



Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Segregation and mortality over time and space

Trevon D. Logan^{a,*}, John M. Parman^b^a Department of Economics, The Ohio State University and NBER, United States^b Department of Economics, College of William and Mary and NBER, United States

ARTICLE INFO

Article history:

Received 14 October 2016

Received in revised form

1 June 2017

Accepted 10 July 2017

Available online xxx

JEL classifications:

I1

J1

N3

N9

Keywords:

United States

Segregation

Race

Mortality

Health disparities

Residential sorting

ABSTRACT

Few studies have been able to measure the evolution of segregation on health disparities or assess whether those disparities existed in rural communities prior to the Great Migration of African Americans to urban centers. We use a newly developed measure of historical racial residential segregation based on individual-level data. The measure exploits complete census manuscript files to identify the races of next-door neighbors. This measure is the first and only measure of historical segregation for rural communities, allowing us to greatly extend the empirical analysis of the effects of racial segregation on health over space and time. Using this comprehensive measure of racial residential segregation, we estimate the historical relationship between racial segregation and mortality. We find that conditional on racial composition, racially segregated environments had higher mortality rates and it was not always the case that the outcomes for blacks were worse than those of whites. These effects of segregation on health differed between urban and rural locations. We conclude by noting how comprehensive measures of segregation can extend the analysis of structural factors in racial health disparities to rural residents and to the historical evolution of health disparities.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Racial residential segregation is thought to be a fundamental determinant of racial inequalities in health. Williams and Collins (2001) provide a review of contemporary evidence suggesting that segregation directly impacts access to education and employment opportunities. These differences in access then translate into racial differences in socioeconomic status (SES) and health outcomes. This is a point reinforced by Phelan and Link (2015) who consider systemic racism as the fundamental cause of racial differences in SES. They point to a large literature identifying both the presence of official policies of residential and school segregation prior to the Civil Rights Act and the continuation of de facto segregation after (Orfield and Eaton, 1997; Feagin, 2014). This segregation has been linked to poor pregnancy outcomes and increased mortality in the black community (Kramer and Hogue, 2009).

Racial residential segregation, or the sorting of different racial

groups into geographic spaces of unequal quality, has been approached through two conceptual levels. The first stems from the seminal work of Massey and Denton (1988, 1993) and considers segregation at the city or MSA level, focusing on the distribution of racial groups across census tracts of a city to get a measure of the overall level of segregation for the city as a whole. This approach aligns with a conceptual model in which segregation modifies social capital for a city as a whole and may do so differently for different racial groups. The second level considers neighborhood segregation, focusing on whether living in a neighborhood segregated from the rest of the city impacts resident health. A growing literature points to segregated neighborhoods impacting health through mechanisms such as access to healthcare, food availability, and walkability (Diez Roux and Mair, 2010; Corral et al., 2012). In this literature, census tract typically serves as a proxy for neighborhood.

There are several challenges for both of these traditions. First, both tend to focus on segregation in urban areas despite 19 percent of the US population living in rural areas and segregation being pronounced in rural communities (Lichter et al., 2007). The meaning of segregation measures based on tract will

* Corresponding author.

E-mail address: logan.155@osu.edu (T.D. Logan).

fundamentally differ between dense urban populations and more dispersed rural populations. Second, many studies do not account for historical processes that segregated groups over time and continue to influence contemporary health. Here again the ability to speak to rural areas is crucial as the majority of the black population resided in rural areas through the 1940s. Finally, the tract-based measures common to both the MSA segregation and neighborhood segregation literatures are subject to politically determined boundaries. As such, they may provide either over- or underestimates of the segregation actually experienced by the black community depending on whether boundaries are drawn to concentrate black voters in a single tract or dilute their votes across several tracts.

In this paper, we address these limitations in the literature to examine the relationship between a novel measure of racial residential segregation in 1880 and health over the subsequent 95 years. The measure of segregation we use focuses the spatial granularity to the individual neighbor-level, eliminating biases introduced by political restructuring of tract boundaries. Furthermore, this measure is not constrained to urban areas; it captures variation in rural contexts as well. This novel measure exploits newly-available digitized federal census data on race at the household level for the complete 1880 census. (These data are also currently available for the 1940 census and are under development for the remaining 1900–1930 censuses).

We use these segregation data in combination with two sources of mortality data. Adult mortality data is obtained from newly-digitized North Carolina death certificates for 1909 to 1975 to demonstrate the importance of historical information when assessing racial health inequalities. We find that rural segregation is well correlated with age at death. Importantly, we find that segregation is related to both black and white mortality. Irrespective of race, individuals from more segregated environments have significantly different mortality rates than those in integrated environments, on average. We then use child mortality data inferred from 1900 to 1910 federal census data to show that there were important differences in rural and urban infant and child mortality with respect to segregation—reinforcing the need to think of the context of segregation in health outcomes. Overall, the results lead to two related conclusions. First, segregation is correlated with health in rural areas in the early twentieth century. This finding suggests that the focus on the relationship between segregation and health should expand to include analysis of rural areas and recognize that segregation influenced health gaps well before the urbanization of the black population. Second, we find that the effects are present for infant and child mortality rates but also for adult mortality, suggesting that historically segregation had both immediate and, perhaps, cumulative effects on health.

We end with both a substantive discussion on the importance of historical segregation processes to contemporary health inequalities and also a brief technical discussion on the current developments in digitizing all census information at the neighbor-level to examine changes in segregation over time in relation to contemporary health inequalities.

2. The connection between segregation and health

2.1. Contemporary

Given the large differences in mortality by geography in the early part of the last century that varied with macrosocioeconomic measures (Crimmins and Condran, 1983), and the unequal distributions of the black and white populations, part of the differences we attribute to “early life circumstances” may be best understood as environmental or social differences rather than socioeconomic

differences per se (see Naidu (2012)). A large literature exists exploring the contemporary relationships between the spatial distribution of racial groups and racial gaps in health outcomes. Kramer and Hogue (2009) review 39 different studies measuring an association between racial segregation and health.

One main pathway explored in the literature is the link between segregation and individual socioeconomic status. Poor black individuals tend to live segregated neighborhoods while poor white individuals tend to live in more integrated neighborhoods (Wilson, 2012). These segregated neighborhoods have worse schooling and employment opportunities, leading to individuals having lower educational attainments and higher rates of unemployment both unconditionally and conditional on educational attainment (Anderson et al., 2003; Howell-Moroney, 2005). This lower socioeconomic status translates into worse health outcomes.

A second pathway involves the impact of neighborhood environment on health. Health can be impacted by the higher crime rates associated with highly segregated neighborhoods (Peterson and Krivo, 1993; Shihadeh and Flynn, 1996; O’Flaherty and Sethi, 2007). Diets are influenced by segregation due to poor, segregated neighborhoods being isolated from healthy food and containing higher numbers of fast-food restaurants (Zenk et al., 2005; Morland and Filomena, 2007; Powell et al., 2007). Studies link residential segregation to physical inactivity leading to poor health outcomes including obesity (Corral et al., 2012; Wilson-Frederick et al., 2014). Researchers have also pointed to the role of increased stress associated with segregated neighborhoods contributing to worse health outcomes (Schulz et al., 2001; Geronimus et al., 2006). Compounding all of these issues is the reduced access to high quality healthcare in segregated neighborhoods (Smedley et al., 2003).

