Segregation and Southern Lynching

Lisa D. Cook[*], Trevon D. Logan[†]and John M. Parman[‡§]

February 19, 2016

Abstract

The empirical relationship between racial segregation and racial violence is unknown. We show that the existing economic, social, and political theories of lynching implicitly contain hypotheses about the relationship between racial segregation and racial violence, which we further note is consistent with general theories of social conflict. Since Southern lynching occurred in rural and urban areas, traditional measures of racial segregation cannot be used to estimate the relationship. Earlier analysis has analyzed the relationship between lynching and racial proportions, a poor proxy for racial segregation. We use a newly developed measure of residential segregation based on individual-level data (Logan and Parman 2015), which exploits complete census manuscript files to derive a measure of segregation based upon the racial similarity of next door neighbors. This new measure distinguishes between the effects of increasing racial homogeneity of a location and the tendency to segregate within a location given a particular racial composition. Using this comprehensive measure of racial residential segregation for every Southern county in the United States, we estimate the relationship between racial segregation and lynching. We find that conditional on racial composition, racially segregated environments were much more likely to experience lynchings and to have more lynchings. In general, a one standard deviation increase in segregation in 1880 resulted in one additional lynching in a county from 1882 to 1930. The result is robust to numerous controls, functional form assumptions, the inclusion of traditional segregation measures, proxies for racial inequality, and a host of potential confounders such as antebellum antecedents of segregation and black political involvement during Reconstruction. Consistent with the hypothesis that segregation is related to interracial violence, we find that segregation is highly correlated with African American lynching, but uncorrelated with white lynching. We conclude by describing how our results call for reformulating theories of Southern lynching to focus on social interactions and interracial proximity.

JEL classifications: I1, J1, N3 Keywords: Segregation, Lynching, Violence, Crime
PRELIMINARY DRAFT

Do Note Cite, Quote, or Circulate Without Permission

[*]Department of Economics, Michigan State University, 110 Marshall-Adams Hall, East Lansing, MI 48824 email: lisacook@msu.edu

[†]Corresponding Author: Department of Economics, The Ohio State University and NBER, 1945 N. High Street, 410 Arps Hall, Columbus, OH 43210 e-mail: logan.155@osu.edu

[‡]Department of Economics, College of William and Mary and NBER, 130 Morton Hall, Williamsburg, VA 23187 e-mail: jmparman@wm.edu

"Our country's national crime is lynching. It is not the creature of an hour, the sudden outburst of uncontrolled fury, or the unspeakable brutality of an insane mob."

- Ida B. Wells, *Lynch Law in America* (1900)

1 Introduction

Ethnic fractionalization is often tied to ethnic conflict that hinders economic development. A growing literature particularly, particularly focused on African nations, has documented both the link between ethnic diversity and civil war (Fearon and Laitin, 2003; Collier and Hoeffler, 2004; Collier and Rohner, 2008) and the subsequent effects of that conflict on growth. In particular, ethnic violence tends to lead to greater distrust which in turn hinders economic performance. These impacts on trust are present in the immediate aftermath of conflict, demonstrated by the work of Rohner et al. (2013) on Uganda's civil conflict and Cassar et al. (2013) behaviorial experiments after the Tajik civil war, but can also persist over several generations. Nunn and Wantchekon (2011) find modern trust levels are lower for individuals whose ancestors were more heavily raided during the slave trade, helping explain the negative relationship between the slave trade and modern economic development (Nunn, 2008). Voigtländer and Voth (2012) find that anti-Semitic attitudes and behaviors from the Black Death in the 1300s persisted into the twentieth century. The memory of ethnic conflict can be quite long with dramatic impacts on trust and economic development.

While violent conflict is a prominent theme in the literature on ethnic fractionalization in developing countries, research on the United States has focused more on the relationships between ethnic diversity, the provision of public goods and the growth of American cities. The relationship between diversity, trust and economic performance within the United States is mixed. Alesina and La Ferrara (2005) demonstrate that increases in ethnic diversity are associated with lower growth rates, consistent with Costa and Kahn (2003) finding that increasing community heterogeneity is associated with declining levels of social capital in the United States. Other studies point to economic benefits of diversity in US cities (Ottaviano and Peri, 2006).

