Journal Pre-proof

Explaining the spatial variation in American life expectancy

Elizabeth A. Dobis, Heather M. Stephens, Mark Skidmore, Stephan J. Goetz

PII: S0277-9536(19)30754-3

DOI: https://doi.org/10.1016/j.socscimed.2019.112759

Reference: SSM 112759

To appear in: *Social Science & Medicine*

Received Date: 6 August 2019

Revised Date: 30 November 2019

Accepted Date: 19 December 2019

EXPLAINING THE SPATIAL VARIATION IN AMERICAN LIFE EXPECTANCY

Elizabeth A. Dobis
Northeast Regional Center for Regional Development
The Pennsylvania State University
University Park, PA, United States
edobis@psu.edu

Heather M. Stephens
Resource Economics and Management
West Virginia University
Morgantown, WV, United States
heather.stephens@mail.wvu.edu

Mark Skidmore
Department of Agriculture, Food, and Resource Economics
North Central Regional Center for Rural Development
Michigan State University
East Lansing, MI, United States
mskidmor@msu.edu

Stephan J. Goetz
Department of Agricultural Economics, Sociology, and Education
Northeast Regional Center for Regional Development
The Pennsylvania State University
University Park, PA, United States
sgoetz@psu.edu

November 29, 2019

Declarations of Interest: none

EXPLAINING THE SPATIAL VARIATION IN AMERICAN LIFE EXPECTANCY

November 29, 2019

Highlights:

- There is statistically significant spatial clustering in county life expectancy.
- Community characteristics are associated with spatial variation in life expectancy.
- Community factors may influence life expectancy in surrounding locations.
- Individual characteristics continue to be strongly correlated with life expectancy.

Abstract:
Since 1980, average life expectancy in the United States has increased by roughly five years; however, in recent years it has been declining. At the same time, spatial variation in life expectancy has been growing. To explore reasons for this trend, some researchers have focused on morbidity factors, while others have focused on how mortality trends differ by personal characteristics. However, the effect community characteristics may play in expanding the spatial heterogeneity has not yet been fully explored. Using a spatial Durbin error model, we explore how community and demographic factors influence county-level life expectancy in 2014, controlling for life expectancy in 1980 and migration over time, and analyzing men and women separately. We find that community characteristics are important in determining life expectancy and that there may be a role for policy makers in addressing factors that are associated with lower life expectancy in some regions.

Keywords: United States; life expectancy; community; spatial variation; spatial analysis

JEL Codes: I12, R12

1. Introduction

In 2016, for the first time since 1962-63, average American life expectancy declined for two consecutive years (*The Economist*, Jan 6, 2018). Although American life expectancy is expected to increase by 2040, if current trends continue, the United States (US) is expected to fall 21 spots in world life expectancy rankings from 43rd to 64th (IHME, 2018) resulting in a life expectancy lower than China's. Kontis et al. (2017) project that the US will fall farther behind in the future due to rising negative health outcomes and lack of universal health insurance. However, focusing on the national average masks the dramatic increase in the variation of life expectancies across the US. From 2010 to 2015, census tract-level average life expectancy ranged from 56 to 97 years; the lower number is on par with Somalia and the higher number exceeds the average for Japan, which has the world's highest overall life expectancy, by 13 years (Arias et al., 2018; *The Economist*, Sep 25, 2018).

Female life expectancy has been higher than that of males since the 19th century (Goldin and Llera-Muney, 2018). However, consistent with the theoretical model in Leung et al. (2004), as more women have entered the workforce, the gap between female and male life expectancy has shrunk. From 1980 to 2014, overall female life expectancy at birth in the US increased from 77.5 to 81.5 years, while overall male life expectancy increased from 70.0 to 76.7 years (IHME, 2016). Concurrently, the variation in county-level life expectancy rose for both sexes, with a 0.57-year increase in the standard deviation for women and a 0.62-year increase for men. The Pearson correlation between life expectancy at birth in 1980 and 2014 is 0.708 for women and 0.815 for men, indicating that approximately 20 to 30 percent of the change in life expectancy cannot be explained by previous life expectancy or that a fundamental change has occurred.

The relative decline in overall American life expectancy and the attendant rise in its variation have spawned numerous recent studies (e.g., Case and Deaton, 2015; Chetty et al., 2016; Currie and Schwandt, 2016; Dwyer-Lindgren et al., 2017). These studies suggest several possible factors may be responsible, including rising obesity rates, opioids abuse, and lack of access to health care, and that these factors may vary by personal characteristics, such as age or race.

To explore this decline, we conduct analysis of the variation in life expectancy using county-level data. We also extend the previous literature in several important ways. First, building on Arora et al. (2016), Chetty et al. (2016), Dwyer-Lindgren et al. (2017), and Frederick et al. (2019), who have examined the spatial variation in county-level American life expectancy, in addition to a full suite of demographic and health characteristics that are associated with life expectancy, we include community-level factors that also may be important to life expectancy in our analysis. Second, we use both descriptive and statistical spatial analysis to illustrate spatial clustering of life expectancy and spatial spillovers from community variables. Finally, we use spatial regression analysis to examine the impact of the community factors on life expectancy at birth, controlling for factors related to personal characteristics and health behaviors and accounting for space. We consider life expectancies for women and men separately. Our approach makes it easier to identify the mechanisms that may have led both to decreases in life expectancies (compared to the rest of the world) and increases in variation across space since 1980, which is critical for making forward-looking policy recommendations.

Consistent with previous work, we find that individual factors are important. Places with more smokers, where residents are less physically active, and with higher rates of obesity are associated with lower life expectancy (for men and women) since 1980, while places with rising incomes and a greater proportion of immigrants have increased both sexes' life expectancies. Recent trends of declining labor force participation are associated with increases in both male and female life expectancies at the county level.

Central to our analysis, we find that community factors also are associated with variation in life expectancy across space. After controlling for lack of physical activity and obesity, increased access to fast food restaurants has led to lower life expectancy, and there are spillovers from nearby areas (for men, especially in rural areas). Perhaps unsurprisingly, places that are growing, have access to better health care, and have more social capital have higher life expectancies. But this is not necessarily an urban benefit, as higher population densities (locally) are associated with lower life expectancy. Additionally, the industry composition of places can affect life expectancy. For example, some rural areas have lower life expectancies if there is a larger share of jobs in mining and oil and gas (both in that county and in surrounding counties).

2. Literature

The decline in overall life expectancy coupled with greater regional variation have led to several recent papers by researchers in economics and the health sciences. These studies either consider the changes in life expectancy, changes in mortality rates (which affect life expectancy), or both.

Examining the variation in mortality rates across space using classifications of regions by how rural or urban they are, James and Cossman (2017) uncover a distinct rural disadvantage. Building on this, James et al. (2018) compare counties and find that the rural South has disproportionately higher mortality rates and this disparity is growing compared to urban areas and even rural areas in the Midwest. Further disaggregating mortality rates based on race and space, James et al. (2019) provide descriptive evidence of clustering of counties with persistently high white mortality rates but no clustering of places with persistently high black mortality rates. The places with persistently high white mortality rates also have distinctly different characteristics: they are generally rural, in the South, and have lower incomes and education levels. Singh et al. (2017) find additional evidence of regional variation in a widening gap between Appalachian infant mortality rates and life expectancy rates and those in the rest of the nation by comparing aggregated data from the federally-designated Appalachian Regional Commission region to the rest of the U.S. They also show that the link between poverty and life expectancy is stronger in Appalachia than elsewhere.

Further examining the reasons for spatial variation in mortality, Cossman et al. (2017) find that expanded health care access at the county level is associated with a decline in black mortality but a rise in mortality rates among white residents. Economic factors such as higher unemployment rates and job loss also appear to be associated with increases in mortality rates (Bloemen et al., 2018; Roelfs et al., 2011; Singh et al., 2013; Strumpf et al., 2017). Using individual-level data, Brodish and Hakes (2016) show that mortality rates fall with household income and that reducing income inequality can have positive effects. Montez et al. (2016) explore the variation in women's mortality between US states, finding that personal characteristics only explain 30 percent of the variation; the rest is due to states' characteristics, including social cohesion and economic conditions. Examining the variation at the county level, Currie and Schwandt (2016) conclude that increasing variance in mortality rates across counties is related to how poverty, income and education affect subgroups of the population differently.

There is also tremendous variation in cause-specific mortality across space. Using a Bayesian model to estimate age- and gender-specific county-level mortality rates, both overall and for specific causes, Dwyer-Lindgren et al. (2016) illustrate clusters of cause-specific mortality in different parts of the US. For example, there are clusters of high respiratory disease mortality in West Virginia and Kentucky, as well as eastern Colorado, while high self-harm and interpersonal violence mortality are clustered in the Southwest and Alaska.

