# RURALITY AND ROBUSTNESS: RURAL COMMUNITIES AND HOUSING INSECURITY RISK

By

Steve Garcia

A.A.S., B.A.

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Data Science and Engineering

(in Spatial Information Science and Engineering)

The Graduate School

The University of Maine

May, 2024

Advisory Committee:

TBD, Advisor

TBD

TBD

# RURALITY AND ROBUSTNESS: RURAL COMMUNITIES AND HOUSING INSECURITY RISK

By Steve Garcia

Thesis Advisor: TBD

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Data Science and Engineering
(in Spatial Information Science and Engineering)
May, 2024

This paper explores rural housing insecurity through Swope and Hernandez's (2019) 4 C's of housing insecurity in rural areas. Rural census tracts are defined by the United States Department of Agriculture (USDA) Rural Urban Code Continuum (RUCA) codes seven through ten to avoid an overly restrictive definition. Little attention has been paid to rural areas in the conversation on housing, to facilitate further discussion this exploratory study uses unsupervised machine learning to group census tracts into risk levels across 7 sectors of ACS data based on housing insecurity factors found in the literature. Multinomial logistic regression is used to determine variation between states based on how well state risk levels can be predicted with the national dataset. Additionally, spatial autocorrelation is used to analyze how spatially clustered the risk levels and housing insecurity risk variables. The results indicate that many rural census tracts have a medium risk of housing insecurity, and the risk levels are hard to predict. The spatial autocorrelation results show that the housing insecurity variables are not as highly spatially clustered as expected.

# RURALITY AND ROBUSTNESS: RURAL COMMUNITIES AND HOUSING INSECURITY RISK

By Steve Garcia

Thesis Advisor: TBD

A Lay Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Data Science and Engineering
(in Spatial Information Science and Engineering)
May, 2024

Keywords: housing insecurity, homelessness, data mining

This paper explores rural housing insecurity through Swope and Hernandez's (2019) 4 C's of housing insecurity in rural areas. Rural census tracts are defined by the United States Department of Agriculture (USDA) Rural Urban Code Continuum (RUCA) codes seven through ten to avoid an overly restrictive definition. Little attention has been paid to rural areas in the conversation on housing, to facilitate further discussion this exploratory study uses unsupervised machine learning to group census tracts into risk levels across 7 sectors of ACS data based on housing insecurity factors found in the literature. Multinomial logistic regression is used to determine variation between states based on how well state risk levels can be predicted with the national dataset. Additionally, spatial autocorrelation is used to analyze how spatially clustered the risk levels and housing insecurity risk variables. The results indicate that many rural census tracts have a medium risk of housing insecurity, and the risk levels are hard to predict. The spatial autocorrelation results show that the housing insecurity variables are not as highly spatially clustered as expected.

## DEDICATION

Dedicated to Steve Mauldin, who taught me to find love in the emptiness of the absurd

## ACKNOWLEDGEMENTS

I would like to thank the following people for helping with this research project: Thank you to my advisor, Dr. Matt Dube for providing guidance and feedback throughout this project, To my wife Katie for encouraging to pursue my dreams, to Dr. Jeremy Mhire for teaching me how to think, and Dr. Dolliann Hurtig for teaching me how to live.

# Table of Contents

# List of Tables

# List of Figures

# CHAPTER 1

# INTRODUCTION

Homelessness research has undergone a significant transformation in recent years. Traditionally, the focus was on categorizing and describing different segments of the homeless population (Lee et al., 2021). A more contemporary approach views homelessness as a spectrum rather than a binary condition (e.g., Desmond et al., 2015; Swope and Hernandez, 2019). This paradigm shift opens opportunities for preventative and reactive services to address homelessness and housing insecurity. To harness these opportunities, four critical areas in literature require attention. First, housing research has concentrated on urban settings resulting in an urban-centric view of social issues like poverty and homelessness. Second, measuring housing insecurityproves challenging due to its dependence on circumstances and obstacles for individuals and communities (Leifheit et al., 2022). Third, housing and homelessness in urban and rural areas necessitates a multi-disciplinary approach to properly capture the aspects that contribute to them, an approach that has been scarcely taken in the extant literature. Finally, the scarcity of identified community-level risk factors in rural areas, coupled with a dearth of rural-specific data and research, makes the study of rural homelessness particularly intricate (Gleason et al., 2021). Addressing these gaps by integrating rural areas into the discourse on homelessness and housing insecurityis essential for creating a just and equitable society with effective policies for preventing and addressing homelessness(O'Regan et al., 2021).

## 1.1 *Rural Areas*

Rural areas differ vastly across regions and even states, often used as a catch-all for non-urban areas. Rurality is often defined simply as not being urban Robertson et al., 2007). Defining rural areas in contrast to urban areas largely excludes the variation between rural areas. The lack of universally accepted definitions of rurality reduces the amount of time and resources that can be dedicated to struggling communities Samadura and Yousey 2018. Rural areas encompass a broad spectrum, including farms, ranches, villages, small towns, and many other characteristics (Cromartie and Bucholtz, 2008). At their core, rural areas are a function of "space, distance, and relative population density" (Castle et al., 2011, ?).Shoup and Houma (2010) group urban areas into three categories: rural ares dependent on nearby urban centers, "destination counties" with natural or artificial amenities that attract non-permanent residents, and production communities that revolve primarily around a single industry. This wide variation between rural areas makes defining and understanding rurality a difficult challenge. Rural areas dominate the land mass of the United States, but with 85 percent of the population living in urban areas, they are often overlooked in the public discussion (Pendall et al. 2016) In the study of housing, rural areas are almost entirely excluded from the conversation (Gkartzios and Ziebarth, 2017). Contributing significantly to this problem is a wide variety of definitions used by governmental organizations, policy makers, and scholars (Samadura and Yousey, 2018; Cromartie and Bucholtz, 2008). Recently, The main policy objective for rural communities has been the promotion of economic development and preservation of the characteristics ascribed to rural areas (Lichter and Johnson, 2007).

*Figure 1.1: Map of Rural areas here vs population density here*

As Figure 1.1 demonstrates, the size of rural areas compared to the population density is vast and deconstructing the urban-centric lens of housing research necessitates a novel approach that can accommodate the differences in rural areas. The size and variation necessitates addressing rural issues differently because there can be no one size fits all policy approach to improving conditions.

## 1.2 *Literal Homelessness*

For decades, scholars have debated if research should focus on the reasons why people become homeless or the structural forces that create homelessness (Shlay and Rossi, 2003).

Prior to significant shifts in the 21st century, research on homelessness focused on identifying and describing categories of homeless people (Lee et al., 2021). In other words, much research has focused on the binary of individuals and families being housed or unhoused and trying to assign them into umbrella categories. This neglects the wide range of individual and societal factors that occur in the phases between when an individual or family is housed and becomes unhoused. Estimating the number of people that are unhoused is notoriously difficult even in urban areas. For measuring homelessness, the most popular mechanism in the United States is the Department of Housing and Urban Development (HUD) point-in-time (PIT) count and housing inventory count. These counts are used for the distribution of federal funds for combating homelessness. as Agans et al. (2014) note that the unhoused frequently relocate and the housed may quickly become unhoused, making it difficult to accurately estimate the number of unhoused people at any given time. When it comes to addressing literal homelessness, public health experts differentiate between preventative services and reactive or emergency services (O'Regan et al., 2021). Preventive services prevent households from becoming homeless, while reactive or emergency services step in after a household becomes homeless. As homelessness is often seen as an urban problem, most intervention occurs in urban areas (Gleason et al., 2021).

### 1.3 *Housing as Health*

A house is far more than four walls, a roof, and some doors. The characteristics and location of a house can make a significant impact on one's life. In the United States, housing is often a family's greatest expenditure, greatest source of wealth, and a place of safety and gathering (Braveman et al., 2011). The federal government has long acknowledged this through legislation like the Housing Act of 1949, and social programs and development goals developed by the U.S. Department of Housing and Urban Development. Housing is often seen as one of the most fundamental determinants of health, and a lack of adequate housing can produce adverse health outcomes and acts as a foundation for "social, psychological, and cultural well-being" (D'Alessandro and Appolloni, 2020, p.17; Leifheit et al., 2022). Part of acknowledging housing as health is moving beyond the housed/ unhoused binary in order to better understand and intervene in households that are at risk of being unhoused. This is often referred to as housing insecurity, a broader term that encompasses a continuum that affects a larger part of the population than simply housed/ unhoused (DeLuca and Rosen, 2022).

### 1.4 *Theoretical Framework*

With little infrastructure for homelessness services in rural areas, the 4 C's approach to housing insecurityproposed by Swope and Hernandez (2020) can highlight areas of critical concern for devoting resources to reactive services and identify areas where preventative services can improve or expand. The pillars of the 4 C's include:

- Conditions: quality of housing
- Cost: housing affordability
- Consistency: residential stability
- Context: neighborhood opportunity

The 4 C's housing as an interconnected web of factors that impact health and encapsulates "this unequal distribution of housing disparities along other axes of inequality, and the historical forces shaping unequal housing opportunities" (Hernández and Swope, 2019, 1). Swope and Hernandez are not the only scholars to design a model encompassing these 4 factors. Metzger and Khare (2017) proposed a similar framework that encompasses stability, affordability, internal housing conditions, and area characteristics.

### 1.5 *Motivation*

There are three primary reasons behind this exploration:

- The first motivation comes from the lack of attention scholars have paid to rural areas as it pertains to housing insecurity.
- The second motivation is to test the efficacy of the 4 C's model of housing insecurityfor the rural United States. As there is an urban lens to housing insecurity, an adequate theoretical model must be capable of adapting to areas often left out of the conversation
- The final motivation is to provide policy makers and researchers with a tool to identify rural areas of housing insecurityin their constituency and create harm reduction approaches and services that can meet the unique needs of their areas.

### 1.6 *Approach*

In order to improve our understanding of rural housing insecurity, this thesis investigates the risk levels of rural census tracts in the United States using the 4 C's model of housing insecurity. Using k-medoids clustering, risk factors across eight different sectors are used to assign housing insecurity risk levels to rural census tracts for the continental United States. Each state is clustered with census tracts from other states within a 15 mile boundary to encapsulate how communities exist across state lines. The cluster medians are analyzed to understand the trends in housing insecurityfactors across states and the clusters are relabeled based on risk factors identified in the literature so that each cluster falls into a low, medium, or high risk level. Using these risk levels, association rules learning is used to identify common patterns between sector risk levels and identify pockets of rural census tractthat are at high risk of housing insecurity. To better understand how factors of housing insecurityrelate to space, Moran's I spatial autocorrelation is used to determine how spatially clustered each housing insecurityfactor is. Finally, a multinomial logistic regression is used to determine how well each states sector risk levels can be predicted and a national model is generated for each sectors risk levels to analyze how each housing insecurityfactor contributes to sector risk levels.

Beginning to understand rural homelessness requires that a number of questions be answered:

- How can risk factors of be used to identify risk levels of housing insecuritywhile accounting for the variation in rural areas?
- Do housing insecurityfactors identified for urban areas exhibit similar characteristics in rural areas?
- When measuring housing insecurityacross different dimensions, how often do the same features arise?
- Are there spatial relations between the different dimensions of housing insecurity?
- To what extent can the housing insecuritydimensions be predicted?

## 1.7 *Major Results*

I have some

## 1.8 *Intended Audience*

Why would anyone read this

This thesis is intended for an audience with a significant interest in rural housing insecurity. such an audience can include, but is not limited to policy makers, economists, political scientists, community psychologists, rural sociology, and many others concerned with housing insecurity.

