Analyzing the Structural Determinants of Homelessness in the United States: A Multilevel Modeling Approach

Braden Baseley
October 2020

Abstract

Homelessness continues to be a pervasive issue across the United States. According to the U.S. Department of Housing and Urban Development's *Annual Homeless Assessment Report to Congress*, more than half a million Americans experienced homelessness on a given night in 2018, with roughly one-third of them living on the streets. Homelessness leads to several sub-optimal social, health and economic outcomes; therefore, it is imperative for policymakers to understand the determinants of homelessness in order to develop interventions for at-risk populations. This study uses linear mixed-effects models to analyze four structural determinants of homelessness–housing characteristics, economic conditions, demographic composition and safety nets–across urban, suburban and rural settings. This study finds that housing dynamics have the strongest association with homelessness, which aligns with findings in earlier studies by Byrne et al. (2013) and Lee et al. (2010).

LITERATURE REVIEW

Definition of Homelessness

Perhaps one of the most difficult parts of researching homelessness is understanding how to operationalize it. Intuitively, we usually understand "homeless" to refer to those without a home. However, social scientists have used various definitions for homelessness over the past several decades. For example, during the tramp (1890s-1920s) and skid row (1940s-1970s) eras, sociologists believed homelessness referred to a degree of disaffiliation with society (Lee et al., 2010). However, researchers found it rather difficult and impractical to measure homelessness in this way, so modern definitions have moved toward "stress[ing] housing hardship linked to extreme poverty" (Lee et al., 2010, p. 502). For example, Rossi's (1989) definition of homelessness, which refers to the absence of "customary and regular access to a conventional dwelling", has been widely cited in contemporary studies (Lee et al., 2010, p. 503). The McKinney-Vento Homeless Assistance Act of 1987 describes homelessness in a similar way "but also specifies physical presence in selected locations— shelters, institutional settings, and places not intended for human habitation—as a sufficient condition to establish one's homelessness" (Lee et al., 2010, p. 503).

Despite its widespread adoption, modern definitions of homelessness are not entirely free of ambiguities. For example, researchers often disagree about what exactly constitutes a "shelter", or whether or not marginally housed persons, such as squatters or those who live in their cars, should be considered homeless. Moreover, recent studies largely ignore the temporal dimensions associated with homelessness. According to Lee et al. (2010), homelessness can span one of three temporal dimensions: (1) temporary or transitional, which suggests a rare or once-in-a-lifetime case of homelessness; (2) episodic, which refers to a short period of time marked by cyclical bouts of homelessness; or (3) chronic, which suggests a permanent condition. The latter dimension is typically overrepresented in cross-sectional studies, even though transitional and episodic homelessness are more common (Lee et al., 2010).

Ultimately, these ambiguities make it somewhat difficult for researchers to accurately count the homeless population in the United States. Nevertheless, according to Lee et al. (2010), the most current national figures come from the U.S. Department of Housing and Urban Development (HUD), which facilitates a point-in-time estimate across the country each year. HUD data has been commonly used for research studies about homelessness (Byrne et al., 2013; Elliott & Krivo, 1991; Hanratty, 2017).

According to HUD's *Annual Homeless Assessment Report to Congress*, more than half a million Americans experienced homelessness on a given night in 2018, with roughly one-third of them living on the streets (U.S. Department of Housing and Urban Development, 2018). Ultimately, homelessness leads to several sub-optimal social, health and economic outcomes (Burt, 2001; Hawkins & Abrams, 2007; Hwang, 2001; Lee & Farrell, 2003; Lee & Greif, 2008; Wolitski, Kidder, & Fenton, 2007; Zerger, 2002). Therefore, if policymakers can better understand what factors are associated with homelessness, they can intervene on at-risk populations more effectively.

Unfortunately, a one-size-fits-all approach to policy intervention is not sufficient for combatting the problem of homelessness, which can be ascribed to a lack of homogeneity among homeless individuals. Depending on the local context, homeless populations have varying demographic, socioeconomic and racial characteristics. Additionally, homelessness is much more prevalent in metropolises or suburban areas rather than rural locales (Burt et al., 2001), which may be attributable to factors like housing supply or other economic characteristics.

Therefore, it is important to control for an assortment of factors when conducting any analysis of homelessness.

Causes of Homelessness

Most researchers believe that homelessness arises from some combination of macro- and micro-level factors (Lee et al., 2010). Macro-level drivers are structural in nature, whereas micro-level determinants embody the individual characteristics that give rise to homelessness. Although Elliott and Krivo (1991) acknowledged the importance of personal factors in contributing to homelessness, they argued that structural issues are often an important precursor to widespread homelessness. As such, this paper focuses more narrowly on the macro-level or structural causes of homelessness. In particular, this paper looks at four structural determinants of homelessness: housing characteristics, economic conditions, demographic composition and safety net features.

Housing Characteristics

Past research has typically shown that housing characteristics, particularly the absence of affordable housing, tend to have the strongest associations with homelessness. Many researchers have consistently found a positive and statistically significant relationship between increased rent and homelessness (Bohanon, 1991; Byrne et al., 2013; Early & Olsen, 2002; Honig & Filer, 1993; Lee & Farrell, 2003; Quigley, 1990; Quigley & Raphael, 2002; Quigley et al., 2001; Troutman et al., 1999). Other widely-used housing measures that were found to have significant associations with homelessness include rental vacancy rates (Byrne et al., 2013; Hanratty, 2017; Quigley, 1990; Quigley & Raphael, 2002; Quigley et al., 2001; Troutman et al., 1999) and the percentage of renter-occupied households (Appelbaum et al., 1991; Elliott & Krivo, 1991; Hanratty, 2017). For example, Elliott and Krivo (1991) found a positive and statistically significant relationship between homelessness and a lack of low-cost housing, which they estimated by looking at the percentage of renter-occupied units renting below $150.

Economic Conditions

The economic conditions in a given area are often believed to have a direct effect on housing affordability. In fact, many researchers have found statistically significant relationships between economic conditions and homelessness. Most studies that have such relationships have either used unemployment rates (Appelbaum et al., 1991; Bohanon, 1991; Burt, 1993; Troutman et al., 1999) or poverty rates (Early & Olsen, 2002; Quigley, 1990; Raphael, 2010; Troutman et al., 1999) for measures of economic conditions. Interestingly, more recent studies like Byrne et al. (2013) and Hanratty (2017) that control for community-level effects have found that unemployment rates and poverty rates do not have a significant association with homelessness.

Demographic Composition

According to an analysis of prior literature by Byrne et al. (2013), no single demographic characteristic has been shown to have a consistent, statistically significant relationship with homelessness. Nevertheless, multiple studies have shown that a link exists between certain demographic groups and homelessness rates. For example, African Americans and Hispanics are overrepresented in homeless populations (Burt et al., 2001; U.S. Department of Housing and Urban Development, 2018). Moreover, Baby Boomers (those born between 1946 and 1964) account for a disproportionately high share of homeless populations, with Byrne et al. (2013)

finding a significant and positive association between Baby Boomers and homelessness rates. Other common demographic measures include the share of single-person households (Burt, 1993; Byrne et al., 2013; Lee & Farrell, 2003) and the proportion of veterans (Hanratty, 2017) in a given area.