There is also a growing body of evidence that individuals living in segregated neighborhoods are subject to greater environmental insults and exposed to greater levels of infectious disease. Even after controlling for neighborhood SES, black neighborhoods are exposed to significantly higher levels of ambient air toxics, increasing cancer risks (Morello-Frosch and Jesdale, 2006). Furthermore, there is evidence that the location of hazardous waste is associated with racial targeting and housing discrimination (Mohai and Saha, 2007). The density and isolation of segregated black populations also contributes to the spread of infectious disease including tuberculosis (Acevedo-Garcia, 2000).

2.2. Historical

Despite this growing literature on segregation and health, it is limited by the use of relatively contemporaneous segregation data. With very limited historical data on health outcomes by race and on levels of segregation, we know little about the relationship between health and segregation in a historical context. However, understanding these historical relationships is crucial to understanding modern health disparities for several distinct reasons. First, with the relatively low levels of intergenerational occupational and income mobility in the US, mobility rates that have declined over the past century (Long and Ferrie, 2013), socioeconomic status of current generations is tied to that of previous generations. Given the persistence of differences in segregation across cities over the past century documented by Cutler et al. (1999), the relationship between segregation and health today could be driven by the impacts of segregation on the socioeconomic status of earlier generations. Second, scholars are beginning to identify links between modern racial health disparities and historical living conditions. An example is the work on contemporary rates of diabetes and cardiovascular disease in the black community and their relationship to historical black poverty in the South

(Steckel, 2013; Steckel and Senney, 2015). Finally, observing the time scale over which historical increases in segregation translated into differences black and white health outcomes provides crucial evidence on how long it may take for health outcomes to respond to recent modest declines in segregation (Logan et al., 2004).

The limited historical data on health outcomes do suggest that racial health disparities have historical roots stretching back to the 19th century. Several studies have looked at the historical relationships between race, socioeconomic status, location and health outcomes. Su (2009), for example, finds occupational gradients in later life mortality for cohorts born in the 1840s favored those who did not leave their earlier occupations as farmers, which is consistent the argument that those of lower socioeconomic status enjoyed a health advantage primarily due to geography. Logan (2009) argues that migration and socioeconomic mobility were related to health and slave status in the late nineteenth century for African Americans, implying that part of the racial differences we observe are due to socioeconomic differences that themselves could be due to health differences.

While these studies consider geography and health, they do not focus specifically on residential segregation. Troesken (2002) is a rare historical study that directly considers segregation's impact on racial health gaps. Troesken argues that it was easier to exclude black households from water and sewerage service in more segregated cities, leaving the black population more exposed to waterborne disease. In this respect, Troesken's work echoes the contemporary literature on segregation and infectious disease (Acevedo-Garcia, 2000). Troesken relies on traditional measures of segregation, the isolation index and the dissimilarity index, which focus on how the racial composition of wards compare to the racial composition of the city overall.

2.3. Limitations of existing measures and data

The nature of traditional segregation measures has dramatically limited the study of historical segregation patterns and precluded any chance of estimating the impacts of those patterns on health. Population counts by race at the ward level essential to calculating traditional segregation indices are only available for a handful of cities – Troesken is limited to just 27 cities. Consequently, segregation and health outcomes in all but the largest American cities have escaped study.

Beyond limiting the number of cities that can be studied, relying on ward-level data (or census-tract level data in contemporary studies) is problematic given political motivations behind the way wards are drawn. A city in which wards are drawn to minimize the voting power of black residents by dispersing their votes across wards may appear to be highly integrated. In this scenario, the reduced political power of black residents could diminish their ability to implement policies directed at improving black health outcomes. If the same city had wards drawn to make it easier to discriminate in the provision of public services relevant to health by placing black residents in a small number of wards it would appear completely segregated according to traditional segregation measures. Thus the way that boundaries are drawn could change the sign of the estimated relationship between segregation and racial health gaps.

A final problem with the reliance on ward-level data is it makes these traditional measures inapplicable to smaller towns and rural areas that lack a comparable political unit, even if ideal population data did exist. In 1870 roughly 90 percent of the black population lived in rural locations. Throughout the first decades of the twentieth century, the majority of the black population remained rural. Therefore traditional segregation measures fail to capture the historical conditions of much of the country.

We address these challenges by exploiting newly available data. Recent advances in the digitization of the federal census make it possible to focus on household-level rather than ward-level data. Logan and Parman (2017) use the recently available 100 percent samples of the 1880 and 1940 censuses to identify the races of each household's next-door neighbors and construct a segregation measure that focuses on the number of black households with white neighbors. This approach to measuring segregation has two distinct advantages for a study of segregation and health. First, the measure of segregation has a consistent interpretation that is not driven by population density or the boundaries of geographic subunits. This makes the neighbor-based measure applicable to smaller cities and rural areas in ways that traditional measures are not. The second key advantage is that this next-door neighbor measure is a better proxy for physical proximity between individuals of different races and interracial interactions than traditional measures based on racial proportions. Physical proximity will be closely related to whether individuals of different races are exposed to the same health shocks, particularly in historical time periods with high rates of infectious disease. Interracial interactions may be crucial for health outcomes particularly if increased interaction makes white populations more likely to extend public health goods to the black population.

We use the 1880 Logan-Parman segregation data and mortality data from the 1900 and 1910 federal censuses as well as death records for North Carolina from 1909 to 1975 as a case study to explore the historical relationship between segregation and black and white mortality rates. We conclude by discussing the application of the Logan-Parman measure to the remaining census years and death records for the entire country.

3. Conceptual model and hypotheses

Historically, segregation will impact African American health outcomes most directly by concentrating the black population in neighborhoods with worse conditions, both in terms of exposure to more environmental hazards and in terms of reduced access to public water and sewerage. As a consequence, we expect mortality rates to be higher for black individuals in more segregated communities. Unlike the modern literature on increased stress in black communities (see, for example, Massey (2004) and Duru et al. (2012)), historically one would expect infectious disease and waterborne illness to drive higher mortality rates as argued by Troesken. These issues are relevant in both rural and urban areas but will be more pronounced in denser populations, leading us to expect larger negative impacts of segregation on African American health outcomes in urban relative to rural areas.

This conceptual framework leads us to the following hypotheses:

H1: Segregation will have a positive correlation with mortality for the black population.

H2: Segregation will have such a relationship with mortality in both urban and rural areas but the marginal effect of segregation on mortality will be larger in urban areas.