## 1.9 *Structure of Thesis*

The thesis is structured into six chapters. Chapter 2 offers a comprehensive theoretical foundation, focusing on the application of the 4 C's of the housing insecuritymodel. This chapter critically reviews pertinent literature pertaining to various facets of the model. Chapter 3 delineates the methodology employed for data processing. It provides an in-depth explanation of the chosen methodology and its execution. The ensuing Chapter 4 presents the study's findings, offering a detailed analysis of the acquired results. Chapter 5 deliberates on the implications of the results, discussing their significance and impact within the scope of the study. This chapter provides a thorough examination of the noteworthy findings. Chapter 6 serves as a synthesis, summarizing the entirety of the work and offering insightful commentary on the major findings. Additionally, it delineates potential avenues for future research and study.

# CHAPTER 2
# FACTORS OF HOUSING INSECURITY

To construct the 4 C's framework, the literature review is divided into one section for each pillar of housing. As each column forms a web rather than separate pieces, the result is a significant amount of overlap between columns.

## 2.1 *Cost*

It is difficult to determine one number that determines when a household is spending too much on housing. A cost-to-income ratio is the most common way of measuring housing affordability. The threshold for housing affordability has ranged between 25 and 50 percent but the current standard is 30 percent (Kropczynski and Dyk, 2012). Housing is considered affordable if the household spends less than 30 percent of its income on housing and 50 percent or more is considered a high-cost burden (Braveman et al., 2011;Swope and Hernandez, 2020;Weicher, 2006). Inherit to any cost-to-income ratio is the understanding that housing is that there are other expenses necessary for survival (Herbert et al., 2018). Housing costs are determined by the rate of household formation and household attrition Pendall et al., 2016. Housing affordability affects individuals, families, and communities while access is largely determined by their demographic characteristics (Braveman et al., 2011;Yadavalli et al., 2020).Housing affordability is directly related to residential stability and has the potential to harm both those being forced to move, the community they are leaving, and the community they are entering (Desmond et al., 2015). Access to affordable housing affects the physical and material comfort of the population as well as the individual (Kumar, 2003). If a household cannot afford to live in their current place, they may be forced to relocate seeking more affordable housing or through eviction and foreclosure.. If too much of a household's money goes to housing, they may be forced to go without other necessities (Herbert et al., 2018). note that a body of research has found that those with high housing costs may also experience food insecurity as food is often considered a flexible expense while housing is a fixed expense (Fletcher et al., 2009;Kropczynski and Dyk, 2012). The shortage of affordable housing drives lower-income families to substandard housing in worse neighborhoods (Braveman et al., 2011). Kang (2021) characterizes housing instability as a by-product of the affordable housing shortage wherein households can be destabilized by minor financial shocks. These factors can create a situation where housing costs lead to residential instability, which is linked to a variety of adverse conditions, especially in children and adolescents (Desmond and Shollenberger, 2015). Part of the blanket construct of rural areas is that they are cheaper to live in. However, Kurre (2003) note that there is relatively little systematic data that supports this presumption. Rural areas face the sale low per capita income and poverty problems faced by urban areas (Castle et al., 2011). Zimmerman (2008) found no consistent pattern of lower princes across all of the rural counties in ? While the dollar amount paid for housing may be lower, given the different socio-economic circumstances of rural areas, housing costs alone may not fully explain the situation (Kropczynski and Dyk, 2012).

### 2.2 *Conditions*

Many scholars have identified internal housing conditions as a significant factor on health (Braveman et al., 2011;Metzger and Khare, 2017; Swope and Hernandez, 2020. In one study, decent housing was found to be a more important determinant of health than education or income (Stefan and Bittschi 2014). Previous environmental health research has identified five broad categories in which housing conditions contribute to adverse health effects: *physical conditions*, *chemical conditions*, *biological conditions*, *building and equipment conditions*, and *social conditions* (Jacobs, 2011). Adequate housing is necessarily related to public health (Matte and Jacobs, 2000). Links to an increase of disease have been tied to poverty, poor housing, and degraded environments reflecting the interconnectedness of housing insecurityissues (Rauh et al., 2008). Stefan and Bittschi (2015) found that the probability of facing a chronic disease increases when housing problems accumulate and that poor housing conditions quickly degrade subjective health. These problems are amplified in the modern world where individuals spend an estimated 90 percent of their time in doors (Palacios et al., 2020). Despite housing conditions playing such a significant role in modern life, there is not a significant since of communal benefit and responsibility when it comes to housing (Jacobs, 2011). Rural areas face unique housing issues because one of the most common housing solutions is mobile homes. Structural problems like poor construction and risks of air pollution and fire create a unique problem (MacTavish et al., 2006). An area of particular concern are marginalized populations who are more likely to be exposed to harmful housing conditions (Swope and Hernandez, 2020). Housing conditions also play a role in residential mobility as Desmond and Bell (2015) place decent housing and affordable housing as fundamentally connected and the previously mentioned rise in housing cost has not brought an increase in housing quality.

### 2.3 *Consistency*

Residential mobility is a complicated subject because residential mobility as a broad concept is conceived as a good thing. That one can pack up and go somewhere with more opportunity is considered a part of the American "mystique" (Molloy et al., 2011). An average of 15 percent of Americans move every year and 25 percent move over the course of two years (Bachmann and Cooper, 2014). Classic urban economic theories explain that households make trade-offs between proximity to jobs and housing prices (Hu and Wang, 2019). This puts low-income households at a particular disadvantage as their access to jobs may be lower than their wealthier counterparts. Consistency or residential stability plays an important role in the physical and social well-being of individuals, families, and communities. Residential instability has been linked to a variety of adverse conditions and affects the neighborhoods being entered and left and it has been identified as a more important predictor of community health than standard factors like poverty and racial composition (Desmond and Shollenberger, 2015;Desmond and Perkins, 2016, Rauh et al., 2008). People move for a variety of reasons, but an important distinction must be made between voluntary and involuntary moves (Siskar, 2019). While most moves are voluntary, there are millions of low-income households that struggle to maintain housing stability (Phinney, 2013;Kang, 2019). Housing is often the biggest expense for low-income families often forcing them to make trade offs between necessities as previously mentioned (Desmond and Bell, 2015). Of particular concern is forced relocation, foreclosure, eviction, and condemnation are all drivers of forced relocation (Phinney, 2013;Siskar, 2019). Forced relocation is linked to an increase in residential instability and households forced to move often end up in places with greater disadvantage and are more likely to face additional moves (Desmond and Shollenberger, 2015). One issue with the study of residential mobility is the limited scope of predictors that have been linked to it (Kang, 2019). In urban areas, renters have been found to be particularly vulnerable to relocating to worse neighborhoods than the one they are exiting (Desmond and Shollenberger 2015). It is yet to be seen how this translates to rural areas, where renting is far less common than urban areas with the exception of mobile homes.

### 2.4 *Context*

Context revolves around neighborhood and community characteristics including demographics, green spaces, education, and healthcare among other things. While it is impossible to capture context in its entirety, this thesis focuses on demographics, economic diversity, housing type, and household wage/ aid factors as these have all been studied as matters related to housing insecuritythat do not fall directly into the other 4 C's. The following sub-sections provide an interdisciplinary review of how these factors affect housing insecurity.

### 2.4.1 *Employment*

In the United States, the labor market is the result of cumulative individual behaviors including geographical migration and educational investments (Wiener, 2020). The demand for labor is driven by firms, which must consider a wide variety of factors in deciding location (Partridge and Rickman, 2007). In recent decades, the United States labor market has entered a risk regime job market where workers hold a greater share of the risk in an employment system without the perceived promise of security and stability, which has become embedded in American social and political institutions (Lowe, 2018). It is agreed that the Fordist regime that brought unprecedented prosperity in the early 20th century came to an end in the 1970s (Stockhammer 2008). Since this shift, productivity of the average worker has increased 64.6 percent while hourly pay has only increased an average of 17.3 percent between 1979 and 2021 Economic Policy Institute). Over this same period, U.S. Housing and Urban Development data shows that the median price of a new single-family home increased from $60,600 ($232,091 adjusted for inflation) in Q1 of 1979 to $369,800 in Q1 of 2021) (U.S. Census Bureau and U.S. Department of Housing and Urban Development, 1963). The great recession has had a lasting impact on the housing market within the United States. As the economy recovery did not benefit all households equally, wealth inequality has grown along both racial and ethnic lines (Lochar, 2014). As wages have failed to keep up with the price of housing, the current economic system under this risk regime places those with low incomes in a precarious situation for housing affordability and residential stability. Thus, employment insecurity and income inequality are two pressing issues the United States is facing that have serious impacts on communities. “housing insecurityhas risen in relative lockstep with employment insecurity” (Desmond and Gershenson, 2016, 48).

Rural communities have been hit hard by economic change, driven by the transition from a production to a consumption based economy (Pendall et al., 2016). what (Bjerke and Mellander, 2019) identified an increasing economic divide between urban and rural areas where over several decades rural areas have lost out. During this shift, employment became increasingly scarce for agricultural workers (Kropczynski and Dyk, 2012). Today, manufacturing is responsible for 21 percent of rural non-agricultural earnings (Low, 2017). Economic development is therefore a fundamental issue to rural areas. While manufacturing has grown, the majority of counties that experienced manufacturing employment growth between 2001 and 2015 had low levels of growth in terms of total employment (Low, 2017). Sherrieb et al. (2010) identify three key elements connected of economic development: the level of economic resources, the level of equality in resource distribution, level of diversity in economic resources. Economic development alongside demographic change in rural areas has been linked to the quality and condition of local housing infrastructure (Barcus, 2011). Thus, how policies shape economic development has a direct affect on the overall housing insecurity risk of rural communities. Demonstrating the interconnectedness of communities, regional economic development in one area can encourage economic stability of its neighboring regions as well (Chen, 2018). Deller and Watson (2016) highlight the importance of economic diversity, a vital aspect of economic development, finding that more diverse economies enhance economic stability. As an insulator against economic instability, employment diversity in rural areas is a key factor that policy makers and scholars should consider as part of a holistic approach to housing insecurity. This may be difficult to achieve for rural areas based on natural amenities, where one industry acts as the lifeblood of the community.

### 2.4.2 *Housing, race, and poverty*

Housing is affected by a variety of social, political, and economic factors. "The ability of residents to access affordable housing, whether renting or buying, is in large part determined by their demographic characteristics, such as income, race, age, and educational attainment" (Yadavalli et al., 2020, 115). While unpredictable events may narrow the disparities, "As a rule, a household's vulnerability to displacement should be shaped in predictable fashion by those characteristics that define its members' position in the [social] stratification system" (Lee and Evans, 2020, 5). Although the federal government took a direct interest in promoting home ownership in 1933, racial discrimination in the housing market was not outlawed until 1968 but enforcement of the law remained difficult until the Fair Housing Act of 1988 (Sharp & Hall, 2014). For example, the practice of redlining made it difficult for Black Americans to receive mortgages under federal aid programs and creating racial segregation that can still be seen today. At the county level, the probability of living in affordable housing decreases as the white population decreases (Brooks, 2022). In addition to racial segregation, income segregation must be considered for a holistic discussion of housing insecurity. Lichter and Brown (2011) found that 40.5 percent of high-poverty places are in high-poverty counties for non-metro areas and the poor and non-poor are becoming increasingly segregated, with higher concentrated poverty among minorities. High concentration of poverty may exacerbate housing condition issues due to a lack of revenue to maintain the necessary services at the household and local government levels. Minorities are also at a disadvantage in income segregation with poor whites being less segregated from their non-poor counterparts (Lichter et al., 2021). As a home is often a household's greatest source of wealth, the disadvantages minorities have in terms of housing are compounded as social and economic inequality are reproduced as these disparities continue (Krivo and Kaufman, 2004).