Safety Net

The last structural determinant of homelessness to be discussed in this paper is the presence of social safety nets. In theory, more robust social safety net programs are associated with lower levels of homelessness. Although prior studies have confirmed the existence of such associations (Burt, 1993; Elliott & Krivo, 1991; Honig & Filer, 1993; Quigley & Raphael, 2002; Quigley et al., 2001; Troutman et al., 1999), Byrne et al. (2013) points out that the safety net measures are largely inconsistent across studies. These measures typically look at recipients of Social Security Income (SSI) and public assistance, as well as spending on mental health services.

Goals of the Study

The primary goal of this study is to verify if the aforementioned structural determinants of homelessness still hold today–a lot of prior studies on homelessness are a bit dated at this point. Moreover, I want to explore whether or not the effects of these structures vary within rural, suburban and urban areas, since earlier research (Burt et al., 2001) has shown considerable differences in homelessness rates in such areas. Lastly, this paper shall add to the existing body of literature by incorporating income inequality, as measured by the Gini index, into the analysis.

Among other things, this paper seeks to understand how income inequality fits into the structural schema of poor economic conditions. Income inequality is an abstract concept that can be difficult to quantify, so it can be operationalized in many different ways. The Gini index is perhaps the most well-known measure of income inequality, which is calculated by looking at how far an income distribution deviates from a perfectly equal distribution. This measure allows for easy comparison across different units of observations, which allows for simple comparisons across different states or metro areas. At the same time, one downside of the Gini index is that the same number can be calculated for different distributions, so researchers need to be careful in ascertaining the underlying distribution. Another popular measure of income inequality is the Atkinson index, which measures the percentage of income that must be given up in order to create more equal shares of income across a population. The strength of the Atkinson index comes from its ability to allow the researcher to make normative statements about society's aversion to inequality by setting a parameter (epsilon) that weights incomes, although this can be difficult in practice.

In the face of so many measures for income inequality, I wanted to understand if one measure is preferable over another. A study by Kawachi & Kennedy (1997) tested the relationship between six different measures of income inequality (Gini index, decile ratio, Robin Hood index, Atkinson index, Theil's entropy, and total proportion of income earned by bottom 50%, 60% and 70% of households) and mortality. The researchers found that the various measures of income inequality all yielded similar results with respect to its relationship with mortality. Moreover, all of the variables were highly correlated with one another. As such, I felt reasonably confident that I may use the Gini index due to its easy interpretation and conventional use in research.

DATA

Data Sources

To begin, I collected data on homelessness populations in the United States in 2018, which I sourced from the HUD's Point-in-Time (PIT) estimates of homelessness. HUD provides a point-in-time count of the total number of homeless people in about 3,000 cities and counties in the United States. The point-in-time data collection process occurs every year on a single night in January, wherein state and local planning agencies (known as Continuums of Care, or CoC) work alongside volunteers to identify homeless individuals and families that live in emergency shelters, transitional housing programs and unsheltered settings. A detailed description of how HUD defines homelessness can be found in Appendix B. HUD provides data either by CoC or state. For this study, I selected the CoC as the unit of analysis in order to avoid well-known statistical issues, such as Simpson's paradox, that may arise from aggregation at a state level.

Next, I collected my independent variables, which were collected for every county in the United States. I used the U.S. Census Bureau's American Community Survey (ACS), which surveys roughly 3.5 million households in the United States every year. The Census Bureau maintains a Master Address File (MAF), a list of all known living quarters and selected nonresidential units in the United States as well as Puerto Rico and island areas, from which the samples are drawn. However, this study excludes Puerto Rico and other territories from the analysis. ACS offers 1-year and 5-year estimates, but for the sake of this paper, I used the ACS 5-year estimates. ACS 1-year estimates are compiled annually for geographic areas that have at least 65,000 people. For areas with populations smaller than 65,000 people, the Census Bureau pools 5 consecutive years of ACS data to come up with the 5-year release. As such, the 5-year

ACS program offers statistically more reliable results compared to the 1-year ACS program. A list of the ACS codes used for the analysis can be found in Appendix B.

Lastly, there were 12 missing values in the raw dataset. After inspecting the pattern of missingness, I imputed the missing values using values from the 2017 ACS 5-year estimates.

Constructing the Research Dataset

As mentioned above, the unit of analysis for the dependent variable is the CoC. However, CoCs typically have irregular geographic boundaries, so CoC-level measures of housing, economic, demographic and safety net characteristics do not exist. As such, I had to rely on county-level measures for my independent variables, which resulted in a mismatch between the units of analysis. To align the units of analysis, I had to construct CoC-level independent variables using statistical adjustments to county-level data. In order to facilitate these conversions, I used a crosswalk created by Byrne (2018) to first map every county to its respective CoC. After being mapped to its relevant CoC, county-level variables were transformed into CoC-level features using statistical adjustments. Byrne et al. (2013) identified three distinct relationship types between counties and CoCs, as outlined below:

1. The boundary for a single CoC and a single county are identical;
2. A single CoC consists of two or more counties; and
3. A single county falls into two or more CoCs.

In the first case, no statistical adjustments were necessary. In the second case, I followed the methodology of Byrne et al. (2013), who “constructed CoC-level [independent] variables from county measures by taking either the sum or a population-weighted average of the county measures from all of the counties within a given CoC" (p. 617). In the third case, I summed the homeless counts for each CoC that fell within a single county to create a series of merged CoCs. As such, the county-level measures did not require any statistical adjustments to transform them into CoC-level features. These mergers ultimately reduced the number of CoCs from 397 to 380 in the final dataset.

VARIABLES

Dependent Variables

Researchers typically look at homelessness rates per 10,000 people (as opposed to counts) in order to account for variation in populations across communities. I looked at two different measures for the homelessness rate, each with a different denominator. These two measures mirror the methodology used by Byrne et al. (2013). The summary statistics for the dependent variables can be found in Table 1 in Appendix A.

1. **Homelessness Rate per 10,000 People in the General Population**: This variable measures the rate of homelessness per 10,000 people in the general population. I calculated this rate by dividing the number of homeless people in each CoC by its total population, and then multiplying the value by 10,000. The values for the denominator were sourced from the ACS 5-year estimates.

2. **Homeless Rate per 10,000 People in Poverty:** This variable measures the rate of homelessness per 10,000 people in the population in poverty, which is defined as

individuals below the 50% poverty level. I calculated this rate by dividing the number of homeless people in each CoC by its respective population of those in poverty, and then multiplying the value by 10,000. The values for the denominator were also sourced from the ACS 5-year estimates.

Independent Variables

I largely mirrored the approach used by Byrne et al. (2013) by grouping the independent variables in terms of housing characteristics, economic conditions, demographic composition and safety net features. I used measures that were widely adopted in prior literature, although this paper adds to the existing body of literature by incorporating the Gini index into the analysis. The summary statistics for the independent variables can be found in Table 1 in Appendix A.

Housing Characteristics

1. **Median Gross Rent**: This variable measures the median monthly housing cost for renters in a given CoC, which can be thought of as a measure of housing affordability. Median gross rent includes the contract rent plus the monthly average cost of fuels and utilities.

