TO WHAT EXTENT IS A RURAL COMMUNITY’S SOCIAL CAPITAL RELATED TO THE LIKELIHOOD OF A HOSPITAL CLOSING?

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Naomi Martha Adaniya, M.P.H., M.A., B.S.

Graduate Program in Public Health

The Ohio State University

2016

Dissertation Committee:

Allard E. Dembe, ScD, Advisor

Linda M. Lobao, PhD,

Sharon B. Schweikhart, PhD,

Thomas Wickizer, PhD
ABSTRACT

BACKGROUND: Rural hospitals provide vital health care services to nearly 54 million Americans. However, when compared to urban hospitals, rural hospitals play an even greater role in their community. As one of the largest employers the area, rural hospitals sustain the local economy. The closure of a rural hospital, therefore, can have a negative impact on the community in multiple dimensions. In the past few years, government programs that provide financial assistance to many rural hospitals, including the Critical Access Hospital (CAH) program, which has been identified by both President Obama and the Office of Management and Budget as a program to cut from the budget. Nationwide, one in four acute care hospitals have attained CAH status. Removing CAH status from just 61 of the over 1,300 Critical Access Hospitals would save the federal government $4 billion.¹ Since social capital has already been found to impact economic development, overall health, and mortality, this study will be the first to analyze whether social capital can be applied to a health system outcome – rural hospital closure. If government support is eliminated or reduced, the social capital of rural communities could be a resource health care administrators can nurture and use to help these facilities remain in operation.

OBJECTIVE: This study aims to develop a comprehensive model of social capital and then apply it nationwide to examine how a rural community’s social capital is associated with rural hospital closure between 2000 and 2008. Closing a hospital is often a controversial decision, and one stemming from long-term pressure. Therefore, after calculating a social capital index, this study will analyze how social capital changes over time (1980 to 2000) are associated with rural hospital closure. Finally, social capital measurement is commonly performed at the county level; however, the use of the county scale can introduce

geographic statistical errors in the analysis. This study will use buffers to more accurately represent a rural hospital's service area and estimate social capital at this scale. The association between social capital at the new hospital service area and hospital closure will be evaluated.

**METHODS:** A comprehensive literature review process, factor analysis, and logistic regression will be used to identify social capital components and categories (or factors) most strongly associated with the likelihood of rural hospital closure. Hospital closure results will be stratified by a county's rurality status and corresponding U.S. Census region. Common social capital variables collected in 1980, 1990 and 2000 surveys will be analyzed for trends related to rural hospital closure between 2000 and 2008. Finally, spatial analysis will be used to create an alternate hospital catchment definition to test for the reliability of the defined geographic areas.

**POLICY IMPLICATIONS:** International organizations and financial institutions like the United Nations and World Bank, as well as national governments have increasingly implemented policies aimed at increasing the level of social capital within communities. Supported by research, in this context, social capital is viewed as a desirable outcome that if facilitated, can positively impact not just healthcare but overall development. If social capital is also found to be associated with hospital closure or other healthcare system outcomes, then policies aimed at building social capital can be adopted not just for economic development and overall community well-being but also for health system benefits.

---

ACKNOWLEDGMENTS

Transformative achievements come from the concerted, and at times unrecognized contributions of many. Belonging to a community and benefitting from the input and support of a wide network of sympathetic actors, I believe are key contributors to what may appear to be a seemingly individual success. Therefore, I attribute and dedicate this successful endeavor to the vast array of significant players that I have had the grand fortune to have entered my life, this achievement is truly OURS. Many, many, feelings of gratitude...

Special note of thanks to my parents for all their love and support, for instilling in me a sense of curiosity about the world, and for putting me through the best education possible. I appreciate their sacrifices, culinary delicacies, and formatting skills. I would not have been able to get to this stage without them, and I will forever be a comensalita.

Para el resto de mi familia estadounidense y colombiana, Hana & Bennett, Mama Carol, mi querida Tía Beatriz, mis Tíos Roberto y Tuto, mi brillante madrina Clemencia, y mis primos (especialmente Adriana, Andres, Juan Manuel y Paola), el apoyo y amor que he recibido de ustedes siempre han sido en momentos oportunos. Les agradezco el suporte que me han brindado.

Also muchísimas gracias to the Jones-Luzzi family (Bruce, Bryce, Carol, Cliff, Kelly, Marge, Ted). The time, wisdom, and unconditional love I receive has and continues to be a huge blessing in my life. Never has mi casa es tu casa felt more true to me than when I visit beautiful St. Joe.

To my ever-patient advisor and academic rabbi, Dr. Allard Dembe, toda rabah! Your specialized and erudite input and careful attention to detail has not gone unnoticed or unappreciated. I would not be here today without you.

To the inspiring Dr. Linda Lobao, agradeço-te! This very dissertation, my rural sociology minor and research interests, my curiosity for geography, my writing skills, and even my
teaching style are all directly a result of you. One of the best decisions I have ever made in graduate school is to take your Diffusion of Innovation course and never looking back.

To the brilliant and compassionate Dr. Sharon Schweikhart, many thanks for all of the intellectual and emotional nourishment. Your unconditional support and clear commitment to all of your students endear you to all of us and is something I aspire to in my teaching.

To the sage Yoda of the division and my committee, Dr. Thomas Wickizer, thank you so much for being a calming force throughout the dissertating process. I hope to keep making you and the Division proud even as Dr. Naomi and not just the “almost Dr. Naomi.”

Thank you to OSU, CPH (especially Dawn Williams, Elizabeth Hope, Renee Watts, and Kynthia Droesch), HSMP, Center for HOPES and Dr. Dembe, Center for Lean Healthcare Research and Dr. Schweikhart, Dr. Eric Seiber, the OSU Department of Geography and Drs. Daniel Sui and Ed Malecki, and ODNR (especially Mike Engle, Jim McDonald, and Madge Fitak), for your financial and other support.

Finally, special thanks to all of my friends. In the words of Anne Shirley, you are all true kindred spirits. I would not be who I am today without all of you. To Lindsey, many well-deserved gracias for the often late-night discussions, suggestions, and input; you were, and are, a wonderful epidemiologist and statistical guru, puzzle solver, and friend. You’re next! To Dr. Sandra Tanenbaum, my deepest appreciation for your endless wisdom and true friendship. You are a treasure, and one I look forward to keeping for years to come. To Brandon Egren, Father Clem Oyafemi, Father Dean Mathewson, Hanna Park, Hillary Appel, Hyeseon Jeong, Jacqueline Mills, Joan McMahon, Mark Temple, Matt Fleming, May Nacion, Mike Zettler (and Ann Schreck), Raluca Ciochina, Sandy Wong (and Johnathan Rush), Vicky Sakr (in alphabetical order) AND all the countless other interesting and unique spirits with whom, and from whom, I have had the good fortune to interact and learn. THANK YOU.
VITA

2002 ................................................................. Westerville North High School

2006 ................................................................. B.S. Economics, Wharton School, University of Pennsylvania

2006-07 ................................................................. Consulting Analyst, Sg2

2008-10 ................................................................. Graduate Research Assistant, Center for HOPES, The Ohio State University

2010 ................................................................. M.P.H. Health Services Management and Policy, The Ohio State University

2010-12, 2015 ................................................................. Graduate Teaching Associate, Division of Health Services Management and Policy, The Ohio State University

2012-14 ................................................................. Senior GIS Intern, Ohio Department of Natural Resources

2013 ................................................................. M.A. Geography in Spatial Analysis & GIS, The Ohio State University

2014-15 ................................................................. Administrative Director, Center for Lean Healthcare Research, Fisher College of Business, The Ohio State University

Fields of Study

Major Field: Public Health
# Table of Contents

Abstract ......................................................................................................................... ii

Acknowledgments ........................................................................................................ iv

Vita ................................................................................................................................. vi

List of Tables .................................................................................................................. xii

List of Figures ................................................................................................................ xiv

Chapter 1  Exploration of Concepts ................................................................................ 1

1.1  Aims ......................................................................................................................... 2

1.1.1  Aim 1 ................................................................................................................... 2

1.1.2  Aim 2 ................................................................................................................... 3

1.1.3  Aim 3 ................................................................................................................... 3

1.2  Conceptual Model ................................................................................................... 3

1.3  Theoretical Background ......................................................................................... 9

1.3.1  Rurality ............................................................................................................... 9

1.3.2  Rural Health ....................................................................................................... 11

1.3.2.1  Rural Hospitals ............................................................................................... 13

1.3.2.2  Critical Access Hospitals .............................................................................. 14

1.3.2.3  Rural Hospital Closure .................................................................................. 16

1.4  Social Capital ......................................................................................................... 17

1.4.1  Development of Social Capital Concept ............................................................ 20

1.4.2  Social Capital & Health ..................................................................................... 21

1.5  Public Health Significance ....................................................................................... 23

Chapter 2  Developing a Social Capital Model for the Health Services Context .............. 25

2.1  Data Sources & Variable Selection ......................................................................... 26
2.2 Variable Definition & Justification ..............................................................................37

2.2.1 Social Capital Variables ..........................................................................................38

2.2.1.1 Domain 1: Population Characteristics .................................................................39

2.2.1.1.1 Population Loss, 1990-2000 (%) .......................................................................40

2.2.1.1.2 Single Parent Family Households, 2000 (%) .........................................................41

2.2.1.1.3 Diversity Index, 2000† (Aim 1 Only) .................................................................42

2.2.1.1.4 Linguistically–Isolated Households, 2000 (%) (Aim 1 Only) .........................42

2.2.1.1.5 Diversity Score, 2000 (%)† ..................................................................................43

2.2.1.2 Domain 2: Housing & Home Ownership .............................................................44

2.2.1.2.1 Vacant Housing, 2000 (%) ..................................................................................45

2.2.1.2.2 Owner-Occupied Housing, 2000 (%)† ...............................................................46

2.2.1.3 Domain 3: Education ............................................................................................47

2.2.1.3.1 High School Graduates, 2000 (%)† .................................................................47

2.2.1.3.2 Some College or Higher Completion, 2000 (%)† .............................................48

2.2.1.4 Domain 4: Economic Indicators ...........................................................................48

2.2.1.4.1 Gini Coefficient, 2000 .........................................................................................49

2.2.1.4.2 Unemployment, 2000 (%) ....................................................................................51

2.2.1.4.3 Per Capita Income, 2000 ($)† .............................................................................52

2.2.1.4.4 Population in Poverty, 2000 (%) .........................................................................52

2.2.1.5 Domain 5: Crime & Violence .................................................................................53

2.2.1.5.1 Total Violent Crimes (#) ......................................................................................54

2.2.1.5.2 Total Property Crimes (#) ....................................................................................55

2.2.1.5.3 Total Substance Abuse Crimes (#) .................................................................55

2.2.1.6 Domain 6: Health ................................................................................................56

2.2.1.6.1 Uninsured, 2000 (%) ...........................................................................................57

2.2.1.6.2 FTE Physicians per 1000, 2000 (#)† ..................................................................58

2.2.1.7 Domain 7: Community Participation .................................................................58

2.2.1.7.1 Penn State Composite Social Capital Index (Aim 1 Only) ....................................59

2.2.1.7.2 Presidential Voter Turnout, 2000 (%)† ..............................................................60

2.2.2 Hospital Variables ..................................................................................................61

2.2.2.1 Defining Hospital Closure .....................................................................................61

2.2.2.2 Describing Hospital Closure ................................................................................62

2.2.2.3 Aggregating Hospital Closure Data by County .....................................................63
Chapter 3 Using the Social Capital Model to Analyze the Key Predictors of Hospital Closings ........ 73

3.1 Abstract ................................................................................................................................. 73

3.2 Introduction ............................................................................................................................ 74

3.3 Methodology .......................................................................................................................... 77

3.3.1 Objective .......................................................................................................................... 82

3.3.2 Model A: Logistic Regression Using Social Capital Indicators ........................................ 82

3.3.3 Model B: Logistic Regression Using Social Capital Factors ............................................ 83

3.3.4 Adjusting Model A and Model B by Possible Confounding Variables ............................ 84

3.3.5 Adjusting Model A: Logistic Regression by Main Interaction Terms .............................. 86

3.4 Results and Interpretation ...................................................................................................... 87

3.4.1 Model A: Logistic Regression Using Social Capital Indicators ...................................... 88

3.4.1.1 Crude Model A: Main Effects ...................................................................................... 88

3.4.1.2 Adjusting Model A for Confounding ........................................................................ 90

3.4.1.3 Adjusting Model A for Possible Interactions ............................................................. 92

3.4.2 Model B: ........................................................................................................................... 94

3.4.2.1 Factor Analysis ......................................................................................................... 94

3.4.2.2 Crude Model B: Main Effects .................................................................................. 97

3.4.2.3 Adjusting Model B for Confounding ........................................................................ 98

3.4.3 Sensitivity analysis ............................................................................................................ 99

3.5 Discussion .............................................................................................................................. 100

3.5.1 Limitations ....................................................................................................................... 106

3.5.2 Conclusion ....................................................................................................................... 109
Chapter 4  In rural Communities, How do Social Capital Trends Over Time between 1980 and 2000 affect the Likelihood of a Hospital Closing? ................................................................. 111

4.1  Abstract ............................................................................................................. 111

4.2  Introduction ........................................................................................................ 112

4.3  Methodology ........................................................................................................ 114
   4.3.1  Objective ...................................................................................................... 117
   4.3.2  Model A: Logistic Regression Using Social Capital Indicators ..................... 117
   4.3.3  Model B: Logistic Regression Using Social Capital Factors ......................... 118

4.4  Results and Interpretation .................................................................................... 118
   4.4.1  Model A: Logistic Regression Using Social Capital Indicators ..................... 119
   4.4.2  Model B: Logistic Regression Using Social Capital Factors ......................... 121
   4.4.3  Sensitivity analysis ...................................................................................... 126

4.5  Discussion ............................................................................................................. 126
   4.5.1  Limitations .................................................................................................. 132
   4.5.2  Conclusion .................................................................................................... 135

Chapter 5  Do Changes to Geographic Scale of Analysis Affect the Association Between Social Capital and Rural Hospital Closure ................................................................. 137

5.1  Abstract ............................................................................................................. 137

5.2  Introduction ........................................................................................................... 138

5.3  Theoretical Background ....................................................................................... 140
   5.3.1  Spatial Data Considerations ........................................................................ 140
       5.3.1.1  Spatial Autocorrelation or Dependency .................................................. 140
       5.3.1.2  Modifiable Areal Unit Problem ............................................................... 142
       5.3.1.3  Boundary Problem ................................................................................ 145
   5.3.2  Scale of Spatial Data .................................................................................... 146
       5.3.2.1  Counties .................................................................................................. 146
       5.3.2.2  Proximal Hospital Areas (PHA) .............................................................. 147

5.4  Methodology ....................................................................................................... 148
   5.4.1  Redefining Geographic Scale: Proximal Hospital Area (PHA) ......................... 149
   5.4.2  Adjusting for Proximal Hospital Areas .......................................................... 151

5.5  Results and Interpretation .................................................................................... 151
LIST OF TABLES

Table 1: RTI Spatial Impact Factor Dataset Description ................................................................. 29
Table 2: Social Capital Variables & Descriptive Statistics (2000) .................................................. 33
Table 3: Rural County Classification ................................................................................................. 36
Table 4: Rural Hospitals by Census Region & Rurality ‡ ................................................................ 37
Table 5: Domain 1 Variables & Descriptive Statistics ...................................................................... 40
Table 6: Domain 2 Variables & Descriptive Statistics ..................................................................... 45
Table 7: Domain 3 Variables & Descriptive Statistics ..................................................................... 47
Table 8: Domain 4 Variables & Descriptive Statistics ..................................................................... 49
Table 9: Domain 5 Variables & Descriptive Statistics ..................................................................... 54
Table 10: Domain 6 Variables & Descriptive Statistics .................................................................... 57
Table 11: Domain 7 Variables & Descriptive Statistics .................................................................... 59
Table 12: Rural Hospitals by Census Regions & Closure Status ....................................................... 63
Table 13: Total Hospitals & Hospital Closures Per County ............................................................... 64
Table 14: Social Capital Variables & Descriptive Statistics by County Hospital Closure Status ........... 66
Table 15: Possible Confounders & Descriptive Statistics by County Hospital Closure Status .............. 69
Table 16: Social Capital Variables & Descriptive Statistics (2000) .................................................... 79
Table 17: County-Level Hospital Data .............................................................................................. 82
Table 18: Odds Ratios in Model A (2000) .......................................................................................... 88
Table 19: Adjusted Odds Ratios in Model A (2000) ......................................................................... 90
Table 20: Crude and Adjusted Odds Ratios for Possible Confounding Variables in Model A (2000) .... 91
Table 21: Crude and Adjusted Odds Ratios for Possible Interaction Terms in Model A (2000) .......... 93
Table 22: Factor Analysis Eigenvalues (2000) .................................................................................. 95
Table 23: Factor Loadings (2000) ..................................................................................................... 96
Table 24: Odds Ratios in Model B (2000) .......................................................................................... 97
Table 25: Crude and Adjusted Odds Ratios for Possible Confounding Variables in Model B (2000) .... 98
List of Figures

Figure 1: Conceptual Model ................................................................. 4
Figure 2: Unemployment & Violent Crime Over Time ........................................... 121
Figure 3: Economic Vulnerability & Crime Risk Over Time ................................. 125
Figure 4: John Snow Cholera Map showing Zoning Effect .................................. 144
Figure 5: Census Region Map ........................................................................ 190
Chapter 1  EXPLORATION OF CONCEPTS

The purpose of this study is to analyze the association between the social capital of rural communities and rural hospital closure. Separately, social capital and rural hospital closure are concepts with a myriad of factors acting upon them including, but not limited to, rurality, accessibility, and economics. For example, rurality is defined in large part by population size and commuting patterns. If a community is becoming more rural and losing population, this means the smaller population living there must become more active in order to maintain the same level of social capital. Population loss and increasing rurality also impacts rural hospitals. Lower population means less demand for healthcare services, which affects tertiary care providers and physician specialists hardest. For example, if an obstetrician is able to deliver 250-300 births per year, this means at least one woman in the community must give birth to a child almost every day. If the overall population in a community is shrinking and/or aging, the community may be unable to sustain an obstetrician practice.

Since conceptually both social capital and rural hospital closure are impacted by a wide-variety of factors, Chapter 1 will provide a foundation to not just social capital and rural hospital closure but also review several prominent concepts crucial to understanding the association between them. Defining social capital is a crucial step in this study. Chapter 2 is dedicated to describing the development of the social capital model to be used in the study. Having established the framework, Chapters 3, 4, and 5, will present one of three overall aims in the study. Finally, after reviewing results, Chapter 6 will further develop the implications of the results to the healthcare system, to the provision of care in rural areas, and even to rural policy.
1.1 AIMS

After describing the study briefly and establishing an understanding of key terms, the purpose of this study is to analyze the association between the social capital of rural communities and rural hospital closure. The development of a new Social Capital Model that defines social capital by factors will be detailed in Chapter 2 since it is a key, novel component of this study. After defining social capital, the actual study is divided into the three aims below. Each aim, with its corresponding background and methodologies, will be described in depth in Chapters 3, 4, and 5.

Aim 1, which will be expounded upon in Chapter 3, is focused on establishing a baseline understanding of the relationship between rural hospital closure and the social capital of the rural communities in the United States where these hospitals are located. Since the decision to close a rural hospital is typically based upon long-term poor indicators, Aim 2, to be presented in Chapter 4, evaluates the association between a rural community’s social capital over a 20 year period (1980 – 2000) and the likelihood of hospitals in these communities to close. For the first two aims, a rural community is defined as the county where the hospital is located in. Although previous research justifies the use of county as a proxy for a rural community, its use introduces potential for geographically-related statistical bias. Therefore, Aim 3, to be described in Chapter 5, will redefine the geographic scale of analysis. In this study, a Proximal Hospital Area (PHA) is calculated by drawing a buffer around each rural hospital with a specified distance as the radius and modifying it to best represent each rural hospital’s service area. The association between the social capital for PHAs and rural hospital closure will be determined.

1.1.1 Aim 1
To what extent is the social capital of a rural community, as defined by county, associated with the likelihood of hospital closure in that community between 2000 and 2008 after adjusting for confounders?

- 1A. Are social capital variables or factors significant in predicting the likelihood of rural hospital closure?

---

1B. Does the relationship between the social capital of a rural community and the likelihood of rural hospital closure vary among the 4 U.S. Census regions?

1.1.2 Aim 2
In rural communities, how do social capital trends over time between 1980 and 2000 affect the likelihood of hospital closure after adjusting for confounders?

- 2A. Are social capital factors in 1980, 1990, and 2000, significant in predicting the likelihood of rural hospital closure between 2000 and 2008?
- 2B. How do social capital factors vary over time? Over time, do changes in social capital factors between 1980 and 2000 affect the likelihood for hospital closure between 2000 and 2008?

1.1.3 Aim 3
In what way does defining a rural community by proximal hospital area, using buffer analysis, enable better estimation of the relationship between a rural community's social capital and hospital closure?

- 3A. To what extent is the social capital of a rural Proximal Hospital Area in 2000, as defined by buffer analysis, associated with the likelihood of rural hospital closure in that community between 2000 and 2008 after adjusting for confounders?
- 3B. In what ways does the association between social capital of rural communities and rural hospital closure differ when a hospital’s community is defined as a county (Aim 1) versus a Proximal Hospital Area?

1.2 Conceptual Model
The main purpose of this study is to analyze the association between the social capital of rural communities and rural hospital closure. This relationship functions within a complex system consisting of individual, community, organizational, and political components; components all of whom are interrelated. The complexity of the social capital and hospital closure relationship warrants a careful transdisciplinary consideration of the appropriate conceptual and methodological approach to analyzing the association. As such, the best conceptual modeling approach is one grounded in systems thinking.
A systems-thinking approach emphasizes both the existence and type of relationships between different types of structures in a system to more fully understand why a problem or event occurs.7 This approach has been found to simply depict a wide variety of healthcare problems and calls for a wide variety of academic disciplines to work together.8 It also facilitates understanding of the study by focusing on the system instead of getting lost in individual-level components.

Because the main goals of this study involve the association between social capital of rural communities and rural hospital closure, each of them are major components of this analysis and they are depicted in the conceptual model as black boxes. However, as this chapter will explain, the relationship between social capital of rural communities and rural hospital closure is complex and does not occur in a vacuum. This relationship is acted upon by a wide variety of variables that fall outside of social capital and rural hospital closure, but yet have been shown to affect social capital, rural hospital closure, or both. Many of these variables measure the impact of federal or state level policies. For simplicity, all arrows in the conceptual model above represent a statistical process or analysis.

The first major component of the conceptual model is the Social Capital Model, as represented by the black box on the right. The development process of this new social

---

8 Ibid
capital model builds upon previous research and uses Putnam’s approach of defining social capital as an aggregation of variables that captures both community and individual characteristics. In this study, 20 components ($v_1, v_2, \ldots v_{20}$) that fall within the following seven domains have been aggregated: population characteristics, housing and home ownership, education, economic indicators, crime and violence, health, and community participation. In Putnam’s words, these 20 variables attempt to measure the degree of “social organization, such as trust, norms, and networks that can improve the efficiency of society by facilitating coordinated actions.”

A detailed description of how this social capital model was developed is provided in Chapter 0.

The second black box in the Conceptual Model represents the dependent or outcome variable – rural hospital closure given a level of Social Capital indicators, independent or input variables. As described in Chapter 2, rural hospital closure data is used to create a dichotomous outcome variable (0/1) reported at the county level that reports whether a rural county experienced any rural hospital closure during the time period of the study (2000 – 2008). In the initial models, the full data set representing social capital and rural hospital closure for all 2,052 rural counties in the United States will be used. However, later for sensitivity analysis, a separate subset of the data set that only includes the rural counties that have at least one rural hospital (n = 1,614) will be used to re-run the models.

To reiterate, all arrows in the conceptual model represent a statistical process or analysis. In Aim 1 and 2 (See Chapter 3), two methodologic approaches will be compared – stepwise logistic regression alone and a combination of factor analysis with logistic regression. The initial approach considers each social capital variable (n = 20) equally and removes the least significant variable, one at a time. The final model (Model A) includes only those social capital variables found to significantly predict rural hospital closure. The second approach first uses a data reduction technique known as factor analysis, a statistical tool commonly used in social capital studies. Factor analysis creates and determines significant factors, which are then run through logistic regression to build a model (Model B). Regardless of the methodological approach taken, both Model A and Model B need to be tested through a process known as model building.

One of the steps in model building is to force possible confounding variables into the model one at a time and determine if each deserves inclusion in the model based on its impact on the odds ratios generated. Confounding variables are variables outside of the independent variables (social capital of rural communities) that may affect the dependent variable (rural hospital closure). Therefore, any confounding variables found to significantly affect the model must be controlled for in the analysis. In this study, six confounding variables will be forced into both Model A and Model B to determine if any requires being controlled for in the study. These six variables represent environmental and policy characteristics of the surrounding area that could have an impact on rural hospital closure. They are the presence of rural health clinics or federally qualified health centers, health professional shortage areas designation, economic dependency classification, degree of rurality, and Census region variation. For example, it is important to check if the presence of additional healthcare providers in a rural hospital’s environment affects the impact a rural hospital closure. All of these variables are based of federal or state level government criteria, which are influenced by policy.

Although not directly depicted in the conceptual model, there are additional factors that can influence the relationship between social capital of rural communities and rural hospital closure. Many of these factors can be classified as being hospital-specific or policy-related, and although they can affect the independent and dependent variables in this study, data collection for these factors is either time or cost prohibitive. Thus, these additional factors fall outside the scope of this study. However, below is a brief description for some of them.

Many hospital-specific factors are associated with rural hospital closure that are not accounted for in the analytical model of social capital, but are nevertheless important. As mentioned previously, much of this data is time or cost prohibitive to collect for a study including every rural hospital in the United States; it is an inherent limitation of this study. Financial factors like days cash on hand or capital expenses can be unrelated to the other parts of the conceptual model but have been found to be quite associated to rural hospital closure. Outside of quantitative characteristics, rural hospital closure is also related to qualitative factors such as the level of executive expertise and type of system affiliation agreement the hospital has entered into. For example, hospital administrators, well versed in the Critical Access Hospital (CAH) program policies, were more likely to know of a loophole that allowed governors to designate a hospital as a CAH, regardless of rurality.
This loophole has since been closed, but dozens of hospitals technically not in a “rural” area received CAH designation and the additional payments from CMS. In certain communities, if the administrators have a good rapport with the community and its leaders, a rural hospital may be able to seek financial support through the form of a county or local tax levy. In the case of a rural hospital struggling to get fast-enough internet service in order to operate an electronic health record, an administrator can seek funding from a designated federal program run by the Federal Communications Commission. Although these hospital-specific factors would be a welcome addition to the analysis, collecting hospital-specific data would be too time intensive or cost prohibitive. Therefore, none of these factors will be included in this study, but the lack of their inclusion will be limitation.

As governments and organizations such as the World Bank have shown and tested, policies adopted by federal (national), state, and local levels can impact (a) social capital and its many components, (b) rural hospital closure, and (c) factors that affect the relationship between social capital and rural hospital closure. Apart from differentiating policies by the level of government, policies that impact social capital, rural hospital closure, and factors affecting both, also can be further categorized by the policy area it targets, which can be healthcare-specific or range from a wide-variety of topics including education, immigration, economic development, and housing.

At the federal level, healthcare-focused enacted policies have had a major impact on rural communities and their hospitals. In 1946, the Hill-Burton Act primarily provided financial assistance for healthcare capital investments that led to hospital openings, but it also prohibited healthcare facilities from discriminating by race, color, national origin, or creed, which could impact social capital. In addition, the creation of federal entitlement or means-tested programs like Medicare and Medicaid or the passage of sweeping reforms like the Patient Protection and Affordable Care Act, qualify rural hospitals for additional federal assistance. Rural hospitals benefits from extra federal financial assistance that acknowledge their critical access to healthcare role and the disproportionate share of Medicaid and uninsured patients they care for. In the United States, policies, even ones not solely focused on healthcare, can impact social capital, rural hospitals, and additional factors. For example, immigration policy changes facilitate the recruitment of foreign-trained physicians or nurses to rural communities. This can affect rural communities’ social capital with the influx of a more diverse, highly-capable population and also help hospitals have the human capital
necessary to remain operational. Therefore, federal policies can act upon social capital and rural hospital closure, as well as factors that impact each individually or the relationship between them.

Although rural communities and their hospitals can be acted upon by federal policy changes, state policies, can significantly impact rural communities' social capital, hospitals, and factors affecting both. Again, health-specific policies have perhaps the clearest impact on all three. The role states play in funding and directing Medicaid affects the percent of the population with health insurance. This directly impacts the financial health of rural hospitals because even low reimbursement from Medicaid plans is preferable to the compensation received from uninsured patients. When compared to uninsured individuals, those with health insurance are better able to contribute both economically and socially to their communities\textsuperscript{11}, thus positively impacting social capital, rural hospitals, and additional factors. Outside of healthcare specific policies, state policies aimed at increasing the effectiveness of public education enable students to better develop into leaders or to more likely pursue higher education, both effects can positively impact social capital, rural hospitals, and additional factors.

In healthcare, care is ultimately delivered locally, which cause many believe that all healthcare is local.\textsuperscript{12,13,14} Not surprisingly, local policies can have a tremendous effect on the local level of social capital, hospitals, and outside factors. In some cases, a county could vote on a levy to help support the local rural hospital, or a hospital’s foundation could organize a fundraiser. Outside of healthcare, local governments also fund and direct primary education and law enforcement. In both these areas, local governments can directly impact social capital and rural hospitals.

1.3 THEORETICAL BACKGROUND

The overall purpose of this study is to analyze the association between the social capital of rural communities and rural hospital closure. As the conceptual model shows, this relationship functions within a complex system consisting of individual, community, organizational, and political components. Each of these components has been analyzed in previous research and is significantly affected by changes in policies. Therefore, before embarking on this work, the study must be grounded by a comprehensive understanding of rurality, rural health and healthcare services, and social capital. This requires an interdisciplinary approach, which is reflected in the methodologies and theories referenced in this study from health services research, rural sociology, and geography. So what is rural? How is it defined? Given the need for defining “rural,” why is the differentiation between rural and urban areas important? How do these differences ultimately affect health and healthcare delivery?

1.3.1 RURALITY

In the United States, “rural” was first used by the U.S. Census Bureau in 1874. In 2010 decennial U.S. Census, 19.3 percent (59.5 million) of the population resided in rural areas.\textsuperscript{15} Yet as people continue to migrate towards metropolitan regions, research is increasingly reflecting a unitary view of development focused on large cities with rural region only being included indirectly or as a contrast to urban regions.\textsuperscript{16} Rural regions are often unstudied or a default category.\textsuperscript{17,18,19} This sense of exclusion continues to the very definition of rural. Latin in origin, “rural” is defined as a land area that is not a city or town.\textsuperscript{20}

Several federal agencies use this definition of exclusion. The U.S. Census Bureau (USCB) classifies all territories, population, and housing units as being located in or outside urban

\begin{footnotes}
\item[18] Byrne, D. (2001). \textit{Understanding the urban}. Houndmills, Basingstoke: Palgrave
\end{footnotes}
clusters. The basis for this classification schema is the combination of population density, relationship to cities, and population size. An “urbanized area” (UA) encompasses a densely settled territory of core U.S. Census block groups or blocks that have a population density of at least 1,000 people per square mile and surrounding U.S. Census blocks with an overall density of at least 500 people per square mile. Building on the USCB definition, the U.S. Office of Management and Budget (OMB) has a similar classification schema that characterizes counties as metropolitan or non-metropolitan based on population size and integration with large cities. Metropolitan counties must have one or more central cities of at least 50,000 inhabitants or have one UA of at least 50,000 inhabitants, and the country must have a total population of at least 100,000 (75,000 in New England counties).

Overall, the dichotomous definition for “rural” may facilitate analysis, but it can make rural areas seem homogeneous. Rural areas do share similarities in (1) having small-scale, low-density settlement, (2) being distant from large, urban centers, and (3) having a specialized economy often dominated by single employer. However, the problem lies in extending the definition of rural to more than just a geographical concept and seeing it as a combination of social practices and structures. The problem of defining “rural” in health care has long been acknowledged by the federal government such as the time following the creation of the “Rural Health Clinics” designation in 1977. Across the United States, rural areas are in fact quite heterogeneous. “The heterogeneity of rural areas makes it almost certain that

---

22 Ibid
broad generalizations about rural conditions and future opportunities are wrong for a large subset of such areas...It must be recognized that rural is both a continuum and a class.”

This dissertation study makes no attempt to assign a generic definition of rural. One method to better differentiate between rural areas could be to use a smaller scale, but unfortunately, not enough data are available nationally at a scale smaller than county. In recognition that rurality falls along a continuum, the U.S. Department of Agriculture’s Economic Research Service developed the Rural-Urban Continuum Codes in 1975, sometimes referred to as “Beale Codes.” This schema still classifies counties down into metropolitan and non-metropolitan. However, metropolitan counties are further differentiated into four categories, non-metropolitan into six categories (See Appendix B). With this schema, rural counties can be classified as urbanized, less urbanized, or thinly populated. Consequently, this study will use the Rural-Urban Continuum classification system to define “rural” but will still use comparisons between rural and urban for illustrative purposes.

1.3.2 Rural Health
In comparison to urban areas, rural areas are particularly vulnerable to hardship and inequality. “Poverty has the three r’s. They are race, region, and rurality.” Factors contributing to a rural community’s vulnerability include: a less-varied industrial mix, fewer employment options, lower quality jobs, a less-educated population, gender, family, poor transportation infrastructure, and a past history of exploitative economic structures or institutional arrangements. These economic, educational, cultural, and social factors, in combination with a lack of recognition by legislators and sheer isolation of living in remote areas, impede rural Americans ability to lead normal, healthy lives.

---

Rural residents in the United States are less likely to receive employer-sponsored health insurance and more reliant on government assistant programs like Medicaid and food stamps. Lack of funding for health behavior and promotion programs contribute to making rural 8th graders twice as likely to smoke cigarettes as their urban counterparts and 40 percent of rural 12th graders to use alcohol while driving compared to 25 percent of urban 12th graders. Although these health statistics begin to describe the lower health outcomes in rural areas, access to health care services and professionals is a well-recognized problem. Although 20 percent of the U.S. population lives in rural areas, less than 10 percent of physicians practice there.33 Over 70 percent of Health Professional Shortage Areas (HPSA) and 87 percent of Mental Health Professional Shortage Areas are in rural and frontier areas of all states and U.S. territories.34 In fact, 20 percent of rural counties lack any mental health services.35 Finally, rural residents have more accessibility issues than their urban counterparts to get access to healthy food and may live in a food desert.36 In Ohio, for example, 24 percent of rural Ohio households do not live within a 10-minute drive of a retail grocery store of any size but may live within closer proximity to a fast food establishment. Of these households, 5 percent of those living a driving distance away from a retail grocery store do not own a car and rural areas often lack reliable public transportation.

Overall, rural populations are vulnerable in many ways that urban ones are not. Lack of accessibility to healthcare is compounded by a weak public transportation infrastructure and geographic remoteness. To add to the health concerns of rural populations, rural residents, on average, are older and have more chronic or serious illnesses, higher injury-related mortality rates, and lower insurance rates than urban residents.37 Consequently, a deeper understanding of the role rural hospitals play in their communities is warranted.

34 Ibid
1.3.2.1 Rural Hospitals

The presence of a hospital is an anchor for rural healthcare services. In a recent study, 94 percent of rural residents living in a community with a closed local hospital want it reopened, 40 percent report having serious problems accessing hospital care, and 17 percent report a family member would decide not to seek medical care on one or more occasions because of the additional travel time to the next closest hospital. Rural hospital closure creates or worsens access problems. Following a rural hospital closure, travel time to the next nearest hospital often increases to over 30 minutes or even up to 3 times longer for rural residents in mountainous regions. Rural hospital closure affects residents not only in the reduced access to facilities but also with the loss of physicians. In the 1980s, rural areas in which their local hospital closed experienced a decrease of 12.8 to 25.4 percent in the number of physicians by 1990.

Although rural hospitals play a key role in supporting rural residents and their health outcomes, those hospitals are symbols of a rural community's identity and sustainability. Rural economies are already vulnerable due to their reliance on a single major employer. Rural hospitals often are the second largest employer in a rural community after the local school system and employ some of the highest paid individuals in the community. A

hospital is a source of economic stability.\textsuperscript{47,48} In a recent Oklahoma study, rural communities would suffer a loss of \$2 to \$6.4 million in retail sales if a local hospital were to close.\textsuperscript{49,50}

Overall, rural hospitals play a crucial role in the survival of rural communities by playing an equally important role economically as well as socially. In recognition of this role, the federal government instituted the Critical Access Hospital (CAH) program to help these front-line hospitals remain operational. CAH designation, with few exceptions, requires a hospital to be serving and based in a rural area. Therefore, one can assume that the great majority of CAHs are in rural areas, but one cannot assume that all rural hospitals are CAHs. Yet, so many rural hospitals have Critical Access Hospital designation that a deeper look at this program is necessary.

1.3.2.2 Critical Access Hospitals

In the late 1980s, the Health Care Financing Administration (HCFA), a precursor to CMS, was concerned with keeping rural hospitals operational. HCFA funded a demonstration program in eight rural states for rural hospitals to receive additional Medicare and Medicaid reimbursement if they complied with specific bed counts and length of stay requirements.\textsuperscript{51} In the 1990s, the introduction of managed care significantly increased the rate of rural hospital closure. Concerned with the health, social, and economic impact a rural hospital closure would have on rural communities, a section of the Balanced Budget Act (BBA) of 1997 introduced the Medicare Rural Flexibility Program (Flex Program). The Flex Program, based on the HCFA demonstrations, helped rural hospitals that were finding it difficult to recover Medicare costs under prospective payment system. The BBA and the Balanced Budget Reconciliation Act of 1999 coined the term "Critical Access Hospitals" or CAH.\textsuperscript{52} Previously, some rural hospitals had already been identified by the Department of Health


and Human Services (HHS) as “Emergency Access Care Hospitals (EACH)” or “Primary Care Hospitals (PCH).” These hospitals converted to CAH status. However, the program was so popular that CAH eligibility was expanded. State governors gained the power to designate new CAHs even if the hospital was in a metropolitan county.

It should be noted that overall, 54 percent of CAHs had to reduce their acute-care capacity, and usually decreased the number of beds, to qualify for CAH designation. By 2001 year end, 545 hospitals in 43 states had received CAH designation, or 1 of every 9 non-federal, short-stay hospitals participating in Medicare. By the end of 2002, there were 723 CAHs, or 1 of every 7 hospitals (1 of every 3 hospitals located in non-metro areas). Most CAHs are clustered in the middle of the country; one-third of CAHs are located in Nebraska, Kansas, Texas, Iowa, and South Dakota. Despite the slightly higher CMS reimbursement rates, 60 percent of CAHs still report negative total margins.

Although rural hospitals report that the CAH program is essential for their survival, in 2003, the Medicare Payment Advisory Commission (MedPAC) estimated that Medicare paid CAHs an average of $850,000 more than the agency normally would have paid. In ongoing deficit reduction talks at the federal level, the Critical Access Hospital program has been identified for cutting at least twice by both the Congressional Budget Office (CBO) and President Barack Obama’s budget proposal. The loophole, allowing state governors to override CAH program requirements and designate additional CAHs, was eliminated. At this point, the CAH program includes over 1,300 hospitals – nearly 1 in 4 acute care hospitals. In 2006, President Obama recommended narrowing the CAH definition to exclude hospitals that are within 10 miles of another hospital. This potential change would affect the CAH status of 61 hospitals and save the federal government $4 billion.

Having established the fact that rural hospitals, especially CAHs, are vulnerable and are “deserving” of federal assistance, the purpose of establishing this program is directly related

55 Ibid
56 Ibid
58 Ibid
59 Ibid
to the goal of avoiding the closure of a rural hospital thus maintaining accessibility to healthcare and support the local economy. This brings to question the very dependent variable of this study – rural hospital closure. What is closure? How often does this occur? What impact does closure ultimately have on rural communities?

1.3.2.3 Rural Hospital Closure

Hospital closure, in general, is a gradual process. Hospitals will experience declines in discharges occurring for a period of time preceding the data the hospital’s license is surrendered.\(^6\) Declines in hospital services are most pronounced in the last 2-3 years before closure.\(^6\) One of the primary causes cited for rural hospital closure is financial distress. Contributing to financial problems are decreases in insurance reimbursement, low occupancy, declining demand for inpatient days due to monitoring by managed care, weak local economies, and lack of capital for reinvestment.\(^6\),\(^6\) In the 1980s, the introduction of the Prospective Payment System (PPS) “may have contributed disproportionately to the financial distress that preceded closure in the smallest rural hospitals”.\(^6\) As a result, 10 percent of all U.S. rural hospitals closed, although nearly half continued operating as a health-related facility.\(^6\),\(^6\) In the 1990s, strict managed plans produced significant cost savings nationally but again many rural hospitals already at risk were unable to cope to the changing healthcare system and closed.