Rural areas face significant consequences for the historical forces that shape housing today. When discussing rural poverty it must be noted that there is an underlying assumption that the dynamics of poverty are fundamentally different from urban areas (Valasik, 2018). Persistent problems faced by the rural poor include "physical isolation and poor public transportation, inadequate schools, and limit access to medical care and other basic public services while institutional support services are frequently limited or simply unavailable" (Lichter and Johnson, 2007, ?). Part of this is driven by the outflow from rural areas to urban areas. Rural areas have seen a reduction in population, reducing the capabilities of public services to accommodate those in need (Bjerke and Mellander, 2019). As mentioned earlier, there are a variety of reasons why households move. In rural areas, a common reason to move is due to the friction that exists when households are too far removed from labor markets that provide adequate employment and income opportunities (Sparks et al., 2013). Breaking the stereotypical images of rural areas, poverty is a significant issue in rural areas. Valasik (2018) found that from 2000 to 2012, increases in poverty were larger in rural counties than urban counties with the highest increases in exposure and the rural black population was by far the most disadvantaged over this time period. Rural areas are not as diverse as the United States overall, and many rural minorities are geographically central in regions tied to historical and economic dynamics (Housing Assistance Council, ?). Another demographic group that is significant to rural areas is Hispanics and Latinos, despite the widespread population decline of rural areas (Lichter and Johnson, 2020). African Americans and Hispanics and Latinos face similar discrimination in the housing market with the benefits of housing are dramatically smaller for these demographics (Krivo and Kaufman, 2004). Thus, the pockets of these groups in rural areas should be considered to be at a higher risk of housing insecurity due to the effects of these historical forces.

### 2.4.3 *Housing Type*

While owning a home is considered a part of the "American Dream," many households rent their housing by choice or by necessity. While the many benefits of home ownership portray it as a means to a better life, renting is not inherently bad and may provide better opportunities for households that can afford it, but there are many potentially destabilizing consequences of high-cost renting (Drew, 2014). Nationally, median rent in a poor neighborhood is $298 compared to $225 in a middle-class neighborhood or $250 in an affluent neighborhood after regular expenses are deducted despite property values typically being much higher in middle-class or affluent neighborhoods (Desmond and Wilmers, 2019). This creates a compounding factor for the previously mentioned disparities in home ownership. Increases in household wealth and secured debt were found to decrease the likelihood of homeowners becoming renters and vice versa (Anderson et al., 2021). Money paid towards a mortgage generates long-term wealth while money paid towards rent generates wealth for the property owner. Renters with high-cost housing are unable to increase household wealth through their means of housing. In addition to whether one rents or owns their home, the type of home can play a significant role in housing insecurity. The housing types considered for this study are single family, multi-family, mobile/ manufactured homes, and "unconventional housing." Unconventional housing includes dwellings not considered for long-term habitation including RVs/ campers, vans, and boats. These unconventional forms of housing may keep people off the streets, but they are not always a stable mode of housing. For RV and camper living, people who are undocumented or are unable to keep up with legal or maintenance costs for vehicles end up losing their housing (Wakin, 2005). Mobile homes also carry a unique set of circumstances that may put households at a greater risk of housing insecurityand are found frequently in rural areas. Mobile homes and the land they are situated on can be either owned or rented. It is common in mobile home parks for households to own their home but not the land it is on. Key issues with mobile homes include their financing: typically done through more expensive but easier obtained means than a mortgage such as personal property or chattel loans; mobile homes do not build wealth in the same way as they typically depreciate rather than appreciate; households on rented land have little control over their length of stay; they also tend to have worse construction and higher risks of air pollution and fire than traditional homes (MacTavish, 2007). Those that live in mobile homes or unconventional housing should be a priority for discussing housing in rural areas.

### 2.4.4 *Household income, aid, and Transportation*

In his first State of the Union address, President Lyndon B. Johnson asked Congress to declare an "unconditional war on poverty... not only to relieve the symptom of poverty but to cure it and, above all, to prevent it." Since then, a patchwork of programs regulated at the federal, state, and local levels have arisen. A large part of the federal government growth in the late 20th century is from the expansion of social welfare spending (Fishback, 2020). As the primary mechanism of income distribution is what Berkowitz and Palakshappa (2023) refer to as the "factor payment system" in which those who work and those who own the means of production and one's relation to this system and the labor market is closely related to one's poverty risk. To alleviate this poverty risk, social programs which utilize different mechanisms are available to those who qualify. These mechanisms can be divided into categorical and income targeted policy designs, alongside decentralization, where some receive benefits based on "demographically defined, categorical eligibility structures" and others enjoy standardized federal assistance through social insurance with some qualifying for income-based or "means-tested" programs (Bruch et al., 2023). Households must fall below certain income and asset thresholds to qualify for means tested programs (Rank and Hirshl, 2002). For housing, there is a wide variety of housing policies and programs aimed at low-income individuals. These take the shape of voucher programs by subsidizing privately held property although some recipients live in public housing (Kim et al., 2017). For rural areas, the U.S. Department of Agriculture (USDA) has a variety of programs aimed at improving living conditions in rural areas including direct or guaranteed loans for single or multi-family housing, and infrastructure programs for water, electricity, and telecommunications ((USDA Rural Development Summary of Programs, 2023). Transportation plays a large role in social and economic life. Access to everything from education to healthcare depends on the infrastructure and ability to use available means of transportation. The expense of owning enough vehicles may prove restrictive, especially for households with high housing costs. Rural areas often do not have public transportation, leading residents to depend more on automobiles. An analysis of 2009 National Household Travel Survey data found that 72 percent of households with a yearly income of $20,000 have access to a household vehicle compared to over 97 percent of households making $50,000 (Blumenberg and Pierce, 2012). Automobile ownership can be a crucial factor in avoiding residential instability (Kang, 2019). Households are twice as likely to be auto-deficient (less than 1 car per driver) than zero-vehicle households where a vehicle is not needed (Blumenberg et al., 2020). This is a harrowing statistic in rural areas without public transportation where distances may be too far to walk or ride a bicycle or it may not be safe due to lacking road infrastructure like bicycle lanes and sidewalks.

### 2.5 *Summary*

Throughout this chapter, the 4 C's of housing insecurity have been covered. It is important to highlight the interconnected nature of the 4 C's. There is significant overlap between each pillar of housing insecurity. Housing costs, housing type, and housing conditions are necessarily linked to the economic conditions of a household. These economic conditions are linked to the household wage/ aid factors that encapsulate their economic status. One's relation to the poverty level and education plays a significant role in housing accessibility and these factors are intrinsically linked to the context that they grew up in. For rural areas where public transportation is scarce, one's access to adequate transportation is highly linked to one's economic opportunity. Any discussion on housing insecurity must consider the historical forces affecting modern day race and poverty, and these forces relate to all aspects of life. When taken as a web, this model encompasses the wide ranging socio-economic factors that surround housing insecurity.

# CHAPTER 3

# ADDRESSING RURAL HOUSING INSECURITY

## 3.1 *Defining Rurality*

Rather than strictly defining rurality, this thesis uses the United States Department of Agriculture (USDA) Rural-Urban Continuum spectrum. The following codes are used to encapsulate rurality:

Table 3.1: RUCA Codes

| **Number** | **Description** |
|---|---|
| 7 | Small town core: primary flow within an Urban Cluster of 2,500 to 9,999 (small UC) |
| 8 | Small town high commuting: primary flow 30 percent or more to a small UC |
| 9 | Small town low commuting: primary flow 10 percent to 30 percent to a small UC |
| 10 | Rural areas: primary flow to a tract outside an urban area or urban cluster |

The range of RUCA codes described in Table 1.1 was chosen to be inclusive than exclusive, including small towns with various levels of commuting to urban clusters and areas classified as rural. We include small towns because they often serve as hubs for rural areas, serving an important role in rural areas. Spatial autocorrelation is used to determine how often similar rates of each variable occurred across each rural census tract in each state. Finally, multinomial logistic regression is used to determine how well the risk levels of a census tract can be predicted based on the nationwide dataset. All analysis was conducted in the R statistical language.

## 3.2 *Data*

Eight different sectors of 2019 ACS 5-year variables are used to capture the 4 Cs of housing insecurity using indicators of housing insecurity identified in the literature. These sectors are housing cost, housing quality, housing type, economic diversity, education mobility, poverty mobility, and household worker/aid. For demographic variables we use seven variables including an "other" variable to account for race/ ethnicity and the number of people over or under 18 by gender. The economic diversity data is the number of people employed across 13 distinct categories. There are It was necessary to create three compound variables: high-cost with a mortgage, high cost without a mortgage, and high-cost rent to use the standard affordability measure of 30 percent. There are four variables for housing conditions which include houses with an incomplete or insufficient kitchen or plumbing for occupied and unoccupied housing units. Two sets of variables account for residential mobility: education and poverty that include those who did not move, those who moved in and out of county and state. Due to the housing affordability and income inequality crises, those below the poverty level and those at 125 percent of the poverty level as high risk for housing instability are included in residential mobility: poverty. For education those with and without a high school diploma are included as those without a college degree may face higher barriers to well-paying and stable employment. Wage/aid data include households without income, households that receive public assistance, households that receive supplemental security income, households with investment income, households with other income, households with 3+ workers and the household Gini index. For housing type, renters and owners of mobile homes, single family homes, small and large multi-family homes, and renters and owners of unconventional housing are included.

### 3.3 *Data processing*

In order to ensure the integrity of the data, census tracts that lacked specific sector-related information were excluded from the analysis. These omitted tracts were assigned a risk level of zero, a measure adopted to preserve the largest possible number of census tracts for subsequent analyses. To mitigate potential biases stemming from differences in population sizes and geographic areas, a standardized approach was employed across each sector. This involved scaling all dataset components to a common base unit. Demographic and economic diversity metrics were adjusted proportionally to the population size. Meanwhile, data pertaining to household expenses and types were scaled based on the counts of homeowners and renters. The household dataset underwent normalization corresponding to the total number of households, whereas housing condition indicators were adjusted relative to the total count of occupied and unoccupied housing units. It's essential to note that all numerical values within the dataset have been uniformly represented as percentages, except for the household Gini Index, which retains its original values.

### 3.4 *Methods*

Supervised and unsupervised machine learning algorithms are used alongside Moran's I spatial autocorrelation and the Queen Contiguity spatial relationship algorithm to form and analyze the housing insecurity risk assignment system.

#### 3.4.1 *Neighbors Algorithm*

Communities often share dependencies across state lines, making it unjust to disregard neighboring communities in a state-based housing insecurity analysis. To address this, the analysis encompasses census tracts within 15 miles of a state's outermost tract. Any census tract sharing a boundary with a tract within this range is considered, ensuring a more inclusive evaluation of rural housing insecurity. This process is repeated for each state in the continental United States.