2. **Percentage of Renters**: This variable measures the percentage of households that are renters in a given CoC. It is calculated by dividing the number of renter-occupied housing units by the total number of occupied housing units and multiplying it by 100.

3. **Rental Vacancy Rate**: This variable measures the percentage of available rental housing units in a given CoC. It is calculated by dividing the number of unoccupied housing units

for rent by the total stock of rental housing units and multiplying it by 100.

Economic Conditions

4. **Gini Index**: This variable measures income inequality in a CoC. It can take on a value between 0 and 100, with 0 representing perfect equality (i.e., every member receives an equal share of income) and 100 indicating perfect inequality (i.e., only one recipient or group of recipients receives all the income). Although the Census Bureau calculates the Gini index on a scale from 0 to 1, I recoded the variable by multiplying each value by 100 in order to make it easier to interpret.

5. **Unemployment Rate**: This variable measures the percentage of unemployed people within a given CoC. It is calculated by dividing the number of unemployed persons by the civilian labor force 16 years and over and multiplying it by 100. A description of how the Census Bureau defines these populations can be found in Appendix B.

6. **Poverty Rate**: This variable measures the poverty rate for a given CoC, which is calculated by dividing the population of persons living below the 50% poverty level by the total population for whom poverty status can be determined and multiplying it by 100. A description of how the Census Bureau defines these populations can be found in Appendix B. N*ote: this variable is excluded for the model where the homelessness rate is given per 10,000 people in poverty, as per Byrne et al. (2013).*

Demographic Composition

7. **Percentage of Blacks**: This variable measures the percentage of Blacks in the total population for a given CoC. It is calculated by dividing the Black population by the total population and multiplying it by 100.

8. **Percentage of Hispanics**: This variable measures the percentage of Hispanics in the total population for a given CoC. It is calculated by dividing the Hispanic population by the total population and multiplying it by 100.

9. **Percentage of Single Households:** This variables measures the percentage of occupied housing units with a single occupant. It is calculated by dividing the number of occupied housing units with a single occupant by the total number of occupied housing units and multiplying it by 100.

10. **Percentage of Veterans:** This variable measures the proportion of veterans within a CoC. It is calculated by dividing the number of veterans by the civilian population 18 years and over and multiplying it by 100.

11. **Percentage of Baby Boomers:** This variable measures the percentage of Baby Boomers within a given CoC. Baby Boomers are those born between 1946 and 1964, which is the definition used by Byrne et al. (2013) and Hanratty (2017). It is calculated by dividing the population aged 50 to 74 by the total population and multiplying it by 100. The

Census Bureau does not explicitly count the number of Baby Boomers, so I had to rely on a slightly larger age range in order to fully measure this demographic.

Safety Net

12. **Public Assistance Rate:** This variable measures the percentage of occupied housing units receiving public assistance. It is calculated by dividing the number of households with public assistance by the total number of occupied housing units and multiplying it by 100.

13. **Supplemental Security Income (SSI) Rate:** This variable measures the percentage of occupied housing units receiving SSI. It is calculated by dividing the number of households receiving SSI by the total number of occupied housing units and multiplying it by 100.

METHODOLOGY

Analytical Approach

Because CoCs are nested within states, the data is considered to have a hierarchical structure. As such, the ordinary least squares (OLS) regression assumption of independence between observations is violated because of the dependencies that exist within the data. In other words, CoCs within the same state are likely to have similar characteristics relative to CoCs in other states. To correct for this violation, I used a multilevel modeling approach (i.e., mixed-effects models), wherein the CoC is the first level and the state is the second level of analysis. This approach mirrors the methodology used by Byrne et al. (2013). Although I could have also run separate models for each state, I decided against this approach because some states had sample sizes that were far too small, thereby increasing the likelihood of a Type I error.

Ultimately, I ran a series of linear mixed-effects models for three different subgroups. In particular, I stratified the data into urban, suburban and rural CoCs and ran separate models for each. This is analogous to the approach used by Byrne et al. (2013), who ran separate models for metropolitan and non-metropolitan CoCs. HUD classifies each CoC in one of four ways–major city, urban, suburban and rural–but I consolidated the major city and urban levels in order to make each of the three subgroups more similar in size. There were 104 urban CoCs, 165 suburban CoCs and 111 rural CoCs.

An initial inspection of the data (see Figures 1-4 in Appendix A) showed a material difference in median homelessness rates between the three subgroups. In particular, rural CoCs had the lowest median rate of homelessness at 7.92 per 10,000 people in the general population, followed by suburban CoCs (10.70 per 10,000 people in the general population) and urban CoCs (16.29 per 10,000 people in the general population). These trends were similar, yet more pronounced, for the median homelessness rate per 10,000 people in poverty, with rural CoCs (142.32) showing lower rates of homelessness compared to suburban CoCs (203.54) and urban CoCs (265.42).

The distributions of the homelessness rates displayed right skewness, so I applied a natural logarithm transformation to make their distributions more normal. Additionally, although I ran initial models without any transformations to the independent variables, the final models presented in Tables 2 and 3 in Appendix A included natural logarithm transformations for every independent variable. The reason for these transformations was twofold. By using a log-log model, the model estimates were easier to interpret. Moreover, these transformations reduced the presence of outliers in several of the independent variables, which resulted in higher R^2 estimates for the models.

The models were all analyzed in RStudio (Version 1.1.383) using the *lme4* and *lmerTest* packages, which are commonly used for mixed-effects models. Random intercepts (grouped by state) were specified for all of the models. The computation of confidence intervals and p-values for mixed-effects models is cumbersome due to the clustering within the dataset, which complicates the reference distributions and degrees of freedom. Therefore, I used the *broom.mixed* package to pull these metrics. The calculation of R^2 is similarly challenging for mixed-effects models, so I used the conditional and marginal R^2 methods presented by Nakagawa and Shielzeth (2012). These were pulled using the *performance* package.

Hypotheses

Based on findings from prior studies, I hypothesize that the following variables have a positive association with homelessness: unemployment rate; poverty rate; median gross rent; percentage of renters; percentage of Blacks; percentage of Hispanics; percentage of single households; percentage of veterans; and the percentage of Baby Boomers. Meanwhile, I hypothesize that the following variables have a negative association with homelessness: rental vacancy rate; public assistance rate; and SSI rate. My hypotheses would be supported if I were to find a statistically significant coefficient at a 5% significance level with the correct sign. On the other hand, my hypotheses would be invalidated if I were to find either a statistically insignificant ($p \geq 0.05$) estimate or a statistically significant estimate with the opposite sign.

Because this paper adds to the existing literature by adding fixed effects for the Gini index, my hypothesis for this variable warrants additional attention. I hypothesize that a positive relationship exists between the Gini index and the homelessness rate. As inequality between the rich and poor widens, lower- and middle-class households become crowded out of better-quality housing and move into lower-quality housing markets. In turn, demand for lower-quality housing grows and places upward pressure on rents for such housing, which invariably pushes the poorest members of society out of the rental market altogether and generates more homelessness. Moreover, as income increasingly concentrates in the hands of the rich, they tighten their grip on the housing supply, thereby giving them greater pricing power over renters that further displaces the poorest citizens from housing.