Apart from financial distress, rural hospitals are at risk for a wide-variety of problems. First, rural areas are often designated Health Professional Shortage Areas. Difficulty in recruiting and retention can lead to burnout among rural-based physicians and all health

\(^6\) Ibid
professionals. Rural physician openings often pay significantly less than urban ones and are less able to support specialized practices.\textsuperscript{68} Second, rural residents themselves threaten the survival of rural hospitals. The rural population continues either migrating towards metropolitan areas or at least commuting for job opportunities not found in rural areas.\textsuperscript{69} Rural residents also may lack confidence in their local hospital and bypass it in favor of an urban hospital further away.\textsuperscript{70} Finally, the relatively weak position rural hospitals have in the healthcare market compared to large, urban-based health care systems fails to give them the ability to respond quickly to the changing healthcare industry.

1.4 SOCIAL CAPITAL

In rural areas, one factor that seems to be protective of local hospitals and helps keep them operational is the level of social capital in the community. It is important to understand what social capital is and how it is measured. First, social capital is “not a single explanation or variable but rather a variety of explanations of how informal social relationships are important for human behavior.”\textsuperscript{71} No single definition has been established. The three most used social capital definitions used by health researchers were conceptualized by Pierre Bourdieu, James Coleman, and Robert Putnam.\textsuperscript{72}

Pierre Bourdieu’s approach to social capital focuses on resources individuals possess through their social networks.\textsuperscript{73} “Social capital is the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition – which provides each of its members with the backing of the collectively-owned capital, a ‘credential’ which

entities them to credit, in various senses of the word.”⁷⁴ Although individuals clearly contribute and benefit from social capital, many researchers argue that social capital is a more encompassing concept that describes an entire social network or community. It extends beyond individuals to include how individuals act upon each other. In fact, the conceptual debate of whether social capital can be measured at an individual or community level is reflected in measurement considerations.⁷⁵

Building upon Bourdieu’s definition, James Coleman’s seminal work defines social capital by its function. “It is not a single entity, but a variety of different entities with two elements in common: they all consist of some aspect of social structures, and they facilitate certain actions of actors – whether persons or corporate actors – within a structure.”⁷⁶

Characterizing social capital as having a structure has directed much effort into the identification of three separate social capital dimensions – structural, cognitive, and behavioral.⁷⁷,⁷⁸,⁷⁹ Structural components of social capital relate to social relations or networks often operationalized as social participation, strength of ties, and organizational affiliation.⁸⁰ The cognitive dimension deals with the quality of social relations often referred to as perceptions of trust and reciprocity.⁸¹ Finally, the behavioral component reflects actions individuals take due to social ties and resources within their community such as voting participation and actual attendance in groups.⁸²

Also noteworthy is the work focused on just the structural component of social capital. Researchers have also begun classifying the types of social networks into bonding or

---

⁸⁰ Ibid
⁸¹ Ibid
⁸² Ibid
bridging varieties of social capital. Bonding ties are strong attachments formed when people already know one another and are frequently brought together, as in close social contacts with family, relatives, and friends. However, they may be exclusionary and may fail to produce society-wide benefits in cooperation and trust. Bridging ties are weaker and refer to connections that cross-cut social groups that may not often interact with one another. Often bridging ties are viewed as more important in helping form widespread solidarity within a community. In the literature, a sub-category of bridging ties is referred to as linking ties. Linking ties are “norms of respect and networks of trusting relationships between people who are interacting across explicit, formal, or institutionalized power or authority gradients in society.”

Reflecting the growing literature and findings, Robert Putnam opted to focus on social capital being a community-level resource or social feature. “Social capital here refers to features of social organization, such as trust, norms, and networks that can improve the efficiency of society by facilitating coordinated actions.” Broadly defined, social capital is the set of rules, norms, obligations, reciprocity, and trust embedded in social relations,

94 Ibid
social structures, and society’s institutional arrangements that enable members to achieve their individual and community objectives.97

1.4.1 DEVELOPMENT OF SOCIAL CAPITAL CONCEPT

Thus, social capital is not a homogeneous concept. It is comprised of various social elements that promote both individual and collective action. Not surprisingly, there are many ways to measure the multifaceted concept of social capital. Three principal ways to measure social capital have been identified from empirical studies.

The first approach focuses on individuals as Bourdieu’s work does and measures social relations directly: assessing the number, structure, or properties of relationships among individuals. Use of this measurement approach is most appropriate for social capital resulting from the transmission of information across overlapping or intense contacts or social networks. Teachman, Paasch, and Carver (1996), for instance, use questions about the frequency of conversations between parents and their eighth grade children about their school experiences and their plans for high school and college to measure parental academically-related social capital.98

The second approach focuses on individuals’ beliefs about their relationships with others, as measured by trust. Paxton (1999), for instance, relies on questions from national survey data about the level of trust in other people and major social institutions to examine changes in levels of social capital over time.99 Although both these approaches are utilized in research, they reflect more of the individual’s interaction with social capital.

The third approach to measuring social capital was originally developed by Robert Putnam’s own work in Italy and the United States. Putnam looks beyond the individual and includes the impact community organizations play on building social capital. He measures membership in certain voluntary organizations because of the social ties these organizations forge within communities. By not depending on individual-level data on social

---

ties within a community, Putnam is able to measure social capital in areas where social tie information is unavailable.\textsuperscript{100}

Since this study requires studying the entire United States, specific social tie information cannot be collected in a timely, cost-effective manner. Furthermore, rural areas are far less studied than urban areas in the field of social capital research. Therefore, this study will build upon Robert Putnam’s work. However, it will expand beyond membership to voluntary organizations and voter participation but also include individual characteristics aggregated at the community-level including home ownership, education, and economic attainment.

1.4.2 SOCIAL CAPITAL & HEALTH

Previous research indicates that education, engagement, and awareness of rural communities can increase their support for the local hospital, including higher utilization of local facilities by not bypassing and continued local financial support through mechanisms such as tax levies.\textsuperscript{101} Social capital can have a number of unanticipated economic and staffing benefits to hospitals.\textsuperscript{102} In addition, higher social capital of rural communities is associated with the presence of rural health professionals.\textsuperscript{103,104,105}

Although there is recent, preliminary evidence of the relationship between social capital and healthcare system outcomes, the association between social dimensions to health and illness has long been a common research theme, such as Durkheim’s work on European suicide rates in the 1950s. In the past decade, social capital has had a “meteoric rise in political, economic, and public health rhetoric”.\textsuperscript{106} Yet, most health-focused social capital

research has focused on epidemiological or health-behavior outcomes. Social capital is linked to higher levels of subjective and self-rated health,107,108,109,110,111 better mental health,112,113 lower mortality,114,115 lower suicide rates,116 lower risky behaviors like binge drinking,117 higher quality of life,118 and greater access to healthcare.119 The positive impact social capital can have on health is even more impressive when researchers have found that higher social capital mediates the effect income equality on health outcomes through psychosocial, political, and social pathways.120,121,122

1.5 **Public Health Significance**

With 81 percent of the country living in urban areas, far more research is focused on healthcare and healthcare delivery in urban settings. However, over 19 percent of the United States, almost 60 million people, continues to live in rural areas. As discussed, this rural population faces many health disparities and struggles with unique vulnerabilities such as access to healthcare or healthy food. Hospitals in rural America are more apt to close than urban hospitals, and rural communities are more profoundly affected by rural hospital closure than urban communities when their local or nearest hospital closes. The unique challenges faced with providing healthcare to nearly 60 million Americans living in rural areas merits deeper understanding because rural is different than urban.

Furthermore, rural hospital closure has been a relevant public health concern for decades. Seminal studies have been published to summarize rural hospital closures in the 1980s following the adoption of prospective payment mechanisms and in the 1990s following the rise of managed care. To date, there is no study focused on rural hospital closure between the years 2000 and 2010. As of October 2014, rural hospital closure is a major concern by HRSA’s Office of Rural Health Policy and is currently being tracked by the North Carolina Rural Health Research Program (NC RHRP).\(^{123,124}\) This study also is significant because it includes the entire country and has the potential to influence policies across the U.S.

This study is focused on the relationship between social capital of rural communities and rural hospital closures. Using social capital in a healthcare systems context instead of an epidemiological one (i.e. mortality, morbidity) is a novel application, as is the development of the new Social Capital Model that builds upon previous work, includes a wider variety of variables, and only requires data that is publically available. If successful, the new Social Capital Model developed for this study could be applied in other healthcare systems analyses.

\(^{123}\) National Center for Rural Health Works Regional Workshop, Terre Haute, IN, October 22, 2014, Presentation: Gerald A. Doeksen & Cheryl St. Clair.

Finally, as Aim 3 will address, public health researchers do not account or recognize geospatial errors introduced by the use of data at a specific geographic scale. Although previous public health researchers have recognizes this problem and addressed it in their studies, most fail to address the problem methodologically. This study will use a Geographic Information System to run multiple analyses and check if the geospatial errors are significant and could affect the result of the overall study.

Before determining the association between social capital of rural communities and rural hospital closures in the United States, a comprehensive social capital definition or model must be determined. The development of the new Social Capital Model, which will be used in all three aims of this study, will be explained in following chapter. A preliminary logistic regression analysis of the relationship between social capital of rural communities and rural hospital closures will also be described.
Chapter 2  DEVELOPING A SOCIAL CAPITAL MODEL FOR THE HEALTH SERVICES CONTEXT

Closing a hospital, especially a rural one, is an outcome that depends on a factors including demographic shifts towards metropolitan centers, financial distress from lack of employment opportunities and poor health insurance reimbursement, and challenges in recruiting and maintaining healthcare professional staffing.\textsuperscript{125,126,127,128} Despite pioneering social capital research by Pierre Bourdieu, James Coleman, and Robert Putnam, as well as studies that find a relationship between social capital to healthcare, defining social capital remains a challenge for researchers. Social capital is often considered a vague concept that requires an interdisciplinary approach in order to apply it in both theoretical and empirical research.\textsuperscript{129}

Since specific social tie information is often unavailable, due in part to an inability to collect data in a timely, cost-effective manner, Putnam and other scholars have developed an innovative approach to measure social capital by using relevant, community-level information collected in national and local surveys and by assessing the level of community involvement.\textsuperscript{130} This study builds on the Putnam approach to defining social capital by identifying as many variables previously used in social capital studies and developing a

\textsuperscript{127} General Accounting Office (1991). Rural hospitals: Federal hospitals should target areas where closures would threaten access to care. GAO. Washington, D.C.
more-inclusive model of social capital. This new social capital model aims to be better suited for applications of social capital and health services research and will be used for all three aims of this study. The new model will contribute to the growing body of social capital literature and may be used in new health services research applications.

As the following conceptual model shows, before any of this study's aims can be accomplished, a robust social capital model must be developed. Additional importance must be placed on ensuring components of the model are particularly sensitive to the measurement of healthcare-related outcomes. In this study, the outcome variable is rural hospital closure.

The main aim of this chapter is to create a comprehensive social capital model, which will serve as the independent variable for the remainder of the study. Data sources, variable selection and inclusion criteria for each of the 20 variables included in the novel social capital model will be described in detail. Since the overall goal of this study is to determine the association between social capital of rural communities and rural hospital closure. A second aim for this chapter is to explain the careful process for defining the dependent variable, rural hospital closure. To finesse the statistic model built to determine the association between the independent and dependent variables, an overview of possible confounding variables will also be provided.

2.1 DATA SOURCES & VARIABLE SELECTION

Before reviewing variable inclusion criteria, a brief overview of data sources and variable selection principles need to be addressed. By following the Putnam approach of social capital modeling, this study requires access to a broad, multi-dimensional database. Data sources need to be vetted for accuracy, precision, and completeness. This is particularly important since this study includes all rural communities in the United States and extends across a 20 year time frame.

2.1.1 SOCIAL CAPITAL DATA

To determine social capital, including health variables, in each rural county, the primary database used for this study is the Research Triangle Institute’s (RTI) Spatial Impact Factor (SIF) database. This database is maintained through collaboration between RTI International and Arizona State University’s (ASU) GeoDa Center for Geospatial Analysis
and Computation. The maintenance of this database is funded by a National Institutes of Health grant through the National Cancer Institute. The SIF database is a collection of 28 different datasets collected by various federal departments and agencies, as well as some individual researchers, including by not limited to Health and Human Services (HHS), CMS, Labor (DOL), Justice (DOJ), and Agriculture (USDA). Upon closer inspection, many of the datasets benefit from the collaboration of several of the same federal departments and agencies. For example, the DOL’s Bureau of Labor Statistics uses data collected by the U.S. Census Bureau (USCB) to calculate unemployment data. Furthermore data includes multiple time periods, and it has also been provided at multiple geographic scales, including county, PCSA, ZCTA, Census tract, and MSSA.\textsuperscript{131} The SIF database is a tremendous resource of publically available data with 9,212 variables. Since USCB data is heavily relied upon in the SIF database, the SIF database was merged with the USCB’s USA counties database in order to provide access to the data from 1980, 1990, and 2000. The USA counties database features over 6,600 variables and combines data collected from U.S. Decennial Censuses, American Community Survey 2005 – 2009, Economic Censuses, Census of Population and Housing, and some federal spending data.\textsuperscript{132} Presidential voter response rate data was calculated by the USCB with some data provided by CQ Press. Between the SIF and USA Counties databases, almost 16,000 variables will be accessed.

**Scale:** Although data is available at multiple levels of geography, the county scale, by far, provides the richest dataset. Consequently, even if a smaller scale could better differentiate rural areas, unfortunately, there is insufficient data available nationally at a scale smaller than county.\textsuperscript{133} For the purpose of this model, data that will help define social capital are aggregated at the county level. According to previous rural hospital research, the county represents the best geographic scale at which to study local, societal impacts on rural hospital closure.\textsuperscript{134,135,136} However, in recognition that rurality falls along a continuum, this

study will define rural using the Rural-Urban Continuum Codes developed by the U.S. Department of Agriculture’s Economic Research Service in 1975, sometimes referred to as “Beale Codes”. This schema still allows for counties to be classified as urban (metropolitan) and rural (non-metropolitan). However, metropolitan urban counties are further differentiated into four categories; rural counties have six categories. With this schema, rural counties can be classified as urbanized, less urbanized, or thinly populated. Consequently, this study will use the Rural-Urban Continuum classification system to define “rural” but will still use comparisons between rural and urban for illustrative purposes.

To begin the process of variable selection, a systematic literature review on social capital and its measurement was conducted to evaluate and identify which variables included in SIF and USA counties databases had been used in prior social capital research. In the process of developing a comprehensive social capital model, variables were selected on the basis of their analytical soundness, measurability (across time), relevance and coverage to social capital and health, and relationship to each other. Although proxy measures have been used in social capital research, the richness of the SIF and USA Counties data included most if not all desired data for inclusion. Additional data would have required changes to existing data collection methods or extensive field work. In addition, since data sources are primarily unbiased government sources, quality and precision of the data can be trusted, although further analysis on accuracy is required since the census collecting process is documented to have its drawbacks. In the end, for the initial social capital model in 2000, 60 variables were extracted from a total of 11 datasets. Data includes nearly complete data for all 2,052 rural counties in the United States. Before reviewing each individual variable selected for inclusion, a detailed description of each of the 11 datasets is provided below in: Table 1.

---

### TABLE 1: RTI SPATIAL IMPACT FACTOR DATASET DESCRIPTION

<table>
<thead>
<tr>
<th>#</th>
<th>Dataset Name</th>
<th>Source Department &amp; Agency</th>
<th>Year</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Atlas of the Spatial Patterning of County-Level Homicide</td>
<td>Department of Commerce (DOC); US Census Bureau (USCB), with Messner, Anselin, et al &amp; National Consortium on Violence Research</td>
<td>1990</td>
<td>Data was collected by Messner, Anselin, et al. through a grant from the National Consortium on Violence Research, which is funded through a NSF and NICHD grants. This dataset was primarily used to access population data collected by the U.S. Census in previous decennial periods prior to 2000.</td>
</tr>
<tr>
<td>2</td>
<td>Bureau of Labor Statistics (BLS)</td>
<td>Department of Labor (DOL); BLS</td>
<td>2000</td>
<td>The BLS uses labor and economic data from the Current Population Survey (CPS), which is conducted by the U.S. Census Bureau. The BLS reports unemployment data on all working-age individuals aged 16 or older.</td>
</tr>
<tr>
<td>3</td>
<td>Centers for Medicare and Medicaid Services (CMS) Provider of Services (POS) File</td>
<td>Department of Health and Human Services (HHS); CMS</td>
<td>2000</td>
<td>The POS file contains data on characteristics of a wide variety of healthcare facilities and staffing. Following the passage of HIPPA, health care providers, health plans, and health care clearings must use a National Provider Identifier (NPI) standard. The NPI is a 10-digit identifier.</td>
</tr>
<tr>
<td>4</td>
<td>Economic Research Service (ERS)</td>
<td>Department of Agriculture (USDA); ERS</td>
<td>2004</td>
<td>The ERS provides data and analysis of economic and policy issues related to agriculture, food, farming, natural resources, and rural development. Variables included define rurality. These terms and measurements were developed by ERS with additional input from the BEA and the U.S. Census Bureau.</td>
</tr>
</tbody>
</table>

---

Table 1 continued

|   | Department of Justice (DOJ); FBI | 2000 | Since 1930, the FBI has managed the Uniform Crime Reports (UCR) database. Since 1998, UCR data is generated by reported crimes from over 18,500 city, university and college, county, state, tribal, federal, and other law enforcement agencies. UCR provides a standardized way of reporting and is accessed by all law enforcement agencies. Incident-level data is collected through the National Incident-Based Reporting System (NIBRS), a component of UCR. NIBRS collects data on each single incident and arrest within 22 offense categories made up of 46 specific crimes called Group A offenses. In addition to the Group A offenses, there are 11 Group B offense categories for which only arrest data is reported. By 2007, UCR Program represents more than 285 million U.S. inhabitants or 94.6% of the total population.
|   | Nancy Lozano-Gracia - Income Disparity | 2000 | The Lozano Smith Labor & Employment Consortium, run by Lozano Smith Attorneys at Law, is a California-based firm that provided RTI and the Geoda Center with data on several income disparity terms calculated using different methodologies. Included in this dataset is the standard Gini Coefficient. Nancy Lozano-Gracia was a postdoctoral fellow at Arizona State's Geoda Center.
|   | François Nielsen and Arthur S. Alderson | 1980, 1990 | The 1980 and 1990 data was compiled and shared by François Nielsen and Arthur Alderson. However, the 1980 and 1990 data was originally collected by the U.S. Census Bureau. The 1980 dataset was prepared by Terry K. Adams of the Inter-University Consortium for Political and Social Research in 1992 from the U.S. Census Bureau Census of Population and Housing 1980 file.

---


<table>
<thead>
<tr>
<th>Table 1 Continued</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>8</strong></td>
</tr>
<tr>
<td><strong>9</strong></td>
</tr>
<tr>
<td><strong>11</strong></td>
</tr>
</tbody>
</table>

---


These 60 variables were collapsed down into 20 social capital variables categorized into seven domains. Certain variables like DIVERSITY INDEX were kept as is for inclusion into the social capital model. Other variables like LINGUISTICALLY-ISOLATED HOUSEHOLD first required the summation of linguistically-isolated Spanish, Indo-European, and Asian-Pacific households and then the division of the summation with total households. This leaves LINGUISTICALLY-ISOLATED HOUSEHOLD as a percent of households. Having selected these 20 variables for this new social capital model, the descriptive statistics show that a nearly complete data set was secured from the SIF and USA Counties databases. The variable with the maximum missing data is Substance Abuse Crimes (4.7% missing data) with most missing values attributed to Florida and Illinois. Due to the completeness of the dataset, no correlations will be made for the missing data.
## TABLE 2: SOCIAL CAPITAL VARIABLES & DESCRIPTIVE STATISTICS (2000)

<table>
<thead>
<tr>
<th>No.</th>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Population Loss, 1990-2000 (%)</td>
<td>2015</td>
<td>0.07</td>
<td>0.14</td>
<td>-0.84</td>
<td>0.90</td>
</tr>
<tr>
<td>2</td>
<td>Single Parent Family Households, 2000 (%)</td>
<td>2049</td>
<td>0.10</td>
<td>0.04</td>
<td>0.01</td>
<td>0.41</td>
</tr>
<tr>
<td>3</td>
<td>Diversity Index, 2000†</td>
<td>2052</td>
<td>0.08</td>
<td>0.08</td>
<td>0.00</td>
<td>0.79</td>
</tr>
<tr>
<td>4</td>
<td>Linguistically-Isolated Households, 2000 (%)</td>
<td>2052</td>
<td>0.02</td>
<td>0.03</td>
<td>0.00</td>
<td>0.34</td>
</tr>
<tr>
<td>5</td>
<td>Diversity Score (%)†</td>
<td>2049</td>
<td>0.13</td>
<td>0.13</td>
<td>0.00</td>
<td>0.50</td>
</tr>
<tr>
<td>6</td>
<td>Vacant Housing, 2000 (%)</td>
<td>2049</td>
<td>0.17</td>
<td>0.10</td>
<td>0.03</td>
<td>0.77</td>
</tr>
<tr>
<td>7</td>
<td>Owner-Occupied Housing, 2000 (%)†</td>
<td>2049</td>
<td>0.62</td>
<td>0.09</td>
<td>0.00</td>
<td>0.80</td>
</tr>
<tr>
<td>8</td>
<td>High School Graduates, 2000 (%)†</td>
<td>2049</td>
<td>0.36</td>
<td>0.06</td>
<td>0.11</td>
<td>0.53</td>
</tr>
<tr>
<td>9</td>
<td>Some College or Higher Completion, 2000 (%)†</td>
<td>2049</td>
<td>0.40</td>
<td>0.10</td>
<td>0.17</td>
<td>0.85</td>
</tr>
<tr>
<td>10</td>
<td>Gini Coefficient, 2000</td>
<td>2024</td>
<td>0.44</td>
<td>0.04</td>
<td>0.34</td>
<td>0.60</td>
</tr>
<tr>
<td>11</td>
<td>Unemployment, 2000 (%)</td>
<td>2050</td>
<td>0.05</td>
<td>0.02</td>
<td>0.02</td>
<td>0.17</td>
</tr>
<tr>
<td>12</td>
<td>Per Capita Income, 2000 ($)†</td>
<td>2031</td>
<td>21349.75</td>
<td>4254.74</td>
<td>7490.00</td>
<td>72844.00</td>
</tr>
<tr>
<td>13</td>
<td>Population in Poverty, 2000 (%)</td>
<td>2049</td>
<td>0.15</td>
<td>0.07</td>
<td>0.00</td>
<td>0.57</td>
</tr>
<tr>
<td>14</td>
<td>Violent Crimes, 2000 (#)</td>
<td>2049</td>
<td>57.84</td>
<td>99.08</td>
<td>0.00</td>
<td>1384.00</td>
</tr>
<tr>
<td>15</td>
<td>Property Crimes, 2000 (#)</td>
<td>2049</td>
<td>522.4</td>
<td>789.6</td>
<td>0.00</td>
<td>8028.00</td>
</tr>
<tr>
<td>16</td>
<td>Substance Abuse Crimes, 2000 (#)</td>
<td>1956</td>
<td>473.1</td>
<td>567.8</td>
<td>0.00</td>
<td>5906.00</td>
</tr>
<tr>
<td>17</td>
<td>Uninsured, 2000 (%)</td>
<td>2050</td>
<td>0.15</td>
<td>0.05</td>
<td>0.04</td>
<td>0.39</td>
</tr>
<tr>
<td>18</td>
<td>FTE Physicians per 1000, 2000 (#)†</td>
<td>2052</td>
<td>0.19</td>
<td>0.64</td>
<td>0.00</td>
<td>15.00</td>
</tr>
<tr>
<td>19</td>
<td>Composite Social Capital Index^†</td>
<td>2023</td>
<td>0.28</td>
<td>1.81</td>
<td>-3.80</td>
<td>15.22</td>
</tr>
<tr>
<td>20</td>
<td>Presidential Voter Turnout (%)†</td>
<td>2049</td>
<td>0.54</td>
<td>0.10</td>
<td>0.19</td>
<td>0.98</td>
</tr>
</tbody>
</table>

^Although this is titled a “Social Capital Index,” it was developed by Penn State University and only measures the degree of community participation.
†As these variables increase, social capital also is expected to increase.

Outside of variables selected for social capital but grounded in the extensive literature review, additional variables were selected from the SIF and USA Counties databases as potential confounders to control for in the three aims of the study. Although quite comprehensive, the SIF and USA Counties databases do not include the rural hospital data necessary for this study. Before delving into variable selection and inclusion criteria for the
social capital model, the data source for rural hospital closure, along with descriptive rural hospital characteristics will be provided.

2.1.2 Hospital Data

In this study, social capital and ultimately the factors that characterize social capital are the independent variables, while rural hospital closure is the dependent variable. Hospital-specific data was purchased through the American Hospital Association’s (AHA) strategic business enterprise known as Health Forum. The AHA is a common source of hospital data, especially in rural hospital research. Since rural hospitals provide care to far fewer patients than urban-located hospitals, they often are excluded from national surveys. For example, many rural hospitals are Critical Access Hospitals and are required to have 25 or fewer beds. As such, exact results from these hospitals are often suppressed and not publicly reported on the Medicare Hospital Compare website. Data purchased from Health Forum includes a hospital’s AHA identifier, CMS provider identifier, name, address, geographic coordinates, Federal Information Processing Standard (FIPS) code, and bed count.

The initial AHA file included all U.S. hospitals (n=5,535). The FIPS code allows for the AHA file to be merged with a federal geographic database to identify the county and state of each hospital. With each hospital’s corresponding county identified, the hospital data file can be merged with the Rural-Urban Continuum (RUC) Code developed by the USDA's Economic Research Service. The RUC Coding scheme will be used to define rural for the entirety of this dissertation. After merging the RUC data and identifying each hospital’s corresponding RUC code, any hospital located in an urban/metro county (as identified by the RUC code 1, 2, or 3) is removed from the database. This removes 2,965 urban/metro based hospitals from the original AHA file. The remaining 2,570 hospitals are all located in rural/non-metro counties (as identified by the RUC code 4, 5, 6, 7, 8, or 9).

To provide a data quality check, the inclusion of two unique identifiers – the AHA identifier and the CMS provider identifier – allows the AHA data to be merged with the CMS provider

---

file. The CMS provider file included hundreds of variables. For the data quality check, the CMS provider identifier, the hospital name, address, hospital type, and bed count data was merged to the AHA file. The additional CMS information matched by three variables (unique identifier, name, and address) facilitated the identification of all hospital not categorized as short-term, acute care hospitals. Since this study is limited to rural short-term, acute care, non-specialty hospital facilities, all hospitals that do not qualify under those criteria must be removed.

In this step, all hospitals with a missing AHA identifier or hospitals identified by CMS as not being a short-term, acute-care hospital were removed (n=444) from the hospital database. Using this process, the 444 hospitals identified for exclusion are operated by the military or the Department of Veterans Affairs and/or are specialty hospitals (children’s, long-term care, Indian Health Service, psychiatric, rehabilitation, correctional). Often, the name of the hospital clearly identified the cause for exclusion, including terms such as “Veterans Affairs,” “VA,” “Children,” “Behavioral Health,” “Rehabilitation,” “Army,” or “Penitentiary.” If a hospital did not include these terms in its name, the accuracy of exclusion for remaining hospitals was checked individually using information from state hospital association or individual hospital websites. Examples of hospitals verified for their exclusion include: The Eastside Hospital (Redmond, WA), Beacon Health Woodlands (The Woodlands, TX), and Hope Hospital (Lockhart, SC).

After removing hospitals located in urban areas and any not considered short-term, acute care, non-specialty hospitals, 2,126 rural hospitals remain for inclusion in the study. If stratified by Census Region (West, Midwest, South and Northeast), the majority of rural hospitals (76.6%) are located in the Midwest (Region 2) and South (Region 3). This is partially explained by historical migration patterns westward in the United States and government-determined, county boundaries. Counties in the western region of the United States are larger, and public, county-level healthcare delivery systems are more common.

To gain a better understanding of rural hospitals included in this study, each hospital will be characterized by its degree of rurality. For this study, rurality will be defined by the 2003 Rural-Urban Continuum (RUC) coding scheme developed by the U.S. Department of Agriculture’s Economic Research Service (USDA ERS). It classifies rural counties into 6
separate categories, which allows for the most flexibility in research. This classification system has been used in previous research.

TABLE 3: RURAL COUNTY CLASSIFICATION

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th>2000 population</th>
<th>Nickname</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Urban population of 20,000 or more, adjacent to a metro area</td>
<td>14,442,161</td>
<td>Bedroom Counties</td>
</tr>
<tr>
<td>5</td>
<td>Urban population of 20,000 or more, not adjacent to a metro area</td>
<td>5,573,273</td>
<td>Lone Stars</td>
</tr>
<tr>
<td>6</td>
<td>Urban population of 2,500 to 19,999, adjacent to a metro area</td>
<td>15,134,357</td>
<td>Home Town Counties</td>
</tr>
<tr>
<td>7</td>
<td>Urban population of 2,500 to 19,999, not adjacent to a metro area</td>
<td>8,463,700</td>
<td>Small Town Counties</td>
</tr>
<tr>
<td>8</td>
<td>Completely rural or less than 2,500 urban population, adjacent to a metro area</td>
<td>2,425,743</td>
<td>Satellite Counties</td>
</tr>
<tr>
<td>9</td>
<td>Completely rural or less than 2,500 urban population, not adjacent to a metro area</td>
<td>2,802,732</td>
<td>Lonely Counties</td>
</tr>
</tbody>
</table>

Rural hospitals are most likely to be located in counties with a population between 2,500 and 19,999, which are either adjacent or not adjacent to a metro area. These counties are coded as a 6 or 7 in the Rural-Urban Continuum Code classification system (Table 4). Rural hospitals located in counties with a larger population and close to a metro area face competition from larger, urban hospitals with more resources. Patients from adjacent, rural counties may opt to bypass the closer rural hospital in favor of the urban hospital, which threatens the financial health of the rural hospital. On the opposite side of the rurality spectrum, rural hospitals in counties with less than 2,500 residents struggle to remain operational due to insufficient patient volume.

---

Although classifying rural hospitals by rurality is useful to understand the environment in which a rural hospital is operating, for the purposes of this study, rurality will remain a dichotomous variable. All counties falling into the six rural, RUC codes will be considered “rural.” This simplifies the study, facilitates interpretation and possible generalizability. In addition, rural hospital closure, the dependent variable in this study, is not a common occurrence. Even with this study including the entire United States, there may be too few cases of hospital closure to allow for any significant findings for each rurality code.

Having described the process of selecting both appropriate datasets and variables for the purpose of this study to measure the association between social capital of rural communities and rural hospital closure, the precise definition and justification for each variable will be discussed. In the end, social capital (consisting of 20 variables in seven domains), rural hospital closure, and six possible confounding variables will be described below. A more detailed description of these variables is available in Appendix C. Ultimately, the extensive literature review and thorough understanding of the concepts strengthens the potential power and generalizability of this social capital model to additional applications.

2.2 VARIABLE DEFINITION & JUSTIFICATION

One of the significant contributions of this study is the development of a new social capital model, and its application to a health systems outcome. Therefore this section provides a
detailed definition of each social capital component and rationale for its inclusion as supported by prior research. After providing justification for each social capital variable, the processes for defining rural hospital closure and linking hospital closure to social capital data will be reviewed.

2.2.1 SOCIAL CAPITAL VARIABLES

The independent variables for this study are the 20 social capital variables or components of rural communities, with the county being used as the definition for a community. As described above, the method to define social capital in this study is similar to the one developed by Robert D. Putnam. Simulating Putnam’s process in defining social capital, this new social capital model will include as many variables from the selected datasets that have been previously used in social capital research. A previous study by Narayan and Cassidy demonstrated this technique by building upon previous social capital studies including World Values, New South Wales, Barometer Social of Capital, and the Index of National Civic Health (INCH). The Narayan and Cassidy social capital model included variables that represent all of domains identified in previous studies including neighborhood connectedness, community participation, subjective well-being, and political engagement.

This study strives to develop a social capital model that is even more comprehensive than Narayan and Cassidy. The richness of the data sources allows for the inclusion of a wide variety of variables that capture all if not more social capital domains or theoretical constructs. The variables included in this study have never been used in this particular combination. In addition, a comprehensive model of social capital has never been constructed for application in rural America or for a healthcare system and not an epidemiological outcome.

Variable selection began by creating a secondary database of any components that can justifiably contribute to a community’s social capital using the RTI Spatial Impact Factor (SIF) database. From the almost 16,000 variables included in the data sources, 673 variables were preliminarily identified as related to social capital. These variables were classified into one of the following seven domains: (1) population characteristic, (2) housing

and home ownership, (3) education, (4) economic indicators, (5) crime and violence, (6) health, and (7) community participation. These seven domains or versions of these categories are commonly used in social capital research. Components were removed if they were duplicates, if no prior social capital research was found to have used them, or if there was insufficient evidence of being strongly associated with social capital. The 673 variables were reduced to 60 variables. From these 60 variables, 20 social capital components were generated for use in this analysis.

2.2.1.1 Domain 1: Population Characteristics

As described previously, social capital is “not a single explanation or variable but rather a variety of explanations of how informal social relationships are important for human behavior.” No single definition has been established. Pierre Bourdieu focuses on the resources individuals possess thanks to their social networks. Conceptually these relationships or networks add a structural component to social capital and are often described as bonding or bridging varieties of social capital. Bonding ties are strong attachments formed when people already know one another and are frequently brought together, as in close social contacts with family, relatives, and friends. However, they may be exclusionary and may fail to produce society-wide benefits in cooperation and trust. Bridging ties are weaker and refer to connections that cross-cut social groups that may not often interact with one another. Often bridging ties are viewed as more

---

important in helping form widespread solidarity within a community.\textsuperscript{175,176} Therefore, variables that help characterize or measure these structural components of social capital are a critical component of any social capital model, as reflected by the selection of these five variables.

\begin{table}
\caption{Domain 1 Variables & Descriptive Statistics}
\begin{tabular}{|c|l|c|c|c|c|c|c|}
\hline
No. & Variable & Obs. & Mean & SD & Min & Max \\
\hline
1 & Population Loss, 1990-2000 (%) & 2015 & 0.07 & 0.14 & -0.84 & 0.90 \\
2 & Single Parent Family Households, 2000 (%) & 2049 & 0.10 & 0.04 & 0.01 & 0.41 \\
3 & Diversity Index, 2000\textsuperscript{†} & 2052 & 0.08 & 0.08 & 0.00 & 0.79 \\
4 & Linguistically-Isolated Households, 2000 (%) & 2052 & 0.02 & 0.03 & 0.00 & 0.34 \\
5 & Diversity Score (%)$\dagger$ & 2049 & 0.13 & 0.13 & 0.00 & 0.50 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{†}As these variables increase, social capital also is expected to increase.

\subsection{Population Loss, 1990-2000 (%)}
Significant population loss means the loss of human capital and a weakening of a community’s interrelationships or network.\textsuperscript{177} In a study on rural fishing villages in Canada, out-migration appeared to be the biggest threat to social capital and can compromise a community’s ability to cope with crisis.\textsuperscript{178} Research has shown that a population loss of 5 percent has a significant impact on the county.\textsuperscript{179} In this study, the average percent population change is actually a gain of 7.2\% in rural counties, but the range goes from a loss of 83 percent to a gain of 90 percent. The percent change in a county’s population is

calculated from two consecutive decennial U.S. Censuses. The data originates from the U.S. Census Bureau reported in the Atlas of the Spatial Patterning of County-Level Homicide dataset, which is available through RTI's Spatial Impact Factor Database. For the partial model, the data source and definition remain consistent using County Population for 1970, 1980, and 1990 data.

2.2.1.1.2 Single Parent Family Households, 2000 (%)

Social capital studies often include single parent households in their statistical models because they are often poorer and report having less time to devote to community building efforts. The well-known monograph *Growing Up with a Single Parent* (1994) examines the negative consequences of single parenthood on social capital because single-parent households lack the supportive benefit of a second at-home parent and tend to change residencies more often thus leading to fewer ties with adults in the community. A previous social capital study based in the United States, the Index of National Civic Health, included the birth rate among single and unmarried women. For this social capital component, different federal government agencies may differ in their definition. For consistency in the full and partial Adaniya Model of Social Capital, the calculation of percent single parent family households will be measured at the household and not individual birth rate level.

On average, 9.5% of family households in rural counties have a single parent head of household. First, the number of single or unmarried parent households with children under 18 years old will be calculated by subtracting the number of family households of married couple with person’s under 18 years from the total number of family households with person’s under 18 years. To calculate percent single or unmarried parent households (with children under 18 years old), the number of single or unmarried parent(s) households will

---

be divided by the total number of family households. For the partial model, the data source and definition remain consistent using USA Counties, the database support by the U.S. Census Bureau, to calculate the percent single or unmarried parent households with children in 1980 and 1990.

2.2.1.1.3 **Diversity Index, 2000† (Aim 1 Only)**

This component is available directly from the RTI SIF database and has a range of [0-1]. This particular diversity index is also known in the literature as the Theil Index.\(^{185}\) It measures the evenness or unevenness of spatial distribution of population subgroups in tracts within a county and takes into account race and ethnicity.\(^{186}\) An index score of 0 means the population subgroups are perfectly distributed throughout the county, while 1 means they are perfectly uneven and more segregated. Although population subgroups may have strong networks, the most developed network would be one where all subgroups interact with one another.\(^{187}\) Even distribution of population subgroups also indicates a level of comfort, trust, and reciprocity between the groups which adds to a county’s social capital.\(^{188}\)

Among rural counties, the average index score is 0.079, and since this is much closer 0 than 1, this score indicates that population subgroups in rural counties are fairly evenly distributed throughout the county. A score of 0 would indicate population subgroups are perfectly distributed throughout the county. Since the variance is low at 0.0057, this indicates that population subgroups in rural counties have low variability, although the diversity index scores do range from 0 to 0.7943. The U.S. Census only began calculating the Diversity Index with the 2000 U.S. Census; therefore, this variable will only be used in the full Adaniya Model of Social Capital.

2.2.1.1.4 **Linguistically–Isolated Households, 2000 (%) (Aim 1 Only)**

Although the Diversity Index measures the distribution of population subgroups, it does not use language in its calculation. Linguistic isolation may not only reflect the spatial

---


\(^{186}\) Iceland, J. (2004). The Multigroup Entropy Index (also known as Theil’s H or the Information Theory Index): University of Maryland.


distribution of certain population subgroups defined by language use but also the degree of assimilation.\textsuperscript{189} The hypothesis in this study is for social capital to decrease as the percent of linguistic isolation increases. Previous literature has not utilized this variable as a component of social capital because the U.S. Census only started preparing for its inclusion in 1990.\textsuperscript{190}

To create this component, the number of linguistically isolated households is calculated by summing the number of Spanish, Indo-Euro, and Asian-Pacific households identified as linguistically isolated. This sum is then divided by the total number of households to get a percent. On average, 1.6\% of households in rural counties can be classified as linguistically isolated. The standard deviation is +/- 3.1\%, which may indicate that there is insufficient variance in the variable to significantly contribute to social capital. Since this is a new variable calculated by the U.S. Census, this variable will only be used in the full new Adaniya Model of Social Capital.

2.2.1.1.5 Diversity Score, 2000 (\%)\textsuperscript{†}

Links between similar people, such as family or those with similar ethnic/cultural backgrounds, can prevent certain population groups from accessing social capital. Therefore, links between people of different backgrounds contribute more towards building intragroup relationships, promoting information diffusion, and promoting social cohesion.\textsuperscript{191} A community's racial diversity is necessary and common component of social capital. In prior studies, social capital definitions or model have not bridged across racial or ethnic groups; results have underscored the importance of racial/ethnic dimensions of equality in the United States and been a major limitation to social capital research.\textsuperscript{192} “The appropriate assessment of social capital’s impact on American civil society and politics shows that it depends on what dimensions of public life we consider, how we define “better off,” whether one is black or white, and whether one lives in a more or less racially

\begin{itemize}
\item \textsuperscript{191} Granovetter, M. S. (1973). The strength of weak ties. \textit{American Journal of Sociology}, 78(6), 1360-1380.
\item \textsuperscript{192} Hero, Rodney E. 2003. Social capital and racial inequality in America. \textit{Perspectives on Politics}, 1(1), 113-122.
\end{itemize}
heterogeneous community.” Although rural communities initially dealing with an influx of a new minority group may initially face challenges in community cohesion and which consequently lowers social capital, Robert Putnam has found that long-term increased diversity and immigration is essential, inevitable, and strengthens communities.