The formula for queen contiguity neighbors is shown in Equation 3.1.

$$
\begin{aligned}
\text{Top-left:} \quad & (x-1, y-1) \\
\text{Top:} \quad & (x, y-1) \\
\text{Top-right:} \quad & (x+1, y-1) \\
\text{Left:} \quad & (x-1, y) \\
\text{Right:} \quad & (x+1, y) \\
\text{Bottom-left:} \quad & (x-1, y+1) \\
\text{Bottom:} \quad & (x, y+1) \\
\text{Bottom-right:} \quad & (x+1, y+1)
\end{aligned}
\tag{3.1}
$$

### 3.4.2 *K-Medoids Clustering*

K-medoids clustering is a partitioning technique aimed at dividing a dataset into $K$ distinct and non-overlapping clusters. Unlike K-means clustering, which utilizes centroids as cluster representatives, K-medoids employs actual data points within the dataset as cluster representatives. The key advantage of K-medoids lies in its robustness to outliers and noise due to its use of real data points. The objective of K-medoids clustering is to minimize the sum of dissimilarities within clusters. Each state, including neighboring census tracts, is clustered. The cluster medians are analyzed to determine which clusters have a high, medium, or low risk of housing insecurity based on the factors previously identified in the literature review. The formula for K-medoids clustering is shown in Equation 3.2.

$$\begin{aligned} \underset{S}{\text{minimize}} \quad & \sum_{i=1}^{K} \sum_{x \in C_i} d(x, m_i) \\ & \text{where:} \\ & \quad S : \text{The set of clusters.} \\ & \quad K : \text{The number of clusters.} \\ & \quad i : \text{Index representing each cluster } (1 \leq i \leq K). \\ & \quad C_i : \text{The } i\text{-th cluster containing data points.} \\ & \quad x : \text{A data point within a specific cluster } (x \in C_i). \\ & \quad m_i : \text{The medoid (representative) of the } i\text{-th cluster.} \\ & \quad d(x, m_i) : \text{The dissimilarity (distance) between data point } x \text{ and medoid } m_i. \end{aligned} \tag{3.2}$$

### 3.4.3 *Association Rules Learning*

Association Rules learning is a data mining technique used to uncover interesting relationships between variables in large datasets. It aims to discover patterns in the form of rules indicating the co-occurrence or association between items within transactions or events.

Association rule learning involves two main metrics:

**Support (s)**: Measures the frequency or occurrence of an itemset in the dataset.

$$\text{Support}(A \to B) = \frac{\text{Transactions containing both A and B}}{\text{Total transactions}}$$

**Confidence (c)**: Measures the conditional probability that an item B appears in a transaction given that item A is present.

$$\text{Confidence}(A \to B) = \frac{\text{Support}(A \cup B)}{\text{Support}(A)}$$

Here, association rules are used to analyze the common occurrences between sector risk levels. Of primary interest are unexpected relationships where a high risk level is association with a low risk level and vice versa.

### 3.4.4 *Moran's I*

The Global Moran's I is a statistical measure used in spatial analysis to detect spatial clustering or dispersion of similar values within a dataset. It quantifies the degree of spatial autocorrelation by assessing whether neighboring locations exhibit similar or dissimilar attribute values. Specifically, Moran's I considers both the values of the locations and the spatial relationship between them, providing a single coefficient that ranges from -1 to 1, with 0 indicating spatial randomness. This measure helps identify patterns in spatial data, highlighting if similar values tend to be close to each other or dispersed across the study area. The Moran's I values for each variable is calculated for each state and nationally in order to analyze how the housing insecurity factors cluster in space. The formula for Moran's I is shown in Equation 3.3.:

$$I = \frac{N}{W} \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{3.3}$$

Where:

$I$ : Moran's I statistic, representing the degree of spatial autocorrelation.

$N$ : Total number of spatial units (e.g., census tracts, regions).

$W$ : Total spatial weight in the dataset.

$w_{ij}$ : Spatial weight between spatial units $i$ and $j$.

$x_i$ : Value of the variable of interest in spatial unit $i$.

$\bar{x}$ : Mean value of the variable of interest across all spatial units.

### 3.4.5 *Multinomial Logistic Regression*

After the clustering is performed and the clusters are analyzed, each sector is assigned a new variable containing the risk levels for each census tract. Cross split validation is used wherein for each state, a new model is trained on all states except the target state. The probability that each census tract is its actual classification is preserved for the analysis. Additionally, to better understand how the housing insecurityfactors contribute to the risk levels, for each sector a national model is trained on the entire dataset so that the model can be analyzed and prediction power can be measured under the best-case scenario. The formula for multinomial logistic regression is shown in Equation 3.4.

$$\log\left(\frac{P(Y=k \mid X)}{P(Y=K \mid X)}\right) = \beta_{0k} + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p \tag{3.4}$$

Where:

log : Natural logarithm function

$P(Y=k \mid X)$ : Probability of the outcome being in category $k$ given predictor variables $X$

$P(Y=K \mid X)$ : Probability of the outcome being in the reference category $K$ given $X$

$\beta_{0k}$ : Intercept for category $k$

$\beta_1, \beta_2, \ldots, \beta_p$ : Coefficients corresponding to predictor variables $X_1, X_2, \ldots, X_p$

$X_1, X_2, \ldots, X_p$ : Predictor variables

$k$ : Specific category being predicted

$K$ : Reference category

# CHAPTER 4

# RESULTS

This analysis focused on a sample of 6,364 rural census tracts with a RUCA code of seven or higher. Four states and Washington D.C. were intentionally excluded from the analysis: Alaska and Hawaii were omitted due to the presence of unique factors, particularly in their rural areas, which may not have been adequately addressed in the existing literature. New Jersey and Rhode Island were excluded from our spatial analysis due to a lack of adequate data. These states are both very urban and once discrepancies in the data were removed, there were not enough observations to include in the analysis. Figure 4.1 shows how the neighbor algorithm changed the state census tract counts.

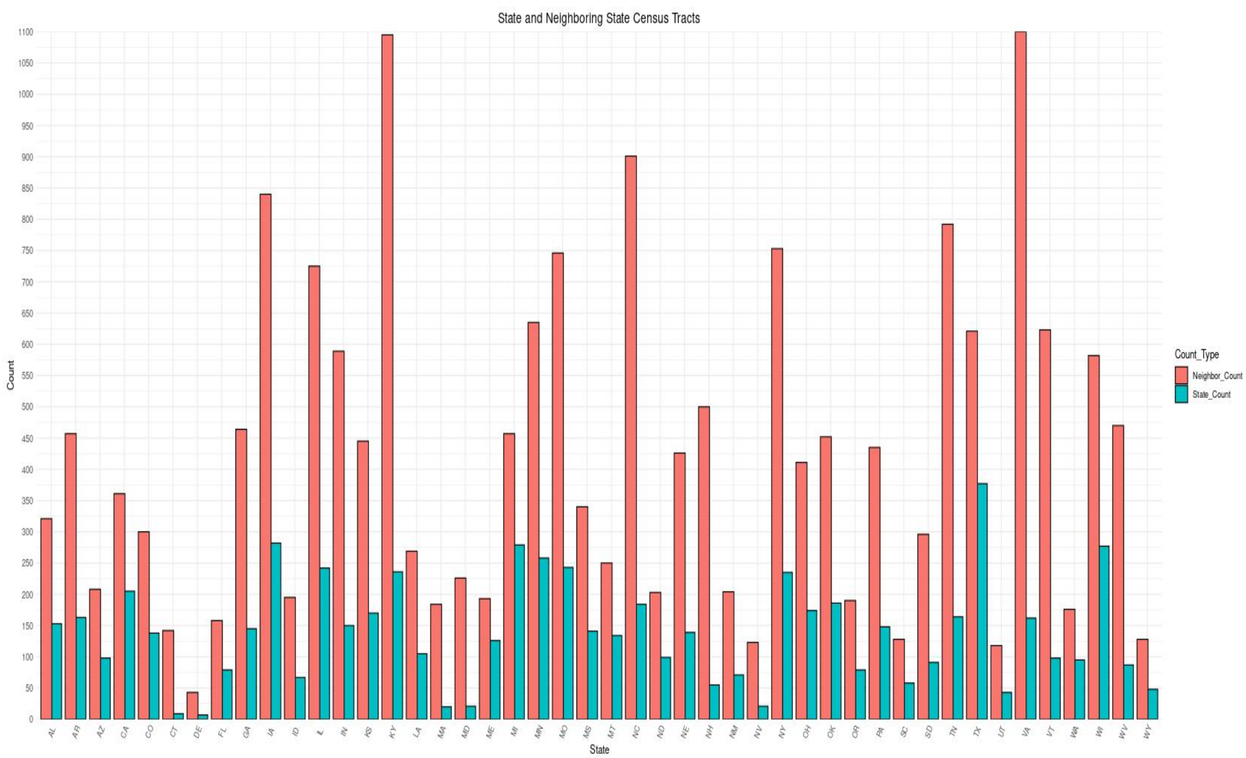


Figure 4.1: State Census Tracts vs. State Neighbors Count

## 4.1 *Cluster Analysis*

Here, the results of the cluster analysis is presented for each sector. The primary mechanism for analyzing the clusters is the average cluster medians for all states. The cluster averages were analyzed as well to ensure that the same trends are found in the dataset under a different descriptive statistic. All values are represented as a percentage corresponding to the base unit each sector is scaled to.

### 4.1.1 *Employment Diversity*

The risk levels for employment diversity are determined based on which clusters have the highest number of highest cluster values compared to the cluster with the lowest number of lowest cluster values. The higher the cluster averages across variables, the better the economic diversity of a cluster. Table 4.1 shows the values for this sector. Cluster 1 had the lowest cluster medians in 61 percent of variables, Cluster 2 has the highest cluster median in 53 percent of variables, and Cluster 3 has the middle value in 69 percent of cases. Based on this analysis Cluster one has the lowest level of economic diversity, cluster two has the lowest level of economic diversity, and cluster three has a medium level of economic diversity. Employment in education, health, and social work has the highest presence across each cluster followed by manufacturing.

Table 4.1: Median Values for Employment Diversity Clusters

| Variable | Cluster 1 | Cluster 2 | Cluster 3 |
|---:|---:|---:|---:|
| ag_for_fish_hunt_mining | 2.54 | 2.24 | 1.87 |
| arts_rec_food | 2.86 | 3.12 | 3.10 |
| construction | 3.16 | 3.06 | 3.07 |
| edu_health_social | 9.38 | 9.73 | 9.46 |
| fin_re_insur | 1.47 | 1.60 | 1.57 |
| information | 0.35 | 0.41 | 0.37 |
| manufacturing | 4.52 | 5.44 | 4.75 |
| othersvcs | 1.76 | 1.88 | 1.91 |
| prof_sci_mgmt_waste | 2.17 | 2.25 | 2.28 |
| public_admin | 1.98 | 1.88 | 1.92 |
| retail_trade | 4.48 | 4.78 | 4.59 |
| trans_warehouse_util | 2.17 | 2.03 | 2.12 |
| wholesale_trade | 0.76 | 0.86 | 0.70 |

### 4.1.2 *Demographics*

Due to the historical forces affecting minorities in both rural and urban areas, the risk levels for demographics are based on which clusters have the highest minority populations and the lowest white populations. Table 4.2 shows the values for this sector. This sector was decided based on the median and average highest, lowest, and medium value counts as clusters two and three had almost the same cluster median counts. Cluster three is the middle value for 90 percent of cluster median variables. Cluster three has the highest number of highest values across means and medians with 55 percent of variables. Cluster two is the lowest for 50 percent of variables. Cluster three alsos has the largest African American and Hispanic and Latino cluster medians. Based on this analysis, cluster one has a medium risk of housing insecurity, cluster two has a low risk of housing insecurity and cluster three has the highest risk of housing insecurity. One notable obeservation is that the cluster medians for the male/ female over 18 years old variables are roughly three times higher than the male/ female under 18 variables. This reflects the ageing population of rural areas noted in the literature.