My hypothesis would be supported if I were to find a positive, statistically significant coefficient for the Gini variable, which shall serve as my approximation for income inequality. My results would be statistically significant if I were to find a p-value less than 0.05 (i.e., a 5% significant level). If my hypothesis were to be true, it would add additional empirical support to the idea that economic conditions contribute to homelessness. However, I have to exercise caution by acknowledging that these results hold only for the year under study (2018). On the other hand, my hypothesis would be invalidated if I were to find either a negative coefficient for the Gini variable, or a positive but statistically insignificant coefficient for the Gini measure.

RESULTS

Homelessness Rate per 10,000 People in the General Population

There were a number of interesting findings when modeling the homelessness rate per 10,000 people in the general population for urban, rural and suburban CoCs. These results can be found in Table 2 in Appendix A.

For the urban CoCs, none of the economic variables were statistically significant. Among the housing characteristics, the percentage of renters was the only statistically significant variable. The p-value is 0.03, which means that there is only a 3% chance that we would get a t-statistic that large (or more extreme) simply due to chance, assuming the null hypothesis is correct. Hence, I can reject the null hypothesis and conclude that the estimate is statistically significant. According to the model, a 1% increase in the percentage of renters leads to an average 1.5% increase in the homelessness rate, all other variables held constant. Among the demographic characteristics, the only significant variables were the percentage of Blacks and percentage of single households. All else equal, a 1% increase in the percentage of Blacks yields an average predicted 0.23% decrease in the homelessness rate. Meanwhile, a 1% increase in the percentage of single households leads to a 1.46% average increase in the homelessness rate, all else held constant. Lastly, among the safety net measures, the coefficient for the percentage of households receiving SSI was statistically significant at a 5% significance level. The coefficient had a positive relationship with the homelessness rate–a 1% increase leads to a predicted 0.55% increase in the homelessness rate, all other variables held constant. Overall, the model explains about 59% (conditional R^2) of the variation in the homelessness rate for urban CoCs.

Next, I shall look at the suburban CoCs model. Similar to the urban CoCs, none of the economic characteristics were statistically significant at a 5% level for this model. Among the housing characteristics, the median gross rent was the only significant variable, with a 1% increase in the median gross rent leading to an average 1% increase in the homelessness rate, all other variables held constant. Meanwhile, many of the demographic characteristics–namely the percentages of Blacks, Hispanics and veterans–had a statistically significant relationship with homelessness rates. However, all of the effect sizes were quite small. In particular, a 1% increase in the percentage of Blacks leads to an average 0.26% decline in the homelessness rate, all else equal. Additionally, a 1% increase in the percentage of Hispanics yields an average 0.17% increase in the homelessness rate, all other variables held constant. Lastly, a 1% increase in the percentage of veterans leads to an average 0.59% increase in the homelessness rate, all else equal. Among the safety net features, the coefficient for the percentage of households receiving SSI was statistically significant at a 5% level. A 1% increase in the percentage of households receiving SSI leads to a predicted 0.79% increase in the homelessness rate, all other variables held constant. Overall, the model explains about 56% of the variation in the homelessness rate for suburban CoCs.

Lastly, for the rural CoCs model, the economic characteristics were all statistically insignificant at a 5% level. However, among housing characteristics, both the median gross rent and percentage of renters variables were statistically significant. All else equal, a 1% increase in the median gross rent coincides with an average predicted 1.73% increase in the homelessness rate. Meanwhile, a 1% increase in the percentage of renter households yields an expected 1.9% average increase in the homelessness rate, all other variables held constant. Looking at the demographic characteristics, the percentage of Blacks and percentage of Baby Boomers were the only statistically significant variables. Similar to the other models, the coefficient for the percentage of Blacks was negative and quite small, with a 1% increase leading to a predicted 0.25% decline in the homelessness rate on average, all else held constant. However, a 1% increase in the percentage of Baby Boomers leads to a robust 3.25% average increase in the homelessness rate, all else equal. None of the safety net features were significant for this model. Ultimately, the model accounts for about 72% of the variation in the homelessness rate for rural CoCs, the largest of the three models.

Homelessness Rate per 10,000 People in Poverty

There were also a number of interesting findings when modeling the homelessness rate per 10,000 people in poverty for urban, rural and suburban settings. These results can be found in Table 3 in Appendix A.

For the urban CoCs, there were no statistically significant economic predictors of homelessness. Among the housing characteristics, the median gross rent was the only statistically significant indicator. In particular, a 1% increase in the median gross rent yields a 1.9% average increase in the homelessness rate, all else equal. The percentage of single households was the only significant variable out of all of the demographic characteristics, with a 1% increase in this variable leading to an average 1.6% upturn in the homelessness rate, all else equal. Among the safety net features, the coefficient for the percentage of households receiving SSI was statistically significant at a 5% level. A 1% increase in the percentage of households receiving SSI leads to a predicted 0.61% average increase in the homelessness rate, all other variables held constant. The model explained about 67% of the variation in the homelessness rate for urban CoCs.

For the suburban CoCs, none of the economic characteristics were statistically significant. The only statistically significant housing characteristic was the median gross rent, whereby a 1% increase in this variable leads to an average 1.53% increase in the homelessness rate, all other variables held constant. Among the demographic features, the percentage of Blacks and percentage of veterans were both statistically significant, although the effect sizes were rather small. In particular, a 1% increase in the percentage of Blacks leads to a predicted 0.26% decline in the homelessness rate on average, all else held constant. However, a 1% increase in the percentage of veterans leads to a predicted 0.4% average increase in the homelessness rate, all else equal. Among the safety net features, the coefficient for the percentage of households receiving SSI was statistically significant at a 5% level. A 1% increase in the percentage of households receiving SSI leads to a predicted 0.61% increase in the homelessness rate on average, all other variables held constant. Overall, the model explained about 50% of the variation in the homelessness rate for suburban CoCs.

For the rural CoCs, none of the economic characteristics were statistically significant. Among the housing variables, the median gross rent and the percentage of renter households were the only significant variables. All else equal, a 1% increase in the median gross rent leads to a 2.07% average upturn in the homelessness rate. Meanwhile, a 1% increase in the percentage of renters leads to a predicted 1.57% average increase in the homelessness rate, all other variables held constant. Looking at demographic features, only the percentage of Blacks and percentage of Baby Boomers were statistically significant. Again, the association between the Black population and the homelessness rate was rather small and negative, with a coefficient of -0.26. On the other hand, a 1% increase in the percentage of Baby Boomers yields an average 3.7% increase in the homelessness rate, all else equal. No safety net variables were statistically significant. The model explained about 72% of the variation in the homelessness rate for rural CoCs.

DISCUSSSION

The major purpose of this study was to verify if the structural determinants of homelessness (housing characteristics, economic conditions, demographic composition and safety net features) still hold for more recent data–in this case, the year 2018. Additionally, the study sought to understand how these effects vary across urban, suburban and rural areas.