4. Data, measures and methods

4.1. Mortality data

We use two data sources to investigate the relationship between segregation and mortality. The first source is a complete set of death certificates from North Carolina that allow us to focus on mortality at later ages. The availability of death records varies by state. North Carolina is the only state with a large black population that has a fully digitized index of the universe of death certificates

with gender and race information extending back to the early 20th century. While other states have digitized indices of death certificates for the early 1900s, they either lack the digitized race information critical to studying health disparities or cover only a fraction of counties in the state, raising serious sample selection issues. The North Carolina data come from the universe of death certificates for individuals who died in the state between the years of 1909 and 1975 and are accessed through [Ancestry.com](#). Before 1909, policies related to death registration were not uniform and therefore we do not consider earlier death certificates representative of the population. After 1975, digitized versions of the death certificates are no longer available online. This set of North Carolina death certificates contains over 1.2 million records. The relevant digitized information from the death certificates includes name, date of birth, date of death, race, and county of birth.

In addition to these death certificate data, we use federal census data that enable us to estimate infant and child mortality. The 1900 and 1910 federal censuses included questions about the number of children ever born to each adult female and the number of children ever surviving. These variables allow us to construct an indicator variable for having ever lost a child and a second variable giving the number of children lost. The children surviving question was not asked in any census years other than 1900 and 1910, restricting us to using these two censuses. We use the 1% sample of the 1910 federal census and the 5% sample of the 1900 federal census available through the Integrated Public Use Microdata Series (IPUMS) ([Ruggles et al., 2009b, 2009c](#)). Given that these data come from before the Great Migration, we restrict our attention to individuals living in the South (Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia and West Virginia). We further restrict the sample to women who have had at least one child and are under the age of 55. This helps ensure that any deaths inferred from the data correspond to children dying when young rather than as adults. These census data also include a rich set of individual characteristics including age, state of birth, current state and county of residence, race, occupational standing, and literacy.

4.2. Segregation data

We use the county-level segregation estimates from [Logan and Parman \(2017\)](#). Logan and Parman utilize the 100% sample of the 1880 federal census available through IPUMS, the first complete census data to be fully digitized and made publicly available, to calculate a neighbor-based measure of segregation ([Ruggles et al., 2009a](#)). The Logan-Parman measure of segregation provides an intuitive approach to measuring residential segregation based on the notion that the location of households in adjacent units can be used to measure the degree of integration or segregation in a community. It exploits the fact that enumerators went door to door when filling out the census manuscript pages. As a result, next-door neighbors appear next to each other on the census manuscript page. Consequently, by looking at the sequence of households on the census manuscript page and the races given in the census for each household head, you can identify all black households living next to white neighbors.

The measure is calculated as

$$\eta = \frac{E(\bar{x}_{b,w}) - x_{b,w}}{E(\bar{x}_{b,w}) - E(\underline{x}_{b,w})}$$

where $x_{b,w}$ is the actual number of black households with a white

next-door neighbor, $E(\bar{x}_{b,w})$ is the expected number of black households with white neighbors under complete integration (household location is independent of race), and $E(\underline{x}_{b,w})$ is the expected number of black households with white neighbors under complete segregation (only the black households on either end of the black neighborhood have white neighbors). This index equals zero for a fully integrated community, increases as black households become more segregated, and equals one in the case of a completely segregated community.

When calculating the measure, only households appearing on the same census manuscript page are considered possible neighbors. This ensures that the households truly are in close physical proximity. The measure is calculated using only household heads. Therefore households in multiunit dwellings are included in the calculation based on their order on the manuscript page but individuals living as lodgers or residents in an institution are not. Full details on the patterns of census enumeration and the calculation of the measure can be found in [Logan and Parman \(2017\)](#).

Currently, the complete count data necessary for calculating the measure is only publicly available for the 1880 and 1940 federal censuses through IPUMS. [Logan and Parman \(2017\)](#) calculate county-level segregation estimates for both of these census years. Our mortality data from the 1900 and 1910 federal censuses capture children dying in the late 1800s and early 1900s. A sizeable number of the deaths in the North Carolina data occur before 1940, and nearly all individuals in the North Carolina death records grew up in the decades prior to 1940. Consequently, we focus on the 1880 segregation estimates from Logan and Parman.

4.3. Empirical strategy

We use the two distinct mortality data sources to uncover the historical correlations between segregation and health outcomes by race. Our goal is to demonstrate that significant variation in both segregation and mortality outcomes existed prior to the Great Migration and that the newly developed Logan-Parman measure of segregation is a useful predictor of racial health gaps in this era both in urban and rural areas. Thus we take the straightforward approach of documenting the correlation between racial residential segregation and health outcomes using basic linear regressions. Our specifications are highly constrained by the available variables in each data source.

For the North Carolina death certificates, our basic specification to test for a relationship between segregation and adult mortality is

$$\text{Age}_{i,s} = \beta_0 + \beta_1 \text{Black}_{i,s} + \beta_2 \eta_s + \beta_3 \eta_s \text{Black}_{i,s} + \beta_4 \theta_s + \beta_5 \theta_s \text{Black}_{i,s} + \beta_6 B_{i,s} + \beta_7 B_{i,s}^2 + \varepsilon_{i,s}$$

where $\text{Age}_{i,s}$ is the age at death for individual i born in county s , η_s is the Logan-Parman measure of segregation for the individual's county of birth, θ_s is the black population share for the individual's county of birth and $B_{i,s}$ is the individual's birth year. $\text{Black}_{i,s}$ is an indicator variable for the race of the individual equal to one if the individual is black and zero if the individual is white. Interacting this indicator with η_s allows the effect of segregation on mortality to differ by race. The goal is to see if β_2 and β_3 are statistically significant, that is whether there is a significant relationship between segregation and mortality overall and whether that relationship differs by race. We include black population share given that Logan and Parman document a positive correlation between black population share and segregation. Controlling for black population share ensures that any estimated effect of segregation is driven by segregation and not black population share. We include the quadratic in birth year to control for secular changes in

longevity over time. Unfortunately, no other relevant information from the death certificates is transcribed to provide for additional controls.

Regressions are run separately for urban and rural counties of birth to account for segregation having differential impacts by population density as discussed in the conceptual model section. We run the regressions separately by gender given that average age at death differs by gender and given the existing literature on sex differences in sensitivity to in utero exposure to environmental insults. Finally, we run the regressions both for all deaths and excluding child deaths (deaths under the age of 5) to focus on the cumulative impacts of segregation on health.