Although these studies focus on the relationship between diversity and peaceful political or economic processes in modern cities, America's past witnessed violent interracial conflict. The most

prominent example is the history of lynchings in the American South. These acts of interracial violence had profound and lasting impacts on the development of the South. As Alesina and La Ferrara (2002) note, America has substantial differences in trust levels by race, with black individuals 24 per cent less likely to trust others than non-black individuals. These low trust levels are most pronounced in the states that high levels of lynchings: Mississippi, Alabama and Arkansas are all among the five states with the lowest levels of trust today Alesina and La Ferrara (2002) and experienced some of the highest levels of lynchings in the early twentieth century.[1] Not only do these states have low trust levels today, they also have among the lowest income per capita levels in the United States. America's violent past suggests that diversity and ethnic tensions in the United States influence American society not just through the modern political and economic processes highlighted by (Alesina and La Ferrara, 2005) but also through the lingering effects of interracial violence more directly related to the literature on ethnic fractionalization and ethnic conflict. Our purpose in this paper is to investigate the links between ethnic diversity and violent conflict within the United States, examining whether residential segregation prevented or contributed to the likelihood of lynchings, whether familiarity bred tolerance or contempt between white and black individuals.

Racial segregation has long been viewed as a factor in explaining racial violence. Lynching is commonly viewed as concurrent with the rise of Jim Crow, the erosion of African American political and economic gains from Reconstruction, and the marker of a period of intense racial hostility. DeFina and Hannon (2011) argue that lynching is linked to contemporary urban segregation measures today. Bailey et al. (2011) argue that lynching victims were socially isolated, and Tolnay and Beck (1992a) argue that racial violence was related to black migration patterns in the early twentieth century. Recent historical scholarship by Loewen (2013), Jaspin (2008), and Kantrowitz (2012) argues that lynching was only one small piece of a larger movement of racial violence in the United States in the late nineteenth century. This movement included the ethnic cleansing of entire counties and the prohibition of African American residence in certain towns (known as "sundown towns" as blacks found to be present after dark would be subject to violence). The full quantitative history of these events is still unknown.

---

[1]Mississippi, Alabama and Arkansas had the first, fourth and sixth highest numbers of lynchings.

The social, political, demographic and economic factors underlying lynching have been a topic of enduring theoretical and empirical interest. Theoretically, scholars have noted that local labor markets, status competition, economic competition and political concerns factored into mob violence in the late nineteenth and early twentieth centuries. The existing theories explicitly hinge on interactions between whites and African Americans over particular spheres of social life. One important missing factor, which would arguably be related to all of the spheres, is racial segregation. The degree to which whites and African Americans perceived themselves to be in competition with one another over status, economic opportunities, or political power would be related to how they or if they interacted with one another. Although intuitively important, it has been difficult to empirically analyze the effect of segregation on lynching. The standard approach in the literature is to use the proportion black in a county, but proportions black tell us little about *segregation*, which could occur with large or small African American populations.[2] The role of racial segregation in lynching has not been empirically estimated.

Further complicating this argument is the fact that the number of lynchings in the United States reached their peak in 1892 (Cook, 2012), far before the rise of urban segregation or significant flows of blacks to urban areas. It is important to note that the effect of segregation on lynching could be seen in locations with large and small African American populations. The social cohesion of a community could be independent of its population shares. Equally important, lynching in rural areas requires a measure of segregation that can be consistently applied in rural and urban communities to estimate the relationship without spatial bias. This necessitates a measure of *segregation* as opposed to the use of racial proportions to estimate the relationship between segregation and lynching.