This study supports the notion that housing dynamics have the strongest association with homelessness, as shown in earlier studies by Byrne et al. (2013) and Lee et al. (2010). In particular, the median gross rent had a positive and statistically significant relationship with the homelessness rate in five of the six models that were ran (the exception being the urban CoCs model where the homelessness rate was given per 10,000 people in the general population, which was positive but not statistically significant), suggesting that housing affordability has a significant association with homelessness. These results are in line with many previous researchers' findings, many of whom have found a positive and statistically significant relationship between increased rent and homelessness (Bohanon, 1991; Byrne et al., 2013; Early & Olsen, 2002; Honig & Filer, 1993; Lee & Farrell, 2003; Quigley, 1990; Quigley & Raphael, 2002; Quigley et al., 2001; Troutman et al., 1999). Nevertheless, there were significant differences in effect sizes across urban, rural and suburban areas, which underscores the point that the homelessness problem is not a one-size-fits-all issue. In particular, although the effect sizes were moderately strong for both urban and suburban CoCs, the estimated coefficient for median gross rent was largest for the rural CoCs models. These findings provide evidence that housing affordability may be a bigger issue in rural areas than previously thought, which may warrant additional safety nets for those living in rural areas. Meanwhile, this study finds that a greater incidence of renters in rural CoCs is positively associated with the rate of homelessness among both the general population and those in poverty. As such, greater protections for renters in rural areas may be of interest to policymakers in order to drive down homelessness in such areas.

Interestingly, none of the economic characteristics were found to be statistically significant for all of the models. These findings generally align with prior research. For example, Elliott and Krivo (1991) found no significant relationship between homelessness and economic factors like poverty rates or unemployment rates. Moreover, Byrne et al. (2013) found no significant association between poverty rates and homelessness rates; moreover, only non-metro CoCs had a significant association between unemployment rates and homelessness rates in the general population. As mentioned earlier, this study adds to the existing literature by incorporating income inequality as a potential economic predictor of homelessness. Although the majority of the models showed a positive association between the Gini index and homelessness rates, the lack of statistical significance for all of the models raises a number of issues with regard to its usefulness as a predictor of homelessness. Perhaps most simply, it may imply that income inequality is simply not a useful predictor of homelessness. However, I believe a more probable explanation is that there may be measurement issues that are obfuscating the analysis. There are many different ways to operationalize income inequality aside from the Gini index, so these other measurements may be useful for future studies. Moreover, it may be worthwhile to see how income inequality interacts with the other structural determinants of homelessness, particularly housing characteristics, in order to get a more nuanced picture of their effect sizes for future studies.

Each of the demographic features were statistically significant in at least one of the models. The percentage of Blacks was consistently significant–five of the six models had a statistically significant estimate, with the exception being the urban CoCs model where the homelessness rate was given per 10,000 people in poverty. Interestingly, the models showed a negative association between the percentage of Blacks and the homelessness rates, which went against my prior. Even so, the 95% confidence interval for this variable is quite narrow for all of the models, which suggests its effect is quite small. In the past, researchers like Elliott and Krivo

(1991) found positive associations between the percentage of Blacks and homelessness rates, although more recent research like Byrne et al. (2013) has shown a negative association. For the other demographic features, their significance largely depended on the geographic region. In particular, the percentage of single households had a positive and statistically significant relationship with homelessness rates (in both the general population and those in poverty) only in urban CoCs. Meanwhile, the percentage of veterans had a positive and statistically significant relationship with homelessness rates (in both the general population and those in poverty) only in suburban CoCs. Lastly, the percentage of Baby Boomers had a positive and statistically significant relationship with homelessness rates (in both the general population and those in poverty) only in rural CoCs–an interesting discovery insofar as it contradicts the findings of Byrne et al. (2013), who found that the prevalence of Baby Boomers was associated with homelessness rates only in metro areas. Ultimately, these findings suggest that interventions based on demographic composition ought to be targeted by region.

Among the safety net features, only the percentage of households receiving SSI was statistically significant. Interestingly, this variable was only significant in urban and suburban settings, and all of the estimated coefficients were positive, which went against my hypothesis. These findings are somewhat consistent with Byrne et al. (2013), who also found a positive (but ultimately insignificant) relationship between the percentage of households receiving SSI and the homelessness rate in metro CoCs. Future research may benefit from incorporating additional measures for safety nets in order to ensure these results are not simply due to bias.

Although there were a number of interesting findings in this study, it suffers from a few limitations. As mentioned earlier, the HUD data does not fully capture the homeless population, especially those experiencing temporary or transitional homelessness. Therefore, these unobserved observations could be biasing the results. Until HUD comes up with new mechanisms for counting these types of homeless people, these issues will continue to be a significant challenge for researchers.

Additionally, because this paper uses the CoC as the unit of observation, there are limitations with regard to the independent variables I could use, which may similarly bias my results. For example, prior research has shown that drug and alcohol abuse disproportionately occurs among homeless persons, which often co-occurs with mental health issues (Burt et al., 2001). Additionally, other researchers have found an association between homelessness and a lack of community mental health facilities, which they measured by looking at both per capita expenditures on beds for the mentally ill and total state health agency expenditures on all mental health services. Unfortunately, these variables are typically only available at a state level. Lastly, this study only looks at 2018 data, so future research may benefit from pooling multiple years of data to see if these findings are consistent over time.

REFERENCES

Appelbaum, R., Dolny, M., Dreier, P., & Gilderbloom, J. (1991). Scapegoating rent control: Masking the causes of homelessness. *Journal of the American Planning Association*, *57*, 153–164.

Bohanon, C. (1991). The economic correlates of homelessness in sixty cities. *Social Science Quarterly*, *72*, 817–825.

Burt, M. R. (1993). *Over the edge: The growth of homelessness in the 1980s*. New York: Russell Sage Foundation.

Burt, M. R., Aron, L. Y., Lee, E., & Valente, J. (2001). *Helping America's Homeless: Emergency Shelter or Affordable Housing?* Washington, D.C.: Urban Institute Press.

Byrne, T. (2018). HUD-CoC-Geography-Crosswalk. Retrieved from https://github.com/tomhbyrne/HUD-CoC-Geography-Crosswalk

Byrne, T., Munley, E. A., Fargo, J. D., Montgomery, A. E., & Culhane, D. P. (2013). New Perspectives on Community-Level Determinants of Homelessness. *Journal of Urban Affairs*, *35*(5), 607–625.

Early, D. W., & Olsen, E. O. (2002). Subsidized housing, emergency shelters, and homelessness: An empirical investigation using data from the 1990 Census. *Advances in Economic Analysis and Policy*, *2*(1), 2, 2–34.

Elliott, M., & Krivo, L. (1991). Structural Determinants of Homelessness in the United States. Social Problems, 38(1), 113-131.

Hanratty, M. (2017). Do Local Economic Conditions Affect Homelessness? Impact of Area Housing Market Factors, Unemployment, and Poverty on Community Homeless Rates. Housing Policy Debate, 27(4), 640–655.

Hawkins, R. L., & Abrams, C. (2007). Disappearing acts: The social networks of formerly homeless individuals with co-occurring disorders. *Social Science and Medicine, 65*, 2031–2042.

Honig, M., & Filer, R. K. (1993). Causes of intercity variation in homelessness. *The American Economic Review*, *83*, 248–255.

Hwang, S. W. (2001). Homelessness and health. *Canadian Medical Association Journal, 164*, 229–233.

Kawachi, I., & Kennedy, B. P. (1997). The relationship of income inequality to mortality: Does the choice of indicator matter? *Social Science & Medicine*, *45*(7), 1121–1127.