For the census data, we employ two different specifications similar in spirit to the one above to test for a relationship between segregation and infant and child mortality. First, we consider whether segregation can predict the likelihood of losing a child. We estimate the following linear probability model

$$\text{Lostchild}_{i,s} = \beta_0 + \beta_1 \text{Black}_{i,s} + \beta_2 \eta_s + \beta_3 \eta_s \text{Black}_{i,s} + \beta_4 \theta_s + \beta_5 \theta_s \text{Black}_{i,s} + \beta_6 A_{i,s} + \beta_7 A_{i,s}^2 + \Gamma X_{i,s} + \varepsilon_{i,s}$$

$\text{Lostchild}_{i,s}$ is an indicator variable for having lost a child, equaling one if the individual has lost a child. We also estimate a second version of the equation above with the same explanatory variables but with number of children lost as the dependent variable. As above, η_s is the estimate of county-level segregation and θ_s is the county-level black population share. Given that only county of residence is observed in the census, segregation and black population share correspond to county of residence at the time of the census. $A_{i,s}$ is the age of the individual and $X_{i,s}$ is a vector of individual characteristics including whether the individual is literate and the individual's occupational income score which is the median earnings for the individual's occupation in 1950. Individual income is not available in the 1900 or 1910 federal census. These are the only individual characteristics available in the 1900 and 1910 federal censuses that allow us to control for socioeconomic status likely correlated with race, segregation and mortality outcomes. Unfortunately, richer controls like measures of concentrated disadvantage are not available in this historical setting. Regressions all include state-year fixed effects.

We caution that this empirical strategy is not establishing a causal link between segregation and mortality by race. Our task here is descriptive, to document the correlation of segregation as opposed to a causal model to establish avenues for future research into historical health gaps.

5. Results

There were large racial inequities in adult mortality in North Carolina during the 1909–1975 period, with white men and women living roughly ten years longer than black men and women (Table 1). While the vast majority of both blacks and whites were from rural locations, blacks, on average, lived in slightly more segregated counties than whites, but the difference is not statistically significant nor large (Table 1). Black men and women tended to be born in counties with greater proportions of black residents compared to white men and women.

As with adult mortality in North Carolina, there were large racial inequities in child mortality in the southern US. Specifically, in 1900–1910, nearly 47% of black women experienced a child death while 39% of white women lost a child (Table 2). As in North Carolina, the vast majority of both black and white women in the entire southern US lived in rural counties during this time and lived in counties with similar levels of segregation (Table 2). Black women

tended to live in counties where about 50% of the residents were black and white women tended to live in counties where about 25% of its residents were black. This suggests that black and white residents lived, on average, with the same degree of racial separation even in areas that had different racial proportions.

Our goal is to show how the effect of race on mortality differs with the inclusion of the segregation measure. As suggested by the descriptive characteristics outlined in Table 1, black men and women experienced earlier age at death compared to white men and women, and these inequities were greater in rural areas (Table 3, left half, top panel). This pattern was the same after accounting for the general increased risk of mortality during childhood (Table 3, right half, top panel). Including controls for segregation and black population share but not allowing their impacts to differ by race does not change these patterns in any substantive way.

Segregation does, however, appear to be differentially related to age at death for black and white adults living in urban and rural counties. For white men and women living in urban areas, there is a suggestive positive relation between segregation and age at death, but the standard errors are relatively large. On the other hand, in rural areas, segregation is related to an earlier age at death for white adults (Table 3, bottom panel, “segregation” coefficient). For black men and women living in urban areas, there is a very strong positive association between segregation and age at death. This suggests that, black adults living in highly segregated urban areas live longer than their counterparts who live in less-segregated urban areas (Table 3, bottom panel, sum of interaction and “segregation” coefficients).

Note, however, that at the mean level of segregation and black population share, black men and women die substantially earlier compared to white men and women. As a better comparison of the isolated effect of segregation, we explicitly control for these factors and allow only segregation to vary. Specifically, we calculated the age at death at the mean of segregation and percent black and also with segregation one standard deviation above and below the mean value. We perform these calculations separately for rural and urban residents by race.

The results in Fig. 1 show the predicted ages at death for urban individuals. They display little variation in white ages at death with segregation but large differences in black longevity. For urban black females, for example, the variation in longevity for a change in segregation from one standard deviation below the mean to above the mean of segregation increases longevity by nearly five years. For black men, the result is even more pronounced and is nearly ten years. These results suggest that racial inequities in adult mortality are particularly high in integrated urban areas and substantially lower in segregated urban areas. Overall, decreasing the level of segregation by one standard deviation *increases* the black-white longevity gap to 6 years for males and 5.2 years for females.

The patterns are different in rural areas. For white adults, segregation is related to a lower age at death (Table 3, bottom panel, “Segregation” coefficient). However, for black adults, segregation is related to a greater age at death (Table 3, bottom panel, sum of interaction and “Segregation” coefficients). At the mean level of segregation and black population share, black men die about 14 years earlier than white men and black women die about 10 years earlier than white women (Table 3, sum of the coefficients for “black” and the interactions of “black” with segregation and percent black multiplied by the means of segregation and percent black). Decreasing the level of segregation by one standard deviation in rural areas increases the black-white gap in longevity to 19.4 years for men and 13.5 years for women. Fig. 2 displays the results graphically. The pattern is similar to that of the urban results in Fig. 1, where segregation has little influence on white longevity but

Table 1
Summary statistics for North Carolina death certificates, 1909–1975.

	Black males	White males	Black females	White females
Age at death	48.36 [25.01]	57.85 [23.14]	47.90 [24.87]	60.22 [24.39]
Year of birth	1895.00 [25.70]	1888.29 [24.20]	1894.44 [24.55]	1883.99 [24.10]
Segregation in county of birth	0.292 [0.079]	0.257 [0.092]	0.293 [0.080]	0.257 [0.092]
Percent black in county of birth	0.435 [0.156]	0.302 [0.173]	0.435 [0.156]	0.300 [0.173]
Born in a rural county (1 = yes)	0.983 [0.128]	0.990 [0.101]	0.984 [0.126]	0.990 [0.101]
Observations	82,164	233,333	79,302	179,811

Notes: Means are reported with standard deviations given in brackets. Samples are restricted to death certificates containing complete information for birth year and county of birth. Segregation and percent black are measured in 1880.

Table 2
Summary statistics for 1900 and 1910 federal census data for the southern states.

	Black females	White females
Lost a child (1 = yes)	0.466 [0.499]	0.390 [0.488]
Number of children lost	1.215 [1.928]	0.760 [1.305]
Age	32.724 [9.932]	34.109 [10.145]
Literate (1 = yes)	0.519 [0.500]	0.901 [0.299]
Segregation in county of residence	0.382 [0.117]	0.312 [0.153]
Percent black in county of residence	0.496 [0.213]	0.252 [0.202]
Living in a rural county (1 = yes)	0.783 [0.412]	0.787 [0.409]
Observations	84,945	163,067

Notes: Means are reported with standard deviations given in brackets. Samples are restricted to females under the age of 55 who have had at least one child. Segregation and percent black are measured in 1880.

is positively related to black longevity.

In Table 4 we use 1900 and 1910 census data described earlier to construct measures of child mortality over the 1900 to 1910 period to consider the effects of segregation on child health. These census data allow us to estimate effects of segregation on mortality for a much larger geographic area and far more urban locations than the North Carolina data, which was primarily rural (only three North Carolina counties in 1880 had at least 20 percent of households living in urban areas).