These differences in cause-specific mortality have been explored to explain the recent decline in US life expectancy. Acciai and Firebaugh (2017) examine national data on men and women separately and find that for men changes in life expectancy are linked to an increased risk of accidental death or suicide, while for women they are due to old-age causes (e.g., heart disease) occurring earlier. This is consistent with the results from Case and Deaton (2015; 2017), who use national data on mortality rates by age groups to highlight the rise in mortality rates of middle-aged non-Hispanic whites and the increases in lifestyle-induced deaths such as suicides, poisonings (e.g., drug overdoses), and chronic liver disease (to which excess drinking can contribute). They posit that an explanation may be the rise of the opioid crisis combined with increasing financial insecurity. Alternatively, Beckfield and Bambra (2016) compare the US to other OECD nations and find that part of the relative decline in American life expectancy can be explained by the lower levels of safety net benefits relative to comparable OECD nations.

To explore variation in life expectancy rates by income and over space, Chetty et al. (2016) use detailed individual data from the Internal Revenue Service (IRS) and Social Security Administration (SSA) to estimate race- and ethnicity-adjusted life expectancy rates by income, sex, and area (county, state, and commuting zone). They find a 10-year difference in life expectancy between the lowest and highest income groups for women and a 15-year difference for men, with these gaps increasing over time. This is consistent with other work using individual data by Clarke et al. (2010) and Crimmins and Saito (2001) who find evidence of racial and income variation in life expectancies and Cristia (2009) and Crimmins and Saito (2001) who show growing differentials over time. Chetty et al. (2016) also find that the geographic variation in life expectancy is being driven primarily by those in lower income groups and their results suggest that most of the variation in income and space appears to be related to differences in health behaviors, such as smoking and obesity, not health access or other economic or community conditions.

The strong association between personal health factors and life expectancy is also apparent in Dywer-Lindren et al.'s (2017) analysis of county-level life expectancy, which combines a wide variety of metabolic and behavioral health variables using principle components analysis, while controlling for race, socioeconomic, and health care characteristics. It identifies the importance of accounting for these factors; thus, in our paper we build on this work by accounting for these factors directly, which can help target policy solutions.

In addition to income and poverty, employment is another economic factor that affects life expectancy. Case and Deaton (2017) explore the tremendous geographic heterogeneity in the "deaths of despair," such as overdoses and suicides, and the relationship with economic factors, such as labor force participation, using mortality data from the US Centers for Disease Control and Prevention. In contrast to Chetty et al. (2016), they find that community factors (especially related to economic security) are tied to changes and variation in life expectancy rates. Consistent with this, Singh and Siahpush (2016) show that counties in the US with higher unemployment rates have lower life expectancies. However, they also find that, despite persistence over time, high unemployment areas narrowed the gap in life expectancy rates between 1990 and 2010.

Arora et al. (2016) seek to incorporate even more community-level factors that may mitigate differences in life expectancy based on race, income, sex, and other demographic and socioeconomic factors. Using composite scores reflecting overall well-being and its six dimensions (physical health, emotional health, healthy behaviors, life evaluation, basic access, and work environment) they assess which factors explain differences in county-level life expectancy. They find that greater well-being is statistically associated with higher life expectancy and that work environment is not statistically important for women but is for men, consistent with the results in Singh and Siahpush (2016). However, because well-being is measured using composite indices, it is difficult to target policy solutions.

Further exploring the interaction among community, socioeconomic, and personal health factors, Allen et al. (2016) consider the effects of income, air pollution, smoking, and obesity on county-level life expectancy. They find positive income effects and negative pollution effects which diminish in size when controlling for smoking and obesity in the population. Frederick et al. (2019) examine similar factors in counties with isolated, mid-size cities and get similar results. Bor (2017) examines the relationship between changes in county-level life expectancy from 1980 to 2014 and voting patterns in the 2008 and 2016 presidential elections. The correlation between votes for Donald Trump in the 2016 elections and lower life expectancy suggests that life expectancy declines may be linked to community-level characteristics associated with factors that were important in the 2016 election. Finally, while Benos et al. (2019) is one of the first papers to consider the impact of spatial spillovers on life expectancy, they consider spillovers in policy only at the state level and do not consider the within-state variation in life expectancy.

Overall, much of the previous research is descriptive, not causal, and leaves many unanswered questions about the effect of specific community-level characteristics on the changes in life expectancy in terms of lower levels and increasing variation, as well as the spatial dependencies in these factors. By including a more complete set of community factors and accounting for spatial relationships in the analysis, our work takes a step toward answering these questions in a way that may be useful for policymakers.

3. Data and Methods

3.1 Data

Our dependent variable is life expectancy in 2014 from the Institute for Health Metrics and Evaluation (IHME) at the University of Washington. The IHME county-level life expectancy at birth data are calculated from the underlying morality rate using a Bayesian small area model that smooths over space, age, and time (Dwyer-Lindgren et al., 2016). The IHME model also incorporates local socioeconomic factors that are expected to predict mortality rates, including the level of education, racial and ethnic composition, median household income, and population density, in addition to the number of residents and deaths. We also include life expectancy from 1980 to control for historic conditions that might affect life expectancy in 2014.

Our overall approach is influenced primarily by Arora et al. (2016), Chetty et al. (2016), and Dwyer-Lindgren et al. (2017), who include community-level characteristics and other personal variables beyond race and socioeconomic characteristics in their analyses. However, we improve upon their work by including a more complete set of county-level or "community" factors that may affect life expectancy to provide a clearer picture of the mechanisms underlying change in county-level life expectancy, which may be useful for policymakers. All variables are measured in 2009, unless otherwise specified.

Our main interest is in looking at the impact of community-level factors as they have mostly been ignored by the previous literature. Community variables describe the aspects of a county that can be accessed by or affect individuals, such as health care access, the built and social environment, and employment by industrial sector. Having access to better health care can positively influence life expectancy. We control for this by including a health care quality index (Chetty et al., 2016), calculated from medical service utilization data from the Dartmouth Atlas of Health Care, and a measure of the number of doctors (non-federal medical doctors) per 1,000 residents, calculated using the Area Health Resources File from the Health Resources and Services Administration. The built and social environment may also affect life expectancy, both positively and negatively. To control for this, we include population change from 1980 to 2009, population density, and net migration from 1990 to 2009, calculated using the

Population Estimates Program (PEP) data from the Census Bureau. Population change and net migration together provide evidence of places that are growing or shrinking and for which the composition of the population is changing over time. We also include the proportion of restaurants offering fast food (NAICS 722211), calculated from the Census Bureau's County Business Patterns (CBP); the proportion of the population with limited access to healthy food from the USDA Economic Research Service's Food Environment Atlas; and the social capital index from the Northeast Regional Center for Rural Development, which includes the rate of social organizations, voter turnout, the decennial census response rate, and the number of domestically-focused nonprofit organizations.

To control for urbanization, we include metropolitan status (using the 2013 definition of metropolitan as specified by the Office of Management and Budget) and the distance from that county to the center of the nearest metropolitan statistical area (MSA). This allows us to distinguish between the urban core and urban fringe, as well as metro-adjacent rural counties and those that are more remote. We also control for the industrial composition of the county, which may affect the types of jobs people are performing (and thus their life expectancy) or the economic vitality. We use employment data from EMSI, Inc., which contains unsuppressed detailed county-level data on industry employment at the four-digit NAICS industry code level, to calculate the proportion of jobs in: agriculture, forestry, fishing, and hunting (NAICS 11); mining, quarrying, and oil and gas extraction (NAICS 21); utilities (NAICS 22); construction (NAICS 23); manufacturing (NAICS 31 to 33); and transportation (NAICS 48).

Since part of the regional variation in life expectancy is due to differences across individuals within those regions, our control variables include personal variables that describe the characteristics of a community's residents. We include the proportion of the population that is non-Hispanic white, calculated from the bridged race files released by the National Center for Health Statistics, as well as several variables calculated from the Census Bureau's American Community Survey (ACS): the proportion of the population who attained a high school or associate's degree, as well as those who attained a bachelor's degree or higher (compared to the omitted category of those who did not complete high school); the proportion of families that are single parents with children; and the proportion of residents that are immigrants. In interpreting these results it is important to note the potential for an ecological fallacy using these and other variables. In particular, a statistically significant coefficient estimate on non-Hispanic whites would not suggest that this ethnic group in itself is associated with or “causes” a change in life expectancy. All we can say is that counties with a higher share of non-Hispanic whites also have a higher or lower life expectancy, all else equal, whether or not that is associated directly with this ethnic group. Likewise, the fact that the share of population that does not exercise physically is associated with a lower life expectancy does not mean that physical exercise allows people to live longer, only that counties in which more individuals exercise also have a higher life expectancy.

As noted above, employment and income can also affect life expectancy. Thus, we include the labor force participation rate for the population from 16 to 64 years of age, calculated using the Local Area Unemployment Statistics (LAUS) from the Bureau of Labor Statistics; the Gini Inequality Index (a measure of income dispersion) from the ACS; and the median household income and all-ages poverty rate from the Census Bureau’s Small Area Income and Poverty Estimates. Finally, to control for health conditions and behaviors, we include the proportion of uninsured residents from the Census Bureau’s Small Area Health Insurance Estimates and several variables calculated from the Behavioral Risk Factor Surveillance System and National Health and Nutrition Examination Survey by the IHME: the incidence of obesity, leisure-time physical activity, smokers, hypertension, diabetes, and alcohol consumption. Summary statistics for all our variables are in Table 1.