In this paper, we exploit a newly developed measure of segregation to estimate the relationship between racial residential segregation and lynching in the United States. The new segregation measure comes from Logan and Parman (2015), who use the availability of the complete (100%) manuscript pages for the federal census to identify the races of next-door neighbors. They measure segregation by

---

[2]More generally, the literature on ethnic fractionalization, conflict and economic development has typically employed the fragmentation index which is a function of the population shares of each ethnic group. It measures the probability that two randomly chosen individuals in an area are of different races but cannot speak to the likelihood that two individuals of different groups will interact with one another, something that crucially depends on the extent to which individuals segregate themselves within a community.

comparing the number of household heads in an area living next to neighbors of a different race to the expected number under complete segregation and under no segregation (random assignment). This measure of segregation is inherently tied to racial dispersion. The measure allows for the distinction between the effects of differences in racial composition and the tendency to segregate given a particular racial composition. A particular advantage is that it can be aggregated to any boundary without losing the underlying properties since it is defined at the household level. Furthermore, the measure is equally applicable to both urban and rural areas. This consistent measure of segregation for rural and urban areas allows us to extend the analysis of lynching to include racial segregation. To our knowledge, this is the first study to comprehensively measure the relationship between segregation and lynching in the South.

A key strength for this approach is that the measure of segregation comes from 1880, while the lynching data come from 1882 onward. This overcomes some concerns about possible reverse-causality in the lynching-segregation relationship. Additionally, Logan and Parman (2015) show that segregation in 1880 is highly correlated with segregation in 1940– communities that were more segregated remained so, and therefore the potential for lynching (post 1880) to influence future segregation variation is limited since the persistence was quite general. Similarly, measurement at the county level allows us to include a host of controls that could presumably explain the relationship. Our main specifications explicitly control for the proportion black in the area and traditional measures of racial segregation, and we therefore are able to estimate the relationship between segregation and lynching while controlling for the overall racial composition of the area, county-level segregation, state fixed effects, and other factors.

Our task in this paper is to determine the relationship between segregation and Southern lynching and to see whether the inclusion of segregation alters the relationship between Southern lynching and other covariates. We match the new measure of segregation to the most comprehensive lynching data available and find that segregation was strongly correlated with lynching. Counties that were more segregated were much more likely to experience a lynching. Conditional on having a lynching, more segregated counties were more likely to experience multiple lynchings. Even when controlling for state and regional effects of lynching, the results show the same relationship between segregation

and lynching. We show that the result is surprisingly robust to alternative specifications and various sample restrictions. As an additional check, we show that segregation was strongly correlated with interracial lynching but uncorrelated with intraracial lynching. Since almost all mob violence at this time involved white perpetrators, the lack of a relationship between segregation and white lynchings suggest that segregation influenced interracial violence but had little effect on intraracial violence. We take this as suggestive evidence that segregation's influence worked through racial isolation as opposed to segregation being a predictor of a generally violent community.

We begin by reviewing the existing theoretical and empirical work on Southern lynching and incorporate segregation more fully into the most prominent hypotheses regarding lynching. There, we show that the predicted effects of lynching depend critically on whether the motivation for lynching was social, economic or political. The theories have clear predictions for the relationship between segregation and racial violence. This is in part due to the fact that all theories of lynching hinge on the determinants and effects of interracial contact. As such, segregation plays a key role in all existing lynching theories, but the theories posit different effects. Ultimately, however, the relationship between segregation and lynching is an empirical question.

We then describe the new measure of segregation and detail how it differs from existing measures and allows for the first empirical analysis of the effect of segregation on lynching. In the subsequent section we present the reduced form empirical results which include segregation as an explanatory variable. We then replicate the most well-known studies of lynching, where we show that the inclusion of segregation alters the previous results. We conclude with a discussion of what the findings imply for the role of social isolation on lynching in particular and racial violence more generally.

2   Theories of Southern Lynching

The academic study of lynchings extends back to the contemporaneous efforts of the NAACP, the Chicago *Tribune*, Tuskegee University and others to document the extent of lynchings and the conditions surrounding them at the turn of the century (National Association for the Advancement of Colored People, 1919; Williams, 1968). Based on these initial data sources describing the character-

istics of lynching victims and lynch mobs, theories of lynching as the product of economic conflict emerged viewing lynching as either an aggressive response to economic frustration (Hovland and Sears, 1940) or a deliberate attempt to improve the economic position of whites relative to blacks (Raper, 1933). The seminal work of Blalock (1967) on minority-group relations emphasized an alternative framework for understanding conflict between groups as a response to threats to the majority group's power and resources posed by the minority group. Corzine et al. (1983) provided some of the first tests of Blalock's proposed power-threat hypothesis in the context of lynchings using the lynching data compiled by the NAACP. They found mixed support for the hypothesized positive and convex relationship between minority concentration and lynching.