These findings justified the addition of a variable that more directly measures the degree to which a rural county is white or non-white. Furthermore, although the Theil Index is a richer measure of diversity, it is unavailable before 2000. To create the diversity score, first, the total white population is subtracted from 1 to calculate the total non-white population. Next, the total non-white population is divided by the total population. For consistency, all 1980, 1990, and 2000 data were collected and released by the U.S. Census Bureau on its USA Counties data portal.

Upon closer review of the social capital variables, the variable measuring the percent of non-white residents in a county needed to be transformed and rescaled. When considering the impact the percent non-white of the population has on social capital, as the non-white population percent approaches either zero or 100, social capital decreases because either whites or non-whites become a small minority. Having a 50 percent non-white population indicates there is at least an equal balance between two or more racial/ethnic groups in a county. After transformation, the non-white population variable becomes a type of diversity score; a score approaching zero indicates only one racial/ethnic group in the county. A score approaching 50 indicates a balance between at least two or more racial/ethnic groups in the county. Now, if the new non-white population variable (Diversity Score) increases so does social capital. This change allows for clearer interpretation of the relationship between racial diversity and social capital. In this study, the average rural county has a diversity score equal to 12.9. The data is skewed more towards 0 than 50, which means Caucasians in rural counties are more likely to be the dominate race/ethnic group.

2.2.1.2 Domain 2: Housing & Home Ownership

Theoretically, social capital is often considered by researchers as a construct resulting from the transmission of information across overlapping or intense contacts or social

networks. Fundamentally, for individual within a community to increase the intensity of these contacts to build social capital, a certain degree of trust must be in place. These interactions are more likely to happen and build as population and housing arrangements remain stable. This is especially important in rural communities because houses of diverse design, size, and price are often in close proximity to each other. Home ownership indicates a certain level of commitment to the community and even a network of neighbors.

### TABLE 6: DOMAIN 2 VARIABLES & DESCRIPTIVE STATISTICS

<table>
<thead>
<tr>
<th>No.</th>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Vacant Housing, 2000 (%)</td>
<td>2049</td>
<td>0.17</td>
<td>0.10</td>
<td>0.03</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>Owner-Occupied Housing, 2000 (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Owner-Occupied Housing, 2000 (%)†</td>
<td>2049</td>
<td>0.62</td>
<td>0.09</td>
<td>0.00</td>
<td>0.80</td>
</tr>
</tbody>
</table>

†As these variables increase, social capital also is expected to increase.

#### 2.2.1.2.1 Vacant Housing, 2000 (%)

Vacant housing visually contributes to the level of physical disorder and decay in a community. It has been linked to fear of crime, high levels of social disorganization, and even negative mental and physical health effects. Data is collected by the U.S. Census Bureau and released as part of the USA Counties database. The percent of vacant housing is defined by dividing the number of vacant housing units by the total housing units in a particular geographic area. For the partial model, the data source and definition remain consistent for percent vacant housing in 1980, 1990, and 2000. The average percent of vacant housing in rural counties in the United States is 16.9% with a standard deviation of +/− 10.2%. The relatively low rate and the geographic clustering of rural residencies may

---


make this variable less significant in determining the relationship between social capital and rural hospital closure.

2.2.1.2.2 Owner-Occupied Housing, 2000 (%)†

Previous research has even found that higher neighborhood social capital among homeowners persists while controlling for overall social capital. Therefore, there is something about owning a home that generates greater within-neighborhood social capital for homeowners. Homeownership, particularly owner-occupied homeownership, has been shown to benefit communities by stabilizing property values, encouraging social participation and maintenance upkeep of properties, and improving social conditions like high school dropout rates or crime rates. Residential stability also strengthens social ties and cohesion with neighbors because there is time to build long-term relationships. Owner-occupied housing information is collected by the U.S. Census Bureau. The percent of owner-occupied housing is calculated by dividing the number of owner-occupied housing units by the total number of housing units. Similar to percent vacant housing, for the partial model, the data source and definition remain consistent for percent owner-occupied housing in 1980, 1990, and 2000. On average 62.0% of housing is owner-occupied. In comparison to vacant housing, there is even less variation in the data (σ = 8.6%). Therefore, the percent owner-occupied housing may be an excellent contributor to social capital theoretically, but it may not significantly contribute to the analysis.

2.2.1.3 Domain 3: Education

Educational achievement has been tied to social capital since the original conceptualization by Bourdieu and Coleman. The social capital, that a student can access, is a result of the involvement of parents, family, and the greater school and neighborhood communities. As a community’s overall level of education increases, it can positively impact social capital. For example, higher educational achievement is associated with greater involvement in civil activities. All of the education data is collected in the decennial U.S. Censuses by the U.S. Census Bureau and released through USA Counties. This data source includes data from the 1980, 1990, and 2000 U.S. Censuses.

### TABLE 7: DOMAIN 3 VARIABLES & DESCRIPTIVE STATISTICS

<table>
<thead>
<tr>
<th>No.</th>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>High School Graduates, 2000 (%)†</td>
<td>2049</td>
<td>0.36</td>
<td>0.06</td>
<td>0.11</td>
<td>0.53</td>
</tr>
<tr>
<td>9</td>
<td>Some College or Higher Completion, 2000 (%)†</td>
<td>2049</td>
<td>0.40</td>
<td>0.10</td>
<td>0.17</td>
<td>0.85</td>
</tr>
</tbody>
</table>

†As these variables increase, social capital also is expected to increase.

2.2.1.3.1 High School Graduates, 2000 (%)†

The percent of high school graduates over the age of 25 is calculated by dividing the number of high school graduates over age 25 by the total number of people over age 25. A high school graduate is defined as any person who has received a high school diploma or equivalent level of education. In this study, 35.9% of rural residents, over the age of 25, have a high school diploma or its equivalence. Although the national high school graduation rate is a little over 70 percent in 2000, the percent high school graduates calculated for this

---


study included the entire rural population, which is skewed towards an older demographic.\textsuperscript{211}

\subsection*{2.2.1.3.2 Some College or Higher Completion, 2000 (\%)†}

It is well established in research and government findings that a large gap exists between rural and urban areas in terms of the percent of adults with a bachelor’s degree or higher. Only 17.4\% of adults in rural areas have a bachelor’s degree or higher compared to 30.2\% in urban areas.\textsuperscript{212} Therefore, a distinction between the percent of high school graduates compared to the percent of those how completed education beyond high school may be relevant. In order to isolate the number of people over age 25 who completed at least some college education (anything beyond high school regardless of degree completion), the number of high school graduates must be subtracted from the total number of people 25 years and over who have completed 12 years or more of school. The latter variable includes high school graduates, those who have completed some college, and those who have graduated from associates, bachelor, master, doctoral, or professional programs. To get a percent, the remaining number of people 25 years and over who have completed at least some college is then divided by the total number of people over age 25. According to the USA Counties data, 40\% of rural residents age 25 or older have completed some kind of education beyond a high school diploma or its equivalence.

\subsection*{2.2.1.4 Domain 4: Economic Indicators}

In 1980, Pierre Bourdieu was the first to systematically analyze social capital in a way that many prominent researchers consider to be the most theoretically refined.\textsuperscript{213} Bourdieu’s work emphasized the connection between different forms of capital (social, human, cultural, economic capital) and concluded that social capital is decomposable into multiple elements.\textsuperscript{214} However, even Bourdieu acknowledged the stronger relationship between

\begin{itemize}
\end{itemize}
social capital and economic indicators.\textsuperscript{215} Economic transactions, for example, require agents to rely on the future actions of others; this is accomplished at a lower cost in environments with higher trust and social capital.\textsuperscript{216} As more recent research confirms the strength of this relationship, the potential for social capital in economic development is being capitalized on by international financial institutions like the World Bank.\textsuperscript{217,218} Multiple World Bank projects, in underdeveloped counties like Sierra Leone, Thailand, and Albania, are striving to further economic development by increasing social capital.\textsuperscript{219}

\begin{table}
\centering
\caption{Domain 4 Variables \& Descriptive Statistics}
\begin{tabular}{|l|l|l|l|l|l|}
\hline
No. & Variable & Obs. & Mean & SD & Min & Max \\
\hline
10 & Gini Coefficient, 2000 & 2024 & 0.44 & 0.04 & 0.34 & 0.60 \\
11 & Unemployment, 2000 (%) & 2050 & 0.05 & 0.02 & 0.02 & 0.17 \\
12 & Per Capita Income, 2000 ($)\textsuperscript{†} & 2031 & 21349.75 & 4254.51 & 7490.00 & 72844.00 \\
13 & Population in Poverty, 2000 (%) & 2049 & 0.15 & 0.07 & 0.00 & 0.57 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{†}As these variables increase, social capital also is expected to increase.

2.2.1.4.1 \textit{Gini Coefficient, 2000}

The Gini coefficient (or index) is a well-known, statistical measure of income equality. Since 1967, household income inequality in the United States has increased by 20 percent.\textsuperscript{220} In a recent National Bureau of Economic Research study, increased concentration of income was

found to be a major contributor for the decrease in social capital between 1952-1998.\textsuperscript{221} In fact, research has found that relative income may be more important for peoples’ health than absolute income.\textsuperscript{222,223,224} Besides the association between income equality and health, income equality, as measured by the Gini coefficient, limits social mobility, consumer spending, educational attainment, and the ability for a community or nation to compete in the global economy.\textsuperscript{225} Earlier, increased ethnic diversity is found to positively impact social capital; income equality, if viewed as measure of income diversity, has the complete opposite effect. Although income inequality exists throughout the United States, increasing inequality creates a bimodal income distribution, characterized by a shrinking middle class and a growing number of people at the top and bottom. In general, marginalized populations are more likely to be in the bottom, poorer group.\textsuperscript{226} This divide hampers the ability for diverse groups within a community to bridge gaps and form intragroup relationships that promote social cohesion and consequently social capital.\textsuperscript{227}

All Gini coefficient data for 1980, 1990, and 2000 is available through the RTI Spatial Impact Factor database. The 1980 and 1990 data was collected for a research study by Nielsen, Francois, and Alderson.\textsuperscript{228} Data for 2000 was provided for the database by Nancy Lozano, a post-doctoral research associate with the GeoDa Center for Spatial Analysis and Geocomputation at Arizona State University. A Gini coefficient of one indicates maximal inequality within a community; a Gini Coefficient of zero expresses perfect equality. On average, rural counties in the United States have a Gini coefficient of 0.44. Although this may

seem like a low Gini coefficient, this average is higher than the overall US Gini coefficient of 0.40, and the range of Gini coefficient values in rural counties goes from 0.34 to 0.60.\textsuperscript{229}

\subsection*{2.2.1.4.2 Unemployment, 2000 (\%)}

The presence and use of social networks is inextricably tied to attaining employment. Previous retrospective surveys have shown that between 25 to 50 percent of workers secure jobs through personal networks or social structures that factor into social capital.\textsuperscript{230,231} Furthermore, unemployment is a destabilizing state for individuals associated with lower self-esteem and increased stress from the job search process, that in turn, has a destabilizing effect on groups of individuals or communities.\textsuperscript{232,233} Unemployed people, therefore, are less able to contribute to building social capital. County unemployment rate data is available for 1980, 1990, and 2000 through the RTI Spatial Impact Factor Database provided by the Bureau of Labor Statistics (BLS). The BLS defines unemployed individuals as those who are jobless, actively seeking work, and available to take a job, which limits the segment to people over the age of 16.\textsuperscript{234} In this study, the average rural unemployment rate was 4.6%, due in part to the strong economy in 2000 and the large percentage of retired or elderly residents in rural counties. However, this also fails to characterize the quality of the jobs. The economically vulnerable position of rural counties is perhaps indicative of the range in unemployment (1.5 – 16.8%).


2.2.1.4.3  *Per Capita Income, 2000* ($)*†

Per capital income is one of the most frequently variable used in social capital modeling. Higher income individuals and families often have more time for civic engagement, and higher income is associated with greater access to help. In fact, families who move from a high-poverty to a low-poverty neighborhood quickly experience improvements in health and education outcomes. Per capita income data for 1980, 1990, and 2000 were generated by the US Department of Commerce Bureau of Economic Analysis, used US Census Bureau data, and is available through RTI Spatial Impact Factor database. It is derived by dividing the total income of all people 15 years old and over in a specific geographic area by the total population in that area. In this study, the average per capita income in rural counties is $21,349 (σ = +/- $4,254), which is almost $7,000 less than the overall average per capita income in the United States. When compared to the Gini coefficient, per capita income is a raw calculation that fails to consider income distribution. In rural areas, for example, lower population means that even the presence of a few extremely wealthy individuals in a particular county may skew the per capita income figure.

2.2.1.4.4  *Population in Poverty, 2000 (%)*

There are countless studies demonstrating that higher social capital will reduce poverty, but even without evidence, such an assumption seems rational. In order to maintain consistency in the data across 1980, 1990, and 2000, population in poverty was calculated using 1979, 1989, and 1999 poverty and population data available on USA Counties. To calculate this variable, the number of persons living below a federally-determined income threshold was divided by the total number of persons for whom poverty status has been

---


239 Ibid

determined. Therefore, the denominator is not the total population in 1979, 1989, or 1999; it does not include the population for which we have no poverty status. Income thresholds are set by the federal government and account for household size, ages of members, and inflation as measured by the Consumer Price Index. In rural America, the average percent of the population in poverty is 15.0%, which predictability is higher than the overall U.S. average of 12.4%. The range extends between 0 to 56.7%. Countries with a poverty rate over 50% are primarily found in countries in the South, in Appalachia, or with American Indian reservations.

2.2.1.5 Domain 5: Crime & Violence

In general, higher levels of social capital are associated with lower crime rates. Social capital builds trust in communities and reinforces norms that dissuade from criminal activities. As Coleman writes, “Effective norms that inhibit crime make it possible to walk freely outside at night in a city and enable old persons to leave their houses without fear for their safety.” In social capital studies, focused on rule enforcement, social capital created by tight community networks is useful to parents, teachers, and police authorities as they seek to maintain discipline and promote compliance among those under their charge. As rule enforcement develops into a norm, solidarity and enforceable trust grows. Although more crime or a greater risk for crime occurring is associated with lower social capital, high crime rates may increase social capital, but only if community members organize to fight crime.

Since the level of motivation for a community to join together and fight crime is not available for the entire United States and likely requires field work, this study, therefore, will assume the more likely directionality of increasing crime risk being associated with lower social capital. In the United States, data are collected and provided by the Federal Bureau of Investigation’s Uniform Crime Reports (UCR) program established in 1930. The

---

243 Ibid
UCR program provides raw counts of reported crimes, clearance, and arrest data. Data for 1981, 1990, and 2000 is available at the country-level on the RTI Spatial Impact Factor database. For the purposes of this study, 1981 will be used as a proxy for 1980 data, since 1980 data is unavailable.

TABLE 9: DOMAIN 5 VARIABLES & DESCRIPTIVE STATISTICS

<table>
<thead>
<tr>
<th>No.</th>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>Violent Crimes, 2000 (#)</td>
<td>2049</td>
<td>57.84</td>
<td>99.08</td>
<td>0.00</td>
<td>1384.00</td>
</tr>
<tr>
<td>15</td>
<td>Property Crimes, 2000 (#)</td>
<td>2049</td>
<td>522.41</td>
<td>789.64</td>
<td>0.00</td>
<td>8028.00</td>
</tr>
<tr>
<td>16</td>
<td>Substance Abuse Crimes, 2000 (#)</td>
<td>1956</td>
<td>473.10</td>
<td>567.81</td>
<td>0.00</td>
<td>5906.00</td>
</tr>
</tbody>
</table>

2.2.1.5.1  Total Violent Crimes (#)

Social capital is linked to lower levels of violent crime. A major line of social capital research focuses on the role of neighborhood collective efficacy on violent crime rates. Collective efficacy is the “shared willingness of residents of a neighborhood to intervene to maintain social order and control crime and delinquency” and demonstrates a level of trust among neighbors. Work by Robert Sampson and colleagues have proven that collective efficacy is a strong predictor of the level of neighborhood violence, even controlling for a number of measures of neighborhood structural conditions. The category of “violent crimes” is defined by the FBI. In 1981, violent crimes included robberies known to police, aggravated assaults known to police, and certain unclassified crimes. In 1990 and 2000, violent crimes include (a) murders and non-negligent manslaughters known to police, (b) forcible rapes known to police, (c) robberies known to police, and (d) aggravated assaults known to police. These four types of crimes are classified as Part I offenses. In rural counties, the average count of violent crimes occurring in 2000 was 57.8.

249 Ibid
250 Ibid
251 Ibid
The large standard deviation indicates a wide range of violent crime values (0 – 1,384), which will require this variable to be normalized using z-scores prior to factor analysis.

2.2.1.5.2 Total Property Crimes (#)

Property crimes merit separation from violent crimes in the analysis due to the severity of the crime and its effect on a community’s social capital. When crime, violent or not, becomes a means of social control in a community, basic trust, among community members, breaks down. The “code of the street” takes hold when even law-abiding residents must project a tough image to avoid victimization. In UCR 1981, 1990, and 2000 data, the FBI classified the following crimes as property crimes: (a) burglaries known to police, (b) larceny-thefts known to police, and (c) motor vehicle thefts known to police. According to UCR statistics, the peak age for property crime arrests in the United States is 16, compared to 18 for violent crime arrests. In rural counties, the average count of property crimes occurring in 2000 was 522.4. Again, a large standard deviation indicates a wide range of property crime values (0 – 8,028), which will require this variable to be normalized prior to factor analysis.

2.2.1.5.3 Total Substance Abuse Crimes (#)

Previous longitudinal and natural-experiment studies have found that alcohol and drug use, as well as participation in delinquency, is related to friend influence or friend selection. In fact, this effect is seen even among roommates, individuals in close proximity to one another who are not necessarily friends. Therefore, alcohol and drug use is influenced by the norms (or laws) accepted or allowed by a community, one well-accepted dimension of social capital. Substance abuse crimes, like violent and property crimes, can have a negative effect to a community’s social capital. The inclusion of total substance abuse crimes is particularly important for studies focused on rural counties because drug and alcohol abuse is at times described as an “epidemic” in rural communities. Isolated, quiet rural

---

communities are ideal for certain drug production like methamphetamines. Although methamphetamine use may be higher in urban areas, rural residents are more like to abuse alcohol, opiate pain relievers, and stimulants than urban ones. Rural drug and alcohol abusers are also more likely to begin their primary substance of abuse at a younger age.

Unlike violent and property crimes, the FBI has not formally defined a “substance abuse” crime category. For this study, substance abuse crimes include: (a-b) opium-cocaine possession or sale/manufacture, (c-d) marijuana possession or sale/manufacture, (e-f) synthetic narcotic possession or sale/manufacture, (g-h) other drug possession or sale/manufacture, (i) drunkenness, and (j) driving under the influence (DUI). The average number of substance abuse crimes reported in rural counties is 473, but the maximum number reaches 5,906 in one rural county. One explanation for the large variance in crime may be due to a county’s geography. The three rural counties with the highest number of reported substance abuse crimes are Humboldt (CA), Mendocino (CA), and Worcester (MD). All three are on an ocean coast and have large footprints with Humboldt having over 4,000 square miles and Mendocino enjoying almost 3,900 square miles.

2.2.1.6 Domain 6: Health

The positive relationship between higher social capital and better health outcomes is well established by the literature. The social capital mechanisms of social support, influence, and engagement, as well as person-to-person contact and access to resources and material goods, influence health through pathways like health-related behavior, psychological processes, and physiological processes. Communities with high social capital may offer health-protective benefits to residents, even mediating the negative impact of living in a disadvantaged neighborhood. As Chapter 1 discusses, the direct link between social

---


257 Ibid

258 Ibid


capital and various health outcomes including mortality, quality of life, self-rated health, and access to healthcare is widely established. Interestingly, positive social capital has also been found to help mediate the effect income inequality has on health outcomes. Although social capital has been found to act upon health outcomes, certain health measures can act upon social capital if they affect the social features, such as trust, norms, and networks. 

Health measures, that enable or hamper a community and its members from achieving objectives, should be considered for inclusion in the social capital model.

<table>
<thead>
<tr>
<th>No.</th>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>Uninsured, 2000 (%)</td>
<td>2050</td>
<td>0.15</td>
<td>0.05</td>
<td>0.04</td>
<td>0.39</td>
</tr>
<tr>
<td>18</td>
<td>FTE Physicians per 1000, 2000 (#)†</td>
<td>2052</td>
<td>0.19</td>
<td>0.64</td>
<td>0.00</td>
<td>15.00</td>
</tr>
</tbody>
</table>

†As these variables increase, social capital also is expected to increase.

2.2.1.6.1 Uninsured, 2000 (%)

Insurance coverage is strongly associated with better health outcomes for both children and adults. Individuals with reporting poor health or even a disability, similar to single parent head of households, have less energy or time to focus on community building activities that positively impact social capital. In addition, these individuals often experience exclusion from spaces or social and economic majority. Uninsured data for 2000 were collected by the U.S. Census and are available through RTI Spatial Impact Factor database. On average, 15 percent of rural residents do not have insurance, although the range extends from a low as three percent and up to 39 percent.

---

Higher community social capital is associated with better access to care likely because social capital improves the functioning and efficiency of community social institutions.\textsuperscript{265,266} The likelihood of physicians to participate in community, politics, and collective advocacy was found to increase with the level of civic mindedness in their community and being in a rural areas.\textsuperscript{267} The number of full-time-equivalent (FTE) physicians per 1000 was collected through the Centers for Medicare and Medicaid Services (CMS) provider file and available at the county-level through the RTI Spatial Impact Factor database. The mean number of FTE physicians per 1000 is 0.19. Since a primary care physician can reasonably see a panel of 2,500 patients, this result reflects the shortage of physicians reported in rural communities.\textsuperscript{268} The range of FTE physicians per 1000 starts with zero and goes to 15. However, only 4 percent of rural counties report having over 1 FTE physician per 1000. The only county reporting a two-figure number of FTE physicians per 1000 is Talbot County (MD) at 15, and although not relevant for this study, even this decreased sharply after 2006 following the acquisition of the local Shore Health System by University of Maryland Medical System.\textsuperscript{269} Due to the low rate of FTE physicians per 1000, the small variance will make this variable less likely to be significant.

\section*{Domain 7: Community Participation}

By definition, social capital focuses on the social ties and/or resources available to all social groups within a community.\textsuperscript{270} These social ties or networks create a valuable resource for a community and are created by community participation. Community participation is a process through which members of a community are involved in and have influence on

\begin{itemize}
\end{itemize}
decisions related to activities that will affect them. This is often facilitated by formal civic, business, religious organizations and the presence or investment in community spaces including parks and community centers. Therefore, community participation is an important component of social capital.

### TABLE 11: DOMAIN 7 VARIABLES & DESCRIPTIVE STATISTICS

<table>
<thead>
<tr>
<th>No.</th>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>Composite Social Capital Index†</td>
<td>2023</td>
<td>0.28</td>
<td>1.81</td>
<td>-3.80</td>
<td>15.22</td>
</tr>
<tr>
<td>20</td>
<td>Presidential Voter Turnout (%)†</td>
<td>2049</td>
<td>0.54</td>
<td>0.10</td>
<td>0.19</td>
<td>0.98</td>
</tr>
</tbody>
</table>

†Although this is titled a “Social Capital Index,” it was developed by Penn State University and only measures the degree of community participation.

†As these variables increase, social capital also is expected to increase.

#### 2.2.1.7.1 Penn State Composite Social Capital Index (Aim 1 Only)

This is a composite variable measuring community participation that used Putnam’s methodological approach. It was created by Penn State University’s Northeast Regional Center for Rural Development (NERCRD), which is funded in part the USDA.²⁷¹ This variable directly captures specific resources individuals have access to in their community that could enable them to benefit from their community’s social capital. It also measures a degree of civic engagement in the form of the population completing their U.S. Census form and voting. All the data was collected in 2005 with the exception of the U.S. Census Mail Response Rate in 2000. To determine a county’s composite social capital index, the following 15 variables are combined:

- The number of (1) bowling centers, (2) physical fitness facilities, (3) public golf courses, and (4) sports clubs per 10,000 people;
- The number of (5) business, (6) civic and social, (7) labor, (8) not-for-profit, (9) political, (10) professional, and (11) religious organizations per 10,000 people;
- The U.S. Census mail response rate in (12) 2000 and (13) 2005;

• The (14) presidential voter response rate as calculated by dividing the number of votes cast in the 2004 presidential election by the total population age 18 and over in 2005; and
• The (15) total resident population.

Despite these data dating to 2005, the richness of the locally-collected data, which Putnam encourages to utilize when available in social capital research, merits its inclusion in the social capital model developed for this study in the year 2000. This level of detail is unavailable for rural communities in 1980 or 1990. Since this is an index, there are no units for this variable, but the range of results (-3.80 – 15.22) will make it more likely to be significant.

2.2.1.7.2  Presidential Voter Turnout, 2000 (%)†
The Composite Social Capital Index is unavailable in 1980 and 1990, yet, as specified previously, community participation and civic engagement is a key component of social capital. Since Aim 2 calls for the use of social capital trends over time, some measure of community participation must be included in the model. From the list of 15 variables used in the Composite Social Capital Index, data for one variable, the Presidential Voter Response Rate, is available through the U.S. Census Bureau's USA counties dataset. In 1980 and 2000, the total number of votes cast for president is divided by the total resident population 18 years and older. To represent 1990, the total number of votes cast for president in 1988 and 1992 were averaged before being divided by the 1990 total resident population 18 years and older. In 2000, the average Presidential voter turnout was 54 percent in rural counties, although the range extends from 19 to 98 percent. Although presidential voter turnout is influenced by the relevance of other issues on the ballot, as well as tight, contentious races, the presidential voter turnout for the years included in the study are fairly steady. The following are the average Presidential voter turnout rates: 52.8% (1980), 50.3% (1988), 55.2% (1992), and 50.3% (2000).272 The Clinton-Bush presidential campaign mobilized more voters, but since this will be averaged with the Bush-Dukakis election, there is little fluctuation in the average presidential voter turnout. Therefore, this study will make not attempt to control these factors and will treat each presidential voter turnout percentage equally.

2.2.2  **Hospital Variables**

Having reviewed each social capital component and rationale for its inclusion in the study, the following section will describe the processes for defining rural hospital closure and linking hospital closure to social capital data. Since rural hospital closure is the dependent variable or outcome being analyzed, establishing clear definition or criteria for what constitutes rural hospital closure is vital for this study.

2.2.2.1 Defining Hospital Closure

As described above, hospital data has already been prepared for closure assessment. The initial AHA data file with 5,535 hospitals was reduced to 2,126 rural, short-term-acute-cares, non-specialty hospitals. Now, these 2,126 remaining hospitals need to be assessed for closure status. The AHA defines closure as cases where hospital data are no longer available for a specific hospital in subsequent years, as identified by an AHA identifier. For this study, rural hospital closure status will only be assessed for the time period between 2000 and 2008. As described in Chapter 1, rural hospitals are at a greater risk for closure than urban ones because rural hospitals are more likely to be in financial distress than their urban counterparts.\(^{273,274,275}\) In addition, rural hospitals are more impacted by changes in federal or state policies and weak economies, since these often directly affect the institution's financial health.\(^{276}\)

Starting in 2008, rural hospitals faced a significant threat to their survival – the Great Recession that began in December 2007, according to the U.S. National Bureau of Economic Research.\(^{277}\) Attributable to the Great Recession, rural counties experienced larger employment losses, felt the recession a full year earlier, and


continue to lag behind urban counties in its economic recovery. Therefore, this study will only consider rural hospital closures that occurred between 2000 and 2008.

To determine hospital closure, AHA data were purchased for three separate years 2000, 2004, and 2009. This will allow rural hospital closure status to be determined for two 4-year periods, 2000 – 2004 and 2004 – 2008. Annual hospital data is available through the AHA, but its acquisition is cost prohibitive. Within the study’s eight year time period, rural hospital closure status was assessed. As an example, Gulf Pines Hospital in Florida reports data in the AHA survey in 2000 and 2004. In 2009, Gulf Pines no longer reports data. According to the AHA, Gulf Pines Hospital can now be classified as closed; its exact closure date is uncertain only the four-year range (2005 – 2008) can be ascertained. However, accuracy and precision in hospital closure data is vital for the integrity of this analysis.

To check the quality of the hospital data, the three AHA datasets (2000, 2004, and 2009) were merged based on the unique AHA identifier with to the CMS Provider File (available through the SIF database). The process enabled the identification of hospitals that changed their name or perhaps were merged with another. In the case of closure, hospitals with a possible name change or merger case were further analyzed to verify open/closure status. In the example of Gulf Pines Hospital, its AHA identifier (6390847) and CMS Provider identifier (100027) were run through the 2009 AHA dataset to see if either identifier reported data. Since no data was reported for either identifier, Gulf Pines Hospital is officially classified as closed.

2.2.2.2 Describing Hospital Closure

After verifying closure status, the 2,126 rural hospitals included in the analysis can be classified into those remaining open (n=2056) and those closed (n=70). As Table 12 shows below, 58.6% of closures (n=41) occur in the South region. Although the South region has the most rural hospitals (n=840), this only accounts for 39.5% of rural hospital nationwide. The Midwest region with 789 rural hospitals, representing 37.1% of all rural hospitals nationwide, accounts for only 22.9% of closures. Rural hospitals, in the Northeast and West regions, account for a smaller percent of overall rural hospitals (7.3%, 16.0%) and an even smaller percent of closures (4.3%, 14.3%). Therefore, based on the simple, descriptive statistics alone, it appears as if the South's rural hospitals are more likely to close than rural

---

hospitals in any of the other three regions. Although insightful, the overall number of hospital closures (n=70) represents only 3.3% of all rural hospitals operational at the start of the study in 2000. With so few hospitals closing, individual analyses at the regional level are likely to not produce statistically significant results.

TABLE 12: RURAL HOSPITALS BY CENSUS REGIONS & CLOSURE STATUS

<table>
<thead>
<tr>
<th>Census Region</th>
<th>1 (Northeast)</th>
<th>2 (Midwest)</th>
<th>3 (South)</th>
<th>4 (West)</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>153</td>
<td>773</td>
<td>799</td>
<td>331</td>
<td>2,056 (96.7%)</td>
</tr>
<tr>
<td>Closed</td>
<td>3</td>
<td>16</td>
<td>41</td>
<td>10</td>
<td>70 (3.3%)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>156 (7.3%)</td>
<td>789 (37.1%)</td>
<td>840 (39.5%)</td>
<td>341 (16.0%)</td>
<td>2,126 (100%)</td>
</tr>
</tbody>
</table>

Although hospital-specific data is useful to characterize them, the geographic scale used in this study is the county. Social capital data is available at the county level. Therefore, hospital closure (dependent variable) data need to be aggregated at the county level in order to merge them with social capital (independent variable) data.

2.2.2.3 Aggregating Hospital Closure Data by County

This study is focused on determining the association between the social capital of rural communities and whether a hospital closure occurred in that community. This calls for the rural hospital closure data to be captured in a dichotomous outcome variable (Yes/No). The dichotomous variable will be marked as 0 if the rural county did not experience a rural hospital closure; a 1 will be reported if the rural county experienced any rural hospital closure.

As Table 13 shows below, the 2,126 rural hospitals are spread out over the 2,052 rural counties in the United States. It is interesting to note that only 1,614 rural counties (78.7%) have at least one hospital, with the majority of these counties (n=1,216) having only one hospital located within its boundaries. In addition, as Table 13 shows, there are rural counties with more than one hospital (n= 398) or even ones that over the time of the study suffer more than one hospital closure (n=8). From Table 12 above, the raw count of rural hospital closures by Census region ranges from 3 in the Northeast to 41 in the South.
Although these differences could be further analyzed, this study is answering whether the social capital of a rural community, as defined by county, is associated with any rural hospital closure. Furthermore, the low rural hospital closure rate calls for rural hospital closure to be aggregated into a single dichotomous outcome variable for Aim 1 and Aim 2 (See Chapter 3 and 4).

**TABLE 13: TOTAL HOSPITALS & HOSPITAL CLOSURES PER COUNTY**

<table>
<thead>
<tr>
<th>Hospitals per County</th>
<th>County (n=2,052)</th>
<th>Counties with 1 Hospital Closure</th>
<th>Counties with 2 Hospital Closures</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>438 (21.3%)</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1,216 (59.3%)</td>
<td>31</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>314 (15.3%)</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>63 (3.1%)</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>13 (&lt; 0.1%)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>7 (&lt; 0.1%)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Yet, as this table shows, using a county-specific scale and a dichotomous rural hospital closure variable for this study limits the analysis and may negate the possible effect of counties with multiple hospitals. For example, a rural county with 3 hospitals that experiences one hospital closure (n = 7) will be less affected by this closure than a rural county where the only hospital in operation closes (n = 14). If the aim of this study is changed to the proportion of rural hospital closure experienced in a county, this would require a separate study and a different methodological approach. Due to the low rural hospital closure rate (3.3%) nationally, the initial aims of the study will be kept, along with the use of a dichotomous outcome variable. To begin addressing the issues highlighted in Table 13, Aim 3 of this study will change the geographic scale from county to Proximal Hospital Area (See Chapter 5).

As stated previously, the richest social capital data is available at the county level, and the county is perhaps the most utilized geographic scale in rural hospital research. It is considered an adequate proxy for the service area a rural hospital typically serves. Therefore, hospital closure (dependent variable) data needs to be aggregated at the county level in order to merge it with social capital (independent variable) data.
2.2.3 **LINKING SOCIAL CAPITAL AND HOSPITAL CLOSURE**

Due to the low rural hospital closure rate and the dichotomous outcome variable aggregated at the county level, the percent of rural counties experiencing any rural hospital closure is lower than the overall closure rate of 3.3 percent. After linking the social capital and hospital closure data at the county scale, the percent of counties experiencing one or more hospital closures is 3.02 percent (n = 62 counties). The percent of counties who did not experience a hospital closure is 96.98 percent (n = 1,990 counties). To begin to characterizing the differences between counties that suffered a rural hospital closure with those that experienced no hospital closure, the second descriptive statistics Table 14 that stratifies results by a county's closure status is provided below.
### TABLE 14: SOCIAL CAPITAL VARIABLES & DESCRIPTIVE STATISTICS BY COUNTY HOSPITAL CLOSURE STATUS

<table>
<thead>
<tr>
<th>Social Capital Indicator</th>
<th>CLOSE (n = 62)</th>
<th>OPEN (n = 1,990)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEAN</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Domain 1: Population Characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Population Loss 2000 (%)</td>
<td>0.07</td>
<td>0.13</td>
</tr>
<tr>
<td>2 Single Parent Family Households 2000 (%)</td>
<td>0.11</td>
<td>0.04</td>
</tr>
<tr>
<td>3 Diversity Index 2000</td>
<td>0.10</td>
<td>0.07</td>
</tr>
<tr>
<td>4 Linguistically Isolated Household 2000 (%)</td>
<td>0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>5 Diversity Score 2000 (%)</td>
<td>0.16</td>
<td>0.14</td>
</tr>
<tr>
<td><strong>Domain 2: Housing &amp; Home Ownership</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Vacant Housing 2000 (%)</td>
<td>0.16</td>
<td>0.10</td>
</tr>
<tr>
<td>7 Owner Occupied Housing 2000 (%)</td>
<td>0.62</td>
<td>0.08</td>
</tr>
<tr>
<td><strong>Domain 3: Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 High School Graduates 2000 (%)</td>
<td>0.35</td>
<td>0.05</td>
</tr>
<tr>
<td>9 Some College or Higher Completion 2000 (%)</td>
<td>0.40</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>Domain 4: Economic Indicators</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 Gini Coefficient 2000</td>
<td>0.45</td>
<td>0.03</td>
</tr>
<tr>
<td>11 Unemployment 2000 (%)</td>
<td>0.05</td>
<td>0.02</td>
</tr>
<tr>
<td>12 Population in Poverty 2000 (%)</td>
<td>0.17</td>
<td>0.07</td>
</tr>
<tr>
<td>13 Per Capita Income, 2000 ($)</td>
<td>21245.50</td>
<td>3813.88</td>
</tr>
<tr>
<td><strong>Domain 5: Crime &amp; Violence</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Violent Crimes, 2000 (#)</td>
<td>130.87</td>
<td>179.85</td>
</tr>
<tr>
<td>15 Property Crimes, 2000 (#)</td>
<td>1053.02</td>
<td>1296.32</td>
</tr>
<tr>
<td>16 Substance Abuse Crimes, 2000 (#)</td>
<td>883.90</td>
<td>1011.19</td>
</tr>
<tr>
<td><strong>Domain 6: Health</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 Uninsured, 2000 (%)</td>
<td>0.16</td>
<td>0.05</td>
</tr>
<tr>
<td>18 FTE Physicians per 1000, 2000 (#)</td>
<td>0.31</td>
<td>0.81</td>
</tr>
<tr>
<td><strong>Domain 7: Community Participation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 Composite Social Capital Index, 2000</td>
<td>-0.21</td>
<td>0.89</td>
</tr>
<tr>
<td>20 Presidential Voter Turnout, 2000 (%)</td>
<td>0.53</td>
<td>0.09</td>
</tr>
<tr>
<td><strong>Hospital Factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Hospitals per County</td>
<td>2.03</td>
<td>0.81</td>
</tr>
<tr>
<td>Total Hospital Closures per County</td>
<td>1.13</td>
<td>0.34</td>
</tr>
</tbody>
</table>
As Table 14 shows, rural counties that experience a hospital closure are more likely to have two hospitals or more in operation. With regards to social capital variables included in the analysis, economic indicator variables show counties with closure to be performing at slightly more negative levels than counties with no hospital closure. By far, the variables in the crime and violence domain – violent crime, property crime, and substance abuse crime – show the biggest difference between rural counties with hospital closure (130.9; 1053.0; 883.9) versus those without closure (55.6; 505.9; 460.8). Besides crime, the social capital index calculated by Penn State University reports a far lower mean of community participation among counties with closure (-0.21) than those without closure (0.01).

2.3 Model Strengths

After creating a new social capital model, clearly defining hospital closure, and merging all of the data at the county level, the process of building a model to describe the relationship between social capital and rural hospital closure can begin. In Aim 1 (See Chapter 3), two methodologic approaches will be compared – stepwise logistic regression alone and a combination of factor analysis with logistic regression. The initial approach includes all variables and considers each social capital variable (n = 20) equally before removing the least significant variable, one at a time. The final model (Model A) includes only those social capital variables found to significantly predict rural hospital closure. The second approach first uses a data reduction technique known as factor analysis, a statistical tool commonly used in social capital studies.279 Factor analysis creates and determines significant factors, which are then run through logistic regression to build a model (Model B). Regardless of the methodological approach taken, both Model A and Model B need to be tested through a process known as model building.

2.3.1 Confounding Variables

One of the steps in model building is to force possible confounding variables into the model one at a time and determine if each deserves inclusion in the model based on its impact on the odds ratios generated. Confounding variables are variables outside of the independent variables (social capital of rural communities) that may affect the dependent variable (rural hospital closure). Therefore, any confounding variables found to significantly affect the

model must be controlled for in the analysis. In this study, six confounding variables will be forced into both Model A and Model B to determine if any requires being controlled for in the study. These six variables represent environmental and policy characteristics of the surrounding area that could have an impact on rural hospital closure. The following section will describe each of the six possible confounding variables tested for.

2.3.1.1 Rural Health Clinics (RHC) & Federally Qualified Health Centers (FQHC)

As described in the glossary, Rural Health Clinics (RHCs) and rural-based Federally Qualified Health Centers (FQHCs) are both healthcare facilities that provide primary care services to rural residents in medically underserved and health professional shortage areas. Clinics or centers designated RHC or FQHC status receive enhanced reimbursement rates for Medicare and Medicaid patients. When compared to urban areas, rural patients are far more likely to be insured through Medicare or Medicaid due to the aging demographic and the lower quality jobs available in rural areas that often do not offer employer sponsored insurance. Since RHCs and FQHCs are recognized as important access points to healthcare in rural areas, their presence in a rural communities may provide rural residents with a sufficient level of care and make them less likely to care if a local hospital closes. As seen in Table 15, the average number of RHCs and FQHCs in counties that experienced a hospital closure (RHC 2.02; FQHC 0.66) is higher than the average in counties that did not experience closure (RHC 1.19; FQHC 0.35).