3.2 Descriptive Analysis

We first explore the variation in life expectancies. As shown in Table 1, in the US, on average, women lived nearly five years more than men at the county level in 2014. This difference is also apparent in the maps of female and male life expectancy in Figure 1. However, disparity in life expectancy is not restricted to differences between the sexes. As noted previously and illustrated in Figure 1, tremendous spatial heterogeneity exists in county-level life expectancy across the US. That the average county-level female and male life expectancies (80.2 years and 75.4 years, respectively) are lower than the overall national life expectancies at birth, 81.5 for women and 76.7 for men, hints at this spatial disparity.

In 2014, life expectancy at birth in American counties ranged from 62.8 years for men and 70.1 years for women in Oglala Lakota County, South Dakota, to 85.5 years for men and 88.5 years for women in Summit County, Colorado. Spatially, the maps of life expectancy in Figure 1 indicate that, while life expectancy varies across space, there are clear regional trends. Perhaps the most striking feature of the map is the extremely low life expectancies of Native American reservations in the Dakotas and Montana juxtaposed against the high life expectancies of surrounding counties. Overall, the South is also noticeably lighter than the rest of the country, indicating lower life expectancy, particularly along the lower Mississippi River and in the Appalachian region in West Virginia and Kentucky.

To assess whether there is spatial correlation among nearby locations, we calculate Moran's I for life expectancy in 2014. Moran's I uses a unit's variation from the mean, as well as the variation in neighboring units, to determine whether a phenomenon (in our case, life expectancy) is clustered or dispersed in space. The null hypothesis is that there is no spatial pattern to the data—that differing life expectancy values are located randomly throughout the country. In this case, the statistically significant estimates ($I = 0.68$ for women and $I = 0.66$ for men) indicate that both female and male life expectancy are highly spatially clustered in the US. We further explore this clustering using local Getis-Ord statistics, which provide information on where clusters of similar life expectancies are located, indicated by hot spots (high life expectancy) or cold spots (low life expectancy). Figure 2 illustrates that there are approximately four clusters of low life expectancy: one in arctic and interior Alaska, one including the Pine Ridge and Rosebud Reservations in South Dakota, one in the Deep South and surrounding the Mississippi River, and one in the Appalachian region in Kentucky and West Virginia. However, there are differences in the intensity and extent of these clusters between the sexes, which is particularly apparent in the addition of a cool spot for men in the Standing Rock Reservation in South Dakota.

Similarly, there are approximately four large spatial clusters of high life expectancy: one from New England to Philadelphia, one in southern Minnesota and the eastern Dakotas into Nebraska, one in

Colorado, and one stretching from central Idaho into the upper Rocky Mountains of the US. The extent and intensity of these clusters is different for women and men, with a larger hot spot in the Dakotas for women but a larger hot spot in the upper Rocky Mountains for men. In addition to these regional clusters, there are also clusters of high life expectancy in several urban areas throughout the country, including in San Francisco and Los Angeles in California; El Paso and an area between Austin and San Antonio in Texas; Miami, Florida; and Washington, DC.

3.3 *Empirical Specification and Estimation*

Our base empirical specification is a county-level cross-sectional model of life expectancy in 2014 in the contiguous US:

$$LifeExp_{14,i} = \alpha + \beta Life\ Exp_{80,i} + \boldsymbol{\varphi Community_i} + \boldsymbol{\gamma Personal_i} + \boldsymbol{\delta_s} + \varepsilon_i, \#(1)$$

where $LifeExp_{14,i}$ is life expectancy at birth in 2014 in county i, $LifeExp_{80,i}$ is life expectancy at birth in 1980, $\boldsymbol{Community_i}$ and $\boldsymbol{Personal_i}$ are groups of explanatory variables describing the counties in 2009, and $\boldsymbol{\delta_s}$ are state-level dummy variables.

We include a deeply lagged measure of life expectancy from 1980 to account for the influence of intangible county characteristics and historic path dependence on county-level life expectancy. This also provides a base observation of life expectancy from which community and personal characteristics will affect change between 1980 and 2014. The correlation between life expectancy in 1980 and 2014 is 0.70 for women and 0.81 for men, which suggests that initial (1980) life expectancy does not fully explain current life expectancy. Therefore, our results indicate which characteristics are positively or negatively correlated with change in life expectancy at the county-level. We also account for changes in the composition of the county over time by including population change from 1980 to 2009 and net migration from 1990 to 2009. Finally, we include state-level dummy variables to account for omitted variables associated with state-specific characteristics, such as policies, that may positively or negatively affect county-level life expectancy. Given that women on average lived nearly five years longer than men at the county level in 2014, we also estimate our models for female and male life expectancy separately.

We use five-year lagged explanatory variables to help establish a causal relationship between life expectancy in 2014 and our key explanatory variables (which are mostly from 2009). Unfortunately, due to data availability limitations, we are unable to test whether this is the appropriate lag. However, we do note that many of the variables we include are not available in earlier years. Additionally, much earlier values of these variables may be quite different than the values of those variables in 2014, making it difficult to know whether they are causing changes in life expectancy. Finally, since the county-level life expectancy values we use are calculated by using current deaths and population characteristics, the shorter lag allows us to be more confident that the relationships between explanatory variables and life expectancy are actually affecting current values. As a sensitivity analysis, we consider the change in life expectancy from 2000 to 2009 by using the lagged 2000 life expectancy (instead of the 1980 life expectancy), and the results are similar. We report those in online appendix Table A1.

Our main variables of interest are the community-level factors, as the previous literature has mostly ignored them and because they are the ones most likely to be affected by policy. As described in Section 3.1, community variables are aspects of a location that can be accessed by or affect both residents living within the community and those living in surrounding areas, such as the types of jobs available or access to healthy food. As controls, we also include personal variables that describe the average characteristics of a community's residents, such as education level or health behaviors. The descriptive spatial analysis in Section 3.2 reveals a spatial component to the level of life expectancy. Thus, we conduct a Moran's I test on the residuals from the OLS estimation of the models for both sexes; the results indicate the presence of spatial clustering and that spatial econometric models better capture the underlying relationships in the data. We discern that a combination of a county's community and personal characteristics and the characteristics of the surrounding communities affects life expectancy, as is the case in a spatial cross-regressive model. Lagrange Multiplier tests also indicate that there may be other unspecified spatial processes affecting life expectancy, as in a spatial error model.

Given the data generating process and our diagnostics, we estimate a spatial Durbin error model:

$$y = \boldsymbol{X}\beta + \boldsymbol{WZ}\theta + u, \qquad u = \lambda \boldsymbol{W}u + \varepsilon, \#(2)$$

where y is our dependent variable (life expectancy at birth in 2014), $\boldsymbol{X}$ is a matrix of the explanatory variables from Equation 1 (community, personal, lagged life expectancy, and state dummy variables), $\boldsymbol{Z}$ is a matrix containing a subset of the explanatory variables in $\boldsymbol{X}$ as discussed above, $\boldsymbol{W}$ is a spatial weights matrix, and u and ε are the error terms. The vector β includes direct coefficient estimates, the vector θ includes spatially-lagged (indirect) coefficient estimates, and λ is the spatial parameter. We define $\boldsymbol{W}$ as a first order, queen contiguity weights matrix, in other words, a matrix that accounts for all counties that touch each other. This captures the effect of factors that may transcend county borders into surrounding communities. The spatially-lagged variables, $\boldsymbol{Z}$**,** include community variables, which may be accessed by residents in neighboring counties. We do not include resident-specific personal variables from neighboring counties as we do not expect them to affect life expectancy in other counties.

Since there is some evidence from the previous literature that factors that affect life expectancy may also vary between urban and rural areas, we also estimate spatial cross-regressive models for metropolitan and nonmetropolitan county subsets of the data. In this case, however, we are unable to model the error term as defined as above, as the subsets do not include all the surrounding counties. We, instead, adjust for spatial correlation in the error term by bootstrapping our standard errors.

Finally, our control variables include a number of factors which are likely highly multicollinear. We test for this and find that, indeed, five of our control variables (the percent of the population with Diabetes, that is Non-Hispanic White, with Bachelor's Degree, with Health Insurance, and with No Leisure Physical Activity) have variance inflation factors, a measure of multicollinearity, of greater than 10. Since these factors are not our main variables of interest, we are not concerned about accurately measuring their impact on life expectancy, just controlling for them. However, we recognize that the presence of multicollinearity may result in less precise estimates of the effect size of certain variables.

4. Results

Thus far, we have demonstrated that the county-level variation in life expectancy is spatially clustered, with several statistically significant clusters of high and low life expectancy occurring throughout the US. Our next step is assessing how life expectancy is affected by a full suite of community- and personal-level characteristics.