With a dramatically improved dataset of Southern lynching victims correcting the errors and omissions of the historical data sources, Beck and Tolnay explored additional economic theories of lynching (Beck and Tolnay, 1990, 1992). Consistent with earlier empirical studies, Beck and Tolnay's improved data confirmed the positive relationship between the size of the black population and the incidence of lynchings. However, they were also able to demonstrate that lynchings were more prevalent when cotton prices were falling and inflationary pressure was increasing, adding support to an economic theory of lynchings. These findings raised the possibility that lynchings could have been an attempt to reduce competition from the black workers for white jobs, either through directly eliminating black workers or through intimidating others thinking about competing for jobs held by white workers.

As these various theories of lynchings have evolved, the role of segregation has been noticeably absent from both theoretical discussions and empirical analyses. Both the theory and empirical tests have focused on the relative size of the black population in a community, not the extent to which that black community is integrated with or segregated from the white community. The reasons for this are not that segregation is assumed to be unrelated to interracial conflict. Instead, it is the product of a lack of adequate data on residential segregation patterns. There are compelling reasons to believe that, conditional on a given racial composition for a community, the degree of segregation within that community matters a great deal in explaining the likelihood and potential extent of interracial violence.

Below, we begin with a general framework drawing on the general theories of social conflict to

describe lynching. We then explore the existing economic, social, and political theories of Southern lynching to identify the role that racial segregation would have in lynching.

2.1   Lynching and Social Conflict

The most general framework for analyzing lynching views lynching as the product of social conflict. Indeed, the economic, social, and political theories further refine the relationship between racial segregation and racial violence by highlighting the mechanisms and channels that would lead to lynching. Segregation plays a role in this conflict by spatially restricting the potential for groups to have interaction with each other. Such spatial restrictions could reflect underlying group animosities, whereby more segregated environments would see more interracial violence as more segregated environments are related to poorer race relations. If segregated environments reflect restrictions on African American mobility and social norms regarding race relations, movements outside of those norms could be met with particularly harsh consequences. On the other hand, racial segregation could be related to fewer incidents of racial violence if the geographic separation minimizes the potential for groups to interact in any fashion. Close interaction between racial groups could also be related to racial violence if sustained interracial interaction breeds social conflict such that racial violence is likely to result. In essence, familiarity could breed contempt and further inflame racial hostility and racial violence.

Irrespective of its motivation, racial violence could be related to racial segregation through the simple fact that the *indirect* effects of lynchings were commonly seen as the key force behind the activity. Indeed, under the NAACP definition of lynching, the lynching activity had to have the purpose of upholding tradition or serving justice. This applied, in many cases, not only to the victims of lynching but also to the surrounding community. In a racially segregated environment, lynching may have particularly large externalities for the perpetrators. That is, the public activity of lynching would be more likely to have its desired group effects if the targeted group was more segregated (and therefore in close contact with one another relative to another group). For example, in many lynchings the bodies of victims were publicly displayed for several days after the lynching and the body parts of victims would be sold and distributed to the mob. Such a public display may have a more chilling effect (larger externalities) in areas where the spacial borders of the black community

8

were less porous. Ultimately, however, the relationship between segregation and racial violence hinges on the relationship between social isolation and racial aggression.