2.3.1.2 Health Professional Shortage Area (HPSA)

Across the country, rural areas consistently report difficulties in recruiting healthcare providers to practice in their area. As a result, the federal government has created specific shortage designations, including the Health Professional Shortage Area (HPSA) program. This program allows healthcare providers in rural areas to qualify for higher reimbursement and facilitates recruitment of healthcare providers by offering foreign-trained physicians specific visas that require them to initially serve HPSAs. The HPSA benefits may affect the likelihood of a rural hospital closing, which therefore merits testing in the model. However, rural areas and providers easily qualify for HPSA designation. This is reflected in Table 15 because the average number of HPSA designations in counties with a hospital closure (1.11) versus those without a hospital closure (0.94) is quite similar. In addition, the HPSA program has existed since Congress approved the National Health
Service Corps in the 1970s. Therefore, in its 40 year history, HPSA designation has not made much of an impact on improving healthcare access to rural communities.\textsuperscript{280}

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|}
\hline
Confounders & \textbf{CLOSE} (n = 62 counties) & & \textbf{OPEN} (n = 1,990 counties) & \\
\hline
 & MEAN & SD & MEAN & SD \\
\hline
Rural Health Clinics, 2000 (#) & 2.02 & 2.69 & 1.19 & 1.59 \\
Federally Qualified Health Centers, 2000 (#) & 0.66 & 0.92 & 0.35 & 0.83 \\
Health Professional Shortage Area, 2000 (#) & 1.11 & 0.89 & 0.94 & 0.82 \\
Rural-Urban Continuum Code, 2000 & 6.10 & 1.66 & 6.84 & 1.54 \\
\hline
\end{tabular}
\caption{Possible Confounders & Descriptive Statistics by County Hospital Closure Status}
\end{table}

\subsection{2.3.1.3 Economic Dependency}
Economic dependency is an economic, county-level classification scheme also developed by the USDA’s Economic Research Service (ERS). There are six non-overlapping categories of economic dependence including: farming, mining, manufacturing, services, Federal/State government, and unspecialized.\textsuperscript{281} Since economically dependent counties are by definition heavily reliant on a specific industry, the county is at risk economically whenever the dependent industry faces a crisis, such as a crop failure, drastic drops in commodities prices, cheap labor availability, etc. This, in turn, could affect the financial health of a rural hospital, especially if the dependent industry directly or indirectly employs a significant percentage of the county and perhaps offers employer-based insurance. Therefore, economic dependency status will be forced into both Model A and Model B to determine if it merits inclusion. It may not require inclusion because the majority of rural counties are economically dependent, but the low rural hospital closure rate does not allow for stratification by type of economic dependency.


2.3.1.4 Rural-Urban Continuum (RUC) Coding Scheme

In recognition that rurality falls along a continuum, the U.S. Department of Agriculture’s Economic Research Service developed the Rural-Urban Continuum Codes in 1975, sometimes referred to as “Beale Codes.” This schema still classifies counties down into metropolitan and non-metropolitan. However, rural or non-metropolitan counties are further differentiated into six categories (See Table 3). As mentioned above, this study uses the RUC Coding Scheme to identify all rural counties and defines “rural” as any county with a RUC Code of 4 or higher. Yet there are clear differences in population and adjacency to urban areas across the categories. The degree of rurality for counties with a rural hospital closure (6.1) and countries without a closure (6.8) differs. This means counties that experienced a rural hospital closure are likely to be more populated and within closer proximity to an urban area, which increases the likelihood of rural residents bypassing their local hospital in favor of a more specialize urban one. 282 Again, the low rural hospital closure rate does not allow for stratification of results by RUC Code, but to improve the model, some consideration could be made. Therefore, three aggregated RUC variables will be created that group rural counties by population. Group 1 will include rural counties with a population of 20,000 or more (RUC Codes 4-5). Group 2 will include rural counties with a population between 2,500 and 19,999 (RUC Codes 6-7). Group 3 will include rural counties with a population less than 2,500 (RUC Codes 8-9).

2.3.1.5 Census Region

The US Census Bureau (USCB) has four official regions, with nine official divisions. The four official regions are the following: Northeast, Midwest, South, and West. The partition of the country into geographic regions is rooted in American history (i.e. New England, Middle Atlantic, and South) and is influenced by colonial governance, slavery, and immigration. This history continues to influence each Census region today. In his research, Robert Putnam, found that the upper Midwest region enjoys some of the highest levels of social capital based on the higher proportion of the population having Scandinavian stock.283 Scandinavian culture and beliefs lend themselves to a strong sense of community. On the other end of the spectrum, the South has, by far, the lowest level of social capital due to its

history of plantation slavery and Jim Crow politics. Putnam writes, "Slavery was, in fact, a social system designed to destroy social capital among slaves and between slaves and freemen." Even after emancipation, dominant wealthy Southerners had a strong interest in inhibiting recently freed slaves from building social networks. Therefore, regional differences may be significant enough to merit the inclusion of Census region into the models. Again, the low rural hospital closure rate (Northeast region only had 3 rural hospital closures) may hamper the effect region could play on the study's results.

### 2.4 Public Health Significance

The overall purpose of this study is to determine the association between social capital of rural communities and rural hospital closure. First, this chapter provided a detailed review of all data sources. Next, the process for creating a novel and comprehensive social capital model with 20 variables was presented. In addition, all social capital variables, hospital closure status, and possible confounders were defined, justified for inclusion, and preliminarily analyzed. Even greater social capital variable descriptive statistics are provided in Appendix C. A concise definition for rural hospital closure was also provided, along with the process for creating the dichotomous dependent variable, rural hospital closure occurring by county. Finally, the merits of testing the resulting models in the study for confounding are discussed. Each possible confounder to be considered in Aim 1 (Chapter 3) was introduced.

As Chapter 2 shows, a great strength to this study is the substantial literature review process undertaken in the development of the social capital model. Each variable included in the study has been carefully considered, selected, and analyzed. Inclusion criteria are grounded in findings from a wide-range of relevant fields including public health, sociology, economics, and political science. The final social capital models builds upon previous studies and includes the most comprehensive group of variables available for every county in the entire United States. Although the number of rural hospital closure is limiting, the data quality check process lends much confidence to the accuracy and precision of which rural hospitals truly closed in the United States between 2000 and 2008. Finally, the review

---

of possible confounders introduces a clear way that the models to be developed in Aim 1 will be analytically strengthened.
Chapter 3  

**USING THE SOCIAL CAPITAL MODEL TO ANALYZE THE KEY PREDICTORS OF HOSPITAL CLOSINGS**

3.1  **ABSTRACT**

**OBJECTIVES:** This study uses a novel social capital model to measure the association of a rural community’s social capital with the likelihood of rural hospital closure.

**METHODS:** Social capital data and rural hospital data for all rural counties in the United States (n = 2,052) were merged. Model A is a multivariable logistic regression model predicting rural hospital closure by backward elimination of social capital variables. Model B combines factor analysis with logistic regression to measure the relationship between social capital and rural hospital closure.

**RESULTS:** In Model A, three social capital components are found to be significant in predicting the likelihood of rural hospital closure – unemployment, substance abuse crime, and violent crime.

After controlling for confounders and interaction terms, performing sensitivity analysis, and adjusting the odds ratios, unemployment (OR=2.04; CI 1.09-2.72; p=0.03) is more strongly associated with the likelihood of rural hospital closure than substance abuse (OR=1.26; CI 1.24-1.29; p=0.01) and violent crime (OR=1.01; CI 1.00-1.02; p=0.02). In Model B, factor analysis determines two social capital factors are significantly associated with rural hospital closure – economic vulnerability (OR=1.42; CI 1.10-1.84; p=0.01) and crime risk (OR=1.47; CI 1.22-1.78; p=0.01).

**CONCLUSIONS:** The traditional focus on institution-specific rural hospital closure risk factors limits our understanding of closure factors. Community-specific rural hospital
closures risk factors such as social capital merit consideration. Factors are a more comprehensive representation of social capital, and they are more strongly associated with rural hospital closure than the social capital components found to be significant in Model A. Findings support the need for further analysis of the relationship between a rural community’s social capital and rural hospital closure.

3.2 Introduction

Although social capital is a challenging concept for researchers to define, social capital is often generally regarded as a community-level resource or social feature, one that a community can find strength and unity to overcome significant periods of duress. Rural communities with higher social capital are better able to overcome problems and survive economic recessions or the loss of a major employer or hospital. Previous work has found strong and clear associations between social capital and a wide-variety of health outcomes including hospital closure. In rural communities, studies have

shown that hospital closure threatens the long-term sustainability of the community and decreases access to healthcare and other social services rural hospitals often provide.\textsuperscript{296} Therefore, preventing rural hospital closure is regarded by both residents and by the government as a desired outcome.\textsuperscript{297} If social capital is one resource rural communities can tap into to prevent a local hospital closure, then a deeper understanding of the relationship between social capital and rural hospital closure is warranted.

The main goal of this study is to determine the association between social capital of rural communities and rural hospital closure. The independent variable is constructed from a large collection of continuous and discrete variables that collectively define social capital (as detailed previously in 1.2). The dependent variable is a dichotomous rural hospital closure variable (closed = 0 or open, operational = 1). In the initial analyses, rural hospital closure will include all rural counties in the United States (n=2,052). Both the independent and dependent variables are depicted in Figure 1 as black boxes. The thicker blue and red arrows, connecting these black boxes, represent the central aim of this study – to determine the association between social capital and rural hospital closure – and reflect the two model building, statistical approaches.

In both Model A and Model B (See Conceptual Model), social capital is considered to be a latent phenomenon that can only be indirectly measured by a combination or subsets of observed, manifest variables. In building Model A, social capital is indirectly measured by 20 identified variables (See 2.2.1) that are treated independently with no interrelationships among variables presumed. Model A, based on those variables will use stepwise multivariable logistic regression to select the set of social capital variables that best estimates the probability of rural hospital closure. Model A is represented in Figure 1 as blue arrows and boxes connecting the independent and dependent variables. The enhanced Model B will initially use factor analysis to reduce the 20 identified variables into social capital factors that incorporate not only the 20 variables but also the interrelations among them. Model B will use the same multivariable logistic regression technique to determine the likelihood of rural hospital closure as predicted by the social capital factors. Model B is


visible in Figure 1 as the red arrows and box connecting the initial social capital model and rural hospital closure.

After performing the initial analyses, both Model A and Model B can be improved by taking into consideration possible confounders. As described in Chapter 2, confounding variables fall outside of social capital and may affect the outcome, rural hospital closure. Confounders can significantly affect Model A and Model B. Therefore, six confounding variables will be controlled for in this study through the model building process. Apart from confounders, plausible interaction variables will also be controlled for in Model A. For example, the interaction or relationship between income inequality (as measured by the Gini coefficient) and diversity is a stronger predictor of rural hospital closure than either variable on its own. Model B, on the other hand, does not need to evaluate interactions because factor analysis already takes into consideration interactions among variables.

To further test the robustness of the Model A and Model B results, both models will be re-run in their entirety with a subset of the rural hospital closure data. In the United States, there are 2,052 rural counties; however, only 1,614 of rural counties have one or more hospitals located within their borders. The goal of this sensitivity analysis is to determine how effective Model A and Model B are in predicting the likelihood of rural hospital closure when the study is limited to just rural counties with a hospital. Perhaps counties without a hospital are fundamentally different than those with at least one hospital, and their inclusion could change the results.

In the end, the overarching aims of this study are three-fold. First, Model A and Model B will determine which social capital variables or factors are significant in predicting the likelihood of rural hospital closure. Second, confounding variables, like US Census regions, and interaction terms will be controlled for in the analysis. Finally, for greater robustness in the findings, sensitivity analysis will be performed to determine if the lack of a hospital in a rural county changes the degree to which social capital predicts rural hospital closure.

3.3 Methodology

In this study, social capital and ultimately the factors that characterize social capital are the independent variables, while rural hospital closure is the dependent variable. Social capital data for this study originated from the robust, nationally-representative Research Triangle Institute’s (RTI) Spatial Impact Factor (SIF) database. It is maintained through collaboration between RTI and Arizona State University’s GeoDa Center for Geospatial Analysis and Computation and work is funded by National Institutes of Health grants. The SIF database includes results from 28 separate datasets that have been linked and geocoded; however, only 11 datasets will be used in this study. As Chapter 2 showed, the 11 datasets utilized are each reported by different federal-level departments or agencies, Health and Human Services (HHS), Labor (DOL), Justice (DOJ), and Agriculture (USDA), or by research teams willing to share their data. Data is aggregated at the county-level and limited to counties classified as rural by the USDA’s Economic Research Service Rural-Urban Continuum or Beale Codes (n=2,052).

A systematic literature review on social capital and its measurement was conducted to evaluate and identify which variables included in SIF database had been previously used in social capital research, as explained in Chapter 2. In the process of developing a comprehensive social capital model, consideration was placed on identifying variables used in prior social capital research, finding data availability across time, and ensuring sampling or surveying techniques provided an accurate description of rural counties. In the end, 60 variables were extracted, in a process grounded in a comprehensive review of the literature, to end up with 20 social capital variables that fall into seven domains. After calculating the descriptive statistics of these 20 variables and running a preliminary factor analysis (Table 2), two of the 20 were removed from the analysis – the number of full-time-equivalent (FTE) physicians per 1,000 rural residents and the percent of rural residents who have completed any form of post-secondary education (some college or higher completion). The number of FTE physicians per 1,000, as expected, is removed due to low variability across counties. With regards to the Some College variable, the descriptive analysis shows an error in the data collection or aggregation. Certain counties had a higher “some college” percentage than “high school graduate” percentage, which logically cannot occur because a high school degree is required for post-secondary enrollment. This discrepancy eliminates this variable from the analysis.
The SIF database provides nearly complete data for all 2,052 rural counties. The variable with the maximum missing data is Substance Abuse Crimes (4.7% missing data) with most missing values attributed to Florida and Illinois. Due to the completeness of the dataset, the minimum amount of analytical data needed for this analysis is easily met. The completeness goal of 100 percent was not expected in order to account for minor losses stemming from issues such as sample collection. Therefore, there is no need to deal with missing data with deletion, single imputation, or model-based methods.299

TABLE 16: SOCIAL CAPITAL VARIABLES & DESCRIPTIVE STATISTICS (2000)

<table>
<thead>
<tr>
<th>No.</th>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Domain 1: Population Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Population Loss, 1990-2000 (%) Single Parent Family Households, 2000 (%)</td>
<td>2015</td>
<td>0.07</td>
<td>0.14</td>
<td>-0.84</td>
<td>0.90</td>
</tr>
<tr>
<td>2</td>
<td>Diversity Index, 2000† Linguistically-Isolated Households, 2000 (%)</td>
<td>2052</td>
<td>0.08</td>
<td>0.08</td>
<td>0.00</td>
<td>0.79</td>
</tr>
<tr>
<td>3</td>
<td>Diversity Score†</td>
<td>2049</td>
<td>0.13</td>
<td>0.13</td>
<td>0.00</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>Domain 2: Housing &amp; Home Ownership</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Vacant Housing, 2000 (%)</td>
<td>2049</td>
<td>0.17</td>
<td>0.10</td>
<td>0.03</td>
<td>0.77</td>
</tr>
<tr>
<td>5</td>
<td>Owner-Occupied Housing, 2000 (%)†</td>
<td>2049</td>
<td>0.62</td>
<td>0.09</td>
<td>0.00</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>Domain 3: Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>High School Graduates, 2000 (%)†</td>
<td>2049</td>
<td>0.36</td>
<td>0.06</td>
<td>0.11</td>
<td>0.53</td>
</tr>
<tr>
<td>7</td>
<td>Some College or Higher Completion, 2000 (%)†</td>
<td>2049</td>
<td>0.40</td>
<td>0.10</td>
<td>0.17</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>Domain 4: Economic Indicators</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Gini Coefficient, 2000</td>
<td>2024</td>
<td>0.44</td>
<td>0.04</td>
<td>0.34</td>
<td>0.60</td>
</tr>
<tr>
<td>9</td>
<td>Unemployment, 2000 (%)</td>
<td>2050</td>
<td>0.05</td>
<td>0.02</td>
<td>0.02</td>
<td>0.17</td>
</tr>
<tr>
<td>10</td>
<td>Per Capita Income, 2000†</td>
<td>2031</td>
<td>21349.75</td>
<td>4254.51</td>
<td>7490.00</td>
<td>72844.00</td>
</tr>
<tr>
<td>11</td>
<td>Population in Poverty, 2000 (%)</td>
<td>2049</td>
<td>0.15</td>
<td>0.07</td>
<td>0.00</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>Domain 5: Crime &amp; Violence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Violent Crimes, 2000</td>
<td>2049</td>
<td>57.84</td>
<td>99.08</td>
<td>0.00</td>
<td>1384.00</td>
</tr>
<tr>
<td>13</td>
<td>Property Crimes, 2000</td>
<td>2049</td>
<td>522.41</td>
<td>789.64</td>
<td>0.00</td>
<td>8028.00</td>
</tr>
<tr>
<td>14</td>
<td>Substance Abuse Crimes, 2000</td>
<td>1956</td>
<td>473.10</td>
<td>567.81</td>
<td>0.00</td>
<td>5906.00</td>
</tr>
<tr>
<td></td>
<td>Domain 6: Health</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Uninsured, 2000 (%)</td>
<td>2050</td>
<td>0.15</td>
<td>0.05</td>
<td>0.04</td>
<td>0.39</td>
</tr>
<tr>
<td>16</td>
<td>FTE Physicians per 1000, 2000†</td>
<td>2052</td>
<td>0.19</td>
<td>0.64</td>
<td>0.00</td>
<td>15.00</td>
</tr>
<tr>
<td></td>
<td>Domain 7: Community Participation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Composite Social Capital Index^†</td>
<td>2023</td>
<td>0.28</td>
<td>1.81</td>
<td>-3.80</td>
<td>15.22</td>
</tr>
<tr>
<td>18</td>
<td>Presidential Voter Turnout (%)†</td>
<td>2049</td>
<td>0.54</td>
<td>0.10</td>
<td>0.19</td>
<td>0.98</td>
</tr>
</tbody>
</table>

^Although this is titled a “Social Capital Index,” it was developed by Penn State University and only measures the degree of community participation.
†As these variables increase, social capital also is expected to increase.

Rural hospital data was purchased from Health Forum, the data repository enterprise for the American Hospital Association (AHA). Many prominent hospital closure studies have used AHA data.\(^{300,301,302}\) However, to verify and ensure accuracy of the data, the AHA file was

merged with the CMS Provider File using the CMS provider number available in both the AHA and CMS databases. As the CMS Provider of Service (POS) data became more accessible in recent years, its first use in rural hospital closure was in a 2016 article.\textsuperscript{303}

The initial AHA file included all U.S. hospitals (n=5,535). Removing all urban-based hospitals eliminates 2,965 hospitals, leaving 2,570 rural-based hospitals. Next, this study only includes short-term, acute care (STAC) hospitals located in rural areas. Thus, all non-STAC hospitals (n=444) were removed from the database; this includes specialty, long-term, VA, and military hospitals. The remaining 2,126 hospitals are rural, STAC hospitals, and they need to be assessed for closure status. The AHA defines closure as cases where hospital data is no longer available for a specific hospital, identified by both an AHA and CMS provider number. To ensure accuracy, all hospitals classified as “closed” were then analyzed further to ensure the hospital did not assume a new name or merged with another hospital and actually remained operational.

In the final rural hospital dataset, there are a total of 2,126 rural hospitals, of which 70 (3.3%) hospitals closed between 2000 and 2007; 1,990 (93.6%) hospitals remained operational. However, for this study, rural hospital closure data must be merged with social capital data and hence be reported at the county level. With so few cases of hospital closure (as shown in Table 17), a dichotomous CLOSE variable was created, with 1 indicating that a particular county did experience at least one hospital closure between 2000 and 2007 and 0 otherwise.

The low rural hospital closure rate also makes analysis of the degree of rurality (4-9 on the Rural-Urban Continuum) difficult due to the sparsity of hospital closure data. Therefore, for the initial analysis, homogeneity among rural counties will be assumed. Although researchers acknowledge that using an urban-rural continuum reflects a deeper

understanding of the subtle differences along the range, rural is most often defined as non-urban and part of a binary urban-rural variable.\textsuperscript{304,305}

Sparsity of rural hospital closure data may prevent stratifying results by degree of rurality, however, the models can be refined by determining if the presence of at least one hospital in a rural community affects the association between social capital and rural hospital closure. Rural hospitals are symbols of a community's identity and sustainability, in addition to being a key healthcare services provider.\textsuperscript{306} Rural hospitals are often the second largest employer in the community, employ some of the highest paid and educated individuals, and are a source of economic stability.\textsuperscript{307,308,309,310} Therefore, rural counties with no hospitals (n = 438) may have less of an identity, be more vulnerable to economic downturns that render people unable to contribute towards community building efforts, and consequently have lower social capital compared to rural counties with at least one hospital (n = 1,614) (See Table 17).\textsuperscript{311,312,313} Sensitivity analysis will test the robustness of the results and determine if the presence of a hospital in a rural community changes the initial findings. All analyses were performed with STATA statistical software. A value of P<0.05 was considered statistically significant.

3.3.1 OBJECTIVE

As explained in the conceptual model, the primary goal of this study is to determine if a rural community’s social capital is associated with the likelihood of a rural hospital closing. Mirroring previous research, this study will run two parallel models. Model A tests the relationship between social capital and rural hospital closure using a simple stepwise logistic regression, in which each of the 18 social capital variables is assumed to have an equal effect on rural hospital closure. Model B also uses logistic regression, but social capital is acknowledged to be measured by not just the 18 variables but also the interrelationships among them. Therefore, Model B first runs factor analysis to determine significant social capital factors and then enters these factors in a logistic regression.

3.3.2 MODEL A: LOGISTIC REGRESSION USING SOCIAL CAPITAL INDICATORS

As described previously, Model A is a stepwise logistic regression. All predictors, the 18 social capital variables, enter the model at once with the dichotomous rural hospital closure outcome variable. The removal of predictors one at a time from the model is an iterative process based on statistical criteria of significance (P<0.05). However, after running a simple logistic regression, Model A needs refinement by testing possible confounders and interaction terms. Confounders, which will be discussed later, are additional variables that may act upon both the independent and dependent variables. Interaction terms are created

---

by multiplying two predictor variables and their inclusion in a regression expands the understanding of relationships among the variables in the model.

3.3.3 **MODEL B: LOGISTIC REGRESSION USING SOCIAL CAPITAL FACTORS**

Model B requires a two-level analysis. Initially, a data reduction technique known as factor analysis (FA) will be used to consider and distill 18 social capital variables and their interrelations into significant factors. Then, logistic regression will be used to determine if these social capital factors predict the likelihood of rural hospital closure.

Although the simplicity of Model A could facilitate interpretation of findings, past studies show the variables, which collectively define social capital, can be highly correlated with one another, and this collection of data has an underlying, fundamental subset of components or factors that can reflect social capital in a simpler way, unencumbered by inter-variable correlations. The goals of FA are to summarize patterns of correlation among observed variables, to reduce a large number of observed variables to a smaller number of factors, to provide an operational definition for an underlying process by using observed variables, and to test a theory about the nature of underlying processes.\(^\text{316}\) FA is commonly used by social capital scholars. In this application of FA, it reduces the 18 social capital variables (and accounts for any interrelationships) to a few factors. Since the assumption of normality is required in FA, the following five social capital variables will be normalized in preparation for factor analysis: (1) per capita income, (2) violent crime, (3) property crime, (4) substance abuse crime, and (5) social capital index developed by Penn State University. Normalization is required prior to any data aggregation as these variables in the data set often have different measurement units. It is needed in order to render them comparable.

Mathematically, FA produces several linear combinations of observed variables. Each linear combination is called a factor, and each factor summarizes the patterns of correlation in the observed correlation matrix.\(^\text{318}\) The correlation matrix helps determine how to reduce the dimension of the problem. Only one of two highly correlated variables would be needed to

---


measure variability in the data. Therefore, less variables are needed to be examined. Geometrically, factors form spatial dimensions or axes of social capital that makes social capital a more easily interpretable construct.

In the end, FA derives factors from the data itself that are independent (meaning no variable contributes to more than one factor) and that measure a different, unique dimension of social capital. Since Model B incorporates interrelationships among the 18 social capital variables, this enhanced model only requires testing possible confounding variables and no interaction terms. Although the simplicity of Model A facilitates interpretation of findings, Model B strengths are rooted in its enhanced representation of social capital. The use of factor analysis in Model B relies on the social capital data itself to determine significant dimensions of social capital that each incorporates the effect of multiple variables and correlations between them.

3.3.4 Adjusting Model A and Model B by Possible Confounding Variables

As the conceptual model previously showed, social capital and rural hospital closure can be influenced by the presence of additional variables identified as environmental factors or hospital factors. Confounding is a situation in which a measure of association between social capital and rural hospital closure is distorted by the presence of another variable. Positive confounding is when the observed association is biased away from the null; negative confounding occurs when the observed associated is biased towards the null. Identification of possible confounders or mediator variables was grounded in a thorough understanding of social capital, rural health, and hospital closure, and was directed by data availability. In this study, the following mediator variables were selected for testing: Census region, degree of rurality, economic dependency, Health Professional Shortage Area (HPSA) classification, presence of Federally Qualified Health Centers (FQHC), and presence of Rural Health Clinic (RHC).

Census regions are important to include because Putnam has found social capital variation by region due in part to historical and cultural contexts such as segregation in the South, and the rate of rural hospital closure also varies by Census region (Table 12).

---

2000 and 2007, 2 percent of rural hospitals closed in the Northeast Census region compared to over 5 percent in the South. Rurality could also impact both social capital and possibly rural hospital closure. For example, the recent Great Recession of 2008 and 2009 affected the most rural, remote counties (RUC 8-9) less than rural counties with a higher population density. Employment only fell by 1.3 percent in the most isolated rural counties (RUC 8-9) compared to 5 to 6 percent in the remaining rural counties (RUC 4-7). This is also due to most rural counties being less dependent on manufacturing and more dependent on industries that encourage social cohesion and interrelationships among residents such as the agriculture sector. The inclusion of HPSA classification and the presence of FQHCs and RHCs are necessary since the presence of other healthcare providers and lack of healthcare professionals are both factors previously found to be related to rural hospital closure. These variables are also related to social capital because healthcare professionals are often the highest educated residents in rural areas and have a greater potential to contribute to social capital building efforts through community involvement.

Since many experts have found that hospital characteristics or factors can predict rural hospital closure, identification of additional and more directly associated hospital factors was attempted. The initial AHA dataset did include bed counts; however, bed count data was found to be very inaccurate. In an attempt to better describe rural hospitals, the AHA dataset was merged with the CMS provider file by linking unique CMS provider numbers for each hospital. The CMS provider file also includes bed counts and other variables include number of healthcare providers, but the bed counts differed from the AHA dataset. Initially, rural hospitals that fell outside the standard deviation of average bed count were considered to be inaccurate. Yet, a random sample of rural hospitals was selected and individually researched to determine if a systematic error occurred in bed count tabulation.

No systematic error was identified. Some rural hospitals report bed count as a total of beds within their STAC facility. Others report the total bed count among all facilities they operate including long-term care and nursing homes. Some Critical Access Hospitals report the number of beds they typically use to provide STAC, tertiary care, instead of the total bed count. Certain rural hospitals do not publish bed counts or are required to report it publically to state health departments or hospital associations. Therefore, bed count was not included in the analysis. Apart from bed count data, financial and operational data specific to a particular hospital was unavailable or too expensive to acquire for the 2000 – 2007 time period of the study. Due to the low number of discharges, rural hospitals are often exempt from every publically reporting data; such is the case with Hospital Compare reports.

To test Model A and Model B for confounding, each identified confounding or mediator variable is forced into the initial model individually. Next, the odds ratios from crude or initial logistic regression model (Model A or Model B) are compared to the odds ratios from the adjusted logistic regression model with the single confounding variable included. If the odds ratio changes by 10 percent or more, the variable added into the model (i.e. number of FQHCs) is truly a confounder and must be added to the final Model A or Model B.\textsuperscript{326}

\subsection*{3.3.5 Adjusting Model A: Logistic Regression by Main Interaction Terms}

In addition to testing for confounding, adding interaction terms to a logistic regression model can greatly expand understanding of the relationships among the variables in the model and allows more hypotheses to be tested. Interaction terms test if the relationship between two variables affects the result in a distinguishable way than solely including each variable separately in the model. Similar to the process of selecting possible confounders, selection of possible interaction terms was based on previous research studies, correlation matrix results, and the crude Model A. This preliminary analysis allows the total number of possible interaction terms (n=190), as calculated by the combinations of 2 from the 18 social capital variables in the model, to be reduced to the most likely 6 interaction terms. This process will be discussed later in results. Having selected which interaction terms to calculate, each interaction term is tested in Model A. Testing for interaction terms in Model

B is unnecessary because factor analysis already accounts the effect of interrelationships or correlations between social capital variables. Once again, a 10 percent or higher change in the odds ratio between the crude and adjusted model may justify inclusion of the interaction term to the model.

### 3.4 Results and Interpretation

This study runs two parallel models. The initial or crude Model A uses logistic regression to determine which of the 18 social capital variables are significant in predicting rural hospital closure. Model A will be refined by first testing if possible confounding variables significantly impact the model and secondly by forcing possible interaction terms into the model. If any confounders or interaction terms significantly change the odds ratio, then they will be added to the final adjusted Model A. In the second approach, Model B also uses logistic regression, but social capital is acknowledged to be measured by not just the 18 variables but also the interrelationships among them. Therefore, Model B first runs factor analysis to determine significant social capital factors, second runs the logistic regression model, and finally tests confounding variables in the model. Adjusting Model B with interaction terms is not required. If any are confounders are found to be significant, Model B will be adjusted by the inclusion of these variables.

Assessing the appropriateness of both Model A and Model B requires examining its fit, or how well the model describes the observed data. \(^{327}\) Calibration of the model was assessed by the Hosmer-Lemeshow Goodness of Fit (GOF) statistic for significance (P>0.05). This GOF test creates 10 ordered groups of subjects and then compares the number actually in each group (observed) to the number predicted by the logistic regression model (predicted). \(^{328}\) Thus the Hosmer-Lemeshow GOF statistic is a chi-squared statistic with a desirable outcome of non-significance, indicating that the model prediction does not significantly differ from the observed. For this study, the Hosmer-Lemeshow GOF statistic was calculated for every step of the model building process, and unless otherwise noted, all models were found to have a desirable outcome of non-significance.

---


3.4.1 Model A: Logistic Regression Using Social Capital Indicators

Since the dependent or response variable is dichotomous, rural hospital closure status, logistic regression allows for the prediction of rural hospital closure from a set of social capital variables that may be continuous, discrete, dichotomous, or a mix of all three. The goal of the analysis is to correctly predict rural hospital closure using the most parsimonious model, the one with the least predictors. Since initially little is known about the relationships between variables, backward stepwise regression is a common method of exploratory analysis. The analysis begins with a full model that includes all variables, and is an iterative process, with one variable being removed at a time (P<0.05). The fit of the model is tested after the elimination of each variable to ensure the model adequately fits the data. When no more variables can be removed from the model, the analysis is complete. Logistic regression model results are reported as odds ratios, which quantifies the probability of rural hospital closure over the probability of a rural hospital not closing.

3.4.1.1 Crude Model A: Main Effects

Multivariate backwards logistic regression analysis identified three social capital variables with a significant influence on rural hospital closure (P<0.05) – unemployment, substance abuse crime, and violent crime. The odds ratio (OR) in a multivariate model is a measure of association and estimates the odds of an adverse outcome given a specific exposure level of an independent variable holding all else constant. The 95% confidence interval (CI) around the odds ratio (OR) was obtained with the formula (β±1.96 SE), in which SE is the standard error of β. The CI indicates the level of uncertainty around the measure of effect; one can infer that the true OR lies between the upper and lower confidence limits. The odds ratio estimates from Model A are shown in Table 18.

<table>
<thead>
<tr>
<th>N = 2,052</th>
<th>CRUDE OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment, 2000</td>
<td>1.15</td>
<td>1.02</td>
<td>1.31</td>
</tr>
<tr>
<td>Substance Abuse Crimes, 2000*</td>
<td>1.0005</td>
<td>1.00</td>
<td>1.001</td>
</tr>
<tr>
<td>Violent Crimes, 2000*</td>
<td>1.0023</td>
<td>1.00</td>
<td>1.004</td>
</tr>
</tbody>
</table>

*To clarify, odds ratios were rounded to 4 decimal places and confidence intervals to 3

Although Model A only includes three social capital variables, only the unemployment OR immediately suggests an association between unemployment and rural hospital closure (OR>1 and CI does not contain 1). For every 1 percent increase in unemployment, the odds of a rural hospital closing increased by 1.15 times (or by 15%) holding all else constant. Unemployment often increases by more than 1 percent in rural communities, as the 2007 – 2009 recession proved, if the odds ratio is recalculated for a 5 percent increase in unemployment, the adjusted OR is 2.04 (β=0.14). This means that for every 5 percent increase in unemployment, the odds of a rural hospital closure is doubled.

In the case of substance abuse crime (SAC) and violent crimes (VC), initially the OR is almost equal to 1 and the CI includes 1 (OR_{SAC}=1.00047; OR_{VC}=1.002266). This means that for every additional substance abuse or violent crime, the odds of a rural hospital closing remains unchanged. Crime data is collected by the FBI through the Uniform Crime Reports (UCR) program and provides raw counts of reported crimes. Like unemployment, if the number of crimes changes in a community, it typically does not increase by just one crime. In general, rural areas face growing trends in both substance abuse crime and violent crime. Substance abuse crime, especially prescription drug abuse, is an increasing in rural areas. Furthermore, due to isolation, rural areas are more likely to serve as production and cultivation sites, as well as important transshipment points for drugs destined for cities. Since the average number of reported substance abuse crimes per rural county is 473.10, a conservative 10% increase in crimes is equivalent to around 50 more substance abuse crimes per rural county. If the OR_{SAC} is recalculated to incorporate an increase of 50 substance abuse crimes, the adjusted OR_{SAC} is 1.26 (β=0.0047). This means the odds of a rural hospital closure is increased 1.26 times for an increase in 50 substance abuse crimes. Violent crimes are the most severe, including murder, forcible rape, and robberies, and are classified as Part I offenses by the FBI. Since the average number of violent crimes per rural

county is 57.4, a conservative 10% increase in crimes is a total of 5 violent crimes. If the OR_{VC} is recalculated for an increase in 5 violent crimes, the adjusted OR_{VC} is 1.01 (β=0.0023). This means that for an increase in 5 violent crimes, the odds of a rural hospital closure is increased by 1.01 times holding all else constant. As the adjusted ORs show, both substance abuse crime and violent crimes do have an effect on rural hospital closure albeit small.


<table>
<thead>
<tr>
<th></th>
<th>n = 2,052</th>
<th>Adjusted OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td>Upper</td>
</tr>
<tr>
<td>Unemployment, 2000</td>
<td>2.04</td>
<td>1.09</td>
<td>3.80</td>
<td>0.03</td>
</tr>
<tr>
<td>Substance Abuse Crimes, 2000</td>
<td>1.26</td>
<td>1.24</td>
<td>1.29</td>
<td>0.01</td>
</tr>
<tr>
<td>Violent Crimes, 2000</td>
<td>1.01</td>
<td>1.00</td>
<td>1.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>

The initial, crude Model A therefore is a logistic regression model with unemployment, substance abuse crime, and violent crime. To ensure that any observed association between social capital of rural communities and rural hospital closure is not affected by confounders or interaction terms, the model building process will now force each possible confounding variable and interaction term into the model to check for its effect on the odds ratio.

### 3.4.1.2 Adjusting Model A for Confounding

One of the key steps in model building is to force possible confounding variables into the model one at a time and determine if each deserves inclusion in the model based on its impact on the odds ratios generated. Confounding variables are variables outside of the independent variables that may affect the dependent variable, and they must be controlled for in the analysis. The six variables selected for possible Model A adjustment represent environmental and policy characteristics of rural communities that could have an impact on rural hospital closure. For further information, see Chapter 2.
### TABLE 20: CRUDE AND ADJUSTED ODDS RATIOS FOR POSSIBLE CONFOUNDING VARIABLES IN MODEL A (2000)

<table>
<thead>
<tr>
<th>YEAR: 2000</th>
<th>UNEMPLOYMENT</th>
<th>SUBSTANCE ABUSE CRIMES,</th>
<th>VIOLENT CRIMES,</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Model</td>
<td>Crude OR</td>
<td>1.15</td>
<td>1.00</td>
</tr>
<tr>
<td>Rural Health Clinics (0/1)</td>
<td>Adjusted OR</td>
<td>1.14</td>
<td>1.00</td>
</tr>
<tr>
<td>Federally Qualified Health Centers (0/1)</td>
<td>Adjusted OR</td>
<td>1.14</td>
<td>1.00</td>
</tr>
<tr>
<td>Health Professional Shortage area (hpsa 1, hpsa2)</td>
<td>Adjusted OR</td>
<td>1.19</td>
<td>1.00</td>
</tr>
<tr>
<td>Economic Dependency (0/1)</td>
<td>Adjusted OR</td>
<td>1.14</td>
<td>1.00</td>
</tr>
<tr>
<td>RUC Codes (4-5, 6-7, 8-9)</td>
<td>Adjusted OR</td>
<td>1.16</td>
<td>1.00</td>
</tr>
<tr>
<td>Census Region (1, 2, 3, 4)</td>
<td>Adjusted OR</td>
<td>1.15</td>
<td>1.00</td>
</tr>
</tbody>
</table>

As Table 20 shows, when any of the six possible confounding variables were forced into the multivariate logistic regression analysis (Model A), the OR failed to change by any significant amount. Therefore, the presence of rural health clinics or federally qualified health centers in a rural community does not affect Model A. The designation of a health professional shortage area in the rural community also does not affect the model. The dependence of a rural county’s economy on a specific industry also fails to have an effect on Model A. Next, since rurality has been found to possibly impact the relationship between social capital and rural hospital closure, Model A is tested for the size and proximity of a rural community to an urban center, through three RUC categories. No effect is reported. Finally, although dividing the entire United States into four regions may be oversimplification, some findings do reflect fundamental differences between these four regions. The South, for example, has a long legacy of racial discrimination, lower social capital, and higher rural hospital closure rates, than the other three regions (Northeast, Midwest, and West). However, even adding Census region into the model failed to change the OR. This result gives credence to crude Model A being the best representation of the relationship between social capital of rural communities and rural hospital closure thus far.

---

To complete building Model A, the model must now be tested with possible interaction terms.

### 3.4.1.3 Adjusting Model A for Possible Interactions

In Model A, the model building process also needs to force possible interaction terms into the model one at a time and determine if each deserves inclusion in the model based on its impact on the odds ratios (OR) generated. The inclusion of a significant interaction term indicates that the effect of one predictor variable on the outcome variable is different at different values of the other predictor variable.\(^{335}\) In this study, two social capital variables are identified as being the most likely to interact with another predictor variable to affect rural hospital closure – diversity score and Gini coefficient, which is a measure of income inequality.

Diversity score is fundamentally a measure of racial diversity. In sociology, some experts consider “race” to be a fluid not just a concrete concept based on person’s biology and upbringing. In fact, Omi and Winant, who developed an analytical tool known as racial formation theory, write, “The racial order is organized and enforced by the continuity and reciprocity between micro-level and macro-level of social relations.”\(^{336,337}\) This means that theoretically the effect of race on rural hospital closure is likely to be different at different values of another social capital variable because race is so interrelated with social structures. Thus interaction terms with diversity score merit testing in the model. Similar to the rationale for testing interaction terms with race, interaction terms with the Gini coefficient also merit consideration in Model A. Income inequality, which is measured by the Gini coefficient, has been found to be related to a wide range of predictor variables including unemployment and violent and substance abuse crime.\(^{338,339}\) The effect of the Gini coefficient on rural hospital closure may differ at different levels of unemployment or crime.

---


thus these interaction terms must also be forced into Model A one at a time to compare the crude OR with the adjusted OR.


<table>
<thead>
<tr>
<th>Year: 2000</th>
<th>Unemployment</th>
<th>Substance Abuse Crimes</th>
<th>Violent Crimes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Model</td>
<td>Crude OR 1.15</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>DiversityScore*Unemployment</td>
<td>Adjusted OR 1.15</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>DiversityScore*SubstanceAbuse Crimes</td>
<td>Adjusted OR 1.18</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>DiversityScore*ViolentCrimes</td>
<td>Adjusted OR 1.17</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Gini*Unemployment</td>
<td>Adjusted OR 0.71</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Gini*SubstanceAbuseCrimes</td>
<td>Adjusted OR 1.11</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Gini*ViolentCrimes</td>
<td>Adjusted OR 1.11</td>
<td>1.00</td>
<td>0.99</td>
</tr>
</tbody>
</table>

As Table 21 shows, none of the interaction terms forced into the multivariate logistic regression model (Model A) changed the odds ratio by any significant amount, except for interaction terms formed by multiplying the Gini coefficient and unemployment. However, this interaction term only affected one of the three variables in Model A, so it can remain excluded from the model. In the end, none of the confounding variables and interaction terms tested in Model A was found to be significant. One reasons for this is likely due to the sparsity of data; the limited number of rural hospital closures (n=70) limits the number of factors that can be adjusted for and included in the model. In addition, the confounding variables and interaction terms tested in Model A could be significant for certain rural communities, but over the entire country, this possible local effect is lost.

