean, to -1.941644 SDs below the mean. One reason for this compression of the z-score range may be attributed to the fact that as the number of variables in multivariate indices such as these increases, the skewness between z-scores typically decreases; additional variables “level the playing field”, so to speak.
Figure 31. Histogram of block group index scores without the access component.

Figure 32. Histogram of block group scores when access is included in the index.
5.4.2 Analyzing Block Group Index Rank Change

To further analyze the impact of access on index scores, each block group was given a numeric rank between 1 and 4,496 based on the scores obtained from the initial deprivation index calculations based on socioeconomic indicators—prior to the inclusion of an access variable. A rank of 1 would indicate the highest score of all block groups (i.e., the least deprivation), while a rank of 4,496 would indicate the lowest score (i.e., the most deprived). The initial block group ranks are illustrated in Figure 33. This process was then repeated using scores based on the MDI including the access component. Those ranks are illustrated in Figure 34.
Figure 33. Deprivation index ranks not accounting for the access component.
Figure 34. Deprivation index ranks accounting for the access component.

Block group ranks for each index were then compared and spatially analyzed. First, the percent change in a block group’s numeric rank (i.e., comparing their index rank without the access component to the index rank when access is included) was calculated (see Figure 35). Not surprisingly, results show that the block groups most negatively affected by the inclusion of an access component were largely those situated in more remote, rural areas. A moderate to very large decrease in index
rank (i.e., lower scores and higher deprivation) occurred for 135 block groups, and of those, all can be characterized as rural or remote.

![Map of Missouri showing rank changes](image)

**Figure 35.** Rates of rank increase or decrease when comparing both indices.

The observation regarding rural block groups was reinforced after investigating the block groups’ population densities. Population density is used by the U.S. Census Bureau to define urban areas—while rural areas are simply defined as those areas “not included within an urbanized area or urban cluster” (Ratcliffe et al. 2016). Determining which areas should be considered rural and which urban can be quite complicated, but given this condensed explanation, population density
seemed an appropriate measure to determine the rurality of block groups for this study.

A cluster analysis of all block groups’ population densities and their change in index rank was performed using Tableau Desktop statistical software. Not to be confused with the spatial cluster analysis performed in ArcGIS Pro, this aspatial cluster analysis uses a k-means algorithm to identify the mean center for all data points in each of a specified number of clusters. In this application, two block group variables were analyzed (population density and index rank change) and a cluster number of four was chosen to provide at least one mean division to be identified for each variable, i.e., two partitions with mean centroids for each of the two variables.

The resulting scatter plot cluster analysis shown in Figure 36 reasserts that changes in index rank and population density (rurality) appear to be heavily dependent on one another. Together, clusters 1 and 3 (represented in blue and yellow) comprise nearly 60 percent of all block groups, and has an average rank-change of +175.28 and an average population density of 2,622.5 persons per square mile. In stark contrast, the remaining block groups in clusters 2 and 4 (represented in red and green) have an average rank-change of -246.6 and an average population density of 275.33 per square mile.

In sum, all block groups in which the inclusion of access in the index caused a decrease in rank position of 400 or more also had a population density below 2,000 persons per square mile, and block groups with population densities of 500 people
per square mile or less (the most rural block groups) accounted for 100 percent of index rank decreases of 800 positions or more.

Figure 36. Block group population density & index rank change when including access.

5.4.3 Hot Spot Analysis of Deprivation Score Impact

After the rank change analysis, an optimized hot spot analysis (HSA) was conducted to measure the clustering of significant rank changes among block groups after access was introduced into the index. The value measured by the HSA is the number of positions in rank by which a block group increased or decreased. Red features represent hot spots where increases in block group rank (high values) are
clustered, and blue features represent cold spots where large decreases in block group rank (low values) are clustered (see Figure 37). Nearly half of all block groups in Missouri (2,400) were within hot spot clusters, and most of those were within or adjacent to urban centers.