We focus our discussion on the results from the spatial Durbin error model in Table 2, which accounts for the underlying data generating processes and observed spatial relationships in the data. For comparison, online appendix Table A2 presents results from representative baseline regressions that include characteristics (primarily individual) that are frequently used as determinants of life expectancy: race, income or poverty, education, and health measures. Online appendix Table A3 presents robust OLS regression results for Equation 1 with the full set of explanatory variables. Together, these tables show that including the additional community- and personal-level factors and considering the spatial relationships been counties leads to changes in the magnitude and significance of several of the variables from the representative baseline regressions, implying that these frequently-used factors are capturing the effect of omitted variables with which they are correlated, rather than a direct relationship between themselves and life expectancy. This underlies the importance of including a full suite of factors in our analysis.

The literature also indicates that different factors are affecting life expectancy in urban and rural locations. While the dummy variable indicating metropolitan status is (mostly) insignificant in Table 2, the distance to the nearest urban area is statistically significant, suggesting there are some differences due to urban proximity. Thus, to explore the differences between urban and rural life expectancy, Table 3 presents the results from separate spatial cross-regressive OLS estimates of metropolitan and nonmetropolitan counties.

Turning to the results, the explanatory factor with the strongest effect on expected years of life in 2014 is lagged (1980) life expectancy. A one-year increase in 1980 life expectancy is associated with an increase in average county-level 2014 life expectancy of 0.48 years for women (95% CI: 0.45, 0.52) and

0.62 years for men (95% CI: 0.59, 0.65). However, overall, the coefficients on life expectancy in 1980 suggest that, while there is a high correlation in life expectancy over time, initial life expectancy does not fully explain current life expectancy. We further explore what factors are contributing to the changes in life expectancy by addressing the relevant community-level variables first, as they are our key variables of interest, followed by personal-level variables, our control variables.

Places that are growing (with increases in population) have higher levels of life expectancy; while those with higher levels of net migration (where positive means higher levels of in-migration than out-migration) have lower levels of life expectancy. These conflicting results may be due to the fact that population changes also include migration. At the same time, there may be something about urban areas that is negatively affecting life expectancy rates, as a one percent increase in population density is associated with a decrease in life expectancy in 2014 of 0.10 years for women (95% CI: -0.19, -0.01) and 0.27 years for men (95% CI: -0.38, -0.17). Additionally, urban counties that have neighboring urban counties experience lower levels of life expectancy (for both men and women).

Consistent with this, we find that rural female life expectancy is higher further away from metropolitan areas. However, for very remote counties, there appears to be at least some negative effects on life expectancy rates (perhaps due to access to jobs, health care, and other goods and services). For each additional 100 kilometers a neighboring county is from the center of the nearest metropolitan area, life expectancy decreases by 0.29 years for both men (95% CI: -0.49, -0.10) and women (95% CI: -0.45, -0.12).

Community social capital also has a large impact on life expectancy. A one-unit increase in the social capital index is associated with increases in life expectancy of 0.13 and 0.11 years for men (95% CI: 0.03, 0.23) and women (95% CI: 0.03, 0.20), respectively. However, it is social capital within the same county that appears to increase average life expectancies for women, while average life expectancies for men are higher when there is more social capital in surrounding counties. This suggests that, at the county-level, greater opportunities for social interaction and participation in community decisions (which comprise our measure of social capital) are associated with higher life expectancies. However, the availability of social interactions is partly the result of efforts of leaders or preferences of community members. Thus, the mechanism underlying this relationship is unclear.

Focusing on the community-level variables related to health and nutrition, increased access to fast foods is associated with a decrease in life expectancy of 0.01 years for both men and women (95% CI: -0.02, 0.00). A one percent increase in fast food restaurants in their county of residence has a direct negative influence of 0.004 years on male life expectancy (95% CI: -0.008, -0.004) and of 0.006 years on female life expectancy (95% CI: -0.007, -0.001), with an additional negative impact of 0.008 years (95% CI: -0.015, -0.001) on life expectancy for men from access in surrounding counties. However, this does not mean that more fast food restaurants themselves reduce life expectancy or that policymakers need only to limit fast food establishments in order to increase life expectancy. Rather, it is more likely that the density of fast food restaurants is a reflection of the preferences, patterns, and values of community residents, which in turn leads to lower life expectancies.

There is a positive relationship between the quality of primary care and life expectancy, with a one-point increase in the index (which ranges from 0 to 100) associated with a 0.03 year increase in the county-level life expectancy for women (95% CI: 0.02, 0.04) and a 0.04 year increase for men (95% CI: 0.02, 0.05). Better primary care in surrounding counties is also positively associated with life expectancies for rural women. This is not unexpected, as access to primary care is associated with preventative care which may lead to people living longer.

Turning to community economic factors, specifically employment opportunities, our findings suggest that a greater reliance on mining, quarrying, and oil and gas extraction has a negative impact on life expectancy, as a one percent increase in the sector's employment share decreases average life expectancy by 0.06 years for women (95% CI: -0.08, -0.04) and 0.04 years for men (95% CI: -0.06, -0.01). This negative association, which is particularly strong in neighboring counties and in rural areas, is possibly due to a greater availability of jobs with a challenging and risky work environment or the cyclical nature of these industries. It may also be due to self-selection—people who choose to work in these industries may inherently engage in riskier behaviors (Hersch and Pickton, 1995) which in turn may affect overall life expectancy rates.

Conversely, rural areas with a greater reliance on manufacturing have higher life expectancies, while life expectancy for men in urban areas is negatively associated with relatively more manufacturing jobs. Additionally, a one percent increase in the share of jobs in transportation is associated with a 0.04 year increase in the average county-level life expectancy for women (95% CI: 0.00, 0.08) and a 0.05 year increase for men (95% CI: 0.00, 0.09). This increase in female life expectancy, which is not found in rural locations, further emphasizes how the interactions among factors may be influencing life expectancy differently in urban and rural locations.

We now shift our focus to our control variables, which describe community residents. Starting with personal economic variables, we see that increases in the labor force participation rate are associated with decreases in average life expectancy for both men and women. This makes sense as certain types of work can induce stress and negatively affect health in the population, thus leading to overall lower levels of life expectancy. Higher median household incomes, on the other hand, are associated with increases in life expectancy, particularly among men. Residents in these locations may experience lower levels of financial stress or have work that provides a sense of purpose, which would overcome the stress of working. After controlling for these other factors, increases in the poverty rate only has a negative association with life expectancy for women. Similarly, income inequality, as measured by the Gini coefficient (where a higher Gini indicates more inequality), only has a small, positive association with female life expectancy. This could indicate that poverty rates and relative income inequality across space have not changed much since 1980.

Places with a greater proportion of immigrants have higher life expectancies. This may be due to a combination of restrictive immigration laws, which mean successful immigrants are more likely to be healthy, and settlement in ethnic enclaves, in which they may retain their healthy behaviors (Singh and Hiatt, 2006) or where they may have a higher degree of social cohesion. Since most immigrants locate in urban areas, it is not surprising that the relationship in stronger in urban areas.

Our health and behavior control variables are generally consistent with the previous research. Places where the population engages in less leisure-time physical activity have lower life expectancies, especially urban locations. This makes sense as low levels of physical activity are associated with additional health risk factors, and a population that engages in less physical activity would then be expected to have shorter life expectancies. However, at least part of the reason people in a location may or may not engage in physical activity is due to the quality of the built and natural environment, whose availability may be influenced by policy.

Higher rates of smoking and obesity are also associated with shortened life expectancy, but obesity rates appear to be more of a factor in increasing life expectancy for women in rural areas. While our models include the portion of the population with hypertension and diabetes, overall, neither factor is statistically significant.

Somewhat surprisingly, there is a positive association between the rates of consumption of alcohol and county-level life expectancy. However, this may be because a moderate level of alcohol consumption can improve health, so a population that drinks moderately may be healthier. For example, Ronksley et al. (2011) find better cardiovascular health and mortality outcomes for individuals who consume up to one alcoholic beverage a day than for those to abstain or drink excessively, although the negative effects of excessive drinking far outweigh those of abstaining. Further, older adults (at least 50 years old) who drink moderately have a lower hazard of depressive episodes than those who occasionally drink, excessively drink, or abstain from drinking (Keyes et al., 2019). However, the alcohol content of the beverage (Rehm and Hasan, 2020) and sex of the individual (Ronksley et al., 2011; Keyes et al., 2019) can affect the magnitude of the impact of alcohol on health.

Interestingly, a higher proportion of residents without health insurance is associated with higher life expectancy. We hypothesize that this is because most uninsured individuals are young adults between the ages of 18 and 40; this age group had an uninsured rate of 26.8 percent in 2009 (versus 15.2 percent for adults from age 40 to 64). However, when we split the sample between urban and rural areas, it appears this result is driven by rural areas.

Finally, consistent with the results based on the community-level social capital, there is evidence that social interactions within families are important to well-being, as a higher portion of single parent families is associated with a lower life expectancy among men.