Lee, B. A., & Farrell, C. R. (2003). Buddy, can you spare a dime? *Urban Affairs Review, 38*, 299–324.

Lee, B. A., & Greif, M. J. (2008). Homelessness and hunger. *Journal of Health and Social Behavior, 49*(1), 3–19.

Lee, B., Tyler, K., & Wright, J. (2010). The New Homelessness Revisited. Annual Review of Sociology, 36, 501-521.

Nakagawa, S., & Schielzeth, H. (2012). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142.

Quigley, J. M. (1990). Does rent control cause homelessness? Taking the claim seriously. *Journal of Policy Analysis and Management*, *9*(1), 89–93.

Quigley, J. M., & Raphael, S. (2002). The economics of homelessness: The evidence from North America. *European Journal of Housing Policy*, *1*, 323–336.

Quigley, J. M., Raphael, S., & Smolensky, E. (2001). Homeless in America, homeless in California. *Review of Economics and Statistics*, *83*(1), 37–51.

Raphael, S. (2010). Housing market regulation and homelessness. In I. G. Ellen, & B. O'Flaherty (Eds.), *How to house the homeless* (pp. 110–140). New York: Russell Sage Foundation.

Rossi, P. H. (1989). *Down and Out in America: The Origins of Homelessness*. Chicago: Univ. of Chicago Pr.

Troutman, W. H., Jackson, J. D., & Ekelund, R. B. (1999). Public policy, perverse incentives, and the homeless problem. *Public Choice*, *98*, 195–212.

U.S. Census Bureau. (2016). Glossary. Retrieved from https://www.census.gov/topics/income-poverty/poverty/about/glossary.html

U.S. Department of Housing and Urban Development (HUD). (2013). *Expanding Opportunities to House Individuals and Families Experiencing Homelessness through the Public Housing (PH) and Housing Choice Voucher (HCV)*. Retrieved from https://www.hud.gov/sites/documents/PIH2013-15HOMELESSQAS.PDF

U.S. Department of Housing and Urban Development (HUD). (2018). *The 2018 Annual Homeless Assessment Report (AHAR) to Congress*. Retrieved from https://www.huduser.gov/portal/sites/default/files/pdf/2018-AHAR-Part-2.pdf

Wolitski, R. J., Kidder, D. P., & Fenton, K. A. (2007). HIV, homelessness, and public health: Critical issues and a call for increased action. *AIDS and Behavior, 11*, 167–171.

Zerger, S. (2002). *A preliminary review of literature: Chronic medical illness and homelessness*. Nashville, TN: National Health Care for the Homeless Council.

APPENDIX A
FIGURES & TABLES

Figure 1: Distribution of Homelessness Rate
(per 10,000 People in the General Population)

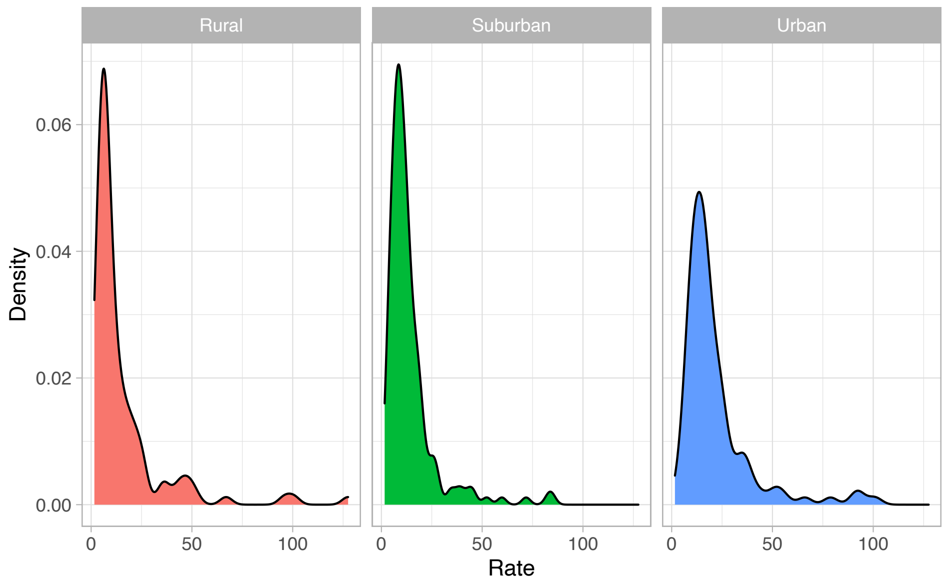

Figure 2: Distribution of Homelessness Rate
(per 10,000 People in Poverty)

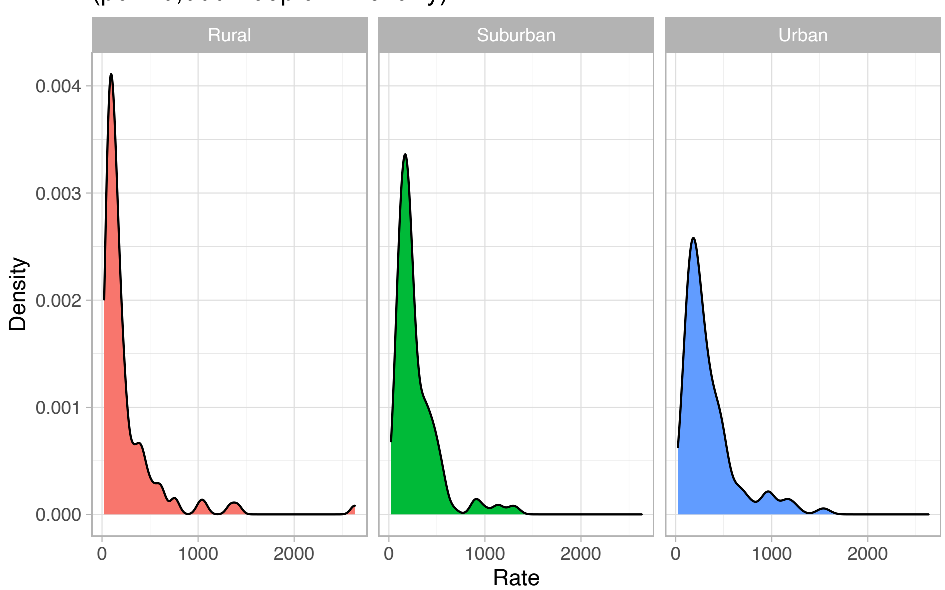

Figure 3: Distribution of Homelessness Rate
(per 10,000 People in the General Population)