We restrict our attention to women in Southern states who have had at least one birth at the time of census enumeration. This restriction allows us to concentrate on the region of the country with the largest share of African Americans at the time. From the census records we derive the number of children who have died for each mother and whether or not a woman has lost a child. Infant mortality was acute at this time; over the 1900 and 1910 period nearly 47% of black women and 39% of white women experienced a child death in the southern US (see Table 2). Given the differences in disease environments and health at this time, we divide the sample

Table 3
Impacts of segregation on age at death for North Carolina, 1909–1975.

Dependent variable:	Age at death				Age at death conditional on living to age of 5			
	Urban males	Rural males	Urban females	Rural females	Urban males	Rural males	Urban females	Rural females
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black coefficient with no segregation controls	−1.599*** [0.445]	−3.374*** [0.0569]	−2.681*** [0.511]	−4.160*** [0.0647]	−1.834*** [0.431]	−3.534*** [0.0545]	−3.240*** [0.496]	−4.812*** [0.0639]
Black (1 = yes)	−25.90*** [6.752]	−5.056*** [0.225]	−17.65*** [5.815]	−5.804*** [0.248]	−23.09*** [6.496]	−5.273*** [0.216]	−24.18*** [5.607]	−6.321*** [0.245]
Segregation	6.011 [7.532]	−1.115*** [0.411]	3.604 [8.396]	−1.995*** [0.501]	3.701 [7.224]	−2.048*** [0.388]	−8.073 [8.311]	−2.562*** [0.488]
Segregation x black	58.89*** [19.57]	5.990*** [0.841]	35.67** [17.83]	9.974*** [0.935]	50.02*** [18.83]	5.001*** [0.804]	53.47*** [17.15]	8.793*** [0.920]
Percent black	−3.150 [3.312]	0.594*** [0.212]	−4.189 [3.626]	2.348*** [0.259]	−1.347 [3.164]	0.390* [0.200]	0.515 [3.504]	2.334*** [0.252]
Percent black x black	−11.21 [7.634]	−0.202 [0.406]	−6.189 [7.787]	−3.439*** [0.459]	−8.628 [7.294]	0.730* [0.390]	−11.72 [7.426]	−2.909*** [0.453]
Birth year	17.38*** [0.968]	17.22*** [0.111]	17.15*** [1.069]	16.02*** [0.128]	17.32*** [1.084]	16.74*** [0.122]	19.04*** [1.261]	16.80*** [0.153]
Birth year squared	−0.00478*** [0.000256]	−0.00475*** [2.94e-05]	−0.00473*** [0.000282]	−0.00444*** [3.38e-05]	−0.00475*** [0.000287]	−0.00461*** [3.24e-05]	−0.00522*** [0.000334]	−0.00464*** [4.05e-05]
Constant	−15,714*** [916.3]	−15,526*** [105.5]	−15,451*** [1011]	−14,339*** [120.8]	−15,696*** [1024]	−15,129*** [115.6]	−17,262*** [1190]	−15,134*** [143.9]
Observations	3775	311,722	3143	255,970	3613	291,624	2992	239,233
R-squared	0.650	0.671	0.656	0.656	0.593	0.604	0.591	0.569

Notes: OLS estimates with standard errors given in brackets. Percent black and segregation correspond to the county of birth and are measured in 1880. The regression samples include all North Carolina death certificates that include birth year and birth county. ***p < 0.01, **p < 0.05, *p < 0.1.

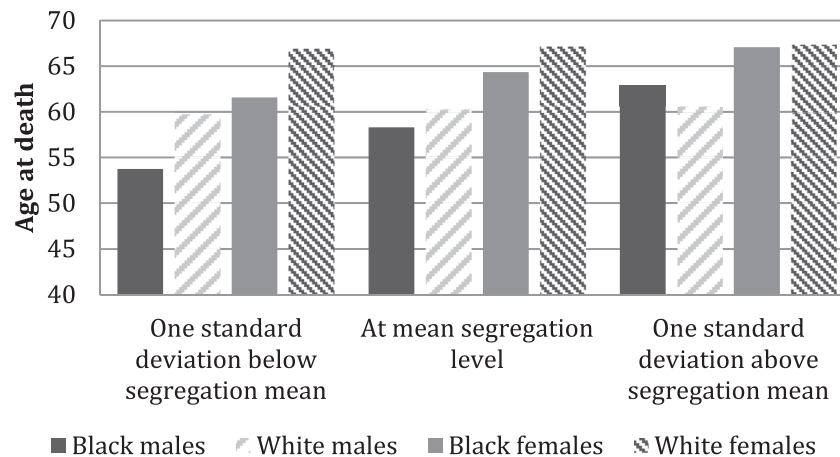


Fig. 1. Predicted age at death for urban individuals in North Carolina at differing levels of segregation. Predicted ages are based on regression coefficients in columns 1 and 3 of Table 3 with percent black and birth year evaluated at their means.

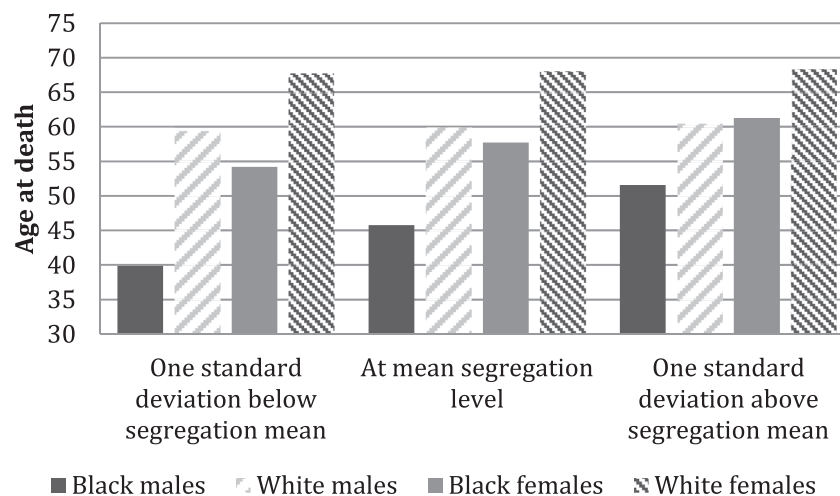


Fig. 2. Predicted age at death for rural individuals in North Carolina at differing levels of segregation. Predicted ages are based on regression coefficients in columns 2 and 4 of Table 3 with percent black and birth year evaluated at their means.

by rural and urban areas and estimate the relationship for rural and urban counties separately, similar to the procedure in the North Carolina data.

The racial inequality in child mortality in the southern states is shown in multivariate models in Table 4. When segregation is included in the model, the racial inequality is reduced by about half in urban counties (and to a much lesser degree in rural counties). We consider the impacts of segregation on the extensive margin of child and infant mortality in Columns 1 and 2 of Table 4. The results for rural and urban mothers show that there is not a strong relationship between segregation and the overall propensity to have experienced an infant death. The comparison, however, reveals that the inclusion of segregation greatly reduces the racial effect (the first row of Table 4). As before, the racial effect in the lower panel is the effect of race and the interacted effect of race with segregation evaluated at the mean. More interesting, the racial effect lowers for both rural mothers and urban mothers, which suggest that segregation may play a large role in infant mortality in both areas. At the mean level of segregation, the racial effect is reduced by more than 35% in urban areas and by more than 30% in rural areas.