Overall, our results suggest that while the characteristics of the population do affect county-level life expectancies, there are important community-level factors that also need to be considered, along with the spatial relationships between these factors, in order to address the growing disparities in life expectancy.

5. Conclusion

The increasing variation in life expectancy across space, combined with the relative decline of overall life expectancy in the US, has spurred a number of recent studies. As a result, the link between life expectancy and personal characteristics (age, education, health status, etc.) has been well-established. However, as Frederick et al. (2019) detail, if policymakers are to address lower levels of life expectancy in certain regions, it is also important to understand how community-level characteristics are associated with life expectancy.

Our study fills this gap in several important ways. First, we delineate and account for the clear spatial relationships among counties using spatial analysis and regression techniques. Second, we include a variety of community-level factors in our analysis, in addition to data on the demographics and other individual characteristics of the population. We also account for path dependency and historic factors related to life expectancy by including a deep lag of life expectancy in a county using 1980 data. Our results allow us to see what factors have led to the increasing variation in life expectancy across space.

We find that community-level factors are important in explaining the changes in life expectancy (both positive and negative) since 1980. Places that have a growing population, better quality primary care, and greater social cohesion have higher levels of life expectancy. Those with more fast food restaurants, greater population density, and a greater proportion of jobs in mining, quarrying, oil and gas have lower levels of life expectancy. Most of these factors appear to be important only locally, but there is evidence of spatial spillovers from nearby places. However, the previously studied demographic and health factors cannot be omitted from the analysis, as they are critical to understanding the variation in life expectancy.

Policymakers should consider how they can help those places with a disparity in life expectancy. That is not to say that it will be easy to use policy levers, but there may be some role for policy. For example, policymakers could help create a built environment that encourages social interactions and physical activity—perhaps through new public transportation, public parks, or other infrastructure. There is also a lot still unknown. Why is life expectancy lower in large, densely populated urban areas? Are jobs in mining leading to lower life expectancies or is it something about the area in which mining jobs are located? Future research could try to tackle these questions in order to help identify other ways to raise life expectancies and shrink the variation in life expectancies across the US.

At the same time, the population for which life expectancy is calculated in a county has changed since 1980, through net migration and net births and deaths. Places that have gained population have enjoyed increases in life expectancy, perhaps due to some selectivity characteristics of those who moved (in general, they may have had more resources and been healthier). Future research may want to consider using data that would allow for tracking individuals over their lifespans in order to control for this sorting.

We recognize that one of the limitations of our study is that we do not have individual-level data but instead use county-level rates of individual characteristics, which requires careful interpretation to avoid ecological fallacies as noted above. Another weakness arises from the long time lags involved and data limitations that forced us to use explanatory variables primarily from 2009. At that same time, a key advantage of our approach using county-level data to proxy individual characteristics is that we can draw more robust conclusions for the entire nation, and we are able to estimate direct and indirect spillover effects of the different variables over space. This would not be possible with individual data alone. While we find important mechanisms that are driving the variation of life expectancy over space, we also recognize that it may be difficult to change some of these factors. Future studies may want to consider what types of policies could help improve outcomes for those places where life expectancy is the lowest.

References

Acciai, F., and G. Firebaugh 2017. “Why Did Life Expectancy Decline in the United States in 2015? A Gender-Specific Analysis.” *Social Science & Medicine* 190: 174-180. https://doi.org/10.1016/j.socscimed.2017.08.004

Allen, R.T., N.M. Hales, A. Baccarelli, M. Jerrett, M. Ezzati, D.W. Dockery, and C.A. Pope. 2016. “Countervailing Effects of Income, Air Pollution, Smoking, and Obesity on Aging and Life Expectancy: Population-Based Study of U.S. Counties.” *Environmental Health* 15(1): 86. https://doi.org/10.1186/s12940-016-0168-2

Arias, E., L.A. Escobedo, J. Kennedy, C. Fu, and J. Cisewski. 2018. “U.S. Small-area Life Expectancy Estimates Project: Methodology and Results Summary” National Center for Health Statistics *Vital Health Stat* 2(181). Retrieved from https://www.cdc.gov/nchs/data/series/sr_02/sr02_181.pdf

Arora, A., E. Spatz, J. Herrin, C. Riley, B. Roy, K. Kell, …, and H.M. Krumholz. 2016. “Population Well-Being Measures Help Explain Geographic Disparities in Life Expectancy at the County Level” *Health Affairs* 35(11): 2075-82. https://doi.org/10.1377/hlthaff.2016.0715

Beckfield, J. and C. Bambra. 2016. “Shorter Lives in Stingier States: Social Policy Shortcomings Help Explain the US Mortality Disadvantage.” *Social Science & Medicine* 171: 30-38. https://doi.org/10.1016/j.socscimed.2016.10.017

Bellos, S., P. Skapinakis, D. Rai, P. Zitko, R. Araya, G. Lewis, …, and V. Mavreas. 2013. “Cross-Cultural Patterns of the Association Between Varying Levels of Alcohol Consumption and the Common Mental Disorders of Depression and Anxiety: Secondary Analysis of the WHO Collaborative Study on Psychological Problems in General Health Care.” *Drug and Alcohol Dependence* 133(3):825-831. https://doi.org/10.1016/j.drugalcdep.2013.08.030

Benos, N., S. Karkalakos, and S. Zotou. 2019. “Spatial and economic patterns in life expectancy among US States.” *Applied Economics*. https://doi.org/10.1080/00036846.2019.1630706

Bloemen, H., S. Hochguertel, and J. Zweerink. 2018. “Job Loss, Firm-Level Heterogeneity and Mortality: Evidence from Administrative Data.” *Journal of Health Economics* 59: 78-90. https://doi.org/10.1016/j.jhealeco.2018.03.005

Bor, J. 2017. “Diverging Life Expectancies and Voting Patterns in the 2016 US Presidential Election.” *American Journal of Public Health*, 107(10), 1560-62. DOI: 10.2105/AJPH.2017.303945

Brodish, P.H., and J.K. Hakes. 2016. “Quantifying the Individual-Level Association Between Income and Mortality Risk in the United States Using the National Longitudinal Mortality Study.” *Social Science & Medicine* 170: 180-187. https://doi.org/10.1016/j.socscimed.2016.10.026

Case, A. and A. Deaton. 2015. “Rising Morbidity and Mortality in Midlife among White Non-Hispanic Americans in the 21st Century.” *Proceedings of the National Academy of Sciences* 112(49): 15078-83. https://doi.org/10.1073/pnas.1518393112

Case, A. and A. Deaton. 2017. “Mortality and Morbidity in the 21st Century.” Brookings Papers on Economic Activity, Spring 2017: 397-476.

Chetty, R., M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, …, and D. Cutler. 2016. “The Association Between Income and Life Expectancy in the US, 2001-2014.” *Journal of the American Medical Association* 315(16): 1750-66. DOI:10.1001/jama.2016.4226

Clarke, C. A., T. Miller, E.T. Chang, D. Yin, M. Cockburn, and S.L. Gomez. 2010. "Racial and Social Class Gradients in Life Expectancy in Contemporary California." *Social Science & Medicine* 70(9): 1373-1380. https://doi.org/10.1016/j.socscimed.2010.01.003

Cossman, J., W. James, and J.K. Wolf. 2017. "The Differential Effects of Rural Health Care Access on Race-Specific Mortality." *SSM - Population Health*, 3: 618-23. https://doi.org/10.1016/j.ssmph.2017.07.013

Cristia, J.P. 2009. "Rising Mortality and Life Expectancy Differentials by Lifetime Earnings in the US." *Journal of Health Economics* 28(5): 984-95. https://doi.org/10.1016/j.jhealeco.2009.06.003

Crimmins, E. M., and Y. Saito. 2001. "Trends in Healthy Life Expectancy in the United States, 1970–1990: Gender, Racial, and Educational differences." *Social Science & Medicine* 52(11): 1629-1641. https://doi.org/10.1016/S0277-9536(00)00273-2

Currie, J. and H. Schwandt. 2016. "Mortality Inequality: The Good News from a County-Level Approach." *Journal of Economic Perspectives* 30(2): 29-52. DOI: 10.1257/jep.30.2.29

Dwyer-Lindgren L., A. Bertozzi-Villa, R.W. Stubbs, C. Morozoff, M.J. Kutz, C. Huynh, …, and C.J.L. Murray. 2016. "US County-Level Trends in Mortality Rates for Major Causes of Death, 1980-2014" *Journal of the American Medical Association* 316(22): 2385-2401. DOI:10.1001/jama.2016.13645

Dwyer-Lindgren, L., A. Bertozzi-Villa, R.W. Stubbs, C. Morozoff, J.P. Mackenbach, F.J. van Lenthe, …, and C.J.L. Murray. 2017. "Inequalities in Life Expectancy Among US Counties, 1980 to 2014: Temporal Trends and Key Drivers" *Journal of the American Medical Association Internal Medicine* 177(7): 1003-11. DOI:10.1001/jamainternmed.2017.0918

Frederick, C., A. Hammersmith, and J.H. Gilderbloom. 2019. "Putting 'place' in its place: Comparing place-based factors in interurban analyses of life expectancy in the United States." *Social Science & Medicine* 232: 148-155. https://doi.org/10.1016/j.socscimed.2019.04.047

Goldin, C. and A. Llera-Muney. 2018. "XX>XY?: The Changing Female Advantage in Life Expectancy." NBER Working Paper No. 24716, The National Bureau of Economic Research. DOI:10.3386/w24716

Hersch J. and T.S. Pickton. 1995. "Risk-taking Activities and Heterogeneity of Job-risk Tradeoffs." *Journal of Risk and Uncertainty* 11(3): 2015-217. https://doi.org/10.1007/BF01207786

"How Life Expectancy Varies across America: Buying Time." 2018, September 25. *The Economist (Online).*

Institute for Health Metrics and Evaluation (IHME). 2016. Life Expectancy & Mortality Risk by County 1980-2014. [Dataset]. Retrieved from http://ghdx.healthdata.org/record/ihme-data/united-states-life-expectancy-and-age-specific-mortality-risk-county-1980-2014.