Results show that MDI scores for both urban and more remote, rural areas were significantly affected when the access to healthcare component was factored into the index. Block groups located in more remote, rural areas that scored very low using the initial index (those with higher levels of deprivation), account for a very large proportion of the block groups with a strong positive association to
access, and consequently accounted for most of the block groups with negatively-affected scores after the access component was incorporated into the deprivation index.

In urban and sub-urban areas, index scores were largely improved once the access component was incorporated. This is to be expected due to the proximity and clustering of hospitals in those areas that provide emergency or trauma care services. In areas situated in between the sub-urban and the more remote rural areas there was far less clustering of index score change when access was included. It is worth noting that these areas of no significant clustering appear to correspond largely with the 15- to 30-minute access areas zones.
CHAPTER 6: CONCLUSIONS

6.1 Summary of Findings

The socioeconomic index developed through this research reveals clear patterns of deprivation in Missouri—in both urban and rural settings. The index also shows that clustering of deprivation at the block group level can occur independently of census tract and county geographies. Moreover, by analyzing access independently of deprivation, this research identifies the areas and populations in Missouri for which driving access to emergency healthcare services are most limited.

After incorporating access measures into the index as an additional variable, the relationship between socioeconomic deprivation and access to emergency healthcare services was further explored. Specifically, analyses compared the impact on block group index scores, before and after, locating spatial patterns and clustering among block groups whose scores were most strongly affected by high deprivation and very limited access to emergency healthcare services. As expected, urban areas are less affected by driving-time access to emergency healthcare services than rural areas, however, urban areas also comprise the highest levels of deprivation. On the other hand, some rural, more remote areas experience the most limited access to emergency healthcare in addition to deprivation levels that are moderately high to very high. Overall, clustering of high and low scores for both indices tends to decrease—and is insignificant—in between urban areas of higher population density and remote areas of very low population density.
Although this study only analyzed healthcare access in the form of driving time to hospitals providing emergency care, findings show that failure to include a healthcare access component in deprivation measures can grossly underestimate differences in deprivation levels between rural and urban areas. In addition, incorporating measures of access to healthcare may be essential when using socioeconomic deprivation indices to identify areas that may be at greater risk for poor health outcomes.

There are other implications regarding the access data generated in this study. For life-threatening medical events, such as a traumatic injury or acute myocardial infarction (i.e., a heart attack), mortality rates increase with increasing driving time to capable, emergency healthcare services (Balamurugan et al. 2016; Karrison et al. 2018). If the effect of very limited access to emergency healthcare services equates to higher probability of mortality due to greater travel times to such services, insurance costs for individuals in those areas could be impacted. Higher insurance rates could potentially increase deprivation for those living farthest from emergency healthcare services.

6.2 Strengths

One key strength of this study can be attributed to the construction of a distinct, block group level deprivation index for the state of Missouri using publicly-available data. At the time that this research was completed the only other research found that had created a similar index for Missouri was that of Sheets et al. (2017). Sheets et al. used a refined version of the 2013 ADI, developed by the University of Wisconsin School of Medicine and Public Health (as mentioned in section 2.1.5).
Other research that has produced socioeconomic indices for Missouri but used different areal units include Nagasako (2018), which utilized previously developed methods to construct a deprivation index at the zip code level, and Lian, Struthers, and Liu (2016), which constructed a deprivation index for Missouri at the census tract level.

In addition to establishing a Missouri deprivation index, this research has also demonstrated how access measures can be calculated using fundamental, distance-based network analysis, then—using a relatively novel approach, incorporate those access measures into such an index. The data generated here can provide policymakers as well as researchers with a much greater understanding of the spatiality of both socioeconomic deprivation and access to emergency healthcare across the state of Missouri.