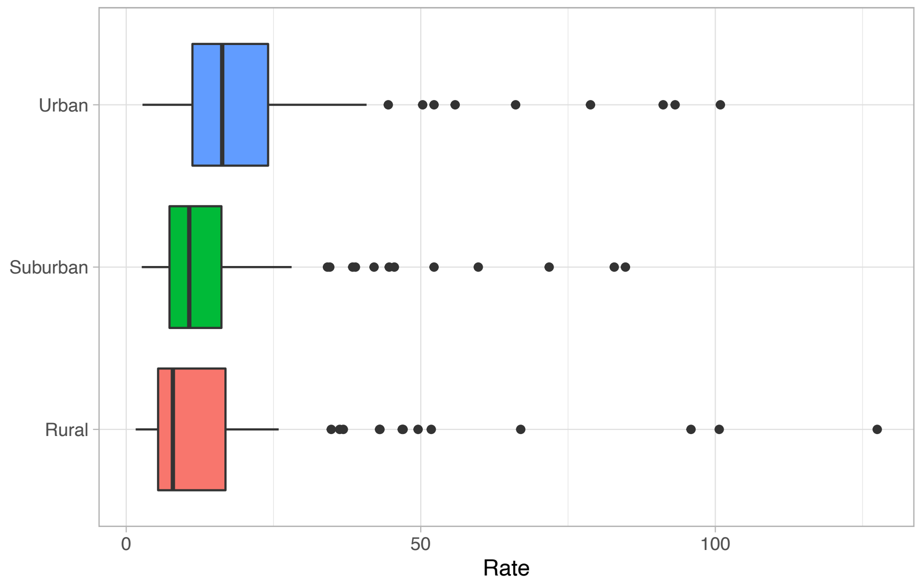

Figure 4: Distribution of Homelessness Rate
(per 10,000 People in Poverty)

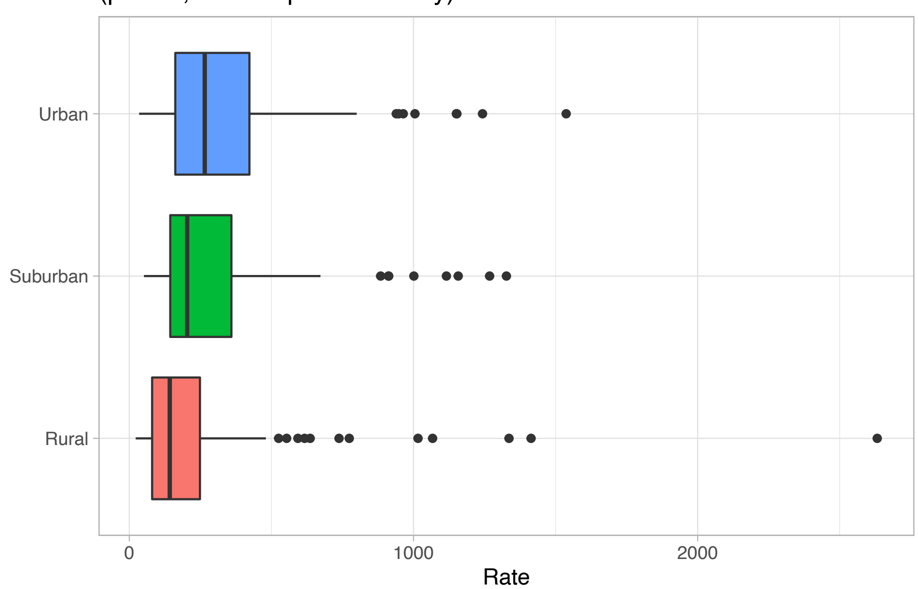

Table 1: Descriptive Statistics

	Urban CoCs (N=104)					Suburban CoCs (N=165)					Rural CoCs (N=111)				
Variable	Mean	Med	SD	Min	Max	Mean	Med	SD	Min	Max	Mean	Med	SD	Min	Max
HOMELESSNESS RATE															
Per 10,000 in General Population	21.75	16.29	17.91	2.76	100.86	14.08	10.7	12.99	2.64	84.74	15.46	7.92	19.84	1.61	127.49
Per 10,000 in Poverty	341.98	265.42	277.86	34.49	1537.31	279.75	203.54	225.71	51.5	1327	248.57	142.32	338.34	22.44	2632
ECONOMIC															
Gini Index	47.22	46.82	2.72	41.72	54.32	45.36	45.42	3.08	38.13	55.29	44.97	44.7	2.05	40.36	51.04
Unemployment Rate	6.07	5.85	1.72	2.94	10.32	5.83	5.58	1.62	3.11	15.25	5.88	5.81	1.53	2.83	9.78
Poverty Rate	7.14	6.73	2.35	3.48	16.89	5.35	5.28	1.8	1.78	11.91	6.82	6.7	1.77	3.54	12.55
HOUSING															
Median Gross Rent	1052.45	961.33	292.24	670	2126	1125.36	1090.66	284.63	682	2158	820.19	775.86	165.7	612.19	1616
Renters	41.93	40.48	7.5	30.21	67.35	32.4	32.24	7.73	14.78	68.53	30.7	30.42	4.86	19.12	47.41
Rental Vacancy Rate	6.22	6	2.75	2.12	21.89	5.78	5.37	2.26	1.96	15.92	7.52	6.95	3.76	2.08	22.25
DEMOGRAPHIC															
Blacks	17.72	12.26	14.88	1.06	62.46	11.79	8.84	10.53	0.59	62.95	6.99	3.05	8.69	0.43	38.69
Hispanics	17.07	10.55	15.95	2.22	82.4	13.32	9.31	12.14	1	83.8	8.85	5.01	9.87	1.05	58.43
Single Households	30.06	30.17	4.91	18.15	44.54	26.81	27.36	4.32	12.78	37.78	28.61	28.69	2.53	20.61	38.41
Veterans	7.73	7.02	3.17	2.32	19.83	7.86	7.56	2.73	2.29	19.81	9.1	8.92	1.41	4.72	13.56
Baby Boomers	26.38	26.4	2.48	19.31	33.01	29.81	29.85	3.7	16.15	44.2	31.13	30.98	2.83	24.62	40.13
SAFETY NET															
Public Assistance	2.61	2.32	1.4	0.93	8.7	2.41	2.16	1.11	0.69	7.09	2.63	2.51	0.88	1.13	6.54
Supplemental Security Income	5.38	5.11	1.71	1.53	10.49	4.91	4.65	1.63	1.62	11.78	5.99	5.83	1.57	2.99	10.76

Table 2: Summary of Mixed-Effects Model for Variables Predicting Homelessness per 10,000 People in General Population