Turning to the number of children lost (Columns 3 and 4), we see that segregation was related to the number of children lost in rural areas, but that this effect did not vary by race. As with the

urban results, however, the inclusion of segregation lowers the racial health disparity in the number of children lost. For rural and urban women, the racial effect is reduced more than 25%. These results suggest that a relatively large amount of the racial health disparity may be driven by segregation. They also imply that rural segregation may play a particularly large role in racial health disparities historically.

As with the North Carolina results, to aid in interpretation of the regression results we isolate the effect of segregation by calculating the predicted number of children lost at the mean of percent black and at the mean of segregation, one standard deviation below the mean of segregation, and one standard deviation above the mean of segregation. Fig. 3 displays the predicted number of children lost for both rural and urban mothers. Segregation appears to be similarly related to child mortality in urban and rural counties. For those in urban areas, there is a suggestive positive relation between segregation and child mortality, but it is very slight. For example, the predicted number of children dying is positively related to segregation for both black and white mothers, but the relationship is stronger for black mothers than white mothers. For example, the predicted number of additional children dying for a one standard deviation increase in segregation would be 0.03 for black mothers and 0.01 for white mothers. In rural areas, segregation is also

Table 4
Impacts of segregation on number of children lost in the 1900 and 1910 federal census data.

Dependent variable:	Lost a child (1 = yes)		Number of children lost	
	Urban females	Rural females	Urban females	Rural females
	(1)	(2)	(3)	(4)
Black coefficient with no segregation controls	0.0830*** [0.00688]	0.0746*** [0.00620]	0.400*** [0.0286]	0.419*** [0.0285]
Black (1 = yes)	0.0481* [0.0246]	0.0674*** [0.0169]	0.250** [0.114]	0.331*** [0.0727]
Segregation	−0.00187 [0.0427]	0.0356 [0.0260]	0.0728 [0.226]	0.198** [0.0914]
Segregation x black	0.00807 [0.0488]	−0.0423 [0.0435]	0.132 [0.222]	−0.0629 [0.138]
Percent black	−0.0674** [0.0238]	−0.0479** [0.0179]	−0.205** [0.0727]	−0.0874 [0.0571]
Percent black x black	0.0880*** [0.0248]	0.0573*** [0.0174]	0.259** [0.103]	0.236*** [0.0584]
Age	0.0291*** [0.00141]	0.0408*** [0.00188]	0.0407*** [0.00541]	0.0559*** [0.00399]
Age squared	−0.000207*** [2.18e-05]	−0.000360*** [2.49e-05]	0.000139* [7.21e-05]	−5.20e-05 [4.34e-05]
Occupational income score	−0.00108*** [0.000339]	−0.00163*** [0.000243]	−0.00364** [0.00149]	−0.00340*** [0.000890]
Literate (1 = yes)	−0.0747*** [0.00926]	−0.0503*** [0.00589]	−0.444*** [0.0466]	−0.260*** [0.0210]
Constant	−0.298*** [0.0373]	−0.479*** [0.0386]	−0.424* [0.200]	−0.857*** [0.112]
Observations	53,132	194,844	53,132	194,844
R-squared	0.095	0.118	0.125	0.146

Notes: OLS estimates with standard errors given in brackets. Regression sample is restricted to woman in the South (KY, MD, DE, VA, NC, SC, WV, GA, FL, AL, MS, TN, AR, OK, TX, LA) under the age of 55 who have had at least one child. Segregation and percent black correspond to the county of residence and are measured in 1880. All regressions include state-year fixed effects. ***p < 0.01, **p < 0.05, *p < 0.1.

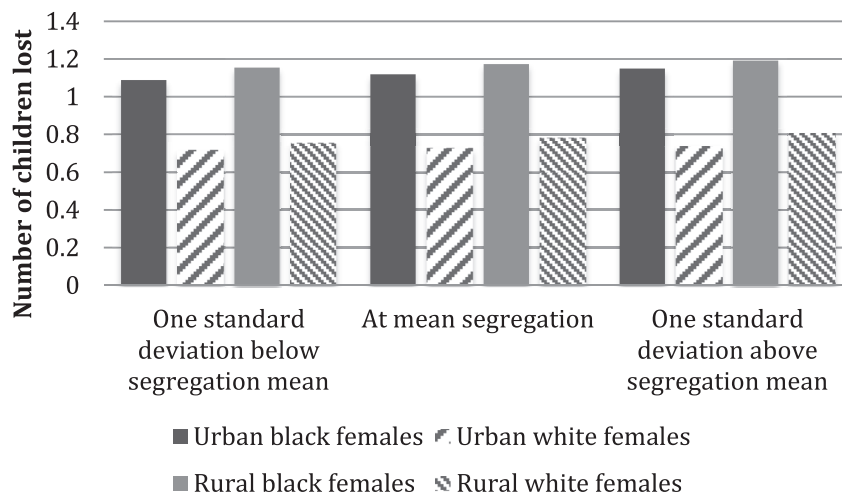


Fig. 3. Predicted number of children lost at differing levels of segregation for women under the age of 55 in the South having at least one child. Predicted number is based on regression coefficients in columns 3 and 4 of Table 4 with all variables other than segregation and race evaluated at their means.

positively related to child mortality for both black and white mothers. For example, the predicted change in the number of children dying with a one standard deviation increase in segregation would be 0.018 for black mothers and 0.026 for white mothers (Fig. 3). For both those in rural and urban areas, the racial differences in child mortality on the extensive margin of experiencing a child death vary insignificantly with segregation in both rural and urban areas.

Fig. 4 displays the effects for the predicted likelihood of losing a child at various segregation levels. As with Fig. 3, we calculated the probability at the mean of percent black and the mean of segregation, one standard deviation above the mean of segregation, and

one standard deviation below the mean of segregation. The figure shows that there is no strong relationship between segregation and the likelihood of losing a child for rural and urban women of either race. The changes in the racial difference in the likelihood of experiencing a child death with changes in segregation are statistically and substantively insignificant.