6.3 Limitations

6.3.1 Transportation Data for Urban Access Measures

As discussed in Chapter Four, there are many ways to measure access to healthcare services. This particular research was only concerned with driving time access to hospitals providing emergency healthcare services. As opposed to rural areas where driving is typically the only practical mode of transportation, urban areas often have more variety in modes of transportation (e.g., public transportation such as bus and rail lines). The decision not to generate access measures that include public transit options may be seen as a limitation, however, the lack of current accurate and reliable data for public transportation times in relevant metropolitan areas of Missouri is itself a very limiting factor.
6.3.2 Rural Clinics and Urgent Care Facilities

There are many other medical facilities across Missouri that are not designated as hospitals. According to the Missouri DHSS, there are 237 Federally Qualified Health Centers (FQHC) as designated by the U.S. Health Resources & Services Administration (HRSA), and 344 Rural Clinics as defined by the DHSS. However, these facilities and clinics are not typically considered capable of providing the same emergency and trauma healthcare services as facilities that are designated as hospitals. As this study was only concerned with access as it relates to emergency and trauma care services, clinics and facilities not designated as hospitals were not included.

6.3.3 Driving Times

Driving times for any given route are not static, they fluctuate in response to an infinite number of factors. There will always be extenuating circumstances that may impede an individual’s ability to travel at a normal rate of speed and arrive within a timeframe that is typical for a given route (e.g., traffic accidents, traffic congestion, road construction, etc.). The network analysis tools used to establish access areas in this study use average travel times to generate isochrone polygons—two-dimensional areal representations. Those averages are updated regularly and supplemented with actual drive-time data from mobile device data collection, however, they are still only estimates.

6.3.4 Qualitative Data

There is an ongoing conversation in the discipline of geography concerning the use of qualitative versus quantitative approaches. With regard to the discipline's
ability to thrive—and to find synthesis among other disciplines conducting similar research, comprehension of multiple approaches and methods seems critical. There also seems to be an increase in the proportion of mixed-methods approaches used in geography research. Perhaps this is out of necessity, and is to the benefit of researchers that are reluctant to explore GIS and quantitative methods. However, one could argue that the success of applying a mixed-methods approach hinges on the researcher’s desire to understand—or at the very least, consider both quantitative and qualitative approaches when designing their research.

Although this particular study does not directly make use of qualitative data, there is a multitude of ways in which qualitative approaches may be used to provide additional insight or clarification into the observations made herein. For example, questions surrounding the use block groups and census tracts to define neighborhoods have been explored using qualitative approaches in other research (Campbell et al. 2009). A study such as this may also benefit from descriptive respondent data regarding available means of transportation, notions of acceptable driving times and costs, or the personal experiences of individuals from both rural and urban settings with respect to healthcare access. The author recognizes the importance of this type of information and hopes this study provides a basis from which both quantitative and qualitative research may benefit from.

6.4 Future Research Opportunities

6.4.1 Accessibility and Future Infrastructure Development

Although the access measures used in this study were based only on the locations of designated hospitals providing emergency care and calculated at the
block group level, much more insight could be gained from the findings generated here. Given that mortality rates for many life-threatening medical events increases with longer transportation times (Balamurugan et al. 2016; Karrison et al. 2018), decisions concerning where and how to expand transportation infrastructure could greatly benefit from access data such as the data produced here. For example, in Figure 10, both the western portion of Howard County and the southern portion of Chariton County contain access areas that fall outside the 30 to 45-minute drive time to the nearest hospital (in this case, Fitzgibbon Hospital located in Saline County). The primary impedance to access in this example is the Missouri River. Future analysis could use these methods to determine the ideal site selection for a new bridge that could drastically reduce hospital access times in this area.