	Urban CoCs (N=104)			Suburban CoCs (N=165)			Rural CoCs (N=111)		
Variable	B	95% CI	p	B	95% CI	p	B	95% CI	p
Intercept	-16.02	(-28.3, -3.74)	0.01	-16.22	(-22.97, -9.47)	0	-29.85	(-42.04, -17.66)	0
ECONOMIC									
Gini Index	0.94	(-2.14, 4.02)	0.55	1.19	(-0.84, 3.22)	0.25	0.4	(-3.25, 4.05)	0.83
Unemployment Rate	0.08	(-0.63, 0.79)	0.82	-0.24	(-0.79, 0.31)	0.39	0.35	(-0.39, 1.09)	0.35
Poverty Rate	-0.63	(-1.34, 0.08)	0.08	0.13	(-0.33, 0.59)	0.57	0.05	(-0.8, 0.9)	0.91
HOUSING									
Median Gross Rent	0.61	(-0.21, 1.44)	0.15	1	(0.35, 1.64)	0	1.73	(0.6, 2.87)	0
Renters	1.5	(0.18, 2.81)	0.03	0.43	(-0.27, 1.13)	0.23	1.9	(0.49, 3.31)	0.01
Rental Vacancy Rate	0.05	(-0.31, 0.4)	0.8	0.04	(-0.27, 0.35)	0.8	0.01	(-0.26, 0.27)	0.96
DEMOGRAPHIC									
Blacks	-0.23	(-0.39, -0.07)	0.01	-0.26	(-0.37, -0.14)	0	-0.25	(-0.38, -0.13)	0
Hispanics	-0.02	(-0.21, 0.16)	0.81	0.17	(0.02, 0.33)	0.03	0.12	(-0.04, 0.28)	0.16
Single Households	1.46	(0.44, 2.48)	0.01	0.49	(-0.55, 1.52)	0.36	0.41	(-1.26, 2.09)	0.63
Veterans	-0.01	(-0.43, 0.41)	0.98	0.59	(0.23, 0.96)	0	-0.54	(-1.31, 0.22)	0.17
Baby Boomers	0.31	(-1.24, 1.86)	0.69	0.5	(-0.74, 1.74)	0.43	3.25	(1.21, 5.29)	0
SAFETY NET									
Public Assistance	0.24	(-0.04, 0.52)	0.1	0.26	(-0.04, 0.56)	0.09	-0.03	(-0.42, 0.37)	0.9
Supplemental Security Income	0.55	(0.05, 1.05)	0.04	0.79	(0.28, 1.3)	0	0.35	(-0.43, 1.14)	0.38
SD of Random Intercepts	0.19			0.13			0.24		
Conditional R^2	0.59			0.56			0.72		
Marginal R^2	0.51			0.52			0.65		

Table 3: Summary of Mixed-Effects Model for Variables Predicting Homelessness per 10,000 People in Poverty

	Urban CoCs (N=104)			Suburban CoCs (N=165)			Rural CoCs (N=111)		
Variable	B	95% CI	p	B	95% CI	p	B	95% CI	p
Intercept	-10.65	(-24.19, 2.89)	0.13	-11.83	(-18.7, -4.95)	0	-20.17	(-31.72, -8.62)	0
ECONOMIC									
Gini Index	-1.8	(-5, 1.4)	0.27	0.15	(-1.88, 2.19)	0.88	-1.61	(-4.86, 1.64)	0.34
Unemployment Rate	-0.6	(-1.31, 0.11)	0.1	-0.47	(-1.02, 0.08)	0.1	-0.1	(-0.73, 0.53)	0.75
HOUSING									
Median Gross Rent	1.9	(1.21, 2.58)	0	1.53	(0.92, 2.14)	0	2.07	(0.95, 3.19)	0
Renters	0.11	(-1.21, 1.42)	0.87	0.12	(-0.58, 0.83)	0.73	1.57	(0.16, 2.97)	0.03
Rental Vacancy Rate	-0.04	(-0.42, 0.34)	0.84	0.09	(-0.23, 0.41)	0.58	0.05	(-0.21, 0.31)	0.7
DEMOGRAPHIC									
Blacks	-0.16	(-0.33, 0.01)	0.07	-0.26	(-0.38, -0.15)	0	-0.26	(-0.39, -0.14)	0
Hispanics	0.11	(-0.08, 0.31)	0.25	0.12	(-0.04, 0.28)	0.14	0.12	(-0.05, 0.28)	0.17
Single Households	1.6	(0.47, 2.74)	0.01	0.32	(-0.76, 1.4)	0.56	0.06	(-1.61, 1.72)	0.95
Veterans	0	(-0.46, 0.46)	0.99	0.4	(0.05, 0.75)	0.03	-0.53	(-1.31, 0.25)	0.19
Baby Boomers	1.31	(-0.34, 2.95)	0.12	1	(-0.25, 2.26)	0.12	3.7	(1.68, 5.72)	0
SAFETY NET									
Public Assistance	0.19	(-0.12, 0.5)	0.24	0.3	(0, 0.61)	0.05	-0.06	(-0.46, 0.34)	0.77
Supplemental Security Income	0.61	(0.05, 1.16)	0.03	0.61	(0.1, 1.12)	0.02	0.38	(-0.41, 1.18)	0.35
SD of Random Intercepts	0.26			0.1			0.24		
Conditional R^2	0.67			0.5			0.72		
Marginal R^2	0.54			0.47			0.65		

APPENDIX B
VARIABLE DEFINITIONS & CODES

I. Definition of Homelessness

HUD classifies homelessness into four categories:

“(1) Individuals and families who lack a fixed, regular, and adequate nighttime residence and includes a subset for an individual who is exiting an institution where he or she resided for 90 days or less and who resided in an emergency shelter or a place not meant for human habitation immediately before entering that institution; (2) Individuals and families who will imminently lose their primary nighttime residence; (3) Unaccompanied youth and families with children and youth who are defined as homeless under other federal statutes who do not otherwise qualify as homeless under this definition; or (4) Individuals and families who are fleeing, or are attempting to flee, domestic violence, dating violence, sexual assault, stalking, or other dangerous or life-threatening conditions that relate to violence against the individual or a family member” (U.S. Department of Housing and Urban Development, 2013, p. 1).

II. Definition of Poverty

The U.S. Census Bureau describes the population in poverty as follows:

“For some persons, such as unrelated individuals under age 15, poverty status is not defined. Since Census Bureau surveys typically ask income questions to persons age 15 or older, if a child under age 15 is not related by birth, marriage, or adoption to a reference person within the household, we do not know the child's income and therefore cannot determine his or her poverty status. For the decennial censuses and the American Community Survey, poverty status is also undefined for people living in college dormitories and in institutional group quarters. People whose poverty status is undefined are excluded from Census Bureau poverty tabulations. Thus, the total population in poverty tables--the poverty universe--is slightly smaller than the overall population” (U.S. Census Bureau, 2016).

III. Definition of Unemployment

The U.S. Census Bureau describes how it determines unemployment as follows:

“All civilians 16 years old and over are classified as unemployed if they (1) were neither "at work" nor "with a job but not at work" during the reference week, and (2) were actively looking for work during the last 4 weeks, and (3) were available to accept a job. Also included as unemployed are civilians who did not work at all during the reference week, were waiting to be called back to a job from which they had been laid off, and were available for work except for temporary illness” (U.S. Census Bureau, 2016).

IV. Codes

The following codes were used to pull data from the U.S. Census Bureau API:

Populations/Demographics
B01003_001 = Total Population
B02001_003 = Black Population
B03001_003 = Hispanic Population
S0101_C01_012 = Population 50 to 54
S0101_C01_013 = Population 55 to 59
S0101_C01_014 = Population 60 to 64
S0101_C01_015 = Population 65 to 69
S0101_C01_016 = Population 70 to 74
S2101_C01_001 = Civilian population 18 years and over
S2101_C03_001 = Veteran Population

Households
B25003_001 = Total Occupied Housing Units
B25003_003 = Total Renter Occupied Housing Units
B25004_002 = Vacancy Status (Total, for rent)
B25004_003 = Vacancy Status (Total, rented, not occupied)
S2501_C01_002 = Single Households
B19057_002 = Households with Public Assistance
B19056_002 = Households with Supplemental Security Income (SSI)

Economic
B19083_001 = Gini Index
B25064_001 = Median Gross Rent
B23025_003 = Civilian Labor Force (16+)
B23025_005 = Unemployed
S1701_C01_001 = Population for whom poverty status is determined
S1701_C01_038 = Population below 50 percent of poverty level