Thus none necessitate inclusion in the model, so Model A requires no adjustment. The initial or crude Model A best predicts the association between a rural community’s social capital and the likelihood of a rural hospital closing there. However, the change in the odds ratios caused by each interaction term, although not significant enough to warrant inclusion, does reflect the interrelationships between social capital variables. The interrelationships between social capital variables are commonly recognized by social capital scholars as a problem that needs addressing. Analytically, this is most often resolved by the use of factor analysis prior to running the regression model. This approach directs Model B.
3.4.2 **MODEL B:**

As mentioned earlier, the second analytical approach, identified as Model B, first performs factor analysis to determine social capital factors, before running the logistic regression model. Similar to Model A, this model will also test confounding variables and adjust Model B if any confounders are found to significantly change the odds ratios.

3.4.2.1 **Factor Analysis**

Factor analysis (FA) is a data reduction, statistical technique that summarizes a set of individual indicators while preserving the maximum possible proportion of the total variation in the original data set. FA analyses can be classified into either, a theory-testing approach, known as exploratory factor analysis, or a theory-generating method, known as confirmatory factor analysis. While exploratory factor analysis is largely directed at identifying a relatively small set of underlying clusters or dimensions that subsume a larger number of inter-correlated variables, confirmatory factor analysis (CFA) statistically tests these clusters as well as the predictive validity of the factor structure. It is appropriate in CFA to test the relationship between various theoretical models. This CFA approach, in which several, increasingly more detailed models are evaluated, will be used in this study and is consistent with commonly advocated procedures.

The factor analysis technique like backwards logistic regression is an iterative process. After each run of FA, the model is refined based on the results. This is repeated until no more variables need to be removed from the analysis and the number of factors is determined. From the initial FA run, three of the 18 variables were removed from the analysis because they did not load highly in any of the factors and failed to meet the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy threshold (KMO>0.60). The

---


The following three variables were removed: percent population change, vacant housing, and owner-occupied housing. With the 15 social capital variables remaining, factors were extracted from this analysis and orthogonally rotated with the promax and varimax methods. Results for two to four factors were examined. Ultimately, two factors were extracted. The selection of two factors was based on examination of eigenvalues, scree plots, and the interpretability of the solution. The estimated factor scores were computed using a regression method based on 15 social capital variables. All factor loading ≥ |0.32| were presented.


<table>
<thead>
<tr>
<th>Factors</th>
<th>Economic Vulnerability</th>
<th>Crime Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenvalue</td>
<td>4.82</td>
<td>2.41</td>
</tr>
<tr>
<td>Proportion</td>
<td>0.57</td>
<td>0.28</td>
</tr>
<tr>
<td>Cumulative</td>
<td>0.57</td>
<td>0.85</td>
</tr>
</tbody>
</table>

A factor with an eigenvalue greater than 1.0 is accounting for a greater amount of variance than was contributed by one variable. Therefore, these factors are accounting for a meaningful amount of variance and merit inclusion in the analysis. In this study, factor 1 is best classified as economic vulnerability and factor 2 as crime risk. The third factor in this model has an eigenvalue less than 1.0, which accounts for less variance than had been contributed by one variable, and it is eliminated from the study. This commonly used rule is known as the latent-root or Kaiser criterion. As Table 22 shows, the first factor, economic vulnerability, explains the maximum variance in the data (57% in the promax rotation). The second factor explains the second greater amount of variance in the data. The inclusion of

---

both factor 1 and factor 2 in the analysis signifies that 85 percent of the (cumulative) variance in the data is explained.

TABLE 23: FACTOR LOADINGS (2000)

<table>
<thead>
<tr>
<th>Factors</th>
<th>Economic Vulnerability</th>
<th>Crime Risk</th>
<th>KMO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Parent Family Households, 2000</td>
<td>0.69</td>
<td></td>
<td>0.79</td>
</tr>
<tr>
<td>Diversity Index, 2000†</td>
<td></td>
<td>0.33</td>
<td>0.87</td>
</tr>
<tr>
<td>Linguistically-Isolated Households, 2000</td>
<td>0.32</td>
<td></td>
<td>0.65</td>
</tr>
<tr>
<td>Diversity Score, 2000†</td>
<td>0.63</td>
<td></td>
<td>0.81</td>
</tr>
<tr>
<td>High School Graduates, 2000‡</td>
<td>-0.32</td>
<td></td>
<td>0.78</td>
</tr>
<tr>
<td>Gini Coefficient, 2000</td>
<td>0.69</td>
<td></td>
<td>0.80</td>
</tr>
<tr>
<td>Unemployment, 2000</td>
<td>0.63</td>
<td></td>
<td>0.88</td>
</tr>
<tr>
<td>Per Capita Income, 2000‡*</td>
<td>-0.68</td>
<td></td>
<td>0.79</td>
</tr>
<tr>
<td>Population in Poverty, 2000</td>
<td>0.94</td>
<td></td>
<td>0.78</td>
</tr>
<tr>
<td>Violent Crimes, 2000*</td>
<td></td>
<td>0.84</td>
<td>0.77</td>
</tr>
<tr>
<td>Property Crimes, 2000*</td>
<td></td>
<td>0.94</td>
<td>0.69</td>
</tr>
<tr>
<td>Substance Abuse Crimes, 2000*</td>
<td></td>
<td>0.83</td>
<td>0.83</td>
</tr>
<tr>
<td>Uninsured, 2000</td>
<td>0.83</td>
<td></td>
<td>0.82</td>
</tr>
<tr>
<td>Composite Social Capital Index†*</td>
<td>-0.56</td>
<td></td>
<td>0.88</td>
</tr>
<tr>
<td>Presidential Voter Turnout, 2000‡</td>
<td>-0.46</td>
<td></td>
<td>0.80</td>
</tr>
<tr>
<td>Overall Kaiser-Meyer-Olkin Measure</td>
<td></td>
<td></td>
<td>0.80</td>
</tr>
</tbody>
</table>

†As these variables increase, social capital also is expected to increase.
*These variables were normalized prior to FA to fulfill normality assumption.

In the final factor extraction, 15 social capital variables were reduced into two factors – economic vulnerability and crime risk. As the factor loadings show (Table 23), economic vulnerability is composed of the following eleven variables from 2000: (1) single parent households (%), (2) linguistically-isolated households (%), (3) diversity score (%), (4) high school graduates (%), (5) Gini coefficient, (6) unemployment (%), (7) population in poverty (%), (8) per capita income, (9) uninsured (%), (10) composite social capital index, and (11) presidential voter turnout (%). Crime risk is composed of the following four variables: (1) diversity index, (2) violent crime, (3) property crime, and (4) substance abuse crime. Economic vulnerability factor is a combination of population characteristics plus economic indicators. Crime risk is essentially limited to variables classified as domain 5 in Chapter 2 crime and violence. For the purpose of the analysis, the promax method of orthogonal
rotation was selected due to clarity and independence of each factor, which facilitates interpretability.

The final factor extraction also yields a Kaiser-Meyer-Olkin value of 0.80. An overall adequacy of 80 percent confirms that these two factors – economic vulnerability and crime risk – are good measures or dimensions of social capital. The FA process has now created two factors that measure social capital for each county. To complete the initial, crude Model B, a logistic regression analysis tests the relationship between these two social capital factors and the likelihood of rural hospital closure.

**3.4.2.2 Crude Model B: Main Effects**

In Model B, two social capital factors were identified and run through a multivariate logistic regression model, where the dichotomous outcome variable is rural hospital closure. Since these social capital factors are composed of 15 social capital variables and more significantly represent the social capital construct, it is not surprising that the odds ratios (OR) reflect a stronger association between social capital factors and rural hospital closure than was found in Model A (Table 18). Despite differences in the odds ratios, it is noteworthy that both factors – economic vulnerability and crime risk – align with the three social capital variables found to significantly predict rural hospital closure in Model A (unemployment, substance abuse crime, and violent crime).


<table>
<thead>
<tr>
<th></th>
<th>Crude OR</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 2,052</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economic Vulnerability (f7)</td>
<td>1.42</td>
<td>1.10 1.84</td>
<td>0.01</td>
</tr>
<tr>
<td>Crime Risk (f8)</td>
<td>1.47</td>
<td>1.22 1.78</td>
<td>0.01</td>
</tr>
</tbody>
</table>

The Model B odds ratio estimations, along with the calculated confidence intervals, suggest a strong association between rural hospital closure and both economic vulnerability and crime risk (OR>1 and CI does not contain 1). Since factors have no units, for every one unit increase in the economic vulnerability factor, the odds of a rural hospital closing increased
by 1.42 times holding all else constant. Regarding crime risk, for every one unit increase in the crime risk factor, the odds of a rural hospital closing increased by 1.47 times holding all else constant.

### 3.4.2.3 Adjusting Model B for Confounding

Similar to the previous model, testing Model B by forcing possible confounding variables into the model is necessary to determine if each deserves inclusion in the model. The same six possible confounding variables tested for in Model A were forced into Model B one at a time. For further information, see Chapter 2.

<table>
<thead>
<tr>
<th>TABLE 25: CRUDE AND ADJUSTED ODDS RATIOS FOR POSSIBLE CONFOUNDING VARIABLES IN MODEL B (2000)</th>
</tr>
</thead>
</table>
| **Year: 2000** \[\begin{array}{|c|c|c|c|} \hline
<table>
<thead>
<tr>
<th>Variable</th>
<th>Crude OR</th>
<th>Adjusted OR</th>
<th>Adjusted OR</th>
<th>Adjusted OR</th>
<th>Adjusted OR</th>
<th>Adjusted OR</th>
<th>Adjusted OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic Vulnerability</td>
<td>1.42</td>
<td>1.38</td>
<td>1.37</td>
<td>1.41</td>
<td>1.38</td>
<td>1.31</td>
<td>1.48</td>
</tr>
<tr>
<td>Crime Risk</td>
<td>1.47</td>
<td>1.43</td>
<td>1.45</td>
<td>1.47</td>
<td>1.38</td>
<td>1.48</td>
<td></td>
</tr>
<tr>
<td>Rural Health Clinics (0/1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Federally Qualified Health Centers (0/1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Professional Shortage Area (hpsa1, hpsa2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economic Dependency (0/1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RUC Codes (4-5, 6-7, 8-9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Census Region (1, 2, 3, 4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After testing all six possible confounding variables, the difference between the crude and adjusted odds ratios are insignificant. Thus no confounding variables are added to Model B. This result indicates that the crude Model B is best at representing the association between social capital of rural communities and rural hospital closure. Although confounding variables are tested for in Model B, unlike the previous model, no interaction terms need to be forced into the model because the interrelationship among social capital components are already incorporated into the factor analysis.
3.4.3 Sensitivity Analysis

Sensitivity analysis is a method "to assess whether altering any of the assumptions made leads to different final interpretations or conclusions." The goal is to assess the robustness or consistency of the results. In this study, every rural county in the United States (n=2,052) is included in the analysis. However, as Table 17 shows, 21.3 percent of rural counties have no hospital operating within their region at the start of the study (n=438). The impact of an operational hospital on a rural county is substantial from an economic, social, or healthcare perspective. Therefore, the assumption that all rural counties all equally likely to have a hospital close is simply not true, which warrants a sensitivity analysis.

In this sensitivity analysis, both Model A and Model B are rerun with a subset of the original dataset; only counties with at least one operational hospital at the start of 2000 will be included (n=1,614). In both cases, the results remain unchanged. In Model A, instead of three variables only two were found significant to include; violent crime is no longer included in the model. In Model B, the exact two factors – economic vulnerability and crime risk – are identified in factor analysis. In fact, economic vulnerability and crime risk consist of virtually the same social capital variables as Model B. There are only two minor differences. First, linguistically-isolated households no longer contribute towards economic vulnerability; and secondly, diversity index is no longer contributing towards crime risk. Using logistic regression, the odds ratios calculated using the subset of data (n=1,614 counties) are indiscriminate from the odds ratios calculated using the entire dataset (n=2,052 counties). This result makes the findings more robust.

3.5 DISCUSSION

The findings suggest that a rural community’s social capital is related to the likelihood of a rural hospital closing. The first analysis (Model A) found that a community’s levels of unemployment, one of three significant social capital variables, has a statistically significant association with odds of rural hospital closure. The second analysis (Model B), which combined factor analysis and the statistical method used in Model A, found economic vulnerability and crime risk to be statistically significant in measuring the likelihood of rural hospital closure. The similarity in the domains of social capital variables or factors found to be significant in the parallel analyses provides researchers with two areas to concentrate future research efforts. This finding also mirrors the results from more general social capital research, which often concludes that crime but especially economic variables are the most significant predictors of social capital. Therefore, the development of the social capital model for a health services research application, as well as finding a reliable and consistent set of social capital variables or factors associated with rural hospital closure, advances theory development and assists empirical research.

Despite similarities in the social capital variables and factors found to be significant in Model A and Model B, the differences in approaches have implications on interventions that rural communities, governments, and rural hospitals could implement to affect social capital. Even after adjustment, the resulting odds ratios from the simpler Model A show that unemployment is by far the most significant social capital component in predicting the likelihood of rural hospital closure. Substance abuse and violent crime, although found to be

significant social capital variables, have smaller odds ratios, implying they are more weakly associated with rural hospital closure. This result is not surprising. Previous social capital studies have shown that economic variables are the strongest predictors of a community's social capital.\textsuperscript{358,359} Unemployment is also a measure that reflects a community's ability to collect taxes, dedicate time to community efforts instead of work production, have health insurance and seek care, and support public education.\textsuperscript{360,361} Unemployment has a tremendous destabilizing effect on groups of individuals and communities. Based on Model A results, unemployment is the most important social capital component in predicting rural hospital closure. Again, this is a logical and well-known finding. A community's economic downturns has a direct effect on the ability for the community to access and pay for healthcare services.\textsuperscript{362} The primary cause of rural hospital closure is financial distress, which would increase if uninsured rates increase as unemployment increases in a community.\textsuperscript{363,364,365} The payer mix rural hospitals rely on is less favorable than urban hospitals because rural residents are more likely to be insured by Medicare and Medicaid.\textsuperscript{366} In fact, rural populations are more likely to be disabled and unable to work. If unemployment increases, this means rural residents, capable of work, are becoming


\textsuperscript{364} General Accounting Office (1991). Rural hospitals: Federal hospitals should target areas where closures would threaten access to care. GAO. Washington, D.C.


uninsured or perhaps qualifying for Medicaid benefits. Therefore, maintaining or increasing employment steady is important for rural communities and rural hospitals. To accomplish this goal, rural communities could benefit from decreasing their dependence on any one industry and try to attract additional employers through tax benefits or by the quality of the local work force. They need to also try attracting higher quality employment opportunities that offer health benefits.

However, these interventions or goals take the time and effort of community leaders and are dependent on a wide-variety of external factors such as the global economy, weather patterns in the major agricultural areas, and effectiveness of elected representatives in attracting investment in rural areas. Strengthening the local economy is a goal any rural community desires, but sustainable successes are hard to achieve. In addition, rural policies, including economic development ones, recommended by elected leaders may be unfeasible in all rural communities and are not prioritized in the political agenda, as Agriculture Secretary Tom Vilsack highlighted in a 2013 speech. In the Obama administration, for example, the pillars of its rural policy has focused on improving education and expanding opportunities for small businesses, tourism and recreation, and clean energy. Small businesses in rural communities are difficult to sustain during economic recessions and are threatened by international trade and the entrance of large corporate chains like Wal-Mart to the area. Some rural counties may lack the capital to develop tourism, be unable to compete with other rural areas with an established tourism industry, or be too remote to attract tourists. Others may not have energy resources. Although policies and interventions can help some rural communities, many rural communities economically struggle for decades and are classified by the federal

---

370 Doering, Christopher. "As more move to the city, does rural America still matter?" USA Today 13 Jan. 2013.
government as a persistent poverty county, one where 20 percent or more of its population has lived in poverty for the past 30 years.\textsuperscript{373,374}

Model A provides a baseline understanding of social capital, and the results highlight the importance of a strong economy on rural communities. Yet, Model A fails to fully capture the interrelationships between all social capital variables, which Model B does address with factor analysis. Social capital experts have confirmed that social capital is a latent variable, and it is best measured by a group of variables, often highly correlated.\textsuperscript{375} Model B determined there are two significant social capital factors – economic vulnerability and crime risk. In factor analysis, each factor is created from a combination of variables. Economic vulnerability includes single parent family households, linguistically-isolated households, diversity score, high school graduates, uninsured, social capital index, presidential voter turnout, along with all four economic variables (Gini coefficient, unemployment, per capital income, and population in poverty). Crime risk consists of diversity index and all three crime variables (violent, property, and substance abuse crime). Even without adjustment, the resulting odds ratios find both factors to be more strongly associated with the likelihood of rural hospital closure than any of the variables in Model A. Since the economic vulnerability and crime risk factors are created from a combination of variables from the Adaniya model of social capital developed in Chapter 2, this model recognizes the interrelationship within social capital, and the results have significant rural policy implications.

This means that any policy or program attempting to intervene on one of these social capital variables would be expected to have an impact on the rest of the variables, especially those within the same factor.

Model B results imply that policies or programs attempting to intervene on one of these social capital variables would be expected to have an impact on the remaining variables, especially those within the same factor. This means that a policy aimed at increasing the high school graduation rate will impact overall economic vulnerability, including the other variables.


outcomes included in this factor. Since economic development policies are so difficult to implement successfully, Model B results indicate that a greater variety of interventions or policies can be adopted in rural communities that positively affect social capital. Previous studies conclude that many of these interventions can be implemented through building and strengthening relationships and social networks within rural communities. A community’s social capital, as Putnam reports, can be impacted by churches, political organizations, Parent Teacher Associations, and even the presence of community resources like a community center, library, or hospital.376 These groups or entities are valuable resources that can be harnessed by community leaders.

Since rural areas struggle to increase the high school graduation rate, a recent report recommends rural school principals and school boards to think creatively and utilize all of their community’s resources to improve graduation rates.377 Recommendations include reaching out to community organizations like the YMCA and United Way to support out-of-school programs. To resolve transportation issues to school, another recommendation was to establish partnerships with community organizations like churches that may have vans or buses. To help support and attract talented personnel, principals and school boards are advised to utilize the PTA as a resource, partner with 4H and Extension offices, form cooperative agreements with neighboring school districts to share resources, and develop strong school-family-community partnerships to invest in family education and encourage parental involvement. Besides the leadership role local schools can have on building social capital in rural communities, extension programs, both in the United States and throughout the world, have also been shown to galvanize community-wide efforts to improve educational opportunities, economic output, and social capital.378,379 Thus, these findings encourage rural communities that they are able to implement program that can positively impact social capital and perhaps prevent their local hospital from closing. As these examples show, social capital building policies require tapping into existing community

resources. However, these policies require individual leaders or champions to implement and maintain these connections.

In rural communities, even one of these individual leaders can have a tremendous positive impact the level of social capital. For example, in Eastern Ohio, one Quaker couple, Richard and Mary Sidwell, are recognized community leaders. First, they creatively raised $6.5 million dollars to save the Olney Friends School and opened the school up to the public for free education and cultural events. Then, the Sidwells, along with 19 others from their church, started and ran a volunteer-run Christmas tree farm for over a decade in order to raise funds to buy a property known as Raven Rock, a sacred American Indian ceremonial site that had historically been a popular tourist attraction. Even after retiring, the Sidwells remain well-known and active community leaders with multiple roles, including Vice President of the Belmont County Community Development Corp. and Secretary of the Chamber of Commerce. Although rural community leaders come from diverse backgrounds, rural hospitals and their highly-educated professional workforce often serve as individual leaders in rural communities. In fact, physicians serving rural areas are more far more committed to participate in community, political, and collective advocacy than physicians in urban areas. This reinforces the importance rural hospital play in a rural community. One of the best examples of a rural champion is Dr. Regina Benjamin, the former Surgeon General, whose community building efforts through her non-profit the Bayou La Batre Rural Health Clinic, were recognized by the MacArthur Foundation. In the end, rural communities do face many challenges and need good leaders. The policy implications of this study are encouraging for rural communities. Although economic outcomes are important, social capital can be positively impacted by adopting policies or interventions that are not just focused on economic development. Community-led efforts that recognize local resources and circumstances may be far easier to achieve and acknowledge rural heterogeneity.

Besides the policy implications of the findings, this study has additional strengths. Rural hospital closure continues to be a growing trend in the United States, although 19 percent of the U.S. population, almost 60 million people, still lives in rural areas. A deeper understanding of rural hospital closure is needed but remains a far less researched topic.

than urban hospital closure.\textsuperscript{381} Next, this study is strengthened by national scope and inclusion of all rural counties in the United States. From a time perspective, there are several studies analyzing rural hospital closure in the 1990s and one, thus far, that covers the period 2010–2014. \textsuperscript{382,383,384} This study is the first rural hospital closure study that focused on a time period in the first decade of the 21\textsuperscript{st} century.

Next, the association between social dimensions of health has been a common research theme, but this study is one of the first to consider a relationship between social capital and a healthcare system outcome variable.\textsuperscript{385} This study is only the second to associate social capital with hospital closure, but it is the first to study the link between a community’s social capital and rural hospital closure.\textsuperscript{386} Finally, a major strength of the study is its design that includes the utilization of two models of analyses, tests for confounding and interaction terms, and sensitivity analysis. The robust findings reflect this. Overall, this study adds to the growing body of social capital, rural health, and rural hospital closure literature. However, this study still faced many limitations.

3.5.1 Limitations

First, social capital data in this study is limited to a single year of analysis (2000). This means a rural community’s social capital is solely a measure for one year. Since studies have found that closing a hospital has far more negative impacts on the community served by that hospital than positive ones, hospital closure requires hospital administrators to justify the decision. Consequently, the process leading to hospital closure occurs over a longer period of time. A measure of social capital over time may better characterize or represent the community than one at a single point in time.

\textsuperscript{385} Lynch, J., et al. (2000). "Social capital--is it a good investment strategy for public health?" \textit{J Epidemiol Community Health} 54(6): 404-408.
Secondly, the county was used as the geographic unit of analysis because insufficient data is available at smaller scales of analysis. This assumes the county is the most available and feasible scale to represent a rural hospital’s market, rural hospitals are centrally located within a county, and equally distribution of the population. However, this can bias the results. Consider, for example, a rural hospital located on the border between two or more rural counties; if only one county’s social capital can be linked to that hospital, which county best represents that hospital’s market?

Third, many hospital-specific characteristics, such as hospital ownership, executive leadership effectiveness, market competitiveness, and fiscal stains, are linked to hospital closure. However, many of these measures were not available for inclusion in this study; much of this information would require in-person, qualitative data collection. Some information is available through sources such as the American Hospital Association (AHA), but it is cost prohibitive for an analysis with a nationwide scope. Unlike the costly AHA data, some data is available through government agencies, but the accuracy of this data is of concern. In this study, the variable representing hospital bed count was not included for this reason. In this case, a small sample of hospitals was selected for bed count analysis. No systematic errors in the data were identified, and thus the bed count data errors could not be addressed. In addition, rural hospitals are often too small and not required to publically share or sometimes report data, as is the case with the government’s Hospital Compare quality program. Finally, there is some support for the idea that additional hospital-specific data may be unnecessary to include in rural hospital closure studies. A recent study found that many of the social capital variables included in this study are related to a hospital’s financial state including population, families living below poverty, number of full-time employees, and nearby facilities. Overall, this analysis could be strengthened by including more hospital-specific variables, but the results are robust with the variables included.

Fourth, rural hospitals do receive support from their community. Hospital boards, for example, always include members that represent the community and its interests. Rural residents often volunteer in their local hospital, and some rural communities even financially support their local hospital through tax bonds. Data on this type of support rural

communities provide their local hospital would strengthen the study. However, given the national scope of this study, collecting this type of data was cost and time prohibitive.

Finally, variables commonly associated with social capital can conceptually be categorized into specific types of social capital – bridging, bonding, and linking. Some studies have utilized or developed a more refined social capital model that distinguishes types of social capital (bridging, bonding, linking) and distinguishes between resources and products of social capital.\(^{389,390}\) However, there is a high degree of variability in both conceptual application and measurement of social capital.\(^{391}\) With little agreement as to what measures should be included in a social capital model, one pioneering study recommends that social capital scales be tailored to the outcome under question because these items can vary within health services and policy.\(^{392}\) In fact, variables that represent different types of social capital may not be available for the study area.\(^{393}\) Furthermore, a recent study analyzing the association of social capital on urban, public hospital closure found that distinguishing social capital into types improves the model, but this may only be the case in the urban context.\(^{394}\) Ko discusses the role strong professional networks in sharing best practices or collaborative efforts to address community health concerns. But in rural areas, strong professional networks, particularly networks consisting of healthcare professionals, may not be possible. Rural areas are widely known for shortages of healthcare professionals. Hence, distinguishing social capital into types in the analysis may be most relevant in the urban context but its inclusion in future rural hospital closure analyses may be warranted.


3.5.2 CONCLUSION

Despite the limitations of this study, both analyses (Model A and Model B) have similar results with regards to the type of social capital variables or factors found to be statistically significant. Economic and crime variables or factors characterizing rural communities, are the most statistically significant at predicting the likelihood of rural hospital closure in a county. However, the odds ratios results from these parallel analyses reflect differences in the models. Model A reaffirms that social capital variables in the economic domain, such as unemployment, may be the most important social capital determinant in all rural counties. However, policies aimed at economic development in rural areas are tough to successfully implement and sustain. Model B proves that social capital has two significant factors or dimensions – economic vulnerability and crime risk. However each factor can be acted upon by policies that affect any social capital variable making up that factor. Regarding economic vulnerability, for example, policies that look to increase the high school graduate rate will positively influence economic vulnerability and thus social capital. The findings reported here support further analysis of the relationship between a rural community's social capital and rural hospital closure, in particular the predictive strength of particular social capital variables or factors.

In addition, political scientists and economic development organizations like the World Bank have long argued that social capital plays a significant, positive role in community building and sustainability. They have found that social capital is associated with a wide variety of positive social, economic, and healthcare outcomes, and consequently, interventions or programs that support social capital should be developed. Although international development organizations such as the World Bank and the International Monetary Fund have funded and reported positive results from social capital development projects, public funding for these type of interventions in the United States are less common. Further analysis is needed to determine the degree to which the healthcare system benefits from social capital building efforts. Social capital interventions may affect certain healthcare outcomes more strongly than hospital closure, or these programs may benefit non-healthcare outcomes more strongly than healthcare ones. This speculation deserves empirical attention, but if future studies continue to support these findings, it would be important for policymakers to understand how community social capital variables or factors affect healthcare. Since social capital is associated with many positive social,
economic, and healthcare outcomes, policies intervening and building social capital could help promote greater vitality of rural communities in the United States.
Chapter 4  IN RURAL COMMUNITIES, HOW DO SOCIAL CAPITAL TRENDS OVER TIME BETWEEN 1980 AND 2000 AFFECT THE LIKELIHOOD OF A HOSPITAL CLOSING?

4.1 ABSTRACT

OBJECTIVES. Having found a relationship between a rural community’s social capital and the likelihood of a rural hospital closing, this study will determine the social capital trends over time. Specifically, the relationship between social capital in 1980, 1990, and 2000 and the likelihood of a rural hospital closing between 2000 and 2007 will be analyzed.

METHODS. Social capital and rural hospital data for all rural counties in the United States (n = 2,052) were merged. The study includes decennial data from 1980, 1990, and 2000, which limits the number of social capital variables to include in the analysis (n=10). Model A is a multivariable logistic regression model predicting rural hospital closure by backward elimination of social capital variables. Model B combines factor analysis with logistic regression to measure the relationship between social capital and rural hospital closure.

RESULTS. In Model A, the odds ratios for unemployment and violent crime, two social capital components previously found to be significant in predicting the likelihood of rural hospital closure, were calculated for the years 1980, 1990, and 2000. In Model A, unemployment over time increasingly becomes a more significant predictor of the likelihood of a rural hospital closing between 2000 and 2008 (OR1980=1.35; OR1990=1.57; OR2000=2.13). Due in part to low variability over time, violent crime over time maintains the same association with rural hospital closure (OR1980=1.03; OR1990=1.01; OR2000=1.02). For Model B, factor analysis, on the 10 social capital variables, determined two social capital factors are significantly associated with rural hospital closure – economic
vulnerability (OR1980=1.25; OR1990=1.43; OR2000=1.42) and crime risk (OR1980=1.63; OR1990=1.39; OR2000=1.39).

**CONCLUSIONS.** Introducing a temporal element to the relationship between a rural community’s social capital and rural hospital closure warrants future analysis. Certain measures of social capital seem to be more temporally sensitive than others, such as unemployment compared to violent crime. Results support the hypotheses that a relationship between a rural community’s social capital and rural hospital closure exists and that social capital components at a time of rural hospital closure is more significant in predicting closure than components collected ten or 20 years earlier. The findings also lend support to policy makers needing to better understand and supporting social capital building initiatives.

### 4.2 Introduction

Social capital can be defined as the set of rules, norms, obligations, reciprocity, and trust embedded in social relations, social structures, and society’s institutional arrangements that enable members to achieve their individual and community objectives.\(^{395}\) Recognized as a latent variable, social capital is most often defined by an aggregation of variables representing demographic, housing, education, economic, crime, health, and community participation domains (Chapter 2).\(^{396}\) A rural community’s social capital measured at a single point in time may fail to accurately characterize or represent the community. For example, if a community experienced a major natural disaster, a political scandal, or layoffs at their largest employer, then the social capital variables reported immediately following the event may reflect a temporary shock instead of the true level of social capital. Furthermore, social capital is a community-level resource that takes time to develop and can fluctuate over time. The strengthening of social ties and cohesion with neighbors


requires building long-term relationships. Therefore, a single measure of social capital at a single point in time may not fully capture a rural community’s social capital.

Temporal factors do not singularly affect social capital. Time also plays a role in the dichotomous, dependent variable rural hospital closure. Although declines in hospital services are most pronounced in the last 2-3 years before closure, rural hospital closure is a result of long-term financial and operational challenges facing rural hospital administrators. In addition, rural hospital closure often increases healthcare accessibility problems and threatens the long-term sustainability of a rural community. Therefore the process of closing a rural hospital typically requires hospital administrators to justify the closure in a public forum and to notify the local, state, and federal governments about their intent to close.

Since time considerations affect both social capital and rural hospital closure, any policies or interventions implemented by rural communities or rural hospitals that target social capital measures must also consider time. With social capital measures reported at three separate time points (1980, 1990, and 2000), any sudden spike in social capital will not entirely skew the trend. By analyzing the impact a rural community’s social capital in 1980 or 1990 has on predicting the odds of rural hospital closure between 2000 and 2008, the long-term impact of social capital can be revealed. Thus, the goal for policymakers should be to design policies that will positively affect a rural community’s social capital long-term because any short-term gains may not affect the likelihood of rural hospital closure. Policies need to be sustainable in order to make a lasting impact on rural communities.

399 Ibid
4.3 Methodology

Building on previous findings, this study further investigates temporal factors of the relationship between a rural community’s social capital and rural hospital closure. In this study, social capital variables or factors in 1980, 1990, and 2000 are the independent variables, while rural hospital closure between 2000 and 2008 remains the dependent variable. Social capital data for this study originated from the robust, nationally-representative Research Triangle Institute’s (RTI) Spatial Impact Factor (SIF) database. It is maintained through collaboration between RTI and Arizona State University’s GeoDa Center for Geospatial Analysis and Computation and work is funded by National Institutes of Health grants. The SIF database provides an aggregation of county-level data collected and maintained by federal-level departments and agencies or university-based researchers.

A systematic literature review on social capital and its measurement was conducted to evaluate and identify which variables included in the SIF database had been previously used in social capital research, as explained in Chapter 2. Since three decennial census years were selected for this study, only 10 social capital variables, reported at the county scale, were selected for this analysis (Table 26). Five fewer social capital variables were included in this analysis than in the analysis in the previous chapter. First, in recognition of the impact social networks and relationships have on communities, the U.S. Census Bureau began collecting data on certain relevant social capital measures in 2000, such as the linguistic-isolation variable. Secondly, many other government agencies including the Federal Bureau of Investigation (FBI) began reporting additional data in 2000, like substance abuse crime. Finally, over time, agencies collecting data may make definitional changes to certain variables. For example, the FBI reclassified or clarified the types of crimes falling in the violent crime category, after the initial 1981 survey. Although violent crime has been kept in this analysis because the change is expected to be insignificant, certain social capital variables in 2000 were lost for prior years of analysis. The SIF database provides nearly complete data for all 2,052 rural counties. The variables with the maximum missing data are population change and unemployment (1.8% missing data). Due to the completeness of the dataset, no adjustments will be made for the missing data. All rural counties were included in the analysis. Social capital data was included for the years 1980, 1990, and 2000. Data from 2010 was considered for inclusion, but significant changes in Census
collection and estimation occurred between 2000 and 2010, making inclusion of 2010 data infeasible.\textsuperscript{404}

Despite the decrease in social capital variables to include in the model, the dependent variable remains a dichotomous rural hospital closure variable (closed = 0 or open, operational = 1). In the initial analyses, rural hospital closure will include all rural counties in the United States (n=2,052). Rural hospital data was purchased from Health Forum, the data repository enterprise for the American Hospital Association (AHA) and merged with the CMS Provider File to verify and ensure accuracy. From the initial AHA dataset of all U.S. hospitals (n=5,535), only 2,126 rural, short-term, acute-care hospitals were identified. In the United States, 70 of the 2,126 rural hospitals between 2000 and 2007 (3.3\%) closed. With so few cases of hospital closure, a dichotomous \textit{CLOSE} variable was created, with 1 indicating that a particular county did experience at least one hospital closure between 2000 and 2008. Rural hospital closure is often studied over a ten year period, but the rural hospital closure time period was cut off in 2008 to prevent introducing the impact of the Great Recession of 2008-2009. Although researchers acknowledge that using an urban-urban continuum reflects a deeper understanding of rural heterogeneity, the low rural hospital closure rate prevents stratification of results by the degree of rurality.\textsuperscript{405,406} Although degree of rurality was not be considered, sensitivity analysis was run to determine if any significant differences exist between rural counties with at least one hospital (n=1,614) and rural counties with no hospital (n=438).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Obs</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>1</td>
<td>Single Parent Family Households (%)</td>
<td>2043</td>
<td>0.06</td>
<td>0.03</td>
</tr>
<tr>
<td>2</td>
<td>Diversity Score†</td>
<td>2043</td>
<td>0.10</td>
<td>0.13</td>
</tr>
<tr>
<td>3</td>
<td>High School Graduates (%)†</td>
<td>2043</td>
<td>0.35</td>
<td>0.08</td>
</tr>
<tr>
<td>4</td>
<td>Gini Coefficient</td>
<td>2045</td>
<td>0.38</td>
<td>0.04</td>
</tr>
<tr>
<td>5</td>
<td>Unemployment (%)</td>
<td>2015</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>6</td>
<td>Per Capita Income ($)†</td>
<td>2015</td>
<td>7710.84</td>
<td>1797.09</td>
</tr>
<tr>
<td>7</td>
<td>Population in Poverty (%)</td>
<td>2021</td>
<td>0.17</td>
<td>0.07</td>
</tr>
<tr>
<td>8</td>
<td>Violent Crimes (#)*</td>
<td>2049</td>
<td>42.37</td>
<td>68.45</td>
</tr>
<tr>
<td>9</td>
<td>Property Crimes (#)*</td>
<td>2049</td>
<td>582.75</td>
<td>781.50</td>
</tr>
<tr>
<td>10</td>
<td>Presidential Voter Turnout (%)†</td>
<td>2043</td>
<td>0.59</td>
<td>0.12</td>
</tr>
</tbody>
</table>

*These variables have been normalized, so as to allow factor analysis. The normalization technique used was standardization (z-scores) by subtracting the mean from the variable and then dividing it by the standard deviation. Without normalization, indicators with extreme values have a greater effect on the composite indicator.

†Although this is titled a “Social Capital Index,” it was developed by Penn State University and only measures the degree of community participation.

†As these variables increase, social capital also is expected to increase.
4.3.1 **Objective**

This study focuses on understanding temporal factors of the relationship between a rural community’s social capital and rural hospital closure. First, a rural community’s social capital will be analyzed at three separate time points (1980, 1990, 2000). Next, this study will run two parallel models. In the first, the analysis is based on prior research findings (Chapter 3) where, unemployment, violent crime and substance abuse were found significant to predict rural hospital closure. However, this model is less refined because one of the significant variables, substance abuse crime, is not available. The model tests the relationship between unemployment and violent crime in 1980, 1990, and 2000, and rural hospital closure using logistic regression. Substance abuse was found to be significant in previous findings, but it was not reported prior to 2000. Model B acknowledges the interrelationship between social capital variables and consequently performs factor analysis to determine significant social capital factors before running the same logistic regression.

Since confounding variables and interaction terms were not found to significantly affect the association between a rural community's social capital and rural hospital closure, this analysis does not check for confounding or interaction. However, the robustness of the Model A and Model B results were tested in sensitivity analysis, using the subset of rural counties with at least one hospital (n=1,614). In the end, the main goals of this study are two-fold. Initially, it will determine if the same social capital factors predicts the likelihood of rural hospital closure, regardless of the year. Then, the strength of association between social capital and the likelihood of rural hospital closure over time will be analyzed. All analyses were performed with STATA statistical software. A value of P<0.05 was considered significant.

4.3.2 **Model A: Logistic Regression Using Social Capital Indicators**

Using stepwise logistic regression, a previous analysis found that unemployment, substance abuse crime, and violent crime were the three social capital variables significantly

---


associated with rural hospital closure. The removal of predictors one at a time from the model was an iterative process based on statistical criteria of significance (P<0.05). For this analysis, ten social capital variables were included in the model. Unfortunately, substance abuse crime, which was found to be a significant social capital variable in Chapter 3, is not among the ten variables included in this model because it is unavailable prior to 2000. Logistic regression will calculate the odds ratios for unemployment and violent crime in 1980, 1990, and 2000, to determine their association with rural hospital closure occurring between 2000 and 2008.

4.3.3 **MODEL B: LOGISTIC REGRESSION USING SOCIAL CAPITAL FACTORS**

Model B requires a two-level analysis. First, factor analysis will be used to consider and distill ten social capital variables and their interrelations into significant factors. The goals of FA are to summarize patterns of correlation among observed variables, to reduce a large number of observed variables to a smaller number of factors, to provide an operational definition for an underlying process by using observed variables, and to test a theory about the nature of underlying processes.\(^\text{409}\) Since the assumption of normality is required in FA, the following three social capital variables will be normalized in preparation for factor analysis: (1) per capita income, (2) violent crime, and (3) property crime. After determining the factors, logistic regression will be used to determine if these social capital factors can predict the likelihood of rural hospital closure.

4.4 **RESULTS AND INTERPRETATION**

This study runs two parallel models. The initial or crude Model A uses logistic regression to determine how unemployment and violent crime are significant in predicting rural hospital closure. In the second approach, Model B first runs factor analysis to determine significant social capital factors before testing the association between these factors and rural hospital closure with a logistic regression analysis. The appropriateness of both Model A and Model B was assessed by the Hosmer-Lemeshow Goodness of Fit (GOF) statistic for significance (P>0.05).\(^\text{410,411}\) In this study, the Hosmer-Lemeshow GOF statistic was calculated for every


4.4.1 Model A: Logistic Regression Using Social Capital Indicators

The goal of the analysis is to predict rural hospital closure using the most parsimonious model, the one with the least predictors. A model only needs to include statistically significant predictors; the fewer predictors included in a model facilitates interpretation. Model A uses previous findings to identify unemployment and violent crime as two social capital variables with the most influence on rural hospital closure. The association between these social capital variables and the dichotomous rural hospital closure variable is tested with logistic regression. Results are reported as odds ratios, which estimates the odds of rural hospital closure given a certain exposure level of an independent variable holding all else constant. The 95% confidence interval (CI) around the odds ratio (OR) was obtained with the formula ($\beta \pm 1.96 \text{SE}$), in which SE is the standard error of $\beta$. The odds ratio estimates from Model A are shown in Table 27.