Institute for Health Metrics and Evaluation (IHME). 2018. "How Healthy Will We Be in 2040?" [Press Release]. Retrieved from http://www.healthdata.org/news-release/how-healthy-will-we-be-2040.

James, W. and J.S. Cossman. 2017. "Long-Term Trends in Black and White Mortality in the Rural US: Evidence of a Race-Specific Rural Mortality Penalty." *The Journal of Rural Health*, 33(1): 21-31. https://doi.org/10.1111/jrh.12181

James, W., J. Cossman, and J.K. Wolf. 2018. “Persistence of Death in the US: The Remarkably Different Mortality Patterns between America’s Heartland and Dixieland.” *Demographic Research* 39: 897-910. DOI:10.4054/DemRes.2018.39.33

James, W., J. Wolf, and J. Cossman. 2019. “Characteristics of Persistently Unhealthy Counties: A Comparison of White and Black Mortality over Time.” *The Social Science Journal*. https://doi.org/10.1016/j.soscij.2019.01.005

Keyes, K.M., K. Allel, U.M. Staudinger, K.A. Ornstein, and E. Calvo. 2019. “Alcohol Consumption Predicts Incidence of Depressive Episodes across 10 Years among Older Adults in 19 Countries.” *International Review of Neurobiology* 148: 1-38. https://doi.org/10.1016/bs.irn.2019.09.001

Kontis, V., J.E. Bennett, C.D. Mathers, G. Li, K. Foreman, and M. Ezzati. 2017. “Future Life Expectancy in 35 Industrialised Countries: Projections with a Bayesian Model Ensemble.” *The Lancet* 389(10076): 1323-35. https://doi.org/10.1016/S0140-6736(16)32381-9

Leung, M.C.M., J. Zhang, and J. Zhang. 2004. “An Economic Analysis of Life Expectancy by Gender with Application to the US.” *Journal of Health Economics* 23(4): 737-59. https://doi.org/10.1016/j.jhealeco.2003.11.001

“Life Expectancy: Not Great, Again.” 2018, January 6. *The Economist* 426(9073): 23.

Montez, J. K., A. Zajacova, and M.D. Hayward. 2016. “Explaining inequalities in women's mortality between U.S. States.” *SSM - Population Health* 2: 561-571. https://doi.org/10.1016/j.ssmph.2016.07.004

Rehm, J. and O.S.M. Hasan. 2020. “Is Burden of Disease Differentially Linked to Spirits? A Systematic Scoping Review and Implications for Alcohol Policy.” *Alcohol* 82: 1-10. https://doi.org/10.1016/j.alcohol.2019.06.005

Roelfs, D.J., E. Shor, K.W. Davidson, and J.E. Schwartz. 2011. “Losing Life and Livelihood: A Systematic Review and Meta-Analysis of Unemployment and All-Cause Mortality.” *Social Science and Medicine* 72(6): 840–854. https://doi.org/10.1016/j.socscimed.2011.01.005

Ronksley, P.E., S.E. Brien, B.J. Turner, K.J. Mukamal, and W.A. Ghali. 2011. “Association of Alcohol Consumption with Selected Cardiovascular Disease Outcomes: A Systematic Review and Meta-Analysis.” *BMJ* 342: d671. https://doi.org/10.1136/bmj.d671

Singh, G.K. and R.A. Hiatt. 2006. “Trends and Disparities in Socioeconomic and Behavioral Characteristics, Life Expectancy, and Cause-Specific Mortality of Native-Born and Foreign-Born Populations in the US, 1979-2003.” *International Journal of Epidemiology* 35(4): 903-919. https://doi.org/10.1093/ije/dyl089

Singh, G.K., M. Siahpush, and S.F. Altekruse. 2013. “Time Trends in Liver Cancer Mortality, Incidence, and Risk Factors by Unemployment Level and Race/Ethnicity, US, 1969-2011.” *Journal of Community Health* 38(5): 926–940. https://doi.org/10.1007/s10900-013-9703-z

Singh, G.K. and M. Siahpush. 2016. “Inequalities in US Life Expectancy by Area Unemployment Level, 1990-2010.” *Scientifica* 2016: 12 pages. http://dx.doi.org/10.1155/2016/8290435

Singh, G.K., M.D. Kogan, and R.T. Slifkin. 2017. “Widening Disparities in Infant Mortality and Life Expectancy between Appalachia and the Rest of the US, 1990–2013.” *Health Affairs* 36(8): 1423-32. https://doi.org/10.1377/hlthaff.2016.1571

Strumpf, E. C., T.J. Charters, S. Harper, and A. Nandi. 2017. “Did the Great Recession Affect Mortality Rates in the Metropolitan United States? Effects on Mortality by Age, Gender and Cause of Death.” *Social Science & Medicine* 189: 11-16. https://doi.org/10.1016/j.socscimed.2017.07.016

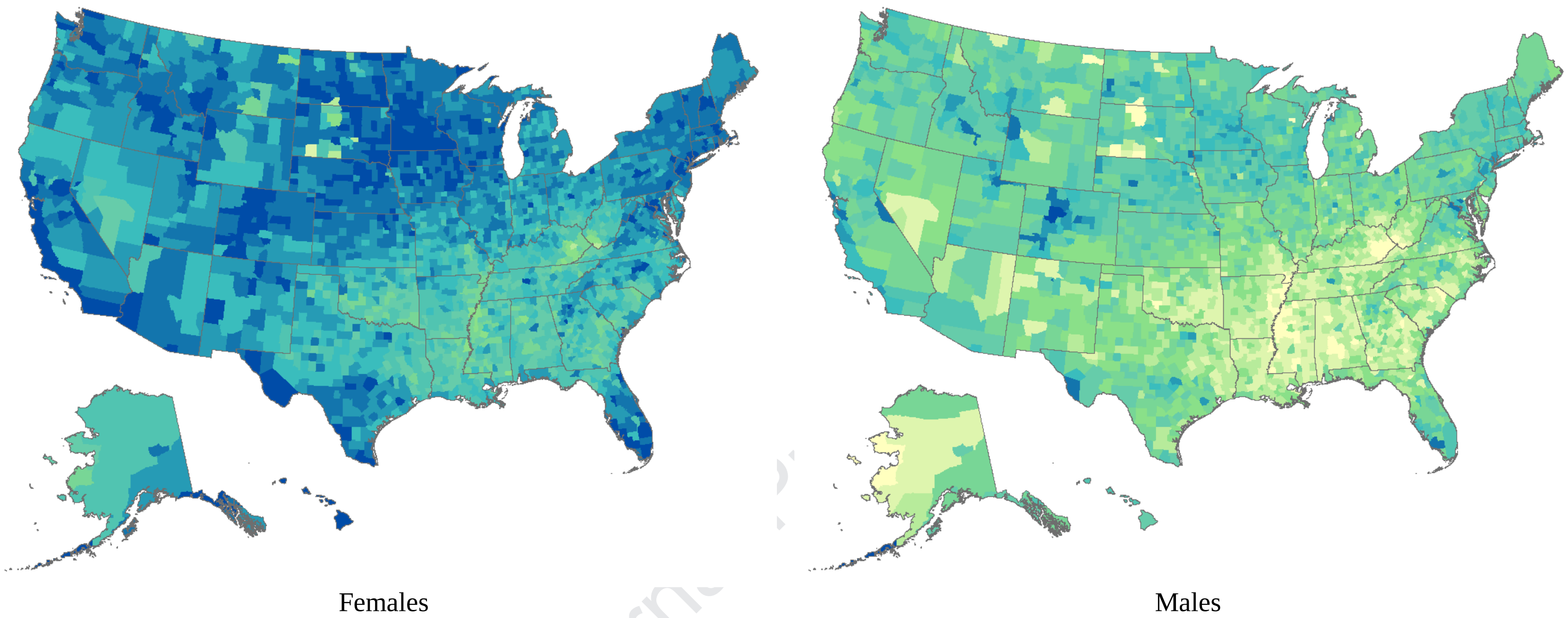

Figure 1. **Life Expectancy in the US, 2014**. This map displays county-level life expectancy at birth for men and women in 2014. The values, which range from 62.759 years to 88.511 years, are divided into 11 categories (62.75–69.88, 69.89–72.22, 72.23–73.88, 73.89–75.26, 75.27–76.52, 76.53–77.68, 77.69–78.81, 78.82–79.95, 79.96–81.19, 81.20–82.69, 82.70–88.51) determined by natural breaks in the data. The color gradation ranges from cream for low values to dark blue for high values.