Table 4.2: Median Values for Demographic Diversity Clusters

| Variable | Cluster 1 | Cluster 2 | Cluster 3 |
| ---: | ---: | ---: | ---: |
| am_in_ala_nat | 0.21 | 0.28 | 0.18 |
| asian | 0.21 | 0.22 | 0.15 |
| black | 0.72 | 0.72 | 0.85 |
| female_o18 | 39.25 | 40.15 | 38.90 |
| female_u18 | 10.77 | 9.79 | 10.98 |
| haw_pac | 0.00 | 0.00 | 0.00 |
| hisp_lat | 2.92 | 2.63 | 3.08 |
| male_o18 | 38.26 | 39.38 | 37.99 |
| male_u18 | 11.44 | 10.27 | 11.70 |
| other | 0.32 | 0.30 | 0.41 |
| white | 94.35 | 93.67 | 92.52 |

### 4.1.3 *Housing Cost*

For housing cost, risk levels are determined by which clusters have the highest cluster medians. Cluster 1 has the highest value for mortgage high cost. Cluster 2 has the lowewest mortgage and rent high cost cluster medians. cluster 3 has the highest no mortgage and rent high cost cluster medians. Table 4.3 shows the values for this sector. Cluster one becomes the medium risk level, cluster two becomes the low risk level, and cluster three becomes the high risk level.

Table 4.3: Median Values for Housing Cost Clusters

| Variable | Cluster 1 | Cluster 2 | Cluster 3 |
| ---: | ---: | ---: | ---: |
| mortgage_high_cost | 5.22 | 4.35 | 4.93 |
| no_mortgage_high_cost | 2.16 | 2.18 | 2.89 |
| rent_high_cost | 15.69 | 14.18 | 16.79 |

### 4.1.4 *Housing Quality*

For housing quality, risk levels are determined by which clusters have the highest values, with preference given to occupied housing as housing conditions in unoccupied housing is of a lesser concern than occupied housing. Table 4.4 shows the values for this sector. Cluster one has the highest values for unoccupied housing with incomplete kitchens and plumbing. Cluster three has the medium value for each variable.Cluster three has the highest values for occupied housing with incomplete kitchens and plumbing. Cluster one becomes the lowest risk level, cluster two becomes the medium risk level, and cluster three becomes the highest risk level.

Table 4.4: Median Values for Housing Quality Clusters

| Variable | Cluster 1 | Cluster 2 | Cluster3 |
|---|---|---|---|
| all_incomplete_kitchen | 25.85 | 25.76 | 19.75 |
| all_incomplete_plumb | 24.00 | 22.73 | 17.28 |
| occ_incomplete_kitchen | 0.46 | 0.52 | 0.64 |
| occ_incomplete_plumb | 0.00 | 0.11 | 0.34 |

### 4.1.5 ***Residential Mobility: Education***

For residential mobility: education, the risk levels are determined primarily by the variables for those who moved with less than a high school education and those in the same house with less than a high school education, and the clusters where more people moved overall will be the highest risk levels. Table 4.5 shows the values for this sector. Cluster one has the medium value for 71 percent of variables including each less than high school education variable. Cluster two has the lowest values for each variable. cluster 3 has the highest values for 71 percent of variables, including each of the less than high school education variables. Cluster one becomes the lowest risk level because it has medium levels of residential mobility but the highest level of residential stability with a high school education. Cluster two becomes the medium risk level, and cluster three becomes the highest risk level.

Table 4.5: Median Values for Residentail Mobility: Education Clusters

| Variable | Cluster 1 | Cluster 2 | Cluster 3 |
|---|---|---|---|
| moved_diff_county_hs | 0.51 | 0.44 | 0.55 |
| moved_diff_county_less_than_hs | 0.13 | 0.10 | 0.18 |
| moved_diff_state_hs | 0.18 | 0.10 | 0.18 |
| moved_diff_state_less_than_hs | 0.00 | 0.00 | 0.00 |
| moved_in_county_hs | 0.95 | 0.84 | 1.36 |
| moved_in_county_less_than_hs | 0.30 | 0.24 | 0.50 |
| same_house_hs | 23.99 | 22.60 | 22.97 |
| same_house_less_than_hs | 7.71 | 6.97 | 7.82 |

### 4.1.6 *Residential Mobility: Poverty*

The residential mobility: poverty sector follows the criteria of residential mobility: education closely with those the variables for those who moved that are below the poverty level are the highest priority. Table 4.6 shows the values for this sector. Cluster one has the lowest values for every variable. Cluster two has the medium value for 57 percent of variables. Cluster three has the highest values for 57 percent of variables including three of the below the poverty level variables. Cluster one becomes the lowest risk level, cluster two becomes the medium risk level, and cluster three becomes the highest risk level. One noteable observation from the cluster medians is that cluster three has the highest value of those living in the same house below and above the poverty measure, indicating high levels of poverty for census tractsin this cluster.

Table 4.6: Median Values for Residential Mobility: Poverty Clusters

| Variable | Cluster 1 | Cluster 2 | Cluster 3 |
|---|---|---|---|
| moved_diff_county_p1 | 0.30 | 0.41 | 0.48 |
| moved_diff_county_p2 | 0.04 | 0.12 | 0.07 |
| moved_diff_state_p1 | 0.05 | 0.10 | 0.08 |
| moved_diff_state_p2 | 0.00 | 0.00 | 0.00 |
| moved_in_county_p1 | 0.74 | 1.00 | 1.06 |
| moved_in_county_p2 | 0.30 | 0.43 | 0.40 |
| same_house_p1 | 9.86 | 10.80 | 12.14 |
| same_house_p2 | 7.79 | 8.55 | 9.04 |

### 4.1.7 *Wage and Household Factors*

For household wage/ aid, the clusters with the highest number of highest cluster medians determines the risk levels with particular attention given to households with no wage and households with three or more workers. Table 4.7 shows the values for this sector. Cluster one has the lowest cluster medians for 89 percent of variables. Cluster two has the medium value for 55 percent of variables. Cluster three has the highest values for 55 percent of variables and the middle value for the other variables. Notablable high values for cluster three include the Gini index, households with no vehicle, households with at least one worker and no vehicle, and households receiving supplemental security income.

Table 4.7: Median Values for Household Wage/ Aid Clusters

| | cluster_1 | cluster_2 | cluster_3 |
|---:|---:|---:|---:|
| gini_index | 42.69 | 42.91 | 44.07 |
| hh_3plus_worker | 1.85 | 1.67 | 1.81 |
| hh_no_investment_income | 32.64 | 33.04 | 33.29 |
| hh_no_other_income | 36.28 | 36.81 | 36.79 |
| hh_no_vehicle | 1.96 | 2.12 | 2.34 |
| hh_no_wage | 13.38 | 14.11 | 14.01 |
| hh_public_assistance | 4.84 | 5.54 | 5.40 |
| hh_ssi | 2.17 | 2.39 | 2.51 |
| hh_worker_no_vehicle | 1.28 | 1.45 | 1.59 |

### 4.1.8 *Housing Type*

For the housing type sector, owner single unit is considered the safest housing while renter and owners of unconventional housing and mobile homes are considered high risk. This sector requires a combination of means and medians for the analysis because for several variables, all cluster medians are zero. Table 4.8 shows the values for this sector. Cluster three has the highest owner mobile and the medium value for renter mobile. Cluster three has the highest renter and owner unconventional cluster averages. Cluster one has the highest owner single and the lowest renter mobile home. Cluster one becomes the low risk level, cluster two becomes the medium risk level, and cluster three becomes the high risk level.

Table 4.8: Median Values for Housing Type Clusters

| Variable | Cluster 1 | Cluster 2 | Cluster 3 |
|---|---|---|---|
| owner_2to4 | 0.00 | 0.00 | 0.00 |
| owner_5plus | 0.00 | 0.00 | 0.00 |
| owner_mobile | 8.29 | 10.06 | 10.41 |
| owner_single | 90.75 | 88.67 | 88.42 |
| owner_unconvent | 0.00 | 0.00 | 0.00 |
| renter_2to4 | 8.29 | 10.57 | 10.59 |
| renter_5plus | 5.78 | 8.50 | 7.70 |
| renter_mobile | 9.25 | 13.04 | 10.79 |
| renter_single | 68.16 | 55.67 | 60.95 |
| renter_unconvent | 0.00 | 0.00 | 0.00 |

## 4.2 *Association Rules*

For 1:1 rules, the rule with the highest confidence and support is between housing type and cost, occurring in 11 percent of transactions with a coverage of 0.29. Wage/ household aid has the most one-one- associations in this group occurring in an average of 1 percent of observations, there is a 36 percent average probability that any of these sectors are in the consequent if wage/ household aid is in the antecedent. In about 9 percent of observations, residential mobility poverty is the consequent if mobility education is the antecedent with a 36 percent probability of a high risk for mobility poverty when there is a high risk of mobility education. mobility education also has an association with employment with 33 percent confidence and 8 percent support.

Of the 3:1 rules with the top 10 highest confidence values, employment diversity has a relationship with every other sector. Support is low at 5 percent with an average confidence of 0.42 and a slightly positive average lift of 1.1. The two other highest rules are housing type to mobility poverty with support of around twelve percent and a confidence interval of 0.41 and employment to residential mobility poverty with support of six percent and a confidence interval of 0.4. Looking at the top ten 3:1 rules with the highest confidence values, the results are significantly different from the 3:1 rules Employment is in the consequent if any other sector is in the antecedent with an average support of 12 percent, and an average confidence interval of 0.32. This relationship between employment diversity having both high-risk-to-low-risk and high-risk-to-low-risk is a surprising relationship.

## 4.3 *Moran's I*

| sector | Morans_I | std_dev | variance | expectation | p_value |
|---|---|---|---|---|---|
| Demographics | 0.32 | 7.02 | 0.00 | -0.01 | 0.00 |
| Employment | 0.25 | 5.38 | 0.00 | -0.01 | 0.01 |
| Household Wage/ Aid | 0.26 | 5.23 | 0.00 | -0.01 | 0.00 |
| Housing Cost | 0.21 | 4.43 | 0.00 | -0.01 | 0.01 |
| Housing Quality | 0.28 | 5.59 | 0.00 | -0.01 | 0.00 |
| Housing Type | 0.25 | 5.32 | 0.00 | -0.01 | 0.01 |
| RUCA | 0.31 | 5.86 | 0.00 | -0.01 | 0.00 |
| Transience: Education | 0.23 | 4.77 | 0.00 | -0.01 | 0.01 |
| Transience: Poverty | 0.19 | 4.11 | 0.00 | -0.01 | 0.01 |

While the association rules dealt exclusively with the housing insecurity risk levels, Moran's I spatial autocorrelation is used to examine how values group together in space for all variables. Moran's I is calculated for every state and the entire dataset. rewrite all of this analysis

### 4.3.1 Moran's I Outliers

While Moran's I values overall are fairly low, there are some outliers that deserve attention. Demographics, household wage/ aid, do not have any outliers based on the IRQ method. mobility poverty has 13 outliers. The mean of all mobility poverty observations is 0.19 while the mean for the outliers is 0.47. 69 percentof these outliers are the same house below the poverty line variable. Connecticut, Nevada, and Arizona have surprisingly high Moran's I statistics for the mobility poverty risk levels variable. The average for these three states is 0.45 compared to 0.21 for the same variable. There are 4 outliers in the mobility education with same house less than high school in Ohio, California, and all states. The final outlier is same house with a high school diploma in Maryland. These outliers have an average of 0.61 while all sector observations have an average of 0.23. For housing type, there are two outliers: owner single unit and owner mobile home, both in the South Dakota. These outliers have an average Moran's I of 0.63 while the sector has an average of 0.25. for Housing quality there are two outliers: occupied incomplete plumbing and occupied incomplete kitchen, both in the state of New Mexico. The sector average is 0.28 while these outliers have an average of 0.66. Housing cost has six outliers: mortgage high cost in Arizona, Maryland, Minnesota, Nevada, and New Mexico. The variable average is 0.28 while these observations have an average of 0.5. For economic diversity there are 16 outliers. 10 of these observations are for manufacturing in all states, Virginia, Florida, Indiana, Kentucky, Mississippi, Ohio, Pennsylvania, South Dakota, and Virginia. The average Moran's I statistic for this sector was 0.47 while these outliers have an average of 0.68. 5 of these outliers are for the agriculture, forestry, fishing, hunting, and mining variable in New Mexico, Oklahoma, Texas, Washington, and all states. The average Moran's I statistic for this variable is 0.36 while these outliers have an average of 0. 61. These outliers provide useful insights into states where conditions do not align with the trends found in the whole dataset and may merit further investigation, especially since all outliers are positive.