<table>
<thead>
<tr>
<th></th>
<th>Crude OR</th>
<th>95% CI</th>
<th>Adjusted OR</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment, 1980</td>
<td>1.06</td>
<td>0.99</td>
<td>1.14</td>
<td></td>
<td>0.07</td>
</tr>
<tr>
<td>Unemployment, 1990</td>
<td>1.09</td>
<td>1.03</td>
<td>1.17</td>
<td>1.57</td>
<td>1.15</td>
</tr>
<tr>
<td>Unemployment, 2000</td>
<td>1.16</td>
<td>1.03</td>
<td>1.31</td>
<td>2.13</td>
<td>1.16</td>
</tr>
<tr>
<td>Violent Crime, 1981</td>
<td>1.01</td>
<td>1.00</td>
<td>1.01</td>
<td>1.03</td>
<td>1.02</td>
</tr>
<tr>
<td>Violent Crime, 1990</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>Violent Crime, 2000</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.02</td>
<td>1.01</td>
</tr>
</tbody>
</table>

Since economic variables are the most significant predictors of social capital, the results do suggest an association between unemployment and rural hospital closure and weak or no association between violent crime and rural hospital closure. In the crude analysis, for every 1 percent increase in unemployment, the odds of a rural hospital closing increased by 1.06 times in 1980, by 1.09 times in 1990, and 1.16 times in 2000, holding all else constant. However, unemployment often increases by more than 1 percent in rural communities, as the 2008 – 2009 recession proved, if the odds ratio is recalculated for a 5 percent absolute
increase in unemployment, the adjusted odds ratios are 1.35 (β=0.06) in 1980, 1.57
(β=0.09) in 1990, and 2.13 (β=0.15) in 2000.⁴¹² These findings show that unemployment in
2000 is more strongly associated with the likelihood of rural hospital closure than
unemployment in either 1990 or 1980 (See Figure 2). In 2000, a 5 percent absolute increase
in unemployment doubles the odds of a rural hospital closure, which mirrors previous
findings (Chapter 3).

In the case of violent crimes (VC), the odds ratios are almost equal to 1 in 1981, 1990, and
2000. This means that for every additional violent crime, the odds of a rural hospital closing
remains unchanged. Crime data is collected by the FBI through the Uniform Crime Reports
(UCR) program and provides raw counts of reported crimes.⁴¹³ Like unemployment, if the
number of crimes changes in a community, it typically does not increase by just one crime.
Violent crimes are the most severe, including murder, forcible rape, and robberies, and they
are classified as Part I offenses by the FBI. Since the average number of violent crimes per
rural county is 57.8 in 2000, a conservative 10% increase in crimes is a total of 5 violent
crimes. If the OR_{VC} is recalculated for an increase in 5 violent crimes, the adjusted odds
ratios are 1.03 (β=0.01) in 1981, 1.01 (β=0.00) in 1990, and 1.02 (β=0.00) in 2000. This
means that for an increase in 5 violent crimes, the odds of a rural hospital closure is
increased by 1.02 times in 2000 holding all else constant. As the adjusted ORs show, violent
crimes have little to no effect on rural hospital closure and its association with rural
hospital closure remains steady over time (See Figure 2).

⁴¹² Hertz, T, L. Kusmin, A. Marré, and T. Parker. (Aug. 2014). Rural Employment Trends in Recession and
Although unemployment and violent crime have been found to be significant social capital variables, unemployment in 1980, 1990, and 2000, best predicts the association between a rural community's social capital and the likelihood of a rural hospital closing between 2000 and 2008. The association between unemployment and the likelihood of rural hospital closure becomes stronger as a rural community’s social capital approaches years hospital closure was analyzed (2000-2008).

4.4.2 **MODEL B: Logistic Regression Using Social Capital Factors**

As mentioned earlier, the second analytical approach, identified as Model B, first performs factor analysis to determine social capital factors, before running the logistic regression model. Factor analysis (FA) is a data reduction, statistical technique that summarizes a set of individual indicators while preserving the maximum possible proportion of the total variation in the original data set. FA is an iterative process that is repeated until no more variables need to be removed from the analysis and the number of factors is determined. All ten variables met the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy threshold (KMO>0.60). Next, factors were extracted from this analysis and orthogonally rotated with the promax method. Results for two to four factors were examined. Ultimately, two factors were extracted. The selection of two factors was based on examination of

---

eigenvalues, scree plots, and the interpretability of the solution.\textsuperscript{415} The estimated factor scores were computed using a regression method based on ten social capital variables. All factor loading $\geq 0.32$ were presented.

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>---------</td>
</tr>
<tr>
<td>Promax Rotation</td>
</tr>
<tr>
<td>Proportion</td>
</tr>
<tr>
<td>Cumulative</td>
</tr>
</tbody>
</table>

A factor with an eigenvalue greater than 1.0 is accounting for a greater amount of variance than was contributed by one variable.\textsuperscript{416} Therefore, these factors are accounting for a meaningful amount of variance and merits inclusion in the analysis. The factor analysis results mirror the findings from Chapter 3. Factor 1 is best classified as economic vulnerability and factor 2 as crime risk. The third factor in this model has an eigenvalue less than 1.0, which accounts for less variance than had been contributed by one variable, and it is eliminated from the study. This commonly used rule is known as the latent-root or Kaiser criterion. As Table 28 shows, the first factor, economic vulnerability, explains the maximum variance in the data (59.7\% in 1980, 65.3\% in 1990, and 64.5\% in 2000). Crime risk, the second factor, explains the second greater amount of variance in the data. The inclusion of both factor 1 and factor 2 in the analysis signifies that 91.8 percent of the (cumulative) variance in the data is explained in 1980, 93.9 percent in 1990, and 93.5 percent in 2000.


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Promax Rotation</td>
<td>Promax Rotation</td>
<td>Promax Rotation</td>
</tr>
<tr>
<td>Single Parent Family Households, 2000</td>
<td>0.674</td>
<td>0.438</td>
<td>0.706</td>
</tr>
<tr>
<td>Diversity Score, 2000†</td>
<td>0.703</td>
<td>0.745</td>
<td>0.651</td>
</tr>
<tr>
<td>High School Graduates, 2000†</td>
<td>-0.737</td>
<td>0.893</td>
<td>-0.502</td>
</tr>
<tr>
<td>Gini Coefficient, 2000</td>
<td>0.769</td>
<td>0.759</td>
<td>0.818</td>
</tr>
<tr>
<td>Unemployment, 2000</td>
<td>0.407</td>
<td>0.541</td>
<td>0.833</td>
</tr>
<tr>
<td>Per Capita Income, 2000**</td>
<td>-0.665</td>
<td>0.822</td>
<td>-0.701</td>
</tr>
<tr>
<td>Population in Poverty, 2000</td>
<td>0.926</td>
<td>0.733</td>
<td>0.949</td>
</tr>
<tr>
<td>Violent Crimes, 2000*</td>
<td>0.818</td>
<td>0.624</td>
<td>0.850</td>
</tr>
<tr>
<td>Property Crimes, 2000*</td>
<td>0.824</td>
<td>0.602</td>
<td>0.867</td>
</tr>
<tr>
<td>Presidential Voter Turnout, 2000†</td>
<td>-0.328</td>
<td>-0.460</td>
<td>0.791</td>
</tr>
<tr>
<td>Overall Kaiser-Meyer-Olkin Measure</td>
<td>0.7295</td>
<td>0.755</td>
<td>0.7179</td>
</tr>
</tbody>
</table>

*These variables were normalized prior to FA to fulfill normality assumption.
†As these variables increase, social capital also is expected to increase.
In the final factor extraction, ten social capital variables were reduced into two factors – economic vulnerability and crime risk. As the factor loadings show (Table 29), economic vulnerability in 1980, 1990, and 2000 is composed of the following six variables: (1) single parent households (%), (2) diversity score (%), (3) Gini coefficient, (4) per capital income, (5) population in poverty (%), and (6) presidential voter turnout (%). Economic vulnerability in 1980 and 1990 also includes high school graduates (%), and this factor includes unemployment in 1990 and 2000. The slight differences in economic vulnerability over time is not concerning because only the factor loadings $\geq |0.32|$ were presented. For example, the factor loading for high school graduates (%) is close to $|0.32|$ in 2000 (-0.26).

Crime risk in 1980, 1990, and 2000 is composed of violent crimes and property crimes and can also include single parent family households (%) in 1980 and presidential voter turnout (%) in 1980 and 1990. By and large, FA derives factors from the data itself that are independent and that measure a different, unique dimension of social capital. For the purpose of the analysis, the promax method of orthogonal rotation was selected due to clarity and independence of each factor, which facilitates interpretability.

The final factor extraction also yields a Kaiser-Meyer-Olkin (KMO) value close to 0.80 ($KMO_{1980}=0.73$; $KMO_{1990}=0.76$; $KMO_{2000}=0.72$). This confirms that these two factors – economic vulnerability and crime risk – are good measures or dimensions of social capital.

The FA process has now created two factors that measure social capital for each county. To complete the initial, crude Model B, a logistic regression analysis tests the relationship between these two social capital factors and the likelihood of rural hospital closure.

In Model B, these two social capital factors, previously identified, run through a multivariate logistic regression model, where the dichotomous outcome variable is rural hospital closure. Since these social capital factors are composed of ten social capital variables and more significantly represent the social capital construct, it is not surprising that the odds ratios (OR) reflect a stronger association between social capital factors and rural hospital closure than was found in Model A (Table 27). Despite differences in the odds ratios, it was expected both factors – economic vulnerability and crime risk – align with previous findings (Chapter 3).

<table>
<thead>
<tr>
<th>Economic Vulnerability, 1980 (f24)</th>
<th>Crude OR</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic Vulnerability, 1990 (f22)</td>
<td>1.43</td>
<td>1.13</td>
<td>1.81</td>
</tr>
<tr>
<td>Economic Vulnerability, 2000 (f20)</td>
<td>1.42</td>
<td>1.12</td>
<td>1.81</td>
</tr>
<tr>
<td>Crime Risk, 1981 (f25)</td>
<td>1.63</td>
<td>1.35</td>
<td>1.98</td>
</tr>
<tr>
<td>Crime Risk, 1990 (f23)</td>
<td>1.39</td>
<td>1.16</td>
<td>1.65</td>
</tr>
<tr>
<td>Crime Risk, 2000 (f21)</td>
<td>1.39</td>
<td>1.16</td>
<td>1.67</td>
</tr>
</tbody>
</table>

The Model B odds ratio estimations, along with the calculated confidence intervals, suggest a strong association between rural hospital closure and both economic vulnerability and crime risk (OR>1 and CI does not contain 1). Since factors have no units, for every one unit increase in economic vulnerability, the odds of a rural hospital closing increased by 1.25 times (1980), 1.43 times (1990), and 1.42 times (2000), holding all else constant. Regarding crime risk, for every one unit increase in crime risk, the odds of a rural hospital closing increased by 1.63 times (1980), 1.39 times (1990), and 1.39 times (2000), holding all else constant.

FIGURE 3: ECONOMIC VULNERABILITY & CRIME RISK OVER TIME

Although crime risk in 1980 seems to be the most significant social capital factor in predicting the likelihood of rural hospital closure between 2000 and 2008, this result is
likely a result of sampling errors. The FBI Uniform Crime Reporting system for crime began in 1981. As discussed in Chapter 2, there are slight differences in the manner crime was reported in 1981 compared to the standard reporting methods used in both 1990 and 2000. Therefore, the odds ratios reported for economic vulnerability in 1980, 1990, and 2000, along with crime risk in 1990 and 2000 will be the ones used for interpretation. First, the odds ratio results indicate that the association between economic vulnerability and the likelihood of rural hospital closure becomes stronger as the rural community’s social capital approaches years hospital closure was analyzed (2000-2008). The 1980 economic vulnerability odds ratio is lower (1.25) than the odds ratio in 1990 (1.43) and 2000 (1.42). Secondly, the 1990 and 2000 odds ratios show that economic vulnerability (1.43 & 1.42) and crime risk (1.39 & 1.39) are almost equally as strong factors in predicting the likelihood of rural hospital closure. This result also indicates that a rural community’s social capital in 1990 is just as significant in its association with rural hospital closure between 2000 and 2008, as social capital in 2000.

4.4.3 Sensitivity Analysis

Although the sensitivity analysis run previously (Chapter 3) found no significant differences between the odds ratios calculated with the entire rural county dataset (n=2,052 counties) and the partial rural county dataset (n=1,614 counties), this exact analysis was repeated for Model A and Model B in this chapter. The purpose of sensitivity analysis is “to assess whether altering any of the assumptions made leads to different final interpretations or conclusions.” The goal is to assess the robustness or consistency of the results. Although previous studies suggest differences could exist between rural counties with at least one hospital versus those without a hospital, the results remains unchanged. Thus, the odds ratio results and established models are robust.

4.5 Discussion

The findings from this temporal analysis of the relationship between a rural community’s social capital and the likelihood for rural hospital closure support the conclusions from Chapter 3. Both Model A and Model B ultimately include economic and crime variables or factors in the logistic regression. This similarity reflects the strength of the parallel analysis

study design commonly used in social capital research. Next, the odds ratios in Model A give additional support to the conclusion that economic variables, like unemployment (OR_{1980} = 1.35; OR_{1990} = 1.57; OR_{2000} = 2.13), are the most significant predictors of social capital.\[418,419,420,421,422,423\] Therefore, maintaining or positively impacting economic outcomes like unemployment is important for rural communities and rural hospitals. This implies that rural communities need to adopt or secure support for policies that directly impact economic outcomes such as diversifying their economy, attracting employers and corporate investment, and improving the quality of their work force.\[424\] As stated previously, these economic goals are challenging for rural communities to achieve, and increasingly rural economic development programs are less prioritized by the state and federal governments.\[425\]

The trend in the unemployment odds ratios over time show that even if a rural community is successful at improving economic measures today, their past economic outcomes still influence the likelihood for a rural hospital to close. Therefore, a rural community would need to maintain any positive economic results over a longer period of time to truly affect the likelihood of a rural hospital closure. This can be extremely challenging in rural communities, who have far more specialized economies than urban areas. In addition, industries rural communities are often dependent on such as agriculture, mining, and low-skilled manufacturing; and they are sensitive to external forces outside of the control of


\[425\] Doering, Christopher. "As more move to the city, does rural America still matter?" USA Today 13 Jan. 2013.
rural communities such as international trade, weather, and government subsidies.\textsuperscript{426} This result implies that rural community leaders should prioritize policies or interventions that are more likely to create long-term, sustainable development. This recommendation is shared by a wide variety of economists and international organizations like the United Nations and World Bank, who recognize that past economic growth has come at the expense of marginalized communities and the environment. In fact, the World Bank Group is vocal on sustainable development, which incorporates economic growth, environmental stewardship, and social inclusion, being reflected in its mission statement and all of its policies.\textsuperscript{427}

Model A results provide a strong illustration of the importance of economic outcomes on social capital. Relevant policy implications include prioritizing interventions that more directly affect economic outcomes and trying to ensure these interventions are sustainable over time. However, Model B, with its use of factor analysis, more fully captures social capital and the interrelationships between social capital variables. Across all three years of data, 1980, 1990, and 2000, the resulting factors are almost identical in their construct and mirror the factor analysis results using the full Adaniya social capital model in 2000 (Chapter 3). The consistency of the results supports the strength of these two factors – economic vulnerability and crime risk – in defining social capital. In addition, each of these factors is constructed from many social capital variables. This implies that policies attempting to intervene on any one of these social capital variables would be expected to have an impact on the rest of the variables, especially significant ones within the same factor. In the case of economic vulnerability, improving high school graduation rates or better supporting single parent households would positively impact this factor and a rural community’s overall social capital. Previous studies have shown that building and strengthening relationships and social networks within rural communities improve high school graduation rates.\textsuperscript{428} There are far more policies or interventions that a rural


community can adopt to target building and strengthening relationships and social networks than ones solely focused on decreasing unemployment.\textsuperscript{429}

Policies that intervene on social capital components can be successfully implemented in rural communities. One successful intervention, known as the Perry Preschool Program, increased the number of high school graduates by 19 for every 100 students in enrolled in the program.\textsuperscript{430} Although initially tested in an urban setting, the program enrolls children (<5 years old) in preschool with an 8:1 child-to-teacher ratio, requires teachers to schedule weekly home visits with the parents of each child, and encourages parental involvement in the school. According to the National Education Association, this programs has been successfully implemented in both urban and rural settings.\textsuperscript{431} Forty years later, the children, who completed the program, were more likely to have graduated from high school, enjoy higher earnings, were more likely to be employed, and had committed fewer crimes.

Another intervention that targets high school students is called Check & Connect, a program designed by the Institute on Community Integration at the University of Minnesota. This program assigns ninth grade students with identified learning, emotional, and/or behavioral disabilities with a “monitor” (e.g. a community member with experience in human services, a graduate student) who works with them year-round as a mentor, advisor, and service coordinator. High risk students who participate in Check & Connect, compared to high risk student not enrolled in a program, are more likely to graduate on time (68\% versus 29\%).\textsuperscript{432} Teachers rated 87 percent of parents with Check & Connect students to be supportive and engaged.\textsuperscript{433} Since policies solely focused on economic development are difficult to successfully implement, these results are encouraging to rural communities. The 1990 and 2000 odds ratios for economic vulnerability and crime risk, for example, are almost equal. This means that policies targeting either factor or any variable within a factor


could increase a rural community’s social capital. Therefore, instead of solely adopting and implementing economic development policies, rural communities can design programs that increase social capital overall. Social capital building policies often focus on nurturing relationships and social networks to increase trust and reciprocity in a community.\textsuperscript{434} Simple actions like organizing a block party, setting up study circles among students, or encouraging street conversations can build social capital, which in turn can improve a wide range of community outcomes including economic ones.\textsuperscript{435} If enough of these social capital building efforts are implemented and maintained in a community, it establishes a social norm of community engagement.\textsuperscript{436} One recent pilot demonstration study, called Save Our Sisters (SOS), implemented in rural North Carolina had a goal of building and nurturing social networks among African-American women to increase cancer screening. This intervention found that rural residents have a willingness and capacity to support their communities. Ultimately this SOS program did not only increase cancer screenings in the community, but it engaged community-based organizations (e.g. church, civic, and social groups), provided a setting for the women to receive support, and motivated the women to establish additional programs to sustain the successes achieve by SOS.\textsuperscript{437} One of these additional programs is called Adopt-A-Sister that assists black women who cannot afford the cost of cancer screening. Perhaps the most important SOS program success has been its sustainability even after the funding period.\textsuperscript{438} In this environment, community engagement and social entrepreneurship can bring together groups within rural communities, which in turn attracts and retains young adults, physicians, and current residents of rural communities.\textsuperscript{439,440} And this type of engaged rural community is less likely to allow their local hospital to close.

\textsuperscript{438} Ibid
Preventing rural hospital closure is an important healthcare outcome rural communities aspire to. To meet this goal, these findings have shown rural communities can design and implement a wide-variety of interventions that target various social capital components. Although rural communities have many successful social capital or community building policies to copy from, rural community leaders must take care to learn from policies that have previously worked in other rural communities. Rural areas are different from urban ones. Leaders cannot simply rescale programs from urban areas and expect them to be successful in rural areas. Rural areas, for example, have social capital strengths to capitalize on by specific programs including “dense social networks, social ties of long duration, shared life experiences, high quality of life, and norms of neighborliness, self-help, and reciprocity.”

Rural communities also face daunting challenges for policy implementation including low population density, transportation issues, lack of access to grant writing, lower public funding levels, difficulties in recruiting staff, and potential fragmentation of scarce resources. Consequently, rural interventions must be tailored to the rural setting to be more likely to succeed.

In addition to the findings of this study and the policy implications, this study has several additional strengths. The study includes every rural county in the United States. Rural hospital closure remains a relevant research interest in rural health. This study is the first rural hospital closure study to include the years 2000 through 2008, and it is one of the first to analyze the association between social capital and health services research. This study is only the second to associate social capital with hospital closure, but it is the first to study the link between a community’s social capital and rural hospital closure. Finally, this analysis addresses a previous limitation of the overall study – time. The analysis presented in this chapter represents a useful addition to the growing body of social capital, rural health, and rural hospital closure literature. However, this study still faced many limitations.

4.5.1 Limitations

First, social capital data from three decades are included in this study (1980, 1990, 2000). Unfortunately, there are many external, environmental factors that could affect social capital variables that were not controlled for in the analysis. For example, between 1980 and 2000, the National Bureau of Economic Research reports three minor recessions. Although these minor recession may not affect social capital too much in rural communities, there was a well-documented farm crisis in the 1980s that did have a profound effect on agriculturally-dependent rural communities. The crisis was caused by domestic financial policies, weather, and a trade embargo prohibiting the shipment of grain to the Soviet Union. The result was a net farm income decline of 30 percent and a land value reduction of 50 percent, which caused one-in-three commercial-sized farms to be unable to pay their bills. Yet, social capital variables most likely to be affected by the farm crisis, like population in poverty (%), remain steady between 1980 and 2000. Perhaps the effect of the farm crisis is limited to agriculturally-dependent counties, and since this study includes the entire United States, the effect is minimized.

Although the inclusion of 2010 was considered for this analysis, significant changes in the U.S. Census occurred in the manner in which the Census was administered and the Census Bureau estimates or projects certain variables. For example, in 2010, the Census Bureau no longer used the long form. The most important difference in 2010, which impacts data reported for rural areas most, is the change to the Census Bureau's American Community Survey. Prior to 2010, community data was collected every 10 years; now, a small percentage of the population will receive the survey every three or five years on a rotating basis. Thus, community data reported each year will be based on estimations or projections. Due to these changes, including the elimination of certain questions previously tested in 2000, social capital data in 2010 was not included in this study.

Furthermore, this study is limited in the time period it can include in the analysis of years prior to 1980. The number of social capital variables included in this analysis was already reduced to ten (from 15) due to problems with data availability and definitional changes. These problems only increase when looking at data prior to 1980. One major change, for

example, was the separation of Hispanics from race, in order to differentiate between white and black Hispanics. In order to perform a well-designed trend analysis, more than three years of data is recommended. Due to changes in the surveys used to gather social capital data, the ideal study design would collect social capital data for rural communities each year over a ten year period. However, rural areas are increasingly ignored in policy. The changes in the American Community Survey, for example, means that a small number of rural American will now be surveyed every five years, compared to annual administration of the survey in urban areas. All these factors make trends in rural America harder to ascertain. Trend analyses in rural areas would almost necessitate the administration of a survey in-person through field work.

Secondly, the county was used as the geographic unit of analysis because insufficient data is available at smaller scales of analysis. This assumes the county is the best scale to represent a rural hospital’s market, rural hospitals are centrally located within a county, and equally distribution of the population. However, this can bias the results. Consider, for example, a rural hospital located on the border between two or more rural counties; if only one county’s social capital can be linked to that hospital, which county best represents that hospital’s market?

Third, many hospital-specific characteristics, such as a hospital ownership, executive leadership effectiveness, market competitiveness, and fiscal stains, are linked to hospital closure. However, many of these measures were not available for inclusion in this study; much of this information would require in-person, qualitative data collection. Some information is available through sources like the American Hospital Association (AHA), but it is cost prohibitive for an analysis with a nationwide scope. Unlike the costly AHA, some data is available through government agencies, but the accuracy of this data is concerning. This is why the variable for hospital bed count was no included. In this case, a small sample of hospitals were selected for bed count analysis. Based on the findings, there are multiple types of errors in data collection, and therefore, the data cannot be corrected. In addition, rural hospitals are often too small and not required to publically share or sometimes report data, as is the case with the government’s Hospital Compare quality program. Finally, there is some support that additional hospital-specific data may be unnecessary to include in

---

rural hospital closure studies. A recent study found that many of the social capital variables included in this study are related to a hospital’s financial state including population, families living below poverty, number of full-time employees, and nearby facilities. Overall, this analysis could be strengthened by including more hospital-specific variables, but the results are robust with the variables included.

Fourth, determining direct local support for a rural hospital through civic engagement or even financially in the form of a tax bond would strengthen the study. However, given the national scope of this study, collecting this type of data was cost and time prohibitive.

Finally, variables commonly associated with social capital can conceptually be categorized into specific types of social capital – bridging, bonding, and linking. Some studies have utilized or developed a more refined social capital model that distinguishes types of social capital (bridging, bonding, linking) and distinguishes between resources and products of social capital. However, there is a high degree of variability in both conceptual application and measurement of social capital. With little agreement as to what measures should be included in a social capital model, one pioneering study recommends that social capital scales be tailored to the outcome under question because these items can vary within health services and policy. In fact, variables that represent different types of social capital may not be available for the study area. Furthermore, a recent study analyzing the association of social capital on urban, public hospital closure found that distinguishing social capital into types improves the model, but this may only be the case in the urban context. Ko et al. discuss the role strong professional networks in sharing best practices or collaborative efforts to address community health concerns. But in rural areas, strong...
professional networks, particularly networks consisting of healthcare professionals, may not be possible. Rural areas are widely known for shortages of healthcare professionals. Hence, distinguishing social capital into types in the analysis may be most relevant in the urban context but its inclusion in future rural hospital closure analyses may be warranted.

4.5.2 **CONCLUSION**

Despite the limitations to this study, both analyses (Model A and Model B) find that rural community’s economic and crime variables or factors are the most significant at predicting the likelihood of rural hospital closure in the United States in 1980, 1990, and 2000. The Model A odds ratios also support the finding that economic outcomes, like unemployment, make the strongest impact on overall social capital. Furthermore, the association between unemployment and the likelihood of rural hospital closure between 2000 and 2008, decreases the further back in time the data was collected. However, the association remains strong enough to show that in order to decrease the likelihood of rural hospital closure, rural communities need to adopt and implement sustainable development policies. Only sustainable positive economic outcomes can ultimately decrease the likelihood of rural hospital closure.

Model B provided robust and compelling findings that proved economic vulnerability and crime risk remain the strongest factors in defining social capital regardless of the year of analysis. The results support the position political scientists and organizations like the UN and World Bank have argued that rural communities can implement policies that focus on social capital, community building efforts. These types of efforts succeed through local leadership and are more likely to be sustainable. They also have a ripple effect in the community and improve outcomes in a wide variety of other social, economic, and healthcare outcomes. Higher social capital has already been found to lower mortality, risky behaviors, violent crime, and suicide rates. It is also linked to improved self-rated

---


Therefore, social capital building efforts have a great potential to impact overall health of rural communities, as well as preventing rural hospital closure. Given the potential positive impact of social capital building policies in rural communities in the United States, as well as the successful implementation of social capital building efforts internationally by the World Bank, future research is warranted to better understand how a rural community's social capital impacts the healthcare sector.
Chapter 5   **DO CHANGES TO GEOGRAPHIC SCALE OF ANALYSIS AFFECT THE ASSOCIATION BETWEEN SOCIAL CAPITAL AND RURAL HOSPITAL CLOSURE**

5.1 **ABSTRACT**

*OBJECTIVES:* Using a geographic information system, this study creates a new geographic scale of analysis, called a Proximal Hospital Area (PHA), to define the rural community served by a specific rural hospital. The three statistically significant social capital variables in the Adaniya social capital model (unemployment, substance abuse crime, and violent crime) were recalculated at the PHA scale to determine if a change in geographic scale would impact the odds of rural hospital closure.

*METHODS:* To partially address geographic analytical challenges, the purpose of this study is to change the geographic scale from counties (n=2,052) to one in which each rural hospital has its own unique Proximal Hospital Area. A PHA is created by georeferencing each rural hospital using a geographic information system (GIS). Next, a circular buffer with a 30-mile radius was drawn around each rural hospital. For every rural county included in each PHA, the percent land area of the county within the PHA was calculated. The unemployment, substance abuse per 1,000, and violent crime per 1,000 rates were calculated one-by-one for each Proximal Hospital Area. This one-by-one process was time intensive. Thus, the PHA social capital recalculation was limited to the South U.S. Census region, which accounts for 39.5 percent of all rural hospitals (n=840) in the United States and has the highest rural hospital closure rate in the United States (4.9 percent compared to the US average of 3.3 percent). The sparsity of rural hospital closure data is least pronounced in the South region. Finally, the impact of the new geographic scale on the association between a rural community's social capital, and the odds of rural hospital closure is estimated.
**RESULTS:** Social capital measures at the PHA scale saw little change from social capital at the county scale. The unemployment rate at the PHA scale is 1.89 percent lower than the rate at the county scale. With regards to crime, the substance abuse crime rate decreased by 3.94 percent at the PHA scale, while the violent crime rate increased by 3.13 percent at the PHA versus county scale. The change in the descriptive statistics is less than 5 percent, which implies that the odds ratios from a logistic regression at the PHA scale falls within the 95 percent confidence interval of the crude Model A results (Chapter 3). The crude odds ratio for unemployment (OR=1.15; CI 1.02-1.31; p=0.03) is the social capital variable most strongly associated to the odds of rural hospital closure. Substance abuse crime (OR=1.00047) and violent crime (OR= 1.002266) odds ratios have no effect on the odds of rural hospital closure.

**CONCLUSIONS:** Our findings indicate that the association between rural community’s social capital and rural hospital closure would not be sensitive to geographic scale. A Proximal Hospital Area may be more representative of the actual community a rural hospital serves, but if the majority of rural hospitals are close to the centroid of a non-oblong shaped county, then the PHA scale is not superior to the county scale in this analysis. The national scope of the overall study, the sparsity of rural hospital closure, and the study limitations may fail to capture subtle differences indicative of rural heterogeneity. Therefore, the odds of rural hospital closure at the PHA scale is no different from the odds at the county scale.

5.2 **INTRODUCTION**

The geographic environment inherently has the characteristics of a hierarchy and introduces spatial relationships between spatial phenomena. Spatial data is data with a spatial component that is connected to a specific place on Earth. For the purposes of this study, the association of social capital of rural communities and rural hospital closure is

---


being determined. Initially, this analysis was conducted at the county (geographic) scale due to data availability and because the county scale has been found to be sufficiently accurate for use in analyses for rural spatial phenomenon. However, county boundaries introduce a man-made, hierarchy of geographic scale.

To demonstrate, for example, states consist of a collection of counties, and counties consist of independent, non-overlapping census tracts. Geographic data thus requires researchers to consider the analytical challenges that spatially referenced data introduces. In the past, social science disciplines outside of geography have focused research on determining spatial patterns and trends but failed to address the analytical challenges of spatial data. More recent work has begun to acknowledge and address them. Special analytical approaches are required to handle problems more commonly referred to as spatial autocorrelation, the modifiable areal unit problem (MAUP), scale and edge effects. These problems will be explained below. However, when properly acknowledged and addressed, the complications stemming from space and error dependence can improve models and make estimates of parameters that are less sensitive to statistical bias, inconsistency, or inefficiency. For the purposes of this study, the major complications caused by spatially referenced data will be described and addressed below.

In this study, spatial autocorrelation remains a limitation but is mitigated in part by the regional scope of the analysis. However, the central purpose of this study is to directly address MAUP and the boundary problems, which includes scale and edge effects, by creating a new geographic scale, known as the Proximal Hospital Area (PHA). The PHA scale is calculated using the georeferencing and buffer analysis tools of a common geographic information system, ESRI's ArcGIS. The main goal of the study is to compare the descriptive

---

statistics for unemployment, substance abuse crime, and violent crime at the county scale versus the PHA scale. If little to no difference (<5% change) exists between social capital at the county compared to social capital at the PHA scales, then the odds ratios measuring the association between social capital and rural hospital closure at the PHA scale fall with the 95 percent confidence intervals of the odds ratios calculated at the county scale. In which case, the relationship between social capital of rural communities and rural hospital closure is considered not sensitive to changes in geographic scale.

5.3 **Theoretical Background**

“Spatial is special.”\(^{471}\) In studies crossing many disciplines, researchers have long asserted that special methods often are necessary for appropriate analysis of spatial data. The nature of spatial data, the mathematics of space, the tools available all contribute in making spatial special. Therefore, use of spatial data requires a baseline understanding of its impact on analyses.

5.3.1 **Spatial Data Considerations**

Overall, four major problems interfering with the accurate estimation of the statistical parameter have been identified in spatial analysis – the scale problem, the pattern problem (more commonly referred to as spatial autocorrelation), the Modifiable Areal Unit Problem (MAUP), and the boundary problem.\(^{472}\) The scale problem was discussed previously, as this study is justifiably using the county scale.\(^{473}\) The remaining three problems – spatial autocorrelation, MAUP, and the boundary problem will be expanded upon below.

5.3.1.1 **Spatial Autocorrelation or Dependency**

Spatially referenced data inherently violates assumptions necessary for standard regression modeling. One is the assumption of independence among observations. In geography, attributes of spatially reference data often are correlated relative to distance, a form of connectivity relationship. This fundamental problem is best expressed by the well-established tenet in geography known as Tobler’s First Law (TFL) which states, “the first law of geography: everything is related to everything else, but near things are more related

---


than distant things.”\textsuperscript{474, 475} If nearby or neighboring areas are more alike, this is known as positive spatial autocorrelation and may violate the independence assumption, thus yielding unstable parameter estimates and unreliable significance tests.\textsuperscript{476} Negative spatial autocorrelation occurs less often and occurs when high values tend to be found by low values (or vice versa) and look like “islands” of dissimilarity or even a checkerboard pattern.\textsuperscript{477}

Since information loss is more severe as spatial autocorrelation becomes more pronounced, methods for measuring the degree of spatial dependency or autocorrelation in a geographic space have been developed.\textsuperscript{478} Spatial autocorrelation can be measured by computing standard spatial autocorrelation indices such as Moran’s \textit{I}, Geary’s \textit{C}, Ripley’s \textit{K}, and Getis’s \textit{G}. The basic premise of these indices is to calculate spatial weights at pairs of locations in a neighborhood that measure the covariance or intensity of the geographic relationship.

In previous studies, the degree of spatial autocorrelation in rural communities across many demographic and socioeconomic variables, as measured by Moran’s \textit{I}, is usually positive but can vary (range of 0.2 – 0.6 in one study).\textsuperscript{479, 480} A zero value indicates random spatial pattern; a one means perfect correlation.\textsuperscript{481} Positive spatial autocorrelation, among geographers, yields the conclusion that “place in space” matters. However, if it does, how

should space be modeled? Regarding this question, there is no standard approach. In addition, one cause of correlation is the physical geography of any area – an impassable mountain, a river with few bridges, lack of roads like in Alaska. Often times, the physical geography is hard to if not impossible to change. Therefore, although spatial autocorrelation is a well-documented and analyzed issue in spatial analysis, the focus of this study is on rural hospitals and their respective communities and is justified in solely listing spatial autocorrelation as a limitation. Furthermore, a major strength of this study comes from including data from the entire South U.S. Census region. There is precedence in health services research literature that directly addressing spatial autocorrelation is unnecessary, particularly for non-geography studies with a national scope.

5.3.1.2 Modifiable Areal Unit Problem

The Modifiable Areal Unit Problem (MAUP) refers to the “geographic manifestation of the ecological fallacy in which conclusions based on data aggregated to a particular set of districts may change if one aggregates the same underlying data to a different set of districts.” MAUP literature was first documented in 1934 when the correlation coefficient of male juvenile delinquency differed by the scale of aggregation with the 252 census tracts in the Cleveland area. In 1979, Openshaw and Taylor coined the term MAUP in a study using the 99 counties in Iowa. MAUP consists of two components: one is the scale problem or aggregation problem and the other is the grouping or zoning problem.

The scale problem occurs when the same data set is aggregated into different spatial resolutions, particularly aggregating small areas into larger ones, thus yielding different statistical inferences and estimates. For example, New Jersey has a relatively low infant mortality rate of 5.3 per 1,000 live births, the seventh best infant mortality rate in the

---


Yet the infant mortality rate in Camden County is 6.9 per 1,000 live births, a rate that is higher than the state average. In the next-smallest geographic scale, when Camden County is divided into city, boroughs, and townships, it becomes clear that over 40 percent of infant mortality cases in Camden County occur in the City of Camden. If infant mortality is only reported at the state or county level, policy makers would either not realize there are areas with high infant mortality or design policies that fail to target the populations truly grappling with high infant mortality. Therefore, aggregating spatial data at different geographic scales can mask trends and introduce errors affecting the validity of results from spatial analysis, which in turn may affect infant mortality policies adopted in the state.

The zoning problem happens when the spatial scale of analysis is fixed but the shape of aggregation units change. Variations in analytic results are found after areal units are grouped in alternative ways at the same spatial scale. The zoning effect can be seen by further analyzing John Snow's cholera outbreak map (Figure 4). In the original analysis, John Snow used a basic form of point pattern analysis when identifying the location of each cholera patient's residence. However if Dr. Snow's study area is aggregated at the same scale into different zones, the conclusion can lead to an erroneous finding. In the top map, the water pump causing the cholera outbreak would still have been identified. In the second map, no strong association between cholera deaths and water pumps is detected in any area. And in the third, two water pumps that did not lead to the outbreak would be incorrectly associated as causing it.


Understanding MAUP is difficult because there is a random aspect to the effects of the MAUP. It may be difficult generalizing how different data sets with different spatial units are effected. Thus, there is no established agreement on a solution to MAUP, although many studies have tried to resolve this problem. Despite the lack of a resolution, researchers must at least recognize the scale and zoning problems. Recently, one claim has been that the only real resolution to MAUP is to use geocoded, individual-level data, but this solution is unrealistic due to confidentiality and privacy issues. With this option being unfeasible, researchers have found techniques to ameliorate the effect of MAUP, which were

adhered to in Chapter 3. First, researchers can avoid making inferences at the individual level when working with aggregated data. Inferences can only be made at the scale the data is collected at. For the purpose of this paper, the spatially referenced data has been collected at the county scale, which in turn is being used as a proxy for the community a rural hospital serves. No inferences are being made at a scale smaller than the county. Finally, when possible, researchers can use scale-independent statistics or methodologies to make inferences, but this analysis requires the inclusion of U.S. Census decennial data that are not scale-independent.

5.3.1.3 Boundary Problem

The boundary problem is often discussed in conjunction with MAUP because MAUP is associated with arbitrary geographic units, which in turn are defined by boundaries. The boundary problem encompasses two parts – the edge and shape effects. The edge effect occurs when patterns of interaction or interdependency across the borders of the bounded region are ignored or distorted. In health geography, this commonly is addressed in accessibility studies when, for example, a hospital or cancer center is located near an existing state border. In those cases, it is unlikely the border would prohibit a patient from seeking care at an institution in a neighboring state. The shape effect happens when the shape imposed on the bounded area affects the perceived interactions between phenomena. A proven example of the shape effect occurs with geographic units that are more elongated, which could occur due to physical geography and history. In point pattern analysis, the elongated shape of a geographic unit tends to cause higher levels of clustering. The shape of a unit can also impact the transportation infrastructure and consequently interaction and flow among spatial entities.

and shape effects of the boundary problem, this study will compare results from the analysis at the county geographic scale and at the Proximal Hospital Area (PHA) scale.

5.3.2 Scale of Spatial Data

5.3.2.1 Counties

As discussed in Chapter 3 and Chapter 4, the geographic scale selected as the base unit for analysis was counties. The county scale does aggregate data from a finer scale and inherently introduces man-made boundaries. In addition, defining a rural hospital’s “community” as the county in which they are located may be inaccurate, particularly if a hospital is located far from the centroid of the county. The county-defined rural community may include areas from which the hospital does not draw patients and also exclude areas from which it does draw patients.\(^{502}\) Another point to consider is the differences in population size between counties. These differences may make rates more unstable\(^{503}\).

Despite MAUP, the boundary problem, and the possible oversimplification of using counties as a proxy hospital service area, the richest data comes at the county level. Federal agencies including the U.S. Census Bureau, the Bureau of Labor Statistics, and the Centers for Disease Control and Prevention all commonly collect data at the county scale thus allowing for the inclusion of a wide range of social, economic, demographic, and political data to this study.\(^{504,505}\) At the county-scale, the number of missing data points of each social capital component ranged from zero to 37, out of a maximum 2,052 counties nationwide. Data available at a finer scale than a county is available in the RTI Spatial Impact Factor Database such as zip code, primary care service area, and census tract, but the percent of missing data significantly increases. Aside from increased data availability at the county scale, previous research has also given credence for the use of the county scale in defining a rural hospital’s market or service area. For example, a seminal study, that compared the predictive ability of


\(^{504}\) Ibid

county versus zip-code scale data in identifying major rural hospital closure risk factors, found no difference between the findings derived from county versus zip code scale data.\textsuperscript{506}

In Chapters 3 and 4, social capital of a rural community is captured at the county scale, in which the county serves as a proxy for the rural hospital’s catchment area. The scale effect of the MAUP is partially addressed by stratifying results by the larger scale Census regions and determining if a statistically significant change in odds ratios occurs. The zoning effect of MAUP and both edge and shape effects of the boundary problem are solely acknowledged as limitations. Therefore, in this aim, this limitation will be addressed by creating a new spatial scale – Proximal Hospital Area (PHA). Social capital of rural communities will be determined at the PHA scale rather than the county. If the observed result for this analysis compared to the findings from Chapter 3 are statistically similar, regarding the relationship between social capital of rural communities and rural hospital closure, then solely using the county scale is justified and sufficient. This result, allowing the analysis to remain at the county level, would simplify further social capital and healthcare system analyses and facilitate generalization.