6. Discussion

This paper exploits a new measure of segregation that is tied to the precise location of households. The measure allows us to estimate the effect of segregation in rural areas on mortality and on the

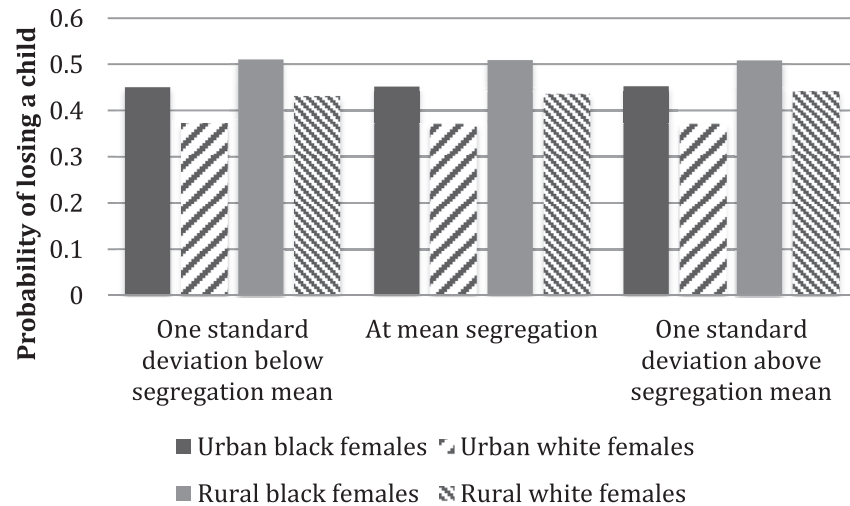


Fig. 4. Predicted probability of losing a child at differing levels of segregation for women under the age of 55 in the South having at least one child. Predicted number is based on regression coefficients in columns 1 and 2 of Table 4 with all variables other than segregation and race evaluated at their means.

effects of this structural factor on historical health inequalities. The results show that both infant and adult mortality were related to the pattern of racial segregation. Further, the results show that segregation in rural areas was an omitted factor in racial health disparities. The results here suggest that the effects of segregation apply not only to the present, but also to the past. Most important, they show that segregation is not only related to outcomes in urban areas, but also related to health outcomes in rural areas and that segregation was influencing racial health gaps prior to the urbanization of the black population during the Great Migration.

The results of Table 3 highlight the fact that segregation is an omitted factor in explaining historical racial health disparities, and its exclusion may in fact lead to underestimates or overestimates of racial health disparities. Even more, segregation is particularly important in explaining variation in longevity by race in rural areas. Surprisingly, the coefficient for the interaction term between segregation and the black indicator variable is consistently positive and statistically significant across all specifications, suggesting that while the impact of segregation on lifespan common to black and white individuals may have been negative, the additional impact specific to black individuals was positive. This suggests that, contrary to our initial hypothesis, segregation had potentially protective effects historically. Overall, the results from Table 3 and Figs. 1 and 2 suggests that segregation has heterogeneous effects that differ by race, sex, and rural/urban location, but that overall increased segregation was related to lower racial disparities in longevity. However, these effects were largely related to adult mortality: the variation in both number of children lost and the likelihood of experiencing a child death varied little with segregation, as displayed in Figs. 3 and 4.

Returning to the hypotheses discussed earlier, we see that segregation is related to adult mortality. Overall, the racial health disparity is partially due to segregation, but in other instances it obscures the extent of racial health disparities. The estimates vary, but across both genders and urban and rural locations a non-negligible portion of the racial effect is correlated with segregation. We hypothesized that segregation would have a negative effect on black health and a positive effect on white health. However, we found that segregation was related to increased longevity for blacks across all specifications while the direction of the effect on white longevity varied. The second hypothesis, that segregation would have a larger effect in urban areas, was not consistently

supported by the data. In the results for infant mortality the effect of segregation was in fact larger in rural areas. The pattern in North Carolina was similar—the effect of segregation was neither confined to nor consistently larger in urban areas.

The different results for adult and child mortality suggest two related effects of segregation. First, segregation may be a factor that plays out over the life course more than an effect that can be estimated immediately, as with an environmental effect. Second, segregation may have had protective effects by reducing the exposure to disease prevalence across races. In a period with a lack of access to quality health care particularly for black individuals, less exposure across races could have had disproportionately large and positive effects on black health.

The results presented here call for renewed research on the effects of segregation on a variety of health outcomes that have intergenerational effects. Doing so would extend the analysis and address the limitations of the existing analysis. For example, the impact of migration on health and human capital must account for the segregation of the sending location, which was not included here. If rural segregation was related to health, as we have shown here, then the evolution of health disparities and racial differences in mortality may be related to a new set of factors that played out over the twentieth century and which require new methods to investigate. At a minimum, these results call for renewed attention to the historical factors which give rise to segregation and the heretofore neglected effect that rural segregation and historical segregation has had on health outcomes and the evolution of health disparities. The complexity of the relationships considered here, both for overall health and racial differences as a function of segregation, are new avenues that could deepen our understanding of segregation and its changing impact on racial disparities in health across space and time. The overall thrust of this exploratory work, however, is to show that the effects of segregation were indeed present in the past and for those in rural areas.

This work sets the stage for future research to explore the effects of early-life segregation on later life health. First, with the rapidly increasing availability of large data sources containing information on childhood location due to the push to digitize historical genealogical records, the role of early life segregation on racial health differences can now be estimated. This is made possible not just because of the types of digitized death and census records we use here but also because of the introduction of a

segregation measure that can apply to rural communities. Second, this work will allow researchers to begin to conceptualize the effects of changes in segregation on health. For example, if segregation impacts health through stress, changes in segregation could intensify or lessen these effects as an individual migrates to different areas and/or experiences changes in their home environment. Additional work can exploit the spatial correlation with segregation to investigate whether rural areas outside of highly segregated cities exhibit similar levels of segregation and if they have similar health disparities. Third, within-race differences in the effect of segregation can be explored. For example, among African Americans in rural or urban areas the differences in the effects of segregation may explain a portion of rural/urban differences in health outcomes. Fourth, specific mechanisms with cause of death can be explored. If segregation is more likely to have an effect on certain conditions than others, cause of death research could help to estimate and test the strength of these relationships. These and other questions about the effects of structural factors such as segregation can be asked and answered now that a comprehensive measure of segregation and large historical datasets with individual-level health outcomes are available to researchers.

Acknowledgements

We thank William D. Biscarri, Nicholas J. Deis, Jackson L. Frazier, Adaeze Okoli, Terry L. Pack, Stephen Prifti and Colin Weinshenker for providing excellent research assistance. The usual disclaimer applies.