this table shows the average local Moran's I for each sector's risk levels

| sector | c1 | c2 | c3 |
|---|---|---|---|
| emp_cluster | 0.83 | 0.00 | 0.89 |
| dem_cluster | 2.09 | 0.14 | 0.58 |
| trans_edu_cluster | 1.28 | 0.10 | 0.25 |
| trans_pov_cluster | 1.16 | 0.05 | 0.48 |
| cost_cluster | 0.79 | -0.00 | 0.57 |
| qual_cluster | 1.06 | 0.00 | 0.74 |
| hhtype_cluster | 0.91 | -0.00 | 0.68 |
| waid_cluster | 1.52 | 0.09 | 0.22 |

### 4.4 *Multinomial Logistic Regression*

The final method applied in this study is a multinomial logistic regression performed on each sector of data and tested on the data for each state. The probability that a predicted risk level is the actual risk level is used to measure how well the data for each state can be predicted based on a model trained on the other states. Average accuracy for all sectors was low: employment diversity, housing quality, residential mobility: poverty and household factors had an average of 0.34; housing type and housing cost had an average of 0.35; residential mobility: poverty had an average of 0.36; demographics had the highest average accuracy at 0.38. Demographics had the highest standard deviation at 14 percent, indicating a high degree of variation in predictability. Looking at average probability for each sector in each state, Utah had the best prediction results with an average of 41 percent and Minnesota was the hardest to predict at 31 percent between sectors. The standard deviation of every state is 14 percent, indicating a high level of variation in each state. While the standard deviation is significant, the mean and median statistics indicate the data is close to symmetrically distributed. In the following subsections, the performance of the state models is evaluated based on the confusion matrices of the actual classification and the predicted classifications. Of primary concern is the misclassification of census tracts as misclassifying risk levels could mislead researchers and policy makers.

#### 4.4.1 *Employment Diversity*

The state models were able to correctly predict 35 percent of census tracts. This model shows a heavy bias towards the low risk levelclassification. Table 4.9 shows that the model struggled to classify the medium risk level, yielding an F1 score of 0.15 compared to 0.44 for the low risk levels and 0.34 for the high risk levels.

When accounting for class imbalance, accuracy increases by 12 to 17 percent.

Table 4.9: Employment Confusion Matrix and Statistics

| | **High Risk** | **Medium Risk** | **Low Risk** |
|---|---|---|---|
| **High Risk** | 708 | 675 | 654 |
| **Medium Risk** | 386 | 232 | 431 |
| **Low Risk** | 987 | 1051 | 1238 |
| Precision | 0.35 | 0.22 | 0.38 |
| Recall | 0.34 | 0.12 | 0.53 |
| F1 | 0.34 | 0.15 | 0.44 |
| Prevalence | 0.33 | 0.31 | 0.37 |
| Detection Rate | 0.11 | 0.04 | 0.19 |
| Detection Prevalence | 0.32 | 0.16 | 0.51 |
| Balanced Accuracy | 0.51 | 0.47 | 0.51 |

Table 4.10: Demographics Confusion Matrix and Statistics

| | **High Risk** | **Medium Risk** | **Low Risk** |
|---|---|---|---|
| **High Risk** | 48 | 95 | 76 |
| **Medium Risk** | 843 | 1975 | 1274 |
| **Low Risk** | 378 | 666 | 1007 |
| Precision | 0.22 | 0.48 | 0.49 |
| Recall | 0.04 | 0.72 | 0.43 |
| F1 | 0.06 | 0.58 | 0.46 |
| Prevalence | 0.2 | 0.43 | 0.37 |
| Detection Rate | 0.01 | 0.31 | 0.16 |
| Detection Prevalence | 0.03 | 0.64 | 0.32 |
| Balanced Accuracy | 0.5 | 0.57 | 0.58 |

### 4.4.2 *Demographics*

### 4.4.3 *Housing Cost*

### 4.4.4 *Housing Quality*

### 4.4.5 *Residential Mobility: Education*

### 4.4.6 *Residential Mobility: Poverty*

### 4.4.7 *Wage and Household Factors*

### 4.4.8 *Housing Type*

Table 4.11: Confusion Matrix and Statistics

| | **High Risk** | **Medium Risk** | **Low Risk** |
|---|---|---|---|
| **High Risk** | 669 | 446 | 422 |
| **Medium Risk** | 660 | 962 | 817 |
| **Low Risk** | 625 | 803 | 938 |
| Precision | 0.44 | 0.39 | 0.4 |
| Recall | 0.34 | 0.44 | 0.43 |
| F1 | 0.38 | 0.41 | 0.41 |
| Prevalence | 0.31 | 0.35 | 0.34 |
| Detection Rate | 0.11 | 0.15 | 0.15 |
| Detection Prevalence | 0.24 | 0.38 | 0.37 |
| Balanced Accuracy | 0.57 | 0.54 | 0.54 |

Table 4.12: Confusion Matrix and Statistics

| | **High Risk** | **Medium Risk** | **Low Risk** |
|---|---|---|---|
| **High Risk** | 285 | 278 | 179 |
| **Medium Risk** | 592 | 296 | 673 |
| **Low Risk** | 954 | 1632 | 1458 |
| Precision | 0.38 | 0.19 | 0.36 |
| Recall | 0.16 | 0.13 | 0.63 |
| F1 | 0.22 | 0.16 | 0.46 |
| Prevalence | 0.29 | 0.35 | 0.36 |
| Detection Rate | 0.04 | 0.05 | 0.23 |
| Detection Prevalence | 0.12 | 0.25 | 0.64 |
| Balanced Accuracy | 0.53 | 0.41 | 0.5 |

Table 4.13: Confusion Matrix and Statistics

| | **High Risk** | **Medium Risk** | **Low Risk** |
|---|---|---|---|
| **High Risk** | 219 | 200 | 197 |
| **Medium Risk** | 427 | 881 | 810 |
| **Low Risk** | 917 | 1169 | 1542 |
| Precision | 0.36 | 0.42 | 0.43 |
| Recall | 0.14 | 0.39 | 0.6 |
| F1 | 0.2 | 0.4 | 0.5 |
| Prevalence | 0.25 | 0.35 | 0.4 |
| Detection Rate | 0.03 | 0.14 | 0.24 |
| Detection Prevalence | 0.1 | 0.33 | 0.57 |
| Balanced Accuracy | 0.53 | 0.55 | 0.53 |

Table 4.14: Confusion Matrix and Statistics

| | **High Risk** | **Medium Risk** | **Low Risk** |
|---|---|---|---|
| **High Risk** | 298 | 394 | 353 |
| **Medium Risk** | 420 | 287 | 324 |
| **Low Risk** | 1182 | 1356 | 1748 |
| Precision | 0.29 | 0.28 | 0.41 |
| Recall | 0.16 | 0.14 | 0.72 |
| F1 | 0.2 | 0.19 | 0.52 |
| Prevalence | 0.3 | 0.32 | 0.38 |
| Detection Rate | 0.05 | 0.05 | 0.27 |
| Detection Prevalence | 0.16 | 0.16 | 0.67 |
| Balanced Accuracy | 0.49 | 0.48 | 0.54 |

Table 4.15: Confusion Matrix and Statistics

| | **High Risk** | **Medium Risk** | **Low Risk** |
|---|---|---|---|
| **High Risk** | 195 | 163 | 208 |
| **Medium Risk** | 492 | 748 | 897 |
| **Low Risk** | 972 | 1352 | 1335 |
| Precision | 0.34 | 0.35 | 0.36 |
| Recall | 0.12 | 0.33 | 0.55 |
| F1 | 0.18 | 0.34 | 0.44 |
| Prevalence | 0.26 | 0.36 | 0.38 |
| Detection Rate | 0.03 | 0.12 | 0.21 |
| Detection Prevalence | 0.09 | 0.34 | 0.58 |
| Balanced Accuracy | 0.52 | 0.5 | 0.48 |

Table 4.16: Confusion Matrix and Statistics

| | **High Risk** | **Medium Risk** | **Low Risk** |
|---|---|---|---|
| **High Risk** | 195 | 163 | 208 |
| **Medium Risk** | 492 | 748 | 897 |
| **Low Risk** | 972 | 1352 | 1335 |
| Precision | 0.34 | 0.35 | 0.36 |
| Recall | 0.12 | 0.33 | 0.55 |
| F1 | 0.18 | 0.34 | 0.44 |
| Prevalence | 0.26 | 0.36 | 0.38 |
| Detection Rate | 0.03 | 0.12 | 0.21 |
| Detection Prevalence | 0.09 | 0.34 | 0.58 |
| Balanced Accuracy | 0.52 | 0.5 | 0.48 |

# CHAPTER 5

# DISCUSSION

RQ1 Across all sectors except employment diversity, the smallest cluster count is the number of high-risk census tracts. In all sectors except housing cost and demographics, there are more low risk-level census tracts than medium-risk census tracts. Demographics has a much lower number of high-risk clusters proportional to the other risk levels compared to other sectors. The high number of high-risk census tracts in the housing cost sector is concerning as the social consequences of high housing costs are well documented. This distribution pattern underscores potential areas for targeted housing insecurity interventions and policy considerations.

RQ2 the association rules show us that while there are some pockets of high-risk-to-high-risk, there are a considerable number of high-risk-to-low-risk and low-risk-to-high-risk relationships in the dataset, contrary to what was expected after reviewing the literature on housing insecurity and related factors. The number of high-risk-to-high-risk relationships that the household wage/ aid sector has with employment, housing costs, mobility poverty, and housing type demonstrates the interconnectedness of housing insecurity However, the high-risk-to-low-risk and low-risk-to-high-risk associations highlight the variation in rural areas demonstrated by the spatial analysis. Based on the 4 C's model, sectors should have similar risk levels. While the results reflect this, the mix of results implies that adjustments are needed to this implementation of it or the theoretical foundation of the 4 Cs model.

RQ3 Numerous studies have shown that poverty clusters around itself in urban areas (Foulkes & Schafft, 2010) as the spatial analysis shows, in rural areas there is little clustering of high levels of poverty for those that moved and those that did not. High housing costs also do not conform to what was expected, with low levels of spatial clustering. One significant observation is the notably higher Moran's I statistic for high-cost housing with a mortgage. In urban areas, renters tend to have higher costs than homeowners. One area where rural and urban areas are similar is in levels of racial segregation. African Americans, Whites, Hispanics, and Latinos have some of the highest spatial autocorrelations in the dataset. The presence of pockets of Hispanics was identified in literature by Lichter (2020) and this analysis supports their findings. Economic diversity results reflect some of the stereotypes of rural areas. Jobs considered to be more rural like mining, agriculture, forestry, and manufacturing had higher spatial autocorrelations than "urban jobs" that primarily serve a consumer rather than a producer community. The results also indicate that the ten previously mentioned states have notably stronger spatial autocorrelations than the overall average. For both residential mobility sectors, those who did not move regardless of high school diploma or poverty status have the highest spatial autocorrelations of their sectors. For all residential mobility variables there are small positive spatial autocorrelations. Desmond and various collaborators have identified patterns of residential mobility, primarily based on a sample of Milwaukee residents. These studies provide great context into residential mobility in urban areas, but no similar studies have been done unique to rural areas. Housing quality was difficult to measure as there are a limited number of variables in the ACS that relate to it. Unoccupied housing with incomplete facilities is more spatially clustered than occupied housing with incomplete facilities. This necessitates a further analysis with a broader range of housing condition features accounted for. Housing cost variables did not have strong spatial autocorrelations except in the previously mentioned states with an average of 0.5 for these states and 0.28 for the entire sector. Research has identified mobile homes as a common housing solution in rural areas, despite their potential health and financial risk. The prevalence of high spatial autocorrelation reflects the widespread use of mobile homes in rural areas. The other variables of interest in this sector are the owner and renter unconventional housing variables which have much lower spatial clustering than single unit and mobile homeowners. Regarding RQ3, there does not appear to be significant spatial clustering of the housing insecurity indicators. This indicates that the urban-centric understanding that indicators of housing insecurity tend to spatially cluster, the same may not be true for rural areas. Alternatively, it may be that urban-centric indicators of housing insecurity do not translate well to rural conditions.