5.3.2.2 Proximal Hospital Areas (PHA)

The constant width buffer zone created around each rural hospital, is represented as a point. A buffer zone, in the shape of a circle, will be created by specifying a Euclidean distance of a specific number of miles from the rural hospital location that captures the area within one-hour access to the rural hospital. “In nearly all locations within the United States, the straight-line distance is an adequate proxy for travel distance...Exceptions are limited to areas located near uncrossable physical features such as lakes, rivers, and mountains and in wilderness areas of the western United States and Alaska.”\textsuperscript{507} Although these uncrossable physical features are more common in rural areas than in urban ones, alternative distance measuring methods are either inadequate or unfeasible. An alternative way to measure distance is known as “Manhattan” distances where movement is restricted to paths parallel to the axes and can be more meaningful than Euclidean distances in urban settings. Travel time would be a more accurate way to capture a rural hospital’s service area. Consequently, a recent study by the National Center for the Analysis of Healthcare Data of West Virginia

rural hospital access used travel times due to the state’s mountainous terrain. However, travel time data for all 50 states is not publicly available. To gain access to this data, researchers either need to have relationships with private industry partners like Google (who use travel times as part of Google Maps capabilities) or with each individual state’s department of transportation.

For this study, the radius for the buffer analysis will be 30 miles, which is based on the 50 kilometer distance defined by Canada’s Ministries of Health Services and Health Planning. The 30-mile buffer also falls within the 15-35 mile range the federal government requires critical access hospitals to be apart from each other. Another Euclidean distance considered for analysis was 10 miles since the Obama administration discussed excluding or cutting support to critical access hospital that are within 10 miles of one another. However, a 10 mile buffer radius often fails to include much outside of the rural hospital’s home county, especially if the rural hospital is located near the county’s centroid.

5.4 METHODOLOGY

In this study, rural hospital closure is the dependent, dichotomous variable, while the following three social capital variables previously found to be statistically significant at measuring social capital (Chapter 3) are the independent variables: unemployment, substance abuse crime, and violent crime. To facilitate the analysis, this study is limited to the South U.S. Census region, a region with 840 rural hospitals and the highest rural hospital closure rate in the United States (4.9 percent compared to a 3.3 percent national closure rate). The South U.S. Census region consists of the following 17 states: Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, Tennessee, Texas, South Carolina, Virginia, and West Virginia. Due to the higher rural hospital closure rate, any geographic effects from using the county

---


scale versus Proximal Hospital Areas would be more pronounced in the South region than in regions where sparsity of rural hospital closure becomes a serious concern.

Social capital data for this study originated from the robust, nationally-representative Research Triangle Institute’s (RTI) Spatial Impact (SIF) database. Data is aggregated at the county-level and limited to counties classified as rural by the USDA’s Economic Research Service Rural-Urban Continuum or Beale Codes (n=873). The SIF database provides nearly complete data for all 873 rural counties in the South census region. The variable with the maximum missing data is Substance Abuse Crimes (3.3% missing data), with most missing values attributed to Florida. Due to the completeness of the dataset, there is no need to address missing data with deletion, single imputation, or model-based methods. Rural hospital data was purchased from Health Forum, the data repository enterprise for the American Hospital Association (AHA). Many prominent hospital closure studies have used AHA data. To verify and ensure accuracy of the data, the AHA file was merged with the CMS Provider File using the CMS provider number available in both the AHA and CMS databases.

5.4.1 Redefining Geographic Scale: Proximal Hospital Area (PHA)

To create the PHA scale, this study will use ArcGIS 10 software engineered by the Environmental Systems Research Institute (ESRI). Within ArcGIS in ArcMap, basemap layers will be added to a Blank Map. Basemap layers contain reference geography that forms a backdrop to the map including state and county boundaries, cities, streets and imagery. Free basemap layers can be downloaded from ESRI’s website and can be stored at multiple scales. The data for basemap layers often includes the metadata, which is a description of what is known about the dataset and vouches for the integrity of data and makes the data searchable. The data for these layers are stored in ArcMap in attribute tables. As long as two separate data sources have a column within the attribute table, the data can be connected through a “union” operation.

After creating the basemap, two data files will be added using the ArcCatalog, a data management application of ArcGIS. In ArcGIS, it is very important for all data files to be stored in the same location, known as a “home folder,” in order to keep map documents, datasets, and related files in the same place.\(^\text{515}\) This keeps the path relationship between layers and their data sources simple and easy to manage or share in the form of a map package or bundle of maps and data.

Next, the location of each rural hospital in the format of (x,y) coordinates from the American Hospital Association database will be used to ensure accuracy and precision. Mapping physical addresses in ArcGIS of rural hospitals using only the ArcGIS (x,y) tool sometimes fails to find a location or marks a hospital in a wrong location, after checking the raster imaging. This occurs more often when mapping physical addresses of rural geographic phenomena than urban phenomena.

Following these initial steps, a buffer analysis will be performed using a radius of 30 miles. The counties included in every buffer will be determined, as well as the percent land area of each county that falls within the buffer boundaries. Any areas of a PHA that are not located in a rural county are removed from the analysis thus modifying the initial buffer. Removal of urban areas from the analysis is necessary because numerous studies have proved that residents living in closer proximity to urban areas are not likely to access health services in a rural hospital when an urban-located hospital is accessible.\(^\text{516}\) In fact for many types of services, distance is not a strong determining factor for health care utilization patterns among rural residents.\(^\text{517}\) Before continuing, the county-level social capital data will be added to ArcCatalog. At this point, the percent land area will be used to weigh each county’s particular social capital data. For example, if 40 percent of a hospital’s PHA, or modified buffer, is in County A and 60 percent is in County B, then the social capital data from County A will be weighted 40 percent (County B’s will be 60 percent). In the end, each PHA will have its own social capital variable measures representing the new scale.


\(^{517}\) Ibid
5.4.2 Adjusting for Proximal Hospital Areas

Having defined social capital at both the county and Proximal Hospital Area (PHA) scales in 2000, the descriptive statistics for unemployment, substance abuse crime, and violent crime will be calculated at both scales. Social capital variables at the PHA scale are calculated by modifying county-level data. Equal distribution of the population in each rural county is assumed. Returning to the example of a PHA including 40 percent of County A and 60 percent of County B, the PHA unemployment rate is calculated by dividing the number of unemployed rural residents in 40 percent of County A and 60 percent of County B by the total population in 40 percent of County A and 60 percent of County B. For both substance abuse crime and violent crime, the crime rate per 1,000 was calculated in a similar manner (Table 31).

<table>
<thead>
<tr>
<th>South Region</th>
<th>Unemployment</th>
<th>Substance Abuse Crime (per 1000)</th>
<th>Violent Crime (per 1000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>County-Scale</td>
<td>0.051</td>
<td>25.089</td>
<td>3.212</td>
</tr>
<tr>
<td>PHA-Scale</td>
<td>0.050</td>
<td>24.100</td>
<td>3.313</td>
</tr>
<tr>
<td>% Mean Change Scales</td>
<td>1.89%</td>
<td>3.13%</td>
<td>3.94%</td>
</tr>
</tbody>
</table>

After calculating the descriptive statistics at the Proximal Hospital Area scale, the average level of these three statistically significant social capital variables (unemployment, substance abuse crime, and violent crime) at the PHA scale is no different from the averages at the county scale. All analyses were calculated and performed by the STATA statistical software. A value of P<0.05 was considered statistically significant.

5.5 Results and Interpretation

This study has two main goals. First, the social capital variables averages for unemployment, substance abuse crime, and violent crime, at two different geographic scales, were compared (Table 31). As the results show, the difference between social capital reported at the county scale versus social capital reported at the Proximal Hospital Area (PHA) scale is negligible (1.9-3.9 percent). A differences of less than 5 percent implies the
The impact of scale would be expected to fall within the initial confidence intervals for the odds ratios (Table 32). The 95% confidence interval (CI) around the odds ratio (OR) was obtained with the formula (β±1.96 SE), in which SE is the standard error of β. The CI indicates the level of uncertainty around the measure of effect; one can infer that the true OR lies between the upper and lower confidence limits.

**Table 32: Odds Ratios at the County Scale (2000)**

<table>
<thead>
<tr>
<th></th>
<th>n = 2,052</th>
<th>Crude OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unemployment, 2000</td>
<td></td>
<td>1.15</td>
<td>1.02</td>
<td>1.31</td>
</tr>
<tr>
<td>Substance Abuse Crimes, 2000*</td>
<td></td>
<td>1.0005</td>
<td>1.000</td>
<td>1.001</td>
</tr>
<tr>
<td>Violent Crimes, 2000*</td>
<td></td>
<td>1.0023</td>
<td>1.000</td>
<td>1.004</td>
</tr>
</tbody>
</table>

*To clarify, odds ratios were rounded to 4 decimal places and confidence intervals to 3 places.

The results imply that the unemployment odds ratio calculated at the PHA scale would lie between 1.02 and 1.31, although it is closer to the crude unemployment OR value of 1.15 than the CI values, since the difference in unemployment between the county and the PHA is 1.89 percent. Since the PHA scale is not expected to change the odds ratios, unemployment remains the most significant social capital variable in predicting the odds of rural hospital closure. For every 1 percent increase in unemployment, the odds of a rural hospital closing increased by 1.15 times (or by 15%) holding all else constant. In the case of both substance abuse crime and violent crime, the odds of a rural hospital closing remains unchanged.

### 5.6 Discussion

Addressing geographic spatial data challenges, in this case the MAUP and the boundary problem, is not common in health services research. This study is the first to define the community or market of a rural hospital using a 30-mile circular buffer. Despite the extensive process undertaken to individually calculate social capital at the PHA scale, the findings suggest the relationship between a rural community’s social capital and the likelihood of rural hospital closure is not sensitive to a change in geographic scale.

---

Initially, studies of hospital markets or the community served by a particular hospital assumed that a rural hospital’s market area coincided with an existing geographic entity. For rural hospitals, the county is considered the most geographically appropriate scale.\textsuperscript{519, 520} However, as the Dartmouth Atlas of Health Care has shown, the community served by any particular hospital, referred to as a Hospital Service Area (HSA), is affected by a wide range of issues including the willingness for both patients and physicians to travel, referral patterns, and even geographic barriers such as mountains or rivers.\textsuperscript{521} Dartmouth’s HSA is based on zip codes and uses Medicare patient origin data, while the PHA, used in this study, is based on a buffer that ignores boundaries. Regardless of the manner in which the community served by a hospital is identified, the same factors characterizing this community make it unlikely that rural hospital markets perfectly coincide with county boundaries.\textsuperscript{522, 523}

To best estimate the community served by a rural hospital, researchers argue that the gold standard would require patient-level data for the address of origin.\textsuperscript{524} Patient-level data introduces significant confidentiality and privacy concerns, even if aggregated. In addition, this type of data is unlikely to be secured from every rural hospital in the United States. Thus, some researchers have proposed using radii around a hospital to delineate market areas or to use a different existing geographic scale of analysis such as zip codes.\textsuperscript{525, 526}

However, this study finds that changing the geographic scale does not affect the association of social capital variables on the likelihood of rural hospital closure. This builds upon a previous rural hospital closure study that similarly found that the geographic scale smaller than a county, zip code level, does not affect socioeconomic and demographic characteristics nor does it affect the estimate of the coefficients for a rural hospital closure

\textsuperscript{521} Dartmouth Atlas of Health Care
Instead of using an existing geographic scale defined by the federal government, this study developed a new scale, the Proximal Hospital Area (PHA). The PHA scale is larger than a county, and it is not restricted to manmade boundaries. Yet, like the previous zip code and rural hospital closure study, the association between social capital and rural hospital closure is not sensitive to geographic scale.

The lack of an impact when changing geographic scale is rooted in a wide variety of factors and study limitations. First, the development of Proximal Hospital Areas requires making many assumptions including equal distribution of the rural population in each county. Although this assumption is commonly accepted in geography, when counties are larger in area, such as counties in the West U.S. Census region, the assumption of equal population distribution becomes more difficult to justify. In fact, if an analysis was limited to the West U.S. Census region, this equal population distribution may not be assumed in the analysis. Alternative, although time consuming, spatial analytical techniques exist to more accurately and precisely determine the population distribution, one of which is dasymetric mapping. This technique uses satellite imagery which can project the population density of an area by identifying land use. Next, Proximal Hospital Areas are created using a circular buffer with a 30-mile radius, which presumes that any rural resident living within the buffer has equal access to the rural hospital located at the centroid of the circle. Buffers fail to consider the road network and geographic barriers which affects travel time or even prohibits access to a rural hospital.

Secondly, a Proximal Hospital Area (PHA) may not be a refined enough geographic scale to capture rural heterogeneity. Calculating the PHA scale, for example, is less useful the more symmetrical counties are in shape. Only counties with elongated shapes or where the rural hospital is not located near the centroid of the county stand to benefit from the PHA. If the majority of rural hospitals are located near the centroid of a county and counties are relatively symmetrical, then the increased precision and accuracy associated with using the PHA scale disappears when the scope of an analysis is the entire United States.

Third, although rural heterogeneity has been shown when comparing rural areas from New England to ones in the South, within a Proximal Hospital Area, the PHA itself may have relative homogeneous social capital characteristics. Again, Tobler’s First Law of Geography

states that “everything is related to everything else, but near things are more related than distant things.”\textsuperscript{528,529} As such, there may be very little difference between counties falling within a PHA.

Fourth, it is possible that variables that are important predictors of closure and able to show important differences between the two market scales are omitted. Many of these variables may not be available for inclusion in this study, such as private hospital-specific financial information patient origin data, or require in-person, qualitative data collection.\textsuperscript{530}

Fifth, the difference between the PHA and a county may be little to none because rural areas with multiple rural hospitals in close proximity were oversampled. If two rural hospitals were within 30 miles of one another, then part of the population captured in the PHA for the first hospital will be included in the PHA for the second hospital. Oversampling, especially of rural areas that are less likely to be isolated, will inherently homogenize results.

Finally, since one-by-one calculation of the social capital variables at the PHA scale was required, this analysis was limited to the South U.S. Census region. Although the South region accounts for 39.5 percent of all rural hospitals in the United States (n=840) and experiences the highest rural hospital closure rate in the country (4.9 percent), limiting the analysis to a specific region can impact the generalizability of the findings. The South region, for example, could have unique characteristics that impact both rural communities and rural hospitals that other regions do not face, such as the historical legacy of slavery and racial discrimination.

\subsection*{5.6.1 Conclusion}

In this study, the findings showing that the association between social capital and rural hospital closure is not sensitive to geographic scale. However, careful consideration must still be given to the unit of analysis in order to address spatial data analytical problems including spatial autocorrelation, the Modifiable Areal Unit Problem, and the boundary

\begin{flushleft}


\end{flushleft}
problem. Further research into the market areas of rural hospitals, particularly ones that have closed, could help develop a more relevant definition of the rural community served by a particular rural hospital.

Despite the limitations of this study, the development of a new geographic scale of analysis contributes to a deeper understanding of hospital markets or the communities each hospital serves. The use of spatial analysis in still relatively innovative in health services research. Although the overall association between social capital and rural hospital closure is not sensitive to the new geographic scale developed, the PHA is expected to be a more accurate geographic scale in representing the community each rural hospital serves. Therefore, if an analysis is limited to a small number of rural hospitals and does not include the entire United States, the PHA may be found to be the most appropriate geographic scale to utilize.531

The PHA scale is perhaps most useful to individual rural hospitals. Rural hospitals, for example, rely on an accurate representation of the market or community they serve to make strategic administrative decisions. There is a well-established research interest focused on how to better determine hospital markets, but the manner in which rural hospital markets are defined is recognized as being quite different from an urban hospital market.532,533 Previous studies have created or used geographic scales at the county level or smaller, such as zip codes, however, rural residents are known to be more willing to travel farther for certain healthcare services than urban residents.534 Therefore, a geographic scale that is larger than a county, such as the PHA scale, may better inform rural hospital administrators on the community their facility serves. Furthermore, rural hospitals are often part of larger healthcare systems and operate healthcare facilities that are not short-term, acute-care hospitals. Although this study focused on just the communities served by short-term, acute-care hospitals, rural healthcare systems run rehabilitation centers, nursing homes, step-down units within their hospital facility, outpatient clinics, etc. These facilities are likely to

not all be located within proximity to each other. A nursing home may be located in another city within the same county. An outpatient facility may be located in an area that is closer to a metropolitan area in order to minimize physician travel time, for physicians commuting from urban areas.

Besides benefiting rural hospitals, the PHA geographic scale could also help rural communities gain a better understanding of the current state of the local area and aid community planners in designing policies. As previous findings show, rural communities could implement policies that directly affect various components of social capital and ultimately impact the odds of rural hospital closure, as well as a number of other outcomes (Chapter 3 and 4). The PHA scale could also inform community and healthcare leaders who are now required to periodically write a Community Health Needs Assessment (CHNA). The Patient Protection and Affordable Care Act requires tax-exempt hospitals to liaise with community leaders and conduct a CHNA every three years in order to maintain their tax-exemption status with the Internal Revenue Service. Analyzing health data at multiple geographic scales can help identify previously undetectable health trends. Therefore, further analysis and application of the PHA scale merits exploration.


Family Environmental and Socially Sustainable Development Network, Social Capital


Intuitive Interpretations and Exploring Analytical Applications. In U. o. T. I. Project (Ed.).
Austin, TX: University of Texas - Austin.

Confounding. Centers for Public Health Preparedness (CPHP). University of Michigan,

Costa, Dora L., & Kahn, Matthew E. (2003). Understanding the American decline in social

County-level Measure of Social Capital. Northeast Regional Center for Rural Development.
http://nercrd.psu.edu


Dartmouth Atlas of Health Care

De Smith, Michael; Goodchild, Michael F.; Longley, Paul A. Geospatial Analysis: A

Department of Agricultural Economics, University of Kentucky, n.d. Web. 15 Oct. 2014.
http://www.uky.edu/


Doering, Christopher. “As more move to the city, does rural America still matter?” USA Today 13 Jan. 2013.


Iceland, J. (2004). The Multigroup Entropy Index (also known as Theil’s H or the Information Theory Index): University of Maryland.


capital and college enrollment: An examination of racial/ethnic group differences. Journal of


and linking aspects of social capital." Health Place 18(2): 286-295.


Knowledge and Social Capital: Foundations and Applications By Eric L. Lesser. Boston:

Principal Components, pg. 294.

Probst Janice C, Samuels Michael E, Hussey James R, Berry David E, Ricketts Thomas C.
“Economic Impact of Hospital Closure on Small Rural Counties, 1984 to 1988:
pgs. 375-390.

Democracy 6: 65-78.

York, Simon and Schuster.

Princeton University Press.


Rowe, Gina. 2014. “Social capital, neighborhood disadvantage, and potentially preventable emergency department visits in metropolitan Baltimore.” Prepared for the 142nd APHA Annual Meeting and Exposition. New Orleans, LA


http://www.icpsr.umich.edu/


http://www.fbi.gov/


## APPENDIX A: ACRONYMS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHA</td>
<td>American Hospital Association</td>
</tr>
<tr>
<td>AHD</td>
<td>American Hospital Directory</td>
</tr>
<tr>
<td>AQS</td>
<td>Air Quality Systems</td>
</tr>
<tr>
<td>ARF</td>
<td>Area Resource File</td>
</tr>
<tr>
<td>ASU</td>
<td>Arizona State University</td>
</tr>
<tr>
<td>BBRA</td>
<td>Balanced Budget Refinement Act of 1999</td>
</tr>
<tr>
<td>BEA</td>
<td>Bureau of Economic Analysis</td>
</tr>
<tr>
<td>BIPA</td>
<td>Benefits Improvement and Protection Act of 2000</td>
</tr>
<tr>
<td>BLS</td>
<td>Bureau of Labor Statistics</td>
</tr>
<tr>
<td>CAH</td>
<td>Critical Access Hospitals</td>
</tr>
<tr>
<td>CFR</td>
<td>Code of Federal Regulations (of the United States)</td>
</tr>
<tr>
<td>CMS</td>
<td>Centers for Medicare and Medicaid Services</td>
</tr>
<tr>
<td>CPS</td>
<td>Current Population Survey</td>
</tr>
<tr>
<td>DOC</td>
<td>Department of Commerce</td>
</tr>
<tr>
<td>DOI</td>
<td>Department of the Interior</td>
</tr>
<tr>
<td>DOJ</td>
<td>Department of Justice</td>
</tr>
</tbody>
</table>

DOL  Department of Labor
EPA  Environmental Protection Agency
ERS  Economic Research Service
ESRI  Environmental Systems Research Institute
FBI  Federal Bureau of Investigation
FQHC  Federally Qualified Health Clinic
GDP  Gross Domestic Product
HHS  Department of Health and Human Services
HRSA  Health Resources and Services Administration
MIPPA  Medicare Improvement for Patients and Providers Act of 2008
MMA  Medicare Modernization Act (2003)
NERCRD  Northeast Regional Center for Rural Development
NIH  National Institutes of Health
NHD  National Hydrography Dataset
NOAA  National Oceanic and Atmospheric Administration
PPACA  Patient Protection and Affordable Care Act (2010)
PRISM  Parameter-elevation Regressions on Independent Slopes Model
OMB  Office of Management and Budget
RHC  Rural Health Center
RTI  Research Triangle Institute
UCR  Uniform Crime Report
USDA  United States Department of Agriculture
APPENDIX B. KEY TERMS

**ACCURACY:** The degree to which information on a map of in a digital database matches the true or accepted values.535

**ATTRIBUTE:** A data item associated with an individual object (record) in a spatial database. Attributes may be explicit, in which case they are typically stored as one or more fields in tables linked to a set of objects, or they may be implicit (sometimes referred to as intrinsic), being either stored but hidden or computed as and when required (e.g. polyline length, polygon centroid). Raster/grid datasets typically have a single explicit attribute (a value) associated with each cell, rather than an attribute table containing as many records as there are cells in the grid.536

**BUFFER:** A zone of a specified distance around coverage features – a point, line, or polygon. In GIS, creating a buffer temporarily groups data from this zone of a specified distance for use in analysis. There are two types of buffers: constant width buffers and variable width buffers.537

**CENSUS REGION:** The United States Census Bureau has four official regions, with nine official divisions. The four official regions are the following: West, Midwest, South, and Northeast. The partition of the country into geographic regions has roots in the colonial period of American history (i.e. New England, Middle Atlantic, and

---


South). However, the U.S. Census officially began reporting data by region with the 1850 U.S. Census after James D. B. DeBow became the Director of the U.S. Census. The nine regional divisions have been in place since the 1910 decennial U.S. Census; the four regions and nine division map has been in place since the 1960 U.S. Census when Alaska and Hawaii were added to the Pacific Division.\textsuperscript{538} Map Source\textsuperscript{539}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{census_region_map.png}
\caption{CENSUS REGION MAP}
\end{figure}

**Community Hospital:** It includes all nonfederal, short-term general (including academic medical centers), and other special hospitals, such as obstetrics and gynecology; eye, ear, nose, and throat; rehabilitation; orthopedic; and other individually described specialty services. Excluded are hospitals not accessible by the general public, such as prison hospitals or college infirmaries.\textsuperscript{540}

**Composite Social Capital Index:** a variable developed by Penn State University’s Northeast Regional Center for Rural Development for research funded by the USDA to describe community participation, which is a component of social capital. It is a composite index created using principal components analysis. The variables used in the composite fall into four categories, which are: (1) total public exercise facilities per 10000 people, (2)  


number of associations or organizations per 10000 people, (3) U.S. Census mail response rate, and (4) presidential voter response rate. [See Appendix]

**CRITICAL ACCESS HOSPITAL:** rural community hospitals that receive cost-based reimbursement and must meet defined criteria that were outlined in the Conditions of Participation 42CFR485 and subsequent legislative refinements to the program through the Balanced Budget Refinement Act of 1999 (BBRA), Benefits Improvement and Protection Act of 2000 (BIPA), the Medicare Modernization Act (MMA) in 2003, the Medicare Improvement for Patients and Providers Act of 2008 (MIPPA), and the Patient Protection and Affordable Care Act (PPACA) in 2010.\(^{541}\)

---

**CAH Eligibility**
The following providers may be eligible to become CAHs:
- Currently participating Medicare hospitals
- Hospitals that ceased operation during the 10 year period from November 29, 1989 through November 29, 1999; or
- Health clinics or centers (as defined by the State) that previously operated as a hospital before being downsized to a health clinic or center.

Unlike facilities such as Medicare Dependent Hospitals or Sole Community Hospitals, CAHs represent a separate provider type with their own Medicare Conditions of Participation as well as a separate payment method.

**CAH Designation**
A hospital must meet the following criteria to be designated a CAH:
- Be located in a state that has established a State Flex Program (as of December 2008, only Connecticut, Delaware, Maryland, New Jersey, and Rhode Island did not have such a program);
- Be located in a rural area or be treated as rural under a special provision that allows qualified hospital providers in urban areas to be treated as rural for purposes of becoming a CAH;
- Furnish 24-hour emergency care services, using either on-site or on-call staff;
- Provide no more than 25 inpatient beds that can be used for either inpatient or swing bed services; however, a CAH may also operate a distinct part rehabilitation or psychiatric unit, each with up to 10 beds;
- Have an average annual length of stay of 96 hours or less (excluding beds that are within distinct part units [DPU]); and
- Be located either more than 35 miles from the nearest hospital or CAH or more than 15 miles in areas with mountainous terrain or only secondary roads OR prior to January 1, 2006 were State certified as a “necessary provider” of health care services to residents in the area.

**Economic Dependency:** A rural area’s economic characteristics have significant effects on its development and need for various types of public programs. One such county-level classification scheme measures economic dependency and has been developed by the USDA’s Economic Research Service (ERS). This typology classifies all rural counties into the following six non-overlapping categories of economic dependence: farming, mining,
manufacturing, services, Federal/State government, and unspecialized. The development of this classification is rooted in dependency theory that views rural areas being kept marginalized by dominant, wealthier urban ones. A rural community's reliance on a single, major industry puts rural areas in a far more precarious economic position.

- Farming-dependent Counties: In these rural counties (n = 403), 15 percent or more of the average annual labor and proprietors’ earnings are derived from farming or 15 percent or more of employed residents worked in farm occupations. The majority of these counties are located in the central plains, Midwest Census region.
- Mining-dependent Counties: In these rural counties (n = 113), 15 percent or more of average annual labor and proprietors’ earnings are derived from mining. These counties are primarily located in the Central Appalachian region, Texas, and the Rocky Mountain region.
- Manufacturing-dependent Counties: In these rural counties (n = 585), 25 percent or more of the average annual labor and proprietors’ earnings are derived from manufacturing. The majority of these counties are found in the Midwest and South Census regions, east of the Mississippi River.
- Federal/State government-dependent Counties: In these rural counties (n = 222), 15 percent or more of average annual labor and proprietors’ earnings are derived from Federal and State government. These counties are scattered across the United States.
- Services-dependent Counties: In these rural counties (n = 114), 45 percent or more of average annual labor and proprietors’ earnings are derived from the following services categories: retail trade, finance, insurance, and real estate. These counties are also spread out across the United States.
- Nonspecialized Counties: These rural counties (n = 615) failed to meet any of the dependence threshold for any of the above industries.

_Euclidean Distance_: Straight-line distance between two points on a plane.

---


**Federally Qualified Health Center:** In rural areas, Federally Qualified Health Centers (FQHC) are important safety net, outpatient care providers. Over one-third of Americans who receive care in an FQHC are rural residents.\(^{545}\) Similar to the Rural Health Clinic program, centers qualified for FQHC status receive higher reimbursement under Medicare and Medicaid because they serve a medically underserved area or population. When compared to the RHC program, however, the FQHC qualification criteria are far more comprehensive. FQHCs must offer provide a minimum level of primary care services as dictated by CMS, which includes maternity, prenatal, preventive, behavioral, dental, emergency, and pharmaceutical services. The center must be open 32 hours per week and make arrangements for after-hour emergency services. FQHCs must be nonprofit or public, are required to have a board of directors with the majority of members composed of their patients, and must have an ongoing quality assurance program. Finally, FQHCs have stricter annual auditing and financial reporting standards.

**Feature:** Frequently used within GIS referring to point, line (including polyline and mathematical functions defining arcs), polygon and sometimes text (annotation) objects (see also, vector).\(^{546}\)

**Geostatistics:** Statistical methods developed for and applied to geographic data. These statistical methods are required because geographic data do not usually conform to the requirements of standard statistical procedures, due to spatial autocorrelation and other problems associated with spatial data (AGI). The term is widely used to refer to a family of tools used in connection with spatial interpolation (prediction) of (piecewise) continuous datasets and is widely applied in the environmental sciences. Spatial statistics is a term

---


more commonly applied to the analysis of discrete objects (e.g. points, areas) and is particularly associated with the social and health sciences.\textsuperscript{547}

\textit{Gini Coefficient}: It is an index of income concentration. It is a statistical measure of income equality ranging from 0 to 1. A measure of 1 indicates perfect inequality; i.e., one person has all the income and rest have none. A measure of 0 indicates perfect equality; i.e., all people have equal shares of income. The U.S. Census Bureau used grouped data to compute all Gini Coefficients.\textsuperscript{548}

\textit{GIS}: A Geographic Information System (GIS) integrates hardware, software, data, and people to capture, manipulate, analyze, and display all forms of geographically referenced information or spatial data. GIS is a tool that allows to visualize, understand, and interpret data to reveal relationships, patterns, and trends.\textsuperscript{549}

\textit{Health Professional Shortage Area (HPSA)}: The Health Professional Shortage Area (HPSA) is a federal health provider shortage designation determined by the Health Resources and Services Administration (HRSA). HRSA primarily uses the number of full-time equivalent health professionals relative to the population and high-need indicators such as high poverty to determine if a geographic area, population, or facility qualifies for HPSA designation. To apply for HPSA designation, a community, individual or facility must apply through the state-level Primary Care Office. This designation helps determine eligibility for a variety of federal and state health workforce programs, such as the Federally Qualified Health Center program. In addition to HPSA status, two other federal shortage designations exist including the Medically Underserved Areas (MUA) and Medically Underserved Populations (MUP), and some states have created state-level shortage designations.

**Hospital Characteristics**: A term used to refer to hospital-level information including number of beds, basic financials, or number of specialties or practitioners.

**Hospital Closure**: This occurs when a healthcare facility is no longer in operation as a hospital, which typically entails maintaining 24/7 emergency services. This study uses hospital data from the American Hospital Association (AHA). As defined by the AHA, closure occurs in cases where hospital data is no longer available for a specific hospital in subsequent years of their annual hospital survey, as identified by both an AHA identifier and CMS Provider identifier. If a hospital changed name or owner, the AHA and CMS Provider identifier remain the same, so close occurs only when a specific identifier ceases to report data (See Chapter 2).

**Hospital Downsizing**: occurs either (1) when two hospitals in proximity to one another close and are replaced by a single hospital with a net loss of at least 10% of beds, or (2) when the number of beds in single hospital decreases by at least 10% between 2000 and 2009.\(^{550,551}\)

**Hospital Replacement**: any case in which a hospital is closed but is replaced by a new hospital which may or may not be in the same location but is within proximity (15-mile radius). Hospital replacement, therefore, cannot be classified as a closure or downsizing.

**Layer**: A collection of geographic entities of the same type (e.g. points, lines or polygons). Grouped layers may combine layers of different geometric types.\(^{552}\)

---


**PER CAPITA INCOME:** a measure of mean income (average income per person) within an economic aggregate like a country, county, or city. In the U.S., it is calculated by the Bureau of Economic Analysis by measuring all sources of income in the aggregate area and dividing it by the total population, from the U.S. Census Bureau’s annual midyear population estimates.553

**PERSISTENT POVERTY:** a term created by the U.S. Department of Agriculture’s (USDA) Economic Research Service that is used to describe counties in which 20 percent or more of the population were below the poverty level in the four previous decennial U.S. Census 1970, 1980, 1990, and 2000.554,555

**POLYGON:** A closed figure in the plane, typically comprised of an ordered set of connected vertices, \( v_1, v_2, \ldots, v_{n-1}, v_n = v_1 \) where the connections (edges) are provided by straight line segments. If the sequence of edges is not self-crossing it is called a simple polygon. A point is inside a simple polygon if traversing the boundary in a clockwise direction the point is always on the right of the observer. If every pair of points inside a polygon can be joined by a straight line that also lies inside the polygon then the polygon is described as being convex (i.e. the interior is a connected point set). The OGC definition of a polygon is “a planar surface defined by 1 exterior boundary and 0 or more interior boundaries. Each interior boundary defines a hole in the polygon.”556,557

---

**Population Loss:** term used to describe a county if its population dropped by 5% or greater between 1990 and 2000.\textsuperscript{558}

**Proximal Hospital Area:** The constant width buffer zone created around each rural hospital, represented as a point. A buffer zone, in the shape of a circle, will be created by specifying a Euclidean distance of a specific number of miles from the rural hospital location that captures the area within one-hour access to the rural hospital. Various distances (in miles) will be used to perform a sensitivity analysis. The first radius for the buffer analysis will be 30 miles, which is based on the 50 kilometer distance defined by Canada’s Ministries of Health Services and Health Planning.\textsuperscript{559} The Critical Access Hospital (CAH) program requires hospitals to be 15 miles apart in mountainous terrain and 35 miles apart otherwise.\textsuperscript{560} The average of 25 miles could be used for a radius distance, which is similar to what has been used in other accessibility studies. Finally, 10 miles could be used as a radius for the buffers since the Obama administration recently discussed excluding or cutting all CAH’s that are within 10 miles of one another at a cost-savings of $4 billion (Gold, 2011).

**Raster/Grid:** A data model in which geographic features are represented using discrete cells, generally squares, arranged as a (contiguous) rectangular grid. A single grid is essentially the same as a two-dimensional matrix, but is typically referenced from the lower left corner rather than the norm for matrices, which are referenced from the upper left. Raster files may have one or more values (attributes or bands) associated with each cell position or pixel.\textsuperscript{561}

**Rural County:** A geographic area (a county, sub-state level) that is identified as non-


metro by the 2003 Rural-Urban Continuum Code classification, which was created by the Office of Management and Budget and is used by the U.S. Census.

- For the purpose of Aim 1 and 2 of this study, every rural hospital’s the county of each rural hospital’s location

**Rural Health Clinic:** The Rural Health Clinic (RHC) status is awarded by the federal government to rural-based, primary care centers serving a patient mix typically insured by either Medicare or Medicaid. These facilities can be public, nonprofit, or for-profit, and must be located in a Health Professional Shortage Area (HPSA). RHCs must use a team approach of care with physicians working with non-physician practitioners including nurse practitioners (NP), physician assistants (PA), and certified nurse midwives (CNM). Clinic with RHC status, as determined by CMS, receive enhanced reimbursed rates for Medicare and Medicaid patients.

**Rural Hospital:** A term for all hospitals located in rural counties which includes Critical Access Hospitals as well as community hospitals.

**Rurality:** It is a measure of the degree of isolation derived from the population in the county and its proximity to a metro area. The 2003 Rural-Urban Continuum Code classification system is on a 1-9 scale, where codes 4-9 are non-metro or “rural.”

---


### TABLE 34: RURALITY CODES

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
<th># Counties</th>
<th>2000 population</th>
<th>Nickname</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Metro Counties: URBAN</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Counties in metro areas of 1 million population or more</td>
<td>413</td>
<td>149,224,067</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Counties in metro areas of 250,000 to 1 million population</td>
<td>325</td>
<td>55,514,159</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Counties in metro areas of fewer than 250,000 population</td>
<td>351</td>
<td>27,841,714</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Nonmetro Counties: RURAL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Urban population of 20,000 or more, adjacent to a metro area</td>
<td>218</td>
<td>14,442,161</td>
<td>Bedroom Counties</td>
</tr>
<tr>
<td>5</td>
<td>Urban population of 20,000 or more, not adjacent to a metro area</td>
<td>105</td>
<td>5,573,273</td>
<td>Lone Stars</td>
</tr>
<tr>
<td>6</td>
<td>Urban population of 2,500 to 19,999, adjacent to a metro area</td>
<td>609</td>
<td>15,134,357</td>
<td>Home Town Counties</td>
</tr>
<tr>
<td>7</td>
<td>Urban population of 2,500 to 19,999, not adjacent to a metro area</td>
<td>450</td>
<td>8,463,700</td>
<td>Small Town Counties</td>
</tr>
<tr>
<td>8</td>
<td>Completely rural or less than 2,500 urban population, adjacent to a metro area</td>
<td>235</td>
<td>2,425,743</td>
<td>Satellite Counties</td>
</tr>
<tr>
<td>9</td>
<td>Completely rural or less than 2,500 urban population, not adjacent to a metro area</td>
<td>435</td>
<td>2,802,732</td>
<td>Lonely Counties</td>
</tr>
</tbody>
</table>

**SCALE:** Scale can be about size (relative or absolute) and involves a fundamental set of issues in geography. Primarily, scale concerns space in geography, which can be referred to as “spatial scale.” However, geographers also consider the domains of temporal and thematic scale. The former deals with the size of time units; the latter is concerned with the groupings of entities or attributes like people or weather variables.\(^{564}\)

- **Analysis Scale:** This refers to the size of the unit at which some problem or phenomena is analyzed, such as at the count or state level. It is essentially the scale of understanding of geographic phenomena.
- **Cartographic Scale:** This refers to the depicted size of a feature on a map relative to its actual size in the world. The cartographic scale is traditionally expressed in one

---

of three ways – a verbal scale, a graphic scale bar, or the representative fraction (RF). The RF expresses scale as a numerical ratio of map distance to earth distance.

- **Phenomenon Scale:** This refers to the size at which human or physical earth structures or processes exist, regardless of how they are studied or represented. It is the ‘true’ scale of geographic phenomena

**SOCIAL CAPITAL:** For the purposes of this study, social capital will be an aggregated construct consisting of 20 factors in 7 general domains describing population characteristics, housing and home ownership, education, economic indicators, crime and violence, health, and community participation within a geographical area (e.g., a county). The 19 social capital factors were derived from 73 variables using 9 different data sources [See Appendix]. This definition of social capital and its components is consistent with recurring underlying dimensions of social capital.

**SPATIAL AUTOCORRELATION:** The degree of relationship that exists between two or more (spatial) variables, such that when one changes, the other(s) also change. This change can either be in the same direction, which is a positive autocorrelation, or in the opposite direction, which is a negative autocorrelation (AGI). The term autocorrelation is usually applied to ordered datasets, such as those relating to time series or spatial data ordered by distance band. The existence of such a relationship suggests but does not definitely establish causality.