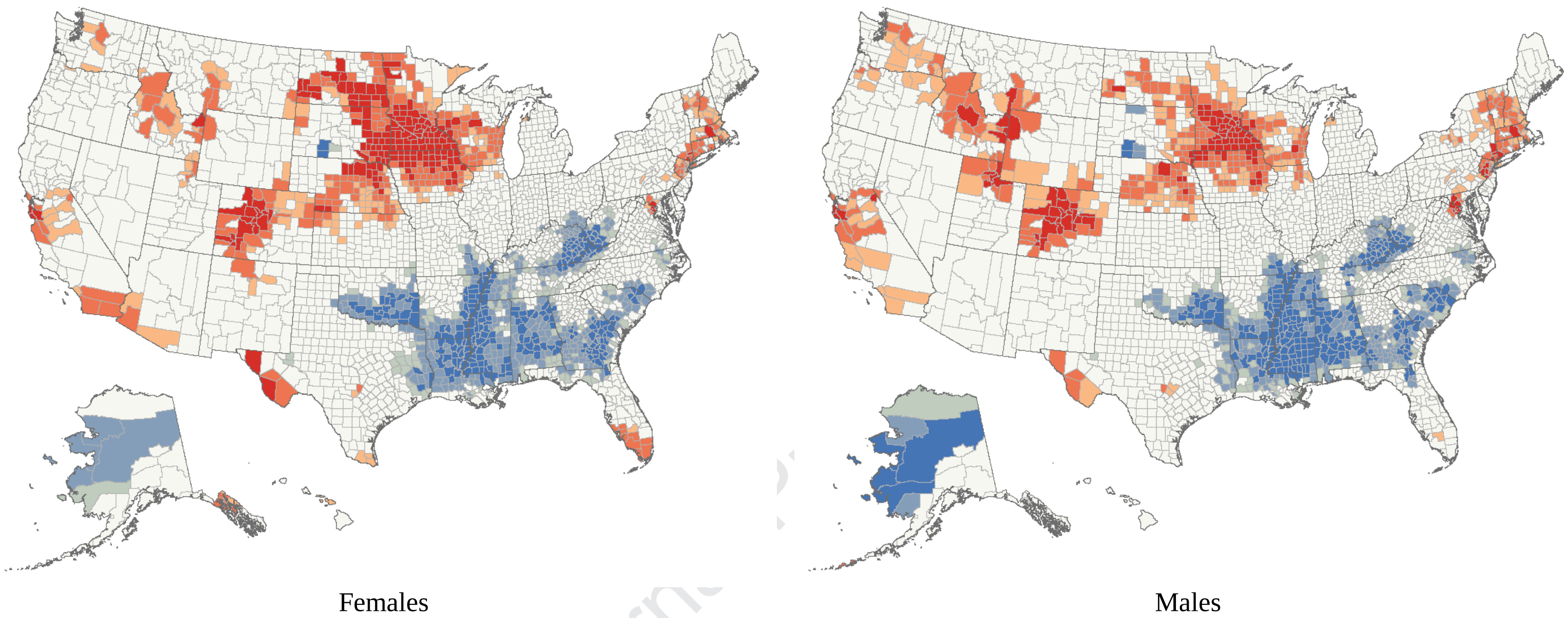

Figure 2. Hot and Cold Spots for Life Expectancy in the US, 2014. This map displays hot and cold spots, clusters of similar values, for county-level life expectancy at birth for men and women in 2014. Red values indicate a statistically significant cluster of high life expectancy levels (90% level for light red, 95% for medium red, and 99% for dark red), white values indicate counties that are not part of a statistically significant cluster, and blue values indicate a statistically significant cluster of low life expectancy levels (90% for light blue, 95% for medium blue, and 99% for dark blue).

Table 1. Descriptive Statistics

	Females				Males			
	Mean	*Std. Dev.*	*Min.*	*Max.*	*Mean*	*Std. Dev.*	*Min.*	*Max.*
Life Expectancy †, 1980	77.813	1.527	66.844	82.606	70.005	2.065	56.601	76.816
Life Expectancy †, 2014	80.205	2.092	70.952	88.511	75.351	2.688	62.759	85.459
Population Change, 1980 to 2009, %	26.540	59.307	-55.358	1,018.698	26.540	59.307	-55.358	1,018.698
Population Density, per sq. km	95.993	668.914	0.044	26,985.40	95.993	668.914	0.044	26,985.40
Total Net Migration, 1990 to 2009, %	8.208	26.057	-68.468	282.951	8.208	26.057	-68.468	282.951
Fast Food Restaurants, %	34.951	14.935	0.000	100.000	34.951	14.935	0.000	100.000
Limited Access to Healthy Food, %	8.408	8.091	0.000	72.300	8.408	8.091	0.000	72.300
Social Capital Index	0.002	1.338	-3.925	17.441	0.002	1.338	-3.925	17.441
Primary Care Quality Index	71.994	6.611	27.700	95.833	71.994	6.611	27.700	95.833
Physicians, per 1,000 residents	1.195	1.480	0.000	30.663	1.195	1.480	0.000	30.663
Jobs in Ag, Forestry, Fishing, Hunting, %	4.967	7.263	0.000	71.170	4.967	7.263	0.000	71.170
Jobs in Mining, Quarrying, Oil, Gas, %	1.411	4.168	0.000	86.910	1.411	4.168	0.000	86.910
Jobs in Utilities, %	0.695	1.536	0.000	41.040	0.695	1.536	0.000	41.040
Jobs in Construction, %	6.518	3.228	0.400	66.460	6.518	3.228	0.400	66.460
Jobs in Manufacturing, %	10.403	8.436	0.000	64.040	10.403	8.436	0.000	64.040
Jobs in Transportation, %	2.580	2.100	0.000	32.590	2.580	2.100	0.000	32.590
Metropolitan County	0.371	0.483	0.000	1.000	0.371	0.483	0.000	1.000
Distance to Center of Nearest MSA, km	70.135	54.019	0.003	411.121	70.135	54.019	0.003	411.121
Labor Force Participation Rate, %	62.032	8.618	29.132	100.000	62.032	8.618	29.132	100.000
Gini Income Inequality Index	0.434	0.036	0.200	0.671	0.434	0.036	0.200	0.671
Poverty Rate, %	16.365	6.442	3.100	62.000	16.365	6.442	3.100	62.000
Median Household Income, $1000	42.730	10.972	18.860	114.200	42.730	10.972	18.860	114.200
High School Degree Only †, %	65.313	7.030	23.166	86.512	63.631	7.713	16.514	90.244
Bachelor's Degree or Above †, %	19.530	8.204	0.000	69.356	18.792	9.309	0.000	78.698
Non-Hispanic White Population †, %	80.313	19.625	2.559	99.565	79.368	19.396	2.551	100.000
Single Parent Family with Children †, %	25.120	9.147	0.000	68.700	10.488	4.643	0.000	51.600
Immigrants †, %	4.166	5.267	0.000	52.600	4.497	5.622	0.000	55.700
No Health Insurance †, %	18.199	5.642	3.000	43.000	18.199	5.642	3.000	43.000
Obesity †, %	38.109	5.473	17.180	59.320	35.965	3.465	17.810	45.570
No Leisure Physical Activity †, %	28.564	5.900	10.810	49.880	25.447	4.773	9.870	44.350
Smokers †, %	22.774	4.353	5.870	37.440	26.693	4.269	9.600	40.180
Hypertension †, %	40.709	3.722	28.520	57.880	38.200	3.636	26.530	54.430
Diabetes †, %	12.685	2.794	6.580	28.280	15.342	2.045	9.390	25.170
Any Alcohol †, %	41.603	12.953	8.300	76.400	57.829	11.691	13.000	83.400

Note: Some variables are available by sex (†), while others are only for the entire population.