RQ4 The multinomial logistic regression reveals how different rural areas are. Figure Z shows the distribution of probabilities across sectors. While most of the sectors are close to being symmetrically distributed, there are a lot of outliers across most sectors. The same trend exists when looking at state averages. The symmetrical distribution indicates similar levels of variation across states as values are similar on each side of the mean. A more advanced classification algorithm such as K-Nearest Neighbors or a Bayesian classifier may be more accurate but based on the multinomial distribution, the answer to RQ4 is that states cannot be predicted very accurately based on the other states in the data set. The implication is another reason a greater understanding of rural-specific housing insecurity indicators is needed and that researchers should find methods of studying rural housing insecurity that can accommodate the differences within and between rural areas. *Figure 5 about here* The risk assessment system is represented with a map where each census tract's color is based on a combination of its risk level and cluster probability based on the multinomial logistic regression (see figure 6). Risk levels are denoted as red for one, orange for two, and green for three. Many rural areas fall somewhere in the middle in terms of housing insecurity risk. There are pockets of high risk and low risk census tracts that can be seen in the national map or the regional breakdown. The light colors on the map indicate that the multinomial logistic regression models were not able to predict the risk level of census tracts very well. The trends reflected in the relative risk assignment system signify a wide dispersion of housing insecurity in rural areas.

Due to the urban-centric lens towards housing insecurity, there is little previous research to compare this study to. Gleason et al. 2021 applied similar spatial techniques to census tracts in Maine and found that poverty, unemployment, and high housing costs are common in rural and urban areas of Maine but found these results to be inaccurate in a later study (Gleason et al 2022?). Lichter and Johnson 2007 did a nationwide county-level analysis on poverty levels specific to rural areas. Insert authors that did research specific to rural areas. This is the first study of this scale to employ data mining techniques on rural housing insecurity. Further, it provides researchers with something to compare to as they begin to conduct more research specific to rural areas. Policy makers at all levels of government can use this as a tool for delegating resources available to their community. One significant limitation is the lower rate of ACS sampling in rural areas, the accuracy of the data is limited in how well it represents the real-world. While the estimates are "likely reasonable approximations of the populations they represent", small area estimates like census tracts used here have issues with attribute uncertainty (Spielman 2014, p#). Despite this, it is currently the most detailed source of data available for rural areas.

Future research should use this study as a starting point for giving housing insecurity and homelessness adequate attention. The most important direction is to identify community level risk factors unique to rural areas. Further studies should also use a wider range of data sources to capture sectors with few available variables such as housing conditions. From our spatial analysis there are several observations worthy of further investigation. Subsequent investigations should focus on examining rural housing insecurity at a localized level. This will enable the refinement and enhancement of the existing model, providing more precise insights into the unique challenges faced by rural communities. Future endeavors should prioritize a closer examination of regions exhibiting unexpected high-risk-to-low-risk and low-risk-to-high-risk relationships, as identified through association rules. Understanding the underlying factors contributing to these unexpected relationships is essential for targeted interventions and policy recommendations. There is a need for in-depth research to discern how levels and trends in income inequality differ between urban and rural areas, shedding light on the specific socio-economic dynamics impacting housing insecurity in each setting. Future research should scrutinize the distinctions in poverty and housing cost dynamics between rural and urban areas, aiming to gain a deeper understanding of the factors at play in each context. By addressing these research gaps, researchers can better inform evidence-based policies and interventions that mitigate housing insecurity and advance the well-being of rural populations. As we strive to enhance housing security and social equity in both rural and urban landscapes, interdisciplinary collaboration and persistent research efforts will remain pivotal in driving meaningful societal change. We encourage scholars and practitioners to join hands in this endeavor, working collectively towards a future where safe and stable housing is a fundamental right for all.

# CHAPTER 6

# CONCLUSION

Future research should use this study as a starting point for giving housing insecurity and homelessness adequate attention. The most important direction is to identify community level risk factors unique to rural areas. Further studies should also use a wider range of data sources to capture sectors with few available variables such as housing conditions. From our spatial analysis there are several observations worthy of further investigation. Subsequent investigations should focus on examining rural housing insecurity at a localized level. This will enable the refinement and enhancement of the existing model, providing more precise insights into the unique challenges faced by rural communities. Future endeavors should prioritize a closer examination of regions exhibiting unexpected high-risk-to-low-risk and low-risk-to-high-risk relationships, as identified through association rules. Understanding the underlying factors contributing to these unexpected relationships is essential for targeted interventions and policy recommendations. There is a need for in-depth research to discern how levels and trends in income inequality differ between urban and rural areas, shedding light on the specific socio-economic dynamics impacting housing insecurity in each setting. Future research should scrutinize the distinctions in poverty and housing cost dynamics between rural and urban areas, aiming to gain a deeper understanding of the factors at play in each context. By addressing these research gaps, researchers can better inform evidence-based policies and interventions that mitigate housing insecurity and advance the well-being of rural populations. As we strive to enhance housing security and social equity in both rural and urban landscapes, interdisciplinary collaboration and persistent research efforts will remain pivotal in driving meaningful societal change. We encourage scholars and practitioners to join hands in this endeavor, working collectively towards a future where safe and stable housing is a fundamental right for all.

## REFERENCES

D. Augustus Anderson, Hye-Sung Han, and John Hisnanick. The Effect of Household Debt and Wealth on Subsequent Housing Tenure Choice. *City & Community*, 20(4):297–325, December 2021. ISSN 1535-6841, 1540-6040. doi: 10.1177/15356841211007757. URL `http://journals.sagepub.com/doi/10.1177/15356841211007757`.

Rüdiger Bachmann and Daniel Cooper. The ins and arounds in the U.S. housing market. Working Paper, Federal Reserve Bank of Boston, 2014.

Holly Barcus. Heterogeneity of rural housing markets. *Rural Housing, Exurbanization, and Amenity-Driven Development: Contrasting the 'Haves' and the 'Have Nots'*, pages 51–73, January 2011.

Seth A. Berkowitz and Deepak Palakshappa. Gaps in the welfare state: A role-based model of poverty risk in the U.S. *PLOS ONE*, 18(4):e0284251, April 2023. ISSN 1932-6203. doi: 10.1371/journal.pone.0284251. URL `https://dx.plos.org/10.1371/journal.pone.0284251`.

Evelyn Blumenberg and Gregory Pierce. Automobile Ownership and Travel by the Poor. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2320:28–36, 2012.

Evelyn Blumenberg, Anne Brown, and Andrew Schouten. Car-deficit households: determinants and implications for household travel in the U.S. *Transportation*, 47(3): 1103–1125, June 2020. ISSN 0049-4488, 1572-9435. doi: 10.1007/s11116-018-9956-6. URL `http://link.springer.com/10.1007/s11116-018-9956-6`.

Paula Braveman, Mercedes Dekker, Susan Egerter, Tabashir Sadegh-Nobari, and Craig Pollack. Housing and Health, 2011.

Sarah K. Bruch, Joseph Van Der Naald, and Janet C. Gornick. Poverty Reduction through Federal and State Policy Mechanisms: Variation over Time and across the United States. *Social Service Review*, 97(2):270–319, June 2023. ISSN 0037-7961, 1537-5404. doi: 10.1086/724556. URL https://www.journals.uchicago.edu/doi/10.1086/724556.

Emery N. Castle, JunJie Wu, and Bruce A. Weber. Place Orientation and Rural–Urban Interdependence. *Applied Economic Perspectives and Policy*, 33(2):179–204, June 2011. ISSN 2040-5790, 2040-5804. doi: 10.1093/aepp/ppr009. URL https://onlinelibrary.wiley.com/doi/abs/10.1093/aepp/ppr009.

Jing Chen. Economic Diversity and Regional Economic Performance: A Methodological Concern from Model Uncertainty. Working Paper, West Virginia University, 2018.

John Cromartie and Shawn Bucholtz. Defining the "Rural" in Rural America. *Amber Waves*, 6(3), 2008.

Steven Deller and Philip Watson. Spatial variations in the relationship between economic diversity and stability. *Applied Economics Letters*, 23(7):520–525, May 2016. ISSN 1350-4851, 1466-4291. doi: 10.1080/13504851.2015.1085630. URL http://www.tandfonline.com/doi/full/10.1080/13504851.2015.1085630.

Stefanie DeLuca and Eva Rosen. Housing Insecurity Among the Poor Today. *Annual Review of Sociology*, 48(1):343–371, July 2022. ISSN 0360-0572, 1545-2115. doi: 10.1146/annurev-soc-090921-040646. URL https://www.annualreviews.org/doi/10.1146/annurev-soc-090921-040646.

Matthew Desmond and Monica Bell. Housing, Poverty, and the Law. *Annual Review of Law and Social Science*, 11(1):15–35, November 2015. ISSN 1550-3585, 1550-3631. doi: 10.1146/annurev-lawsocsci-120814-121623. URL https://www.annualreviews.org/doi/10.1146/annurev-lawsocsci-120814-121623.

Matthew Desmond and Carl Gershenson. Housing and Employment Insecurity among the Working Poor. *Social Problems*, 63(1):46–67, February 2016. ISSN 0037-7791, 1533-8533. doi: 10.1093/socpro/spv025. URL https://academic.oup.com/socpro/article-lookup/doi/10.1093/socpro/spv025.

Matthew Desmond and Kristin L. Perkins. Housing and Household Instability. *Urban Affairs Review*, 52(3):421–436, May 2016. ISSN 1078-0874, 1552-8332. doi: 10.1177/1078087415589192. URL http://journals.sagepub.com/doi/10.1177/1078087415589192.

Matthew Desmond and Tracey Shollenberger. Forced Displacement From Rental Housing: Prevalence and Neighborhood Consequences. *Demography*, 52(5):1751–1772, October 2015. ISSN 0070-3370, 1533-7790. doi: 10.1007/s13524-015-0419-9. URL https://read.dukeupress.edu/demography/article/52/5/1751/169475/Forced-Displacement-From-Rental-Housing-Prevalence.

Matthew Desmond and Nathan Wilmers. Do the Poor Pay More for Housing? Exploitation, Profit, and Risk in Rental Markets. *American Journal of Sociology*, 124(4): 1090–1124, January 2019. ISSN 0002-9602, 1537-5390. doi: 10.1086/701697. URL https://www.journals.uchicago.edu/doi/10.1086/701697.

Rachel Bogardus Drew. Believing in Homeownership: Behavioral Drivers of Housing Tenure Decisions. Technical report, Joint Center for Housing Studies, Harvard Univerity, 2014.