**TOPOLOGY:** The relative location of geographic phenomena independent of their exact position. In digital data, topological relationships such as connectivity, adjacency and relative position are usually expressed as relationships between nodes, links and polygons. For example, the topology of a line includes its from- and to-nodes, and its left and right polygons (AGI). In mathematics, a property is said to be topological if it survives stretching and distorting of

---

VECTOR: Within GIS the term vector refers to data that are comprised of lines or arcs, defined by beginning and end points, which meet at nodes. The locations of these nodes and the topological structure are usually stored explicitly. Features are defined by their boundaries only and curved lines are represented as a series of connecting arcs. 2. In mathematics the term refers to a directed line, i.e. a line with a defined origin, direction, and orientation. The same term is used to refer to a single column or row of a matrix, in which case it is denoted by a bold letter, usually in lower case.⁵⁶⁷

⁵⁶⁷ Ibid
APPENDIX C: SOCIAL CAPITAL VARIABLE CHARACTERISTICS

TABLE 35: POPULATION LOSS

<table>
<thead>
<tr>
<th>Population Loss, 1990-2000 (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain 1: Population Characteristic</td>
<td></td>
</tr>
<tr>
<td>Descriptive</td>
<td></td>
</tr>
<tr>
<td><strong>Obs</strong></td>
<td><strong>Mean</strong></td>
</tr>
<tr>
<td>2015</td>
<td>0.0722</td>
</tr>
<tr>
<td>Percentiles</td>
<td></td>
</tr>
<tr>
<td>1%</td>
<td>0.1893</td>
</tr>
</tbody>
</table>

Graphic

| Correlation |  |
| Variables with HIGH Correlation | ρ | Variables with LOW Correlation | ρ |  |
| Percent Population Loss | 1.0000 | Percent Population in Poverty | -0.0892 |  |
| SC Index | -0.3604 | Percent Single Unmarried Parent | 0.0877 |  |
| Percent Presidential Voter Turnout | -0.2403 | Percent Uninsured | 0.0757 |  |
| Substance Abuse Crime | 0.2247 | Diversity Score | 0.0686 |  |
| Property Crime | 0.2188 | Percent Linguistic Isolation | 0.0575 |  |
| Percent High School Grad | -0.1915 | Percent Unemployed | 0.0548 |  |
| Violent Crime | 0.1530 | Physicians per 1000 | 0.0503 |  |
| Percent Vacant Housing | 0.1488 | Diversity Index | 0.0326 |  |
| Percent Owner Occupied Housing | -0.1484 | Gini Coefficient | 0.0043 |  |
| Percent Some College | 0.1103 | Per Capita Income | 0.0026 |  |
### TABLE 36: FAMILY HOUSEHOLDS

<table>
<thead>
<tr>
<th>Single Parent Family Households, 2000 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain 1: Population Characteristic</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Descriptive</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Obs</strong></td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>2049</td>
</tr>
<tr>
<td><strong>Percentiles</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>1%</strong></td>
</tr>
<tr>
<td>.0363</td>
</tr>
<tr>
<td><strong>Graphic</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><img src="image1.png" alt="Graphic" /></td>
</tr>
<tr>
<td><img src="image2.png" alt="Graphic" /></td>
</tr>
<tr>
<td><strong>Correlation</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Variables with HIGH Correlation</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Percent Unmarried Parent</strong></td>
</tr>
<tr>
<td><strong>Diversity Score</strong></td>
</tr>
<tr>
<td><strong>Percent Population in Poverty</strong></td>
</tr>
<tr>
<td><strong>SC Index</strong></td>
</tr>
<tr>
<td><strong>Percent Unemployed</strong></td>
</tr>
<tr>
<td><strong>Percent Presidential Voter Turnout</strong></td>
</tr>
<tr>
<td><strong>Percent Uninsured</strong></td>
</tr>
<tr>
<td><strong>Gini Coefficient</strong></td>
</tr>
<tr>
<td><strong>Per Capita Income</strong></td>
</tr>
<tr>
<td><strong>Violent Crime</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Variables with LOW Correlation</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Percent Some College</strong></td>
</tr>
<tr>
<td><strong>Diversity Index</strong></td>
</tr>
<tr>
<td><strong>Substance Abuse Crime</strong></td>
</tr>
<tr>
<td><strong>Property Crime</strong></td>
</tr>
<tr>
<td><strong>Percent High School Grad</strong></td>
</tr>
<tr>
<td><strong>Percent Vacant Housing</strong></td>
</tr>
<tr>
<td><strong>Physicians per 1000</strong></td>
</tr>
<tr>
<td><strong>Percent Population Loss</strong></td>
</tr>
<tr>
<td><strong>Percent Linguistic Isolation</strong></td>
</tr>
<tr>
<td><strong>Percent Owner Occupied Housing</strong></td>
</tr>
</tbody>
</table>
TABLE 37: DIVERSITY INDEX

Diversity Index, 2000†

Domain 1: Population Characteristic

Descriptive

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2052</td>
<td></td>
<td>0.0793</td>
<td>0.0756</td>
<td>0.0057</td>
<td>0</td>
<td>0.7943</td>
</tr>
</tbody>
</table>

Percentile

<table>
<thead>
<tr>
<th></th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0.0019</td>
<td>0.0315</td>
<td>0.0638</td>
<td>0.1057</td>
<td>0.1631</td>
<td>0.2059</td>
<td>0.3916</td>
</tr>
</tbody>
</table>

Graphic

Correlation

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation.</th>
<th>$\rho$</th>
<th>Variables with LOW Correlation.</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diversity Index</td>
<td>1.0000</td>
<td>Physicians per 1000</td>
<td>0.1196</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
<td>0.2997</td>
<td>Percent Vacant Housing</td>
<td>-0.1106</td>
</tr>
<tr>
<td>Substance Abuse Crime</td>
<td>0.2846</td>
<td>Percent Some College</td>
<td>-0.0830</td>
</tr>
<tr>
<td>Violent Crime</td>
<td>0.2842</td>
<td>Per Capita Income</td>
<td>-0.0774</td>
</tr>
<tr>
<td>Property Crime</td>
<td>0.2813</td>
<td>Gini Coefficient</td>
<td>0.0485</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>-0.2099</td>
<td>Percent Linguistic Isolation</td>
<td>-0.0330</td>
</tr>
<tr>
<td>Diversity Score</td>
<td>0.1943</td>
<td>Percent Population Loss</td>
<td>0.0326</td>
</tr>
<tr>
<td>Percent Unemployed</td>
<td>0.1772</td>
<td>Percent Owner Occupied Housing</td>
<td>0.0160</td>
</tr>
<tr>
<td>SC Index</td>
<td>-0.1687</td>
<td>Percent Uninsured</td>
<td>0.0070</td>
</tr>
<tr>
<td>Percent Population in Poverty</td>
<td>0.1542</td>
<td>Percent High School Grad</td>
<td>-0.0038</td>
</tr>
</tbody>
</table>

†As these variables increase, social capital also is expected to increase.
TABLE 38: LINGUISTICALLY ISOLATED HOUSEHOLDS

Linguistically – Isolated Households, 2000 (%)

Domain 1: Population Characteristic

Descriptive

<table>
<thead>
<tr>
<th>Obs</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2052</td>
<td>0.0158</td>
<td>0.0314</td>
<td>0.0010</td>
<td>0</td>
<td>0.3389</td>
</tr>
</tbody>
</table>

Percentiles

<table>
<thead>
<tr>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.0013</td>
<td>0.0028</td>
<td>0.0059</td>
<td>0.0136</td>
<td>0.0378</td>
<td>0.0668</td>
<td>0.1670</td>
</tr>
</tbody>
</table>

Graphic

[Graph showing correlation]

Correlation

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>P</th>
<th>Variables with LOW Correlation</th>
<th>ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Linguistic Isolation</td>
<td>1.0000</td>
<td>Per Capita Income</td>
<td>-0.0934</td>
</tr>
<tr>
<td>Percent Uninsured</td>
<td>0.4568</td>
<td>Percent Single Unmarried Parent</td>
<td>0.0845</td>
</tr>
<tr>
<td>Percent High School Grad</td>
<td>-0.3909</td>
<td>Substance Abuse Crime</td>
<td>0.0604</td>
</tr>
<tr>
<td>Percent Population in Poverty</td>
<td>0.2284</td>
<td>Percent Some College</td>
<td>-0.0591</td>
</tr>
<tr>
<td>Diversity Scale</td>
<td>0.2155</td>
<td>Percent Vacant Housing</td>
<td>0.0587</td>
</tr>
<tr>
<td>Percent Unemployed</td>
<td>0.1837</td>
<td>Percent Population Loss</td>
<td>0.0575</td>
</tr>
<tr>
<td>SC Index</td>
<td>-0.1405</td>
<td>Diversity Index</td>
<td>-0.0330</td>
</tr>
<tr>
<td>Percent Owner Occupied Housing</td>
<td>-0.1316</td>
<td>Physicians per 1000</td>
<td>-0.0286</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>-0.1315</td>
<td>Violent Crime</td>
<td>0.0252</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>0.1171</td>
<td>Property Crime</td>
<td>0.0203</td>
</tr>
<tr>
<td>Variables with HIGH Correlation</td>
<td>ρ</td>
<td>Variables with LOW Correlation</td>
<td>p</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----</td>
<td>----------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Diversity Scale</td>
<td>1.0000</td>
<td>Per Capita Income</td>
<td>-0.2797</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
<td>0.7202</td>
<td>Percent Some College</td>
<td>-0.2303</td>
</tr>
<tr>
<td>Percent Population in Poverty</td>
<td>0.5085</td>
<td>Percent Linguistic Isolation</td>
<td>0.2155</td>
</tr>
<tr>
<td>Percent Uninsured</td>
<td>0.4973</td>
<td>Substance Abuse Crime</td>
<td>0.2086</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>0.4631</td>
<td>Diversity Index</td>
<td>0.1943</td>
</tr>
<tr>
<td>SC Index</td>
<td>-0.4310</td>
<td>Property Crime</td>
<td>0.1863</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>-0.4201</td>
<td>Physicians per 1000</td>
<td>0.0865</td>
</tr>
<tr>
<td>Violent Crime</td>
<td>0.3308</td>
<td>Percent Vacant Housing</td>
<td>-0.0746</td>
</tr>
<tr>
<td>Percent Unemployed</td>
<td>0.3302</td>
<td>Percent Owner Occupied Housing</td>
<td>-0.0746</td>
</tr>
<tr>
<td>Percent High School Grad</td>
<td>-0.3271</td>
<td>Percent Population Loss</td>
<td>0.0686</td>
</tr>
</tbody>
</table>

†As these variables increase, social capital also is expected to increase.
### TABLE 40: VACANT HOUSING

**Vacant Housing, 2000 (%)**  
Domain 2: Housing & Home Ownership

#### Descriptive

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vacant Housing, 2000 (%)</td>
<td>2049</td>
<td>0.1696</td>
<td>0.1019</td>
<td>0.0104</td>
<td>0.0346</td>
<td>0.7701</td>
</tr>
</tbody>
</table>

#### Percentiles

<table>
<thead>
<tr>
<th></th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0528</td>
<td>0.0653</td>
<td>0.0769</td>
<td>0.1021</td>
<td>0.1395</td>
<td>0.2033</td>
<td>0.3091</td>
<td>0.3697</td>
<td>0.5530</td>
</tr>
</tbody>
</table>

#### Graphic

![Graph](image)

#### Correlation

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>P</th>
<th>Variables with LOW Correlation</th>
<th>ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Vacant Housing</td>
<td>1.0000</td>
<td>Diversity Index</td>
<td>-0.1106</td>
</tr>
<tr>
<td>Percent Owner Occupied Housing</td>
<td>-0.8057</td>
<td>Percent Unemployed</td>
<td>0.1091</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>0.2949</td>
<td>Percent High School Grad</td>
<td>-0.0886</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
<td>-0.1808</td>
<td>Physicians per 1000</td>
<td>-0.0861</td>
</tr>
<tr>
<td>Substance Abuse Crime</td>
<td>-0.1785</td>
<td>Gini Coefficient</td>
<td>0.0753</td>
</tr>
<tr>
<td>Property Crime</td>
<td>-0.1672</td>
<td>Diversity Scale</td>
<td>-0.0746</td>
</tr>
<tr>
<td>Percent Some College</td>
<td>0.1617</td>
<td>SC Index</td>
<td>0.0633</td>
</tr>
<tr>
<td>Percent Uninsured</td>
<td>0.1552</td>
<td>Percent Linguistic Isolation</td>
<td>0.0587</td>
</tr>
<tr>
<td>Percent Population Loss</td>
<td>0.1488</td>
<td>Per Capita Income</td>
<td>-0.0183</td>
</tr>
<tr>
<td>Violent Crime</td>
<td>-0.1468</td>
<td>Percent Population in Poverty</td>
<td>-0.0004</td>
</tr>
</tbody>
</table>
TABLE 41: OWNER OCCUPIED HOUSING

**Owner-Occupied Housing, 2000 (%)†**

**Domain 2: Housing & Home Ownership**

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>2049</td>
<td>0.6204</td>
<td>0.0863</td>
<td>0.0074</td>
<td>0.00</td>
<td>0.8046</td>
</tr>
</tbody>
</table>

**Percentiles**

<table>
<thead>
<tr>
<th></th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Owner Occupied Housing</td>
<td>0.3118</td>
<td>0.4606</td>
<td>0.5156</td>
<td>0.5809</td>
<td>0.6353</td>
<td>0.6780</td>
<td>0.7099</td>
<td>0.7264</td>
<td>0.7546</td>
</tr>
</tbody>
</table>

**Graphic**


**Correlation**

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>P</th>
<th>Variables with LOW Correlation</th>
<th>ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Owner Occupied Housing</td>
<td>1.0000</td>
<td>Percent Unemployed</td>
<td>-0.0947</td>
</tr>
<tr>
<td>Percent Vacant Housing</td>
<td>-0.8057</td>
<td>Diversity Scale</td>
<td>-0.0746</td>
</tr>
<tr>
<td>Percent Some College</td>
<td>-0.3422</td>
<td>Percent Single Unmarried Parent</td>
<td>-0.0441</td>
</tr>
<tr>
<td>Percent High School Grad</td>
<td>0.3156</td>
<td>Property Crime</td>
<td>-0.0237</td>
</tr>
<tr>
<td>Percent Uninsured</td>
<td>-0.2713</td>
<td>SC Index</td>
<td>-0.0226</td>
</tr>
<tr>
<td>Percent Population Loss</td>
<td>-0.1484</td>
<td>Diversity Index</td>
<td>0.0160</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>-0.1462</td>
<td>Physicians per 1000</td>
<td>0.0094</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>-0.1361</td>
<td>Violent Crime</td>
<td>-0.0080</td>
</tr>
<tr>
<td>Percent Linguistic Isolation</td>
<td>-0.1316</td>
<td>Substance Abuse Crime</td>
<td>-0.0073</td>
</tr>
<tr>
<td>Percent Population in Poverty</td>
<td>-0.1271</td>
<td>Per Capita Income</td>
<td>-0.0054</td>
</tr>
</tbody>
</table>

†As these variables increase, social capital also is expected to increase.
TABLE 42: HIGH SCHOOL GRADUATES

<table>
<thead>
<tr>
<th>High School Graduates, 2000 (%)†</th>
<th>Domain 3: Education</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Descriptive</strong></td>
<td></td>
</tr>
<tr>
<td>Obs</td>
<td>Mean</td>
</tr>
<tr>
<td>2049</td>
<td>0.3588</td>
</tr>
<tr>
<td><strong>Percentiles</strong></td>
<td></td>
</tr>
<tr>
<td>1%</td>
<td>5%</td>
</tr>
<tr>
<td>0.2058</td>
<td>0.2644</td>
</tr>
<tr>
<td><strong>Graphic</strong></td>
<td></td>
</tr>
</tbody>
</table>

| **Correlation**                |                     |
| Variables with HIGH Correlation | P   | Variables with LOW Correlation | P   |
| Percent High School Grad        | 1.0000 | Substance Abuse Crime | -0.1728 |
| Percent Some College            | -0.4557 | Property Crime | -0.1555 |
| Percent Uninsured               | -0.4366 | Violent Crime | -0.1528 |
| Percent Linguistic Isolation    | -0.3909 | Percent Unemployed | -0.1309 |
| Gini Coefficient                | -0.3306 | Percent Vacant Housing | -0.0886 |
| Diversity Scale                 | -0.3271 | SC Index | 0.0862 |
| Percent Owner Occupied Housing  | 0.3156 | Percent Presidential Voter Turnout | 0.0815 |
| Percent Population in Poverty    | -0.3054 | Physicians per 1000 | -0.0573 |
| Percent Single Unmarried Parent  | -0.2104 | Per Capita Income | -0.0304 |
| Percent Population Loss         | -0.1915 | Diversity Index | -0.0038 |

†As these variables increase, social capital also is expected to increase.
### TABLE 43: SOME COLLEGE OR HIGHER COMPLETION

#### Domain 3: Education

**Some College or Higher Completion, 2000 (%)†**

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive</td>
<td>2049</td>
<td>0.4000</td>
<td>0.1001</td>
<td>0.0100</td>
<td>0.1691</td>
<td>0.8539</td>
</tr>
</tbody>
</table>

#### Percentiles

<table>
<thead>
<tr>
<th></th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentiles</td>
<td>0.2110</td>
<td>0.2510</td>
<td>0.2777</td>
<td>0.3239</td>
<td>0.3933</td>
<td>0.4659</td>
<td>0.5251</td>
<td>0.5651</td>
<td>0.6825</td>
</tr>
</tbody>
</table>

#### Graphic

![Graph showing distribution of some college or higher completion](image)

#### Correlation

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>P</th>
<th>Variables with LOW Correlation</th>
<th>ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Some College</td>
<td>1.0000</td>
<td>Diversity Scale</td>
<td>-0.2303</td>
</tr>
<tr>
<td>Per Capita Income</td>
<td>0.4945</td>
<td>Percent Uninsured</td>
<td>-0.1888</td>
</tr>
<tr>
<td>SC Index</td>
<td>0.4649</td>
<td>Percent Vacant Housing</td>
<td>0.1617</td>
</tr>
<tr>
<td>Percent High School Grad</td>
<td>-0.4557</td>
<td>Property Crime</td>
<td>0.1396</td>
</tr>
<tr>
<td>Percent Population in Poverty</td>
<td>-0.4397</td>
<td>Percent Population Loss</td>
<td>0.1103</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>0.4367</td>
<td>Diversity Index</td>
<td>-0.0830</td>
</tr>
<tr>
<td>Percent Unemployed</td>
<td>-0.3699</td>
<td>Substance Abuse Crime</td>
<td>0.0701</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>-0.3453</td>
<td>Percent Linguistic Isolation</td>
<td>-0.0591</td>
</tr>
<tr>
<td>Percent Owner Occupied Housing</td>
<td>-0.3422</td>
<td>Physicians per 1000</td>
<td>0.0357</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
<td>-0.3258</td>
<td>Violent Crime</td>
<td>0.0074</td>
</tr>
</tbody>
</table>

†As these variables increase, social capital also is expected to increase.
### TABLE 44: GINI COEFFICIENT

**Gini Coefficient, 2000**  
**Domain 4: Economic Indicators**

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2024</strong></td>
<td></td>
<td>0.4389</td>
<td>0.0374</td>
<td>0.0014</td>
<td>0.3360</td>
<td>0.6049</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.3645</td>
<td>0.3819</td>
<td>0.3924</td>
<td>0.4125</td>
<td>0.4365</td>
<td>0.4638</td>
<td>0.4878</td>
<td>0.5025</td>
<td>0.5334</td>
</tr>
</tbody>
</table>

**Graphic**  

![Graph showing distribution of Gini Coefficient](image)

**Correlation**

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>P</th>
<th>Variables with LOW Correlation</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gini Coefficient</td>
<td>1.0000</td>
<td>Presidential Voter Turnout (%)</td>
<td>-0.3225</td>
</tr>
<tr>
<td>Percent Population in Poverty</td>
<td>0.7007</td>
<td>Owner Occupied Housing (%)</td>
<td>-0.1462</td>
</tr>
<tr>
<td>Percent Uninsured</td>
<td>0.5187</td>
<td>Violent Crime</td>
<td>0.1255</td>
</tr>
<tr>
<td>Diversity Scale</td>
<td>0.4631</td>
<td>Percent Linguistic Isolation</td>
<td>0.1171</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
<td>0.4570</td>
<td>Physicians per 1000</td>
<td>0.0845</td>
</tr>
<tr>
<td>Percent Unemployed</td>
<td>0.4173</td>
<td>Percent Vacant Housing</td>
<td>0.0753</td>
</tr>
<tr>
<td>SC Index</td>
<td>-0.3742</td>
<td>Substance Abuse Crime</td>
<td>0.0654</td>
</tr>
<tr>
<td>Per Capita Income</td>
<td>-0.3471</td>
<td>Diversity Index</td>
<td>0.0485</td>
</tr>
<tr>
<td>Percent Some College</td>
<td>-0.3453</td>
<td>Property Crime</td>
<td>0.0065</td>
</tr>
<tr>
<td>Percent High School Grad</td>
<td>-0.3306</td>
<td>Percent Population Loss</td>
<td>0.0043</td>
</tr>
</tbody>
</table>
### TABLE 45: UNEMPLOYMENT

**Unemployment, 2000 (%)**

**Domain 4: Economic Indicators**

<table>
<thead>
<tr>
<th>Obs</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2050</td>
<td>0.0464</td>
<td>0.0178</td>
<td>0.0003</td>
<td>0.0150</td>
<td>0.1680</td>
</tr>
</tbody>
</table>

**Percentiles**

<table>
<thead>
<tr>
<th></th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0210</td>
<td>0.0250</td>
<td>0.0270</td>
<td>0.0340</td>
<td>0.0430</td>
<td>0.0550</td>
<td>0.0690</td>
<td>0.0790</td>
<td>0.1020</td>
</tr>
</tbody>
</table>

**Graphic**

#### Correlation

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>ρ</th>
<th>Variables with LOW Correlation</th>
<th>ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Unemployed</td>
<td>1.0000</td>
<td>Percent Linguistic Isolation</td>
<td>0.1837</td>
</tr>
<tr>
<td>Percent Population in Poverty</td>
<td>0.5976</td>
<td>Diversity Index</td>
<td>0.1772</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
<td>0.5044</td>
<td>Percent High School Grad</td>
<td>-0.1309</td>
</tr>
<tr>
<td>Per Capita Income</td>
<td>-0.4443</td>
<td>Violent Crime</td>
<td>0.1179</td>
</tr>
<tr>
<td>Percent Uninsured</td>
<td>0.4367</td>
<td>Percent Vacant Housing</td>
<td>0.1091</td>
</tr>
<tr>
<td>SC Index</td>
<td>-0.4291</td>
<td>Percent Owner Occupied Housing</td>
<td>-0.0947</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>0.4173</td>
<td>Substance Abuse Crime</td>
<td>0.0725</td>
</tr>
<tr>
<td>Percent Some College</td>
<td>-0.3699</td>
<td>Percent Population Loss</td>
<td>0.0548</td>
</tr>
<tr>
<td>Diversity Scale</td>
<td>0.3302</td>
<td>Physicians per 1000</td>
<td>0.0536</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>-0.2253</td>
<td>Property Crime</td>
<td>0.0439</td>
</tr>
</tbody>
</table>
### TABLE 46: PER CAPITAL INCOME

**Per Capita Income, 2000*†**

Domain 4: Economic Indicators

**Descriptive**

<table>
<thead>
<tr>
<th>Obs</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2031</td>
<td>-9.56e-07</td>
<td>1</td>
<td>1</td>
<td>-3.2577</td>
<td>12.1034</td>
</tr>
</tbody>
</table>

**Percentiles**

<table>
<thead>
<tr>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.8984</td>
<td>-1.3902</td>
<td>-1.0755</td>
<td>-0.6158</td>
<td>-0.0506</td>
<td>0.4920</td>
<td>1.0582</td>
<td>1.4557</td>
<td>3.0636</td>
</tr>
</tbody>
</table>

**Graphic**

* These variables have been normalized, so as to allow factor analysis. The normalization technique used was standardization (z-scores) by subtracting the mean from the variable and then dividing it by the standard deviation. Without normalization, indicators with extreme values have a greater effect on the composite indicator.

† As these variables increase, social capital also is expected to increase.
### TABLE 47: POPULATION IN POVERTY

#### Domain 4: Economic Indicators

##### Descriptive

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2049</td>
<td>0.1500</td>
<td>0.0650</td>
<td>0.0042</td>
<td>0</td>
<td>0.5669</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0552</td>
<td>0.0712</td>
<td>0.0809</td>
<td>0.1026</td>
<td>0.1379</td>
<td>0.1820</td>
<td>0.2325</td>
<td>0.2682</td>
<td>0.3609</td>
</tr>
</tbody>
</table>

#### Graphic

![Histogram and Box Plot](image)

#### Correlation

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>$\rho$</th>
<th>Variables with LOW Correlation</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Population in Poverty</td>
<td>1.0000</td>
<td>Percent High School Grad</td>
<td>-0.3054</td>
</tr>
<tr>
<td>Percent Uninsured</td>
<td>0.7475</td>
<td>Percent Linguistic Isolation</td>
<td>0.2284</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>0.7007</td>
<td>Diversity Index</td>
<td>0.1542</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
<td>0.6603</td>
<td>Percent Owner Occupied Housing</td>
<td>-0.1271</td>
</tr>
<tr>
<td>Per Capita Income</td>
<td>-0.6118</td>
<td>Percent Population Loss</td>
<td>-0.0892</td>
</tr>
<tr>
<td>Percent Unemployed</td>
<td>0.5976</td>
<td>Violent Crime</td>
<td>0.0864</td>
</tr>
<tr>
<td>Diversity Scale</td>
<td>0.5085</td>
<td>Property Crime</td>
<td>-0.0564</td>
</tr>
<tr>
<td>SC Index</td>
<td>-0.4561</td>
<td>Physicians per 1000</td>
<td>0.0253</td>
</tr>
<tr>
<td>Percent Some College</td>
<td>-0.4397</td>
<td>Percent Vacant Housing</td>
<td>-0.0004</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>-0.3651</td>
<td>Substance Abuse Crime</td>
<td>0.0001</td>
</tr>
</tbody>
</table>
**TABLE 48: VIOLENT CRIME**

<table>
<thead>
<tr>
<th>Violent Crimes, 2000*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain 5: Crime &amp; Violence</td>
</tr>
<tr>
<td><strong>Descriptive</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>2049</td>
</tr>
<tr>
<td><strong>Percentiles</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>-0.5837</td>
</tr>
<tr>
<td><strong>Graphic</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variables with HIGH Correlation</strong></td>
</tr>
<tr>
<td>Violent Crime</td>
</tr>
<tr>
<td>Property Crime</td>
</tr>
<tr>
<td>Substance Abuse Crime</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
</tr>
<tr>
<td>Diversity Scale</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
</tr>
<tr>
<td>Diversity Index</td>
</tr>
<tr>
<td>SC Index</td>
</tr>
<tr>
<td>Physicians per 1000</td>
</tr>
<tr>
<td>Percent Population Loss</td>
</tr>
</tbody>
</table>

*These variables have been normalized, so as to allow factor analysis. The normalization technique used was standardization (z-scores) by subtracting the mean from the variable and then dividing it by the standard deviation. Without normalization, indicators with extreme values have a greater effect on the composite indicator.*
### TABLE 49: PROPERTY CRIMES

**Property Crimes, 2000***

Domain 5: Crime & Violence

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs</td>
<td>Mean</td>
<td>Std</td>
<td>Variance</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>2049</td>
<td>3.74e-08</td>
<td>1</td>
<td>1</td>
<td>-0.6616</td>
<td>9.5051</td>
</tr>
</tbody>
</table>

**Percentiles**

<table>
<thead>
<tr>
<th></th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.6616</td>
<td>-0.6616</td>
<td>-0.6071</td>
<td>-0.4007</td>
<td>0.1958</td>
<td>1.1696</td>
<td>1.9168</td>
<td>3.8519</td>
<td></td>
</tr>
</tbody>
</table>

**Graphic**

![Graphic showing data distribution]

**Correlation**

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>P</th>
<th>Variables with LOW Correlation</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Property Crime</td>
<td>1.0000</td>
<td>Percent Vacant Housing</td>
<td>-0.1672</td>
</tr>
<tr>
<td>Violent Crime</td>
<td>0.8156</td>
<td>Percent High School Grad</td>
<td>-0.1555</td>
</tr>
<tr>
<td>Substance Abuse Crime</td>
<td>0.7773</td>
<td>Per Capita Income</td>
<td>0.1487</td>
</tr>
<tr>
<td>Diversity Index</td>
<td>0.2813</td>
<td>Percent Some College</td>
<td>0.1396</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>-0.2485</td>
<td>Percent Population in Poverty</td>
<td>-0.0564</td>
</tr>
<tr>
<td>Percent Single Unmarri Parent</td>
<td>0.2362</td>
<td>Percent Uninsured</td>
<td>-0.0523</td>
</tr>
<tr>
<td>Physicians per 1000</td>
<td>0.2338</td>
<td>Percent Unemployed</td>
<td>0.0439</td>
</tr>
<tr>
<td>SC Index</td>
<td>-0.2236</td>
<td>Percent Owner Occup. Housing</td>
<td>-0.0237</td>
</tr>
<tr>
<td>Percent Population Loss</td>
<td>0.2188</td>
<td>Percent Linguistic Isolation</td>
<td>0.0203</td>
</tr>
<tr>
<td>Diversity Scale</td>
<td>0.1863</td>
<td>Gini Coefficient</td>
<td>0.0065</td>
</tr>
</tbody>
</table>

*These variables have been normalized, so as to allow factor analysis. The normalization technique used was standardization (z-scores) by subtracting the mean from the variable and then dividing it by the standard deviation. Without normalization, indicators with extreme values have a greater effect on the composite indicator.
TABLE 50: SUBSTANCE ABUSE CRIMES

Substance Abuse Crimes, 2000*
Domain 5: Crime & Violence

<table>
<thead>
<tr>
<th></th>
<th>Observations</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1956</td>
<td>6.75e-08</td>
<td>1</td>
<td>1</td>
<td>-0.8332</td>
<td>9.5682</td>
<td></td>
</tr>
</tbody>
</table>

Percentiles

<table>
<thead>
<tr>
<th>Percentile</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.8332</td>
<td>-0.8138</td>
<td>-0.7874</td>
<td>-0.6553</td>
<td>-0.3313</td>
<td>0.2860</td>
<td>1.1868</td>
<td>1.9459</td>
<td>3.9959</td>
</tr>
</tbody>
</table>

Graphic

![Graph showing distribution of Substance Abuse Crimes](image)

Correlation

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>P</th>
<th>Variables with LOW Correlation</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substance Abuse Crime</td>
<td>1.0000</td>
<td>Percent Vacant Housing</td>
<td>-0.1785</td>
</tr>
<tr>
<td>Property Crime</td>
<td>0.7773</td>
<td>Percent High School Grad</td>
<td>-0.1728</td>
</tr>
<tr>
<td>Violent Crime</td>
<td>0.6617</td>
<td>Per Capita Income</td>
<td>0.1065</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>-0.2991</td>
<td>Percent Unemployed</td>
<td>0.0725</td>
</tr>
<tr>
<td>SC Index</td>
<td>-0.2917</td>
<td>Percent Some College</td>
<td>0.0701</td>
</tr>
<tr>
<td>Diversity Index</td>
<td>0.2846</td>
<td>Gini Coefficient</td>
<td>0.0654</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
<td>0.2617</td>
<td>Percent Linguistic Isolation</td>
<td>0.0604</td>
</tr>
<tr>
<td>Percent Population Loss</td>
<td>0.2247</td>
<td>Percent Uninsured</td>
<td>-0.0135</td>
</tr>
<tr>
<td>Physicians per 1000</td>
<td>0.2198</td>
<td>Percent Owner Occupied Housing</td>
<td>-0.0073</td>
</tr>
<tr>
<td>Diversity Scale</td>
<td>0.2086</td>
<td>Percent Population in Poverty</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

*These variables have been normalized, so as to allow factor analysis. The normalization technique used was standardization (z-scores) by subtracting the mean from the variable and then dividing it by the standard deviation. Without normalization, indicators with extreme values have a greater effect on the composite indicator.
TABLE 51: UNINSURED

Uninsured, 2000 (%)
Domain 6: Health

Descriptive

<table>
<thead>
<tr>
<th>Obs</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2050</td>
<td>0.1523</td>
<td>0.0498</td>
<td>0.0025</td>
<td>0.0437</td>
<td>0.3856</td>
</tr>
</tbody>
</table>

Percentiles

<table>
<thead>
<tr>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0630</td>
<td>0.0820</td>
<td>0.0928</td>
<td>0.1151</td>
<td>0.1473</td>
<td>0.1830</td>
<td>0.2164</td>
<td>0.2414</td>
<td>0.2899</td>
</tr>
</tbody>
</table>

Graphic

Correlation

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>P</th>
<th>Variables with LOW Correlation</th>
<th>ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Uninsured</td>
<td>1.0000</td>
<td>Percent Owner Occupied Housing</td>
<td>-0.2713</td>
</tr>
<tr>
<td>Percent Population in Poverty</td>
<td>0.7475</td>
<td>Presidential Voter Turnout (%)</td>
<td>-0.2428</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>0.5187</td>
<td>Percent Some College</td>
<td>-0.1888</td>
</tr>
<tr>
<td>Per Capita Income</td>
<td>-0.5086</td>
<td>Percent Vacant Housing</td>
<td>0.1552</td>
</tr>
<tr>
<td>Diversity Scale</td>
<td>0.4973</td>
<td>Violent Crime</td>
<td>0.0783</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
<td>0.4626</td>
<td>Percent Population Loss</td>
<td>0.0757</td>
</tr>
<tr>
<td>Percent Linguistic Isolation</td>
<td>0.4568</td>
<td>Property Crime</td>
<td>-0.0523</td>
</tr>
<tr>
<td>Percent Unemployed</td>
<td>0.4367</td>
<td>Substance Abuse Crime</td>
<td>-0.0135</td>
</tr>
<tr>
<td>Percent High School Grad</td>
<td>-0.4366</td>
<td>Diversity Index</td>
<td>0.0070</td>
</tr>
<tr>
<td>SC Index</td>
<td>-0.3810</td>
<td>Physicians per 1000</td>
<td>0.0012</td>
</tr>
</tbody>
</table>
TABLE 52: FTE PHYSICIANS PER 1000

<table>
<thead>
<tr>
<th>FTE Physicians per 1000, 2000†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain 6: Health</td>
</tr>
<tr>
<td><strong>Descriptive</strong></td>
</tr>
<tr>
<td>Obs</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>2052</td>
</tr>
<tr>
<td><strong>Percentiles</strong></td>
</tr>
<tr>
<td>1%</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td><strong>Graphic</strong></td>
</tr>
<tr>
<td><img src="image_url" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<p>| <strong>Correlation</strong>                |</p>
<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>P</th>
<th>Variables with LOW Correlation</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physicians per 1000</td>
<td>1.0000</td>
<td>SC Index</td>
<td>-0.0815</td>
</tr>
<tr>
<td>Property Crime</td>
<td>0.2338</td>
<td>Per Capita Income</td>
<td>0.0603</td>
</tr>
<tr>
<td>Substance Abuse Crime</td>
<td>0.2198</td>
<td>Percent High School Grad</td>
<td>-0.0573</td>
</tr>
<tr>
<td>Violent Crime</td>
<td>0.1777</td>
<td>Percent Unemployed</td>
<td>0.0536</td>
</tr>
<tr>
<td>Diversity Index</td>
<td>0.1196</td>
<td>Percent Population Loss</td>
<td>0.0503</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
<td>0.1053</td>
<td>Percent Some College</td>
<td>0.0357</td>
</tr>
<tr>
<td>Diversity Scale</td>
<td>0.0865</td>
<td>Percent Linguistic Isolation</td>
<td>-0.0286</td>
</tr>
<tr>
<td>Percent Vacant Housing</td>
<td>-0.0861</td>
<td>Percent Population in Poverty</td>
<td>0.0253</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>0.0845</td>
<td>Percent Owner Occupied Housing</td>
<td>0.0094</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>-0.0836</td>
<td>Percent Uninsured</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

†As these variables increase, social capital also is expected to increase.
## TABLE 53: COMPOSITE SOCIAL CAPITAL INDEX

### Composite Social Capital Index^†

**Domain 7: Community Participation**

### Descriptive

<table>
<thead>
<tr>
<th>Obs</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Variance</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>2023</td>
<td>1.48e-08</td>
<td>1</td>
<td>1</td>
<td>-2.2577</td>
<td>8.2559</td>
</tr>
</tbody>
</table>

### Percentiles

<table>
<thead>
<tr>
<th>%</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>-1.5399</td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>-1.2239</td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>-1.0242</td>
<td></td>
</tr>
<tr>
<td>25%</td>
<td>-0.7128</td>
<td></td>
</tr>
<tr>
<td>50%</td>
<td>-0.1917</td>
<td></td>
</tr>
<tr>
<td>75%</td>
<td>0.5022</td>
<td></td>
</tr>
<tr>
<td>90%</td>
<td>1.2859</td>
<td></td>
</tr>
<tr>
<td>95%</td>
<td>1.8865</td>
<td></td>
</tr>
<tr>
<td>99%</td>
<td>3.2022</td>
<td></td>
</tr>
</tbody>
</table>

### Graphic

![Graphical representation of the data]

### Correlation

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>P</th>
<th>Variables with LOW Correlation</th>
<th>ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC Index</td>
<td>1.0000</td>
<td>Percent Population Loss</td>
<td>-0.3604</td>
</tr>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>0.6022</td>
<td>Substance Abuse Crime</td>
<td>-0.2917</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
<td>-0.5178</td>
<td>Violent Crime</td>
<td>-0.2751</td>
</tr>
<tr>
<td>Percent Some College</td>
<td>0.4649</td>
<td>Property Crime</td>
<td>-0.2236</td>
</tr>
<tr>
<td>Percent Population in Poverty</td>
<td>-0.4561</td>
<td>Diversity Index</td>
<td>-0.1687</td>
</tr>
<tr>
<td>Per Capita Income</td>
<td>0.4370</td>
<td>Percent Linguistic Isolation</td>
<td>-0.1405</td>
</tr>
<tr>
<td>Diversity Scale</td>
<td>-0.4310</td>
<td>Percent High School Grad</td>
<td>0.0862</td>
</tr>
<tr>
<td>Percent Unemployed</td>
<td>-0.4291</td>
<td>Physicians per 1000</td>
<td>-0.0815</td>
</tr>
<tr>
<td>Percent Uninsured</td>
<td>-0.3810</td>
<td>Percent Vacant Housing</td>
<td>0.0633</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>-0.3742</td>
<td>Percent Owner Occupied Housing</td>
<td>-0.0226</td>
</tr>
</tbody>
</table>

^Although this is titled a “Social Capital Index,” it was developed by Penn State University and only measures the degree of community participation.

†As these variables increase, social capital also is expected to increase.
TABLE 54: PRESIDENTIAL TURNOUT

<table>
<thead>
<tr>
<th>Presidential Voter Turnout, 2000 (%)†</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain 7: Community Participation</td>
<td></td>
</tr>
<tr>
<td><strong>Descriptive</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Obs</strong></td>
<td><strong>Mean</strong></td>
</tr>
<tr>
<td>2049</td>
<td>0.5355</td>
</tr>
<tr>
<td><strong>Percentiles</strong></td>
<td></td>
</tr>
<tr>
<td>1%</td>
<td>0</td>
</tr>
</tbody>
</table>

**Graphic**

**Correlation**

<table>
<thead>
<tr>
<th>Variables with HIGH Correlation</th>
<th>P</th>
<th>Variables with LOW Correlation</th>
<th>ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent Presidential Voter Turnout</td>
<td>1.0000</td>
<td>Percent Vacant Housing</td>
<td>0.2949</td>
</tr>
<tr>
<td>SC Index</td>
<td>0.6022</td>
<td>Property Crime</td>
<td>-0.2485</td>
</tr>
<tr>
<td>Percent Single Unmarried Parent</td>
<td>-0.4688</td>
<td>Percent Uninsured</td>
<td>-0.2428</td>
</tr>
<tr>
<td>Percent Some College</td>
<td>0.4367</td>
<td>Percent Population Loss</td>
<td>-0.2403</td>
</tr>
<tr>
<td>Per Capita Income</td>
<td>0.4235</td>
<td>Percent Unemployed</td>
<td>-0.2253</td>
</tr>
<tr>
<td>Diversity Scale</td>
<td>-0.4201</td>
<td>Diversity Index</td>
<td>-0.2099</td>
</tr>
<tr>
<td>Percent Population in Poverty</td>
<td>-0.3651</td>
<td>Percent Owner Occupied Housing</td>
<td>-0.1361</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>-0.3225</td>
<td>Percent Linguistic Isolation</td>
<td>-0.1315</td>
</tr>
<tr>
<td>Violent Crime</td>
<td>-0.3016</td>
<td>Physicians per 1000</td>
<td>-0.0836</td>
</tr>
<tr>
<td>Substance Abuse Crime</td>
<td>-0.2991</td>
<td>Percent High School Grad</td>
<td>0.0815</td>
</tr>
</tbody>
</table>

†As these variables increase, social capital also is expected to increase.
APPENDIX D. HUMAN SUBJECTS PROTECTION

This study is being conducted for a doctoral dissertation and for publication of the results. No primary data collection through surveys, interviews, questionnaires, other tests or instruments will be used. No audio or video recording data will be utilized.

All data being used for this study comes from one of the following three public, secondary data sources: the American Hospital Association (AHA), RTI International’s Spatial Impact Factor Database, and USA Counties (a database run by the U.S. Census Bureau). The AHA data is an existing data set developed by the AHA from its Annual Survey of hospitals. It is available for purchase from the AHA and regularly used in health services research.\(^{568,569}\)

The Spatial Impact Factor Database (SIFD) is a publicly-available database overseen by one of RTI’s Senior Fellows Lee Rivers Mobley through a grant from the National Institutes of Health (R01CA126858-01A1). The latest version of the database, updated in November of 2011, is available online at the following website: rtispatialdata.rti.org. It includes secondary data from a wide variety of government and academic sources. The data is geographically-referenced; however, it is aggregated at a county-level, (U.S. Census) tract-level, or Zip Code Tabulation Area (ZCTA) level. These geographic aggregation levels do not allow individual subjects or participants to be identified because person-specific data is not being collected and the data has been made publically available and in order for this to be the case, the data has to be de-identified. Therefore, there are no individual level identifiers.


This research project is being conducted by Naomi Adaniya, a doctoral candidate in the College of Public Health at The Ohio State University. It is being conducted under the approval and oversight of the Division of Health Services, Management, and Policy and its Chair Allard Dembe, ScD, who is also the Chair of the Dissertation Committee.

According to Michael Donovan, Senior IRB Protocol Analyst at The Ohio State University’s Office of Academic Affairs – Responsible Research Practices, this study not meet the definition of human subjects research since data is aggregated at the county level with no individual identifiers and publicly available. No IRB or IRB exemption application is needed (Response Received Electronically October 21, 2014).

Even if the study was determined to be human subjects research, the study qualifies for exemption from review from The Ohio State University’s Institutional Review Board (IRB) as a study falling under Category #4. All data sources are publicly available, data is aggregated at the county level, and no human subjects are used. Therefore, no data can have personal identifiers.

Category #4 – Research, involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified directly or through identifiers linked to the subjects.