Table 2. Spatial Durbin Error Regression Results

	Females			Males		
	Direct Effects	*Indirect Effects*	*Total Effects*	*Direct Effects*	*Indirect Effects*	*Total Effects*
Intercept	43.868 ***			30.278 ***		
Life Expectancy, 1980	0.482 ***			0.622 ***		
Community Variables						
Population Change, 1980 to 2009, %	0.005 ***	0.002	0.007 ***	0.005 ***	0.004 **	0.009 ***
Population Density, per sq. km	-0.163 ***	0.064	-0.099 **	-0.247 ***	-0.027	-0.275 ***
Total Net Migration, 1990 to 2009, %	-0.008 ***	-0.005	-0.013 ***	-0.004 **	-0.008 **	-0.012 ***
Fast Food Restaurants, %	-0.006 ***	-0.005	-0.011 ***	-0.004 ***	-0.008 **	-0.012 ***
Limited Access to Healthy Food, %	0.006 ***	0.008	0.014 **	-0.001	-0.005	-0.006
Social Capital Index	0.077 ***	0.036	0.113 ***	0.037 *	0.092 **	0.129 **
Primary Care Quality Index	0.022 ***	0.005	0.027 ***	0.027 ***	0.009	0.036 ***
Physicians, per 1,000 residents	-0.009	0.029	0.020	-0.005	0.061	0.056
Jobs in Ag, Forestry, Fishing, Hunting, %	0.006 **	-0.011 **	-0.005	0.004	0.002	0.006
Jobs in Mining, Quarrying, Oil, Gas, %	-0.013 ***	-0.043 ***	-0.056 ***	-0.008	-0.028 ***	-0.036 ***
Jobs in Utilities, %	0.005	0.016	0.021	0.009	0.017	0.026
Jobs in Construction, %	0.016 ***	0.004	0.020	0.010	0.001	0.011
Jobs in Manufacturing, %	0.001	-0.009 *	-0.008	0.001	0.001	0.002
Jobs in Transportation, %	0.021 ***	0.022	0.043 **	0.013	0.035 *	0.047 **
Metropolitan County	0.010	-0.168 *	-0.158	0.040	-0.146	-0.106
Distance to Nearest MSA, 100 km	0.009	-0.296 ***	-0.287 ***	-0.0001	-0.294 **	-0.294 ***
Personal Variables						
Labor Force Participation Rate, %	-0.010 ***			-0.023 ***		
Gini Income Inequality Index	0.972 *			0.640		
Poverty Rate, %	-0.018 ***			-0.009		
Median Household Income, $1000	0.011 ***			0.046 ***		
High School Degree Only, %	-0.015 ***			-0.011 *		
Bachelor's Degree or Above, %	-0.006			0.009		
Non-Hispanic White Population, %	-0.002			-0.003		
Single Parent Family with Children, %	-0.003			-0.008 **		
Immigrants, %	0.034 ***			0.036 ***		
No Health Insurance, %	0.041 ***			0.065 ***		
Obesity, %	-0.030 ***			-0.030 ***		
No Leisure Physical Activity, %	-0.035 ***			-0.025 **		
Smokers, %	-0.098 ***			-0.067 ***		
Hypertension, %	0.012			0.005		
Diabetes, %	0.0002			-0.028		
Any Alcohol, %	0.034 ***			0.032 ***		
State Effects	Y			Y		
λ	0.424 ***	0.024		0.417 ***	0.024	
N	3085			3085		
Log-likelihood	-3,343.58			-3,880.13		

Statistical Significance: '*' 10%, '**' 5%, '***' 1%

Table 3. Metropolitan and Nonmetropolitan Spatial Cross-Regressive Regression Results

	METROPOLITAN COUNTIES				NONMETROPOLITAN COUNTIES			
	Females		Males		Females		Males	
	Direct Eff	*Indirect Eff*	*Direct Eff*	*Indirect Eff*	*Direct Eff*	*Indirect Eff*	*Direct Eff*	*Indirect Eff*
Intercept	53.649 ***		44.185 ***		40.229 ***		26.262 ***	
Life Expectancy, 1980	0.369 ***		0.445 ***		0.520 ***		0.663 ***	
Community Variables								
Population Change, 1980 to 2009, %	0.003 ***	0.001	0.002	0.002	0.005 ***	0.0003	0.010 ***	0.004 *
Population Density, per sq. km	-0.166 ***	0.080	-0.259 ***	0.032	-0.186 ***	-0.020	-0.380 ***	-0.108
Total Net Migration, 1990 to 2009, %	-0.004 *	-0.009 **	0.0004	-0.007	-0.008 ***	0.002	-0.008 **	-0.008
Fast Food Restaurants, %	-0.014 ***	-0.001	-0.012 ***	-0.0003	-0.003 *	-0.006 *	-0.0002	-0.011 ***
Limited Access to Healthy Food, %	-0.002	0.016	-0.003	-0.004	0.007 ***	0.002	-0.001	-0.012
Social Capital Index	0.058 *	-0.058	0.065	-0.026	0.056 *	-0.003	-0.010	0.098 *
Primary Care Quality Index	0.023 ***	0.006	0.035 ***	0.012	0.021 ***	0.014 **	0.023 ***	0.010
Physicians, per 1,000 residents	-0.001	0.046	-0.032 *	0.031	-0.023	-0.011	0.047	0.048
Jobs in Ag, Forestry, Fishing, Hunting, %	0.023 ***	0.002	0.008	0.013	0.004	-0.015 **	0.002	0.002
Jobs in Mining, Quarrying, Oil, Gas, %	0.014	-0.036 *	0.004	-0.021	-0.012 *	-0.046 ***	-0.007	-0.044 ***
Jobs in Utilities, %	-0.009	0.038	-0.015	0.069 *	0.011	0.006	0.010	-0.018
Jobs in Construction, %	0.021	0.036 *	0.009	0.013	0.016 **	-0.030 *	0.010	-0.001
Jobs in Manufacturing, %	-0.006	-0.004	-0.011 **	0.003	0.007 **	-0.007	0.008 **	0.005
Jobs in Transportation, %	0.024 **	0.010	0.005	0.023	0.016	0.003	0.011	0.018
Metropolitan County		-0.384 **		-0.573 ***		-0.037		0.046
Distance to Nearest MSA, 100 km	-0.161	-0.475 *	-0.042	-0.795 ***	0.304 **	-0.567 ***	0.172	-0.432 **
Personal Variables								
Labor Force Participation Rate, %	-0.017 ***		-0.021 ***		-0.009 **		-0.026 ***	
Gini Income Inequality Index	0.901		-1.478		0.580		0.362	
Poverty Rate, %	-0.027 **		-0.013		-0.021 **		-0.009	
Median Household Income, $1000	0.006		0.030 ***		0.006		0.057 ***	
High School Degree Only, %	0.004		-0.005		-0.016 **		-0.011	
Bachelor’s Degree or Above, %	-0.004		0.018		-0.001		0.002	
Non-Hispanic White Population, %	-0.002		-0.002		-0.001		-0.004	
Single Parent Family with Children, %	0.0004		-0.031 **		-0.007 *		-0.007	
Immigrants, %	0.069 ***		0.080 ***		0.010		0.011	
No Health Insurance, %	0.021		0.026		0.051 ***		0.069 ***	
Obesity, %	-0.015		-0.013		-0.034 ***		-0.020	
No Leisure Physical Activity, %	-0.053 ***		-0.040 **		-0.022 **		-0.023 *	
Smokers, %	-0.102 ***		-0.065 ***		-0.116 ***		-0.090 ***	
Hypertension, %	-0.029		-0.034		0.004		0.037 *	
Diabetes, %	0.014		0.015		0.060		-0.018	
Any Alcohol, %	0.037 ***		0.036 ***		0.040 ***		0.035 ***	
State Effects	Y		Y		Y		Y	
N	3085		3085		3085		3085	
R-squared	0.877		0.902		0.871		0.886	

Statistical Significance: '*' 10%, '**' 5%, '***' 1%

EXPLAINING THE SPATIAL VARIATION IN AMERICAN LIFE EXPECTANCY

Elizabeth A. Dobis
Northeast Regional Center for Regional Development
The Pennsylvania State University
University Park, PA, United States
edobis@psu.edu

Heather M. Stephens
Resource Economics and Management
West Virginia University
Morgantown, WV, United States
heather.stephens@mail.wvu.edu

Mark Skidmore
Department of Agriculture, Food, and Resource Economics
North Central Regional Center for Rural Development
Michigan State University
East Lansing, MI, United States
mskidmor@msu.edu

Stephan J. Goetz
Department of Agricultural Economics, Sociology, and Education
Northeast Regional Center for Regional Development
The Pennsylvania State University
University Park, PA, United States
sgoetz@psu.edu

November 29, 2019

Funding Acknowledgement: Dobis and Goetz gratefully acknowledge USDA NIFA funding under grants 2017-51150-27125 and 2018-51150-28696 as well as support from the Pennsylvania State University and NIFA Multistate/Regional Research Appropriations under project #NE1749. The funders were not involved in the research, and any opinions expressed are strictly those of the authors and not of the funders.

EXPLAINING THE SPATIAL VARIATION IN AMERICAN LIFE EXPECTANCY

Author Contributions:

Elizabeth A. Dobis: Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Visualization, Writing – Original Draft, Writing – Review & Editing

Heather M. Stephens: Conceptualization, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing

Mark Skidmore: Conceptualization, Methodology, Writing – Original Draft, Writing – Review & Editing

Stephan J. Goetz: Conceptualization, Methodology, Funding Acquisition, Writing – Original Draft, Writing – Review & Editing