References

- Acevedo-Garcia, Dolores, 2000. Residential segregation and the epidemiology of infectious diseases. *Soc. Sci. Med.* 51 (8), 1143–1161.
- Ancestry.com, 2007. North Carolina, Death Certificates, 1909–1976 [database online]. Ancestry.com Operations Inc, Provo, UT, USA.
- Anderson, Laurie M., et al., 2003. Providing affordable family housing and reducing residential segregation by income: a systematic review. *Am. J. Prev. Med.* 24 (3), 47–67.
- Corral, Irma, et al., 2012. Residential segregation, health behavior and overweight/obesity among a national sample of african american adults. *J. Health Psychol.* 17 (3), 371–378.
- Crimmins, E.M., Condran, G.A., 1983. Mortality variation in U.S. Cities in 1900: a two-level explanation by cause of death and underlying factors. *Soc. Sci. Hist.* 7, 31–60. JSTOR.
- Cutler, David M., Edward, L. Glaeser, Vigdor, Jacob L., 1999. "The rise and decline of the American ghetto. *J. Political Econ.* 107 (3), 455–506. JSTOR.
- Diez Roux, Ana V., Mair, Christina, 2010. Neighborhoods and health. *Ann. N. Y. Acad. Sci.* 1186 (1), 125–145.
- Duru, O.K., Harawa, N.T., Kermah, D., Norris, K.C., 2012. Allostatic load burden and racial disparities in mortality. *J. Natl. Med. Assoc.* 104 (1–2), 89–95.
- Feagin, Joe R., 2014. *Racist America: Roots, Current Realities, and Future Reparations*. Routledge.
- Geronimus, Arline T., et al., 2006. "Weathering" and age patterns of allostatic load scores among blacks and whites in the United States. *Am. J. Public Health* 96 (5), 826–833.
- Howell-Moroney, Michael, 2005. The geography of opportunity and unemployment: an integrated model of residential segregation and spatial mismatch. *J. Urban Aff.* 27 (4), 353–377.
- Kramer, Michael R., Hogue, Carol R., 2009. Is segregation bad for your health? *Epidemiol. Rev.* 31 (1), 178–194.
- Lichter, Daniel T., et al., 2007. National estimates of racial segregation in rural and small-town America. *Demography* 44 (3), 563–581.
- Logan, Trevon D., 2009. Health, human capital, and African-American migration before 1910. *Explor. Econ. Hist.* 46 (2), 169–185.
- Logan, Trevon D., Parman, John M., 2017. "The national rise in residential segregation. *J. Econ. Hist.* 77 (1), 127–170.
- Long, Jason, Ferrie, Joseph, 2013. Intergenerational occupational mobility in Great Britain and the United States since 1850. *Am. Econ. Rev.* 103 (4), 1109–1137.
- Logan, John R., Brian, J. Stults, Farley, Reynolds, 2004. Segregation of minorities in the metropolis: two decades of change. *Demography* 41 (1), 1–22.
- Massey, D.S., 2004. Segregation and stratification: a biosocial perspective. *Du. Bois Rev.* 1 (01), 7–25.
- Massey, Douglas S., Denton, Nancy A., 1988. The dimensions of residential segregation. *Soc. Forces* 67 (2), 281–315.
- Massey, Douglas S., Denton, Nancy A., 1993. *American Apartheid: Segregation and the Making of the Underclass*. Harvard University Press.
- Mohai, Paul, Saha, Robin, 2007. Racial inequality in the distribution of hazardous waste: a national-level reassessment. *Soc. Probl.* 54 (3), 343–370.
- Morello-Frosch, Rachel, Jesdale, Bill M., 2006. Separate and unequal: residential segregation and estimated cancer risks associated with ambient air toxics in US metropolitan areas. *Environ. Health Perspect.* 386–393.
- Morland, Kimberly, Filomena, Susan, 2007. Disparities in the availability of fruits and vegetables between racially segregated urban neighbourhoods. *Public Health Nutr.* 10 (12), 1481–1489.
- Naidu, Suresh, 2012. Suffrage, Schooling, and Sorting in the Post-Bellum U.S. South. Working Paper No. w18129. National Bureau of Economic Research.
- O'Flaherty, Brendan, Sethi, Rajiv, 2007. Crime and segregation. *J. Econ. Behav. Organ.* 64 (3), 391–405.
- Orfield, Gary, Eaton, Susan E., 1997. *Dismantling Desegregation: the Quiet Reversal of Brown v. Board of Education*. The New Press, 1997.
- Peterson, Ruth D., Krivo, Lauren J., 1993. Racial segregation and black urban homicide. *Soc. Forces* 71 (4), 1001–1026.
- Phelan, Jo C., Link, Bruce G., 2015. Is racism a fundamental cause of inequalities in health? *Annu. Rev. Sociol.* 41, 311–330.
- Powell, Lisa M., Frank, J. Chaloupka, Bao, Yanjun, 2007. The availability of fast-food and full-service restaurants in the United States: associations with neighborhood characteristics. *Am. J. Prev. Med.* 33 (4), S240–S245.
- Ruggles, S., Sobek, M., Alexander, T., Fitch, C., Goeken, R., Hall, P., King, M., Ronnander, C., 2009a. Integrated Public Use Microdata Series Sample of the 1880 Federal Census. Minnesota Population Center. Accessed through usa.ipums.org/usa/.
- Ruggles, S., Sobek, M., Alexander, T., Fitch, C., Goeken, R., Hall, P., King, M., Ronnander, C., 2009b. Integrated Public Use Microdata Series Sample of the 1900 Federal Census. Minnesota Population Center. Accessed through usa.ipums.org/usa/.
- Ruggles, S., Sobek, M., Alexander, T., Fitch, C., Goeken, R., Hall, P., King, M., Ronnander, C., 2009c. Integrated Public Use Microdata Series Sample of the 1910 Federal Census. Minnesota Population Center. Accessed through usa.ipums.org/usa/.
- Schulz, Amy, et al., 2001. Social context, stressors, and disparities in Women's health. *J. Am. Med. Women's Assoc.* 56 (4), 143–150.
- Shihadeh, Edward S., Flynn, Nicole, 1996. Segregation and crime: the effect of black social isolation on the rates of black urban violence. *Soc. Forces* 1325–1352.
- Smedley, B.D., Stith, A.Y., Nelson, A.R., 2003. *Unequal Treatment: what Healthcare Providers Need to Know about Racial and Ethnic Disparities in Healthcare*. National Academy Press, Washington, DC.
- Steckel, Richard H., 2013. The hidden cost of moving up: type 2 diabetes and the escape from persistent poverty in the American South. *Am. J. Hum. Biol.* 25 (4), 508–515.
- Steckel, Richard H., Senney, Garrett, 2015. Historical Origins of a Major Killer: Cardiovascular Disease in the American South. Working Paper No. w21809. National Bureau of Economic Research.
- Su, D., 2009. Occupational career and risk of mortality among US Civil War veterans. *Soc. Sci. Med.* 69, 460–468.
- Troesken, Werner, 2002. The limits of Jim Crow: race and the provision of water and sewerage services in American cities, 1880–1925. *J. Econ. Hist.* 62 (3), 734–772.
- Williams, David R., Collins, Chiquita, 2001. Racial residential segregation: a fundamental cause of racial disparities in health. *Public Health Rep.* 116 (5), Association of Schools of Public Health: 404.
- Wilson, William Julius, 2012. *The Truly Disadvantaged: the Inner City, the Underclass, and Public Policy*. University of Chicago Press.
- Wilson-Frederick, Shondelle M., et al., 2014. Examination of race disparities in physical inactivity among adults of similar social context. *Ethn. Dis.* 24 (3), 363.
- Zenk, Shannon N., et al., 2005. Fruit and vegetable access differs by community racial composition and socioeconomic position in Detroit, Michigan. *Ethn. Dis.* 16 (1), 275–280.