6.4.2 Modifying Index Components for Future Applications

As mentioned in section 3.2, the six variables used in the construction of the initial Missouri deprivation index were selected to recreate the preexisting methods considered to be most commonly used, or widely accepted. Perhaps future research could explore modifying these variables to take into consideration factors such as health insurance coverage, or the number of vehicles available per household\textsuperscript{16} as included in both the Carstairs and Townsend deprivation indices. Vehicle availability may also be an ideal variable to use as a limiting factor when considering access measures based on driving times. Future research may also benefit by measuring access in a way that takes into consideration the primary means of transportation for the majority of individuals in a given block group level (e.g., walking, biking, personal vehicle, public transportation, etc.). This may provide
greater insight into cross-area access and differences in access between rural and urban areas, whereby individuals in urban areas may be less affected by vehicle availability than those in rural areas.

6.4.3 Measuring COVID-19 Pandemic Effects on Deprivation in Missouri

The COVID-19 pandemic has had a profound impact on the Missouri economy that has uprooted the financial stability of many individuals and entire communities across the state. Missouri hospitals have not been spared in this respect, as many are still reeling from staffing and resource shortages and some have closed their doors completely (Wheeler 2020; Sable-Smith 2022). The data chosen for this research was intended to be representative of the time period up to—and as close to the pre-COVID-19 condition as possible to avoid the confounding effects the pandemic may have on subsequent data. This study provides an ideal opportunity for future comparative research as a control study; to explore patterns and trends in deprivation, pre and post-COVID. Recreating this study using equivalent data from a post-COVID time period may not only provide insight into deprivation, but it may also provide insight into changes in healthcare accessibility, as the number of hospitals—as well as the transportation infrastructure, many have changed during this time period.
APPENDIX

Additional maps were generated to illustrate the differences that can occur when data is represented using different classification methods, or embodied differently on a map—as discussed in Chapter 1. After calculations for the Missouri Deprivation Index (MDI) were completed, the block group index scores were normalized into quintiles—not of equal distribution as in Figure 14, but of equal interval percentiles relative to the index scores. The MDI scores rendered using quintile normalization are provided in Appendix 1, Appendix 2, and Appendix 3.

The first quintile represents the top 20 percent of block groups by their deprivation index score (i.e., those with the least deprivation), while the fifth quintile represents the bottom 20 percent of block groups by their deprivation index score (i.e., the most deprived). The combination of this normalization method and an equal interval symbolization method results in a very different distribution of block group counts for each classification when compared to equal distribution as shown in Figure 14, Figure 15, and Figure 16. Maps of the index scores when access was included were also generated in the same fashion (see Appendix 4).
Appendix 1. MDI block group scores classified using equal interval quintiles.
Appendix 2. MDI scores for St. Louis, MO classified using equal interval quintiles.
Appendix 3. MDI scores for the Kansas City, MO classified using equal interval quintiles.
Appendix 4. Deprivation when access is included, illustrated using quintile classes.
REFERENCES


Campbell, Elizabeth, JULIA R. HENLY, DELBERT S. ELLIOTT, and Katherin Erwin. 2009. “SUBJECTIVE CONSTRUCTIONS OF NEIGHBORHOOD BOUNDARIES: LESSONS FROM A


Ratcliffe, Michael, Charlynn Burd, Kelly Holder, and Alison Fields. 2016. “Defining Rural at the U.S. Census Bureau.” *U.S. Census Bureau, ACSGEO-1,*.


ENDNOTES


2 Health Resources & Services Administration, Neighborhood Atlas®, Area Deprivation Index, October, 2021, 
https://www.neighborhoodatlas.medicine.wisc.edu/


4 U.S. Census Bureau’s TIGER/Line® Shapefiles download portal, 
https://www.census.gov/cgi-bin/geo/shapefiles


6 U.S. Census Bureau’s American Community Survey Data Portal, 
https://data.census.gov/cedsci/table

7 Tableau Desktop Pro software, 
https://www.tableau.com/products/desktop

8 ESRI ArcGIS Pro, Data classification methods, 


10 Missouri Department of Health and Senior Services, “Data: Geographic Information Systems (GIS) & Maps”, March, 2022, 
https://health.mo.gov/data/gis/index.php


12 TravelTime ArcGIS add-in, 
https://docs.traveltime.com/arcgis/about/overview