D D'Alessandro and L Appolloni. Housing and health: an overview. *Ann Ig*, 32(5):17–26, 2020.

Price V. Fishback. Social Insurance and Public Assistance in the Twentieth-Century United States. *The Journal of Economic History*, 80(2):311–350, June 2020. ISSN 0022-0507, 1471-6372. doi: 10.1017/S0022050720000200. URL https://www.cambridge.org/core/product/identifier/S0022050720000200/type/journal_article.

Jason M. Fletcher, Tatiana Andreyeva, and Susan H. Busch. Assessing the effect of changes in housing costs on food insecurity. *Journal of Children and Poverty*, 15(2): 79–93, September 2009. ISSN 1079-6126, 1469-9389. doi: 10.1080/10796120903310541. URL http://www.tandfonline.com/doi/abs/10.1080/10796120903310541.

MEnelaos Gkartzios and Ann Ziebarth. Housing: A Lens to rural Inequalities. *International Handbook of Rural Studies*, November 2017.

Kristin Gleason, Matthew Dube, and Jennifer Martin. Using Geographic Information Systems to Assess Community-Level Vulnerability to Housing Insecurity in Rural Areas. *Journal of Community Psychology*, 50(4):1993–2012, 2021.

Christopher Herbert, Alexander Hermann, and Daniel McCue. Measuring Housing Affordability: Assessing the 30-Percent of Income Standard, 2018.

Diana Hernández and Carolyn B. Swope. Housing as a Platform for Health and Equity: Evidence and Future Directions. *American Journal of Public Health*, 109(10):1363–1366, October 2019. ISSN 0090-0036, 1541-0048. doi: 10.2105/AJPH.2019.305210. URL https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2019.305210.

Lingqian Hu and Liming Wang. Housing location choices of the poor: does access to jobs matter? *Housing Studies*, 34(10):1721–1745, November 2019. ISSN 0267-3037, 1466-1810. doi: 10.1080/02673037.2017.1364354. URL https://www.tandfonline.com/doi/full/10.1080/02673037.2017.1364354.

David E. Jacobs. Environmental Health Disparities in Housing. *American Journal of Public Health*, 101(S1):S115–S122, December 2011. ISSN 0090-0036, 1541-0048. doi: 10.2105/AJPH.2010.300058. URL https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2010.300058.

Seungbeom Kang. Why Low-Income Households Become Unstably Housed: Evidence From the Panel Study of Income Dynamics. *Housing Policy Debate*, 29(4):559–587, July 2019.

ISSN 1051-1482, 2152-050X. doi: 10.1080/10511482.2018.1544161. URL https://www.tandfonline.com/doi/full/10.1080/10511482.2018.1544161.

Seungbeom Kang. Severe and persistent housing instability: examining low-income households' residential mobility trajectories in the United States. *Housing Studies*, pages 1–26, October 2021. ISSN 0267-3037, 1466-1810. doi: 10.1080/02673037.2021.1982871. URL https://www.tandfonline.com/doi/full/10.1080/02673037.2021.1982871.

Huiyun Kim, Sarah A. Burgard, and Kristin S. Seefeldt. Housing Assistance and Housing Insecurity: A Study of Renters in Southeastern Michigan in the Wake of the Great Recession. *Social Service Review*, 91(1):41–70, March 2017. ISSN 0037-7961, 1537-5404. doi: 10.1086/690681. URL https://www.journals.uchicago.edu/doi/10.1086/690681.

Lauren J. Krivo and Robert L. Kaufman. Housing and wealth inequality: Racial-ethnic differences in home equity in the United States. *Demography*, 41(3):585–605, August 2004. ISSN 0070-3370, 1533-7790. doi: 10.1353/dem.2004.0023. URL https://read.dukeupress.edu/demography/article/41/3/585/170331/Housing-and-wealth-inequality-Racial-ethnic.

Jessica N. Kropczynski and Patricia H. Dyk. Insights into Housing Affordability for Rural Low-Income Families. *Housing and Society*, 39(2):125–148, January 2012. ISSN 0888-2746, 2376-0923. doi: 10.1080/08882746.2012.11430603. URL http://www.tandfonline.com/doi/full/10.1080/08882746.2012.11430603.

Barrett A. Lee and Megan Evans. Forced to move: Patterns and predictors of residential displacement during an era of housing insecurity. *Social Science Research*, 87:102415, March 2020. ISSN 0049089X. doi: 10.1016/j.ssresearch.2020.102415. URL https://linkinghub.elsevier.com/retrieve/pii/S0049089X20300132.

Barrett A. Lee, Marybeth Shinn, and Dennis P. Culhane. Homelessness as a Moving Target. *The ANNALS of the American Academy of Political and Social Science*, 693(1): 8–26, January 2021. ISSN 0002-7162, 1552-3349. doi: 10.1177/0002716221997038. URL `http://journals.sagepub.com/doi/10.1177/0002716221997038`.

Kathryn M Leifheit, Gabriel L Schwartz, Craig Evan Pollack, and Sabriya L Linton. Building health equity through housing policies: critical reflections and future directions for research. *Journal of Epidemiology and Community Health*, 76(8):759–763, August 2022. ISSN 0143-005X, 1470-2738. doi: 10.1136/jech-2021-216439. URL `https://jech.bmj.com/lookup/doi/10.1136/jech-2021-216439`.

Daniel T. Lichter and David L. Brown. Rural America in an Urban Society: Changing Spatial and Social Boundaries. *Annual Review of Sociology*, 37(1):565–592, August 2011. ISSN 0360-0572, 1545-2115. doi: 10.1146/annurev-soc-081309-150208. URL `https://www.annualreviews.org/doi/10.1146/annurev-soc-081309-150208`.

Daniel T. Lichter and Kenneth M. Johnson. The Changing Spatial Concentration of America's Rural Poor Population*. *Rural Sociology*, 72(3):331–358, September 2007. ISSN 00360112. doi: 10.1526/003601107781799290. URL `http://doi.wiley.com/10.1526/003601107781799290`.

Daniel T. Lichter and Kenneth M. Johnson. A Demographic Lifeline? Immigration and Hispanic Population Growth in Rural America. *Population Research and Policy Review*, 39(5):785–803, October 2020. ISSN 0167-5923, 1573-7829. doi: 10.1007/s11113-020-09605-8. URL `https://link.springer.com/10.1007/s11113-020-09605-8`.

Daniel T. Lichter, David L. Brown, and Domenico Parisi. The rural–urban interface: Rural and small town growth at the metropolitan fringe. *Population, Space and Place*, 27(3), April 2021. ISSN 1544-8444, 1544-8452. doi: 10.1002/psp.2415. URL https://onlinelibrary.wiley.com/doi/10.1002/psp.2415.

Sarah A Low. Rural Manufacturing at a Glance, 2017 Edition. *Plastics and rubber*, 2017.

Katherine MacTavish, Michelle Eley, and Sonya Salamon'. POLICY AND PRACTITIONER PERSPECTIVES. *Georgetown Journal on Poverty Law & Policy*, 13 (1):95–117, 2006.

Katherine A. MacTavish. The Wrong Side of the Tracks: Social Inequality and Mobile Home Park Residence. *Community Development*, 38(1):74–91, March 2007. ISSN 1557-5330, 1944-7485. doi: 10.1080/15575330709490186. URL http://www.tandfonline.com/doi/abs/10.1080/15575330709490186.

Thomas D. Matte and David E. Jacobs. Housing and Health- Current Issues and Implications for Research and Programs. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 7(1):7–25, 2000.

Molly W Metzger and Amy T Khare. Fair Housing and Inclusive Communities. *American Academy of Social Work and Social Welfare*, 2017.

Raven Molloy, Christopher L Smith, and Abigail Wozniak. Internal Migration in the United States. *Journal of Economic Perspectives*, 25(3):173–196, August 2011. ISSN 0895-3309. doi: 10.1257/jep.25.3.173. URL https://pubs.aeaweb.org/doi/10.1257/jep.25.3.173.

Katherine M. O'Regan, Ingrid Gould Ellen, and Sophie House. How to Address Homelessness: Reflections from Research. *The ANNALS of the American Academy of Political and Social Science*, 693(1):322–332, January 2021. ISSN 0002-7162, 1552-3349. doi: 10.1177/0002716221995158. URL http://journals.sagepub.com/doi/10.1177/0002716221995158.

Robin Phinney. Exploring Residential Mobility among Low-Income Families. *Social Service Review*, 87(4):780–815, December 2013. ISSN 0037-7961, 1537-5404. doi: 10.1086/673963. URL `https://www.journals.uchicago.edu/doi/10.1086/673963`.

Mark R. Rank and Thomas A. Hirshl. Welfare Use as a Life Course Event: Toward a New Understanding of the U.S. Safety Net. *Social Work*, 47(3):237–248, 2002.

Virginia A. Rauh, Philip J. Landrigan, and Luz Claudio. Housing and Health: Intersection of Poverty and Environmental Exposures. *Annals of the New York Academy of Sciences*, 1136(1):276–288, July 2008. ISSN 00778923. doi: 10.1196/annals.1425.032. URL `http://doi.wiley.com/10.1196/annals.1425.032`.

Kathleen Sherrieb, Fran H. Norris, and Sandro Galea. Measuring Capacities for Community Resilience. *Social Indicators Research*, 99(2):227–247, November 2010. ISSN 0303-8300, 1573-0921. doi: 10.1007/s11205-010-9576-9. URL `http://link.springer.com/10.1007/s11205-010-9576-9`.

Anne Shlay and Peter Rossi. Social Science Research and Contemporary Studies of Homelessness. *Annual Review of Sociology*, 18:129–160, 2003. doi: 10.1146/annurev.so.18.080192.001021.

Lilly Shoup and Becca Houma. Principles for Improving Transportation Options in Rural and Small Town Communities. White Paper, Transportation for America, 2010.

Thomas C Siskar. *Who Is Forced to Move?* Master of Arts, Pennsylvannia State University, 2019.

P. Johnelle Sparks, Corey S. Sparks, and Joseph J. A. Campbell. Poverty Segregation in Nonmetro Counties: A Spatial Exploration of Segregation Patterns in the US. *Spatial Demography*, 1(2):162–177, June 2013. ISSN 2364-2289, 2164-7070. doi: 10.1007/BF03354896. URL `https://link.springer.com/10.1007/BF03354896`.

Carolyn B. Swope and Diana Hernandez. Housing as a determinant of health equity_ A conceptual model. *Soc Sci Med*, 243, 2020. doi: 10.1016/j.socscimed.2019.112571. URL https://reader.elsevier.com/reader/sd/pii/S0277953619305659?token=0BF7CE998C4949B5AD4DB9DB6CE6CA02DF1448E46028A9D722AB36EBD1473C1B404CCF479C8632CC8BEF36 originRegion=us-east-1&originCreation=20230118070028.

U.S. Census Bureau and U.S. Department of Housing and Urban Development. Median Sales Price of Houses Sold for the United States, January 1963. URL https://fred.stlouisfed.org/series/MSPUS. Publisher: FRED, Federal Reserve Bank of St. Louis.

Michele Wakin. Not Sheltered, Not Homeless: RVs as Makeshifts. *American Behavioral Scientist*, 48(8):1013–1032, April 2005. ISSN 0002-7642, 1552-3381. doi: 10.1177/0002764204274197. URL http://journals.sagepub.com/doi/10.1177/0002764204274197.

John C. Weicher. Housing conditions and homelessness. *Gender Issues*, pages 35–53, 2006.

Anita Yadavalli, Brenna Rivett, James Brooks, Christiana K McFarland, and David Hardiman. A Comprehensive Look at Housing Market Conditions Across America's Cities. *Cityscape*, 22(2):111–132, 2020.