2.2 Economic Theories of Lynching

The Theory of Labor Control, proposed by Beck and Tolnay, is tied to the idea of the use of lynching as a form of social control over black workers. The theory is based on the idea that lynchings are tied to the demand for labor. Lynchings were used as a form of social control over the African-American labor force. The Economic Competition model notes that when southern whites became more and more economically disadvantaged as the southern economy stagnated, whites and blacks began to compete for the same jobs and lynchings occurred as a result (Raper, 1933). Two hypotheses exist as to why this occurred, the first being the frustration-aggression model (Hovland and Sears, 1940), the second being lynching used to improve the economic position of whites (Raper, 1933). In the frustration-aggression model, lynching is used as an aggressive response to economic frustration; a vent for labor market competition with blacks. As a way to improve the economic status of whites, it displaces black workers with white workers and/or keeps them in segments where they do not face competition. One might also contend that social control could also be related to the Competition model (Tolnay et al., 1992; Beck and Tolnay, 1992).

Each model shares the idea that whites and African Americans are labor forces and the goal of whites was to control African American labor either by restricting movement (to control the supply of black labor– fewer black workers would result in higher wages) or competition (to discourage African Americans from competing with whites for employment). In either case the theory hinges on the relative sizes of the African American population relative to white. Economic factors such as segmentation of the labor market also play a role as they predict the likelihood of competition between whites and African Americans.

In these economic theories, segregation is not explicitly stated to play a role in lynching itself. As a measure of competition, percent black is normally used along with other economic factors which would be related to competition (the share of the labor force in agriculture, for example). As the proportions of African-Americans grow, the more competition whites would have with them. As

the size of the African American population increases, the desire to control the African American population increases as they are a more numerous factor in the labor market. As early as Raper, the relationship between the share of the community that was black and lynching was noted, and the general relationship has been confirmed in previous empirical work. At very small and large proportions of the African American population lynching is not as likely, but as the population share grows lynching likelihood increases.[3]

The incorporation of segregation into the economic theories of lynching leads to a passive role for segregation, per se. The role of labor control and competition are not changed by the presence or absence of racial segregation. The economic incentive is driven by the supply and demand for black labor relative to white labor. While it could be the case that whites would be more likely to view African Americans as a threat in more segregated environments, the essentials of competition and labor control are not directly influenced by segregation unless one would argue that *integration* would lessen the likelihood of whites viewing seeing African Americans as competing for the same jobs or seeing African Americans as a labor force needing to be controlled.[4] Since economic competition is relatively silent on the effects of segregation itself we would predict a passive role for segregation in economic theories.

2.3   Social Theories of Lynching

Social theories of lynching hinge on class relations. In the Status Competition model of lynching, the idea is that lynchings are a product of competition between the class status of whites and African Americans. When the white population perceived their class status was threatened they responded with violence/disenfranchisement of African Americans (Tolnay and Beck, 1995; Price et al., 2008). A key for this model is the fact that lynchings are reactive– they are responses to the perception of a deterioration of white status and used to reinforce a racial hierarchy.

Segregation plays a critical role in the Status Competition model. Segregation itself is a signal of a racial hierarchy and separation of racial groups. As such, segregated environments by themselves can

---

[3]The relationship is concave, with a positive first derivative and a negative second derivative.

[4]Roediger and Esch (2012) describe the use of racial animus by firms to exert rents from labor in the late nineteenth and early twentieth centuries, but the authors do not claim an independent role for segregation.

be used as a signal of white status. The social separation of the groups would act as a complementary factor in to alleviate the problems caused by status competition. If the social superiority of whites was reinforced via segregation the Status Competition model predicts that there would be fewer lynchings. Since white status is not under threat the likelihood of racial violence would decrease. This is due to the fact that segregation helped to reinforce the idea of racial inequality and is a substitute for it.

Since segregation is a form of social isolation that would serve to reinforce racial inequality, the incorporation of segregation into the theory of Status Competition leads to the hypothesis that segregation would have a negative relationship with lynching. Highly segregated environments would be less likely to use racial violence to reinforce racial status as segregation already plays that role.

2.4   Political Theories of Lynching

Political theories see lynching reflecting fears of greater political participation by blacks. In the Power-Threat hypothesis, the idea is that when two groups coexist with unequal access to political/power resources, the dominant group will engage in a wide variety of methods (including violence/lynching) to secure their privileged access to those resources. The idea being the larger the political threat of African-Americans, the more lynchings that would occur. In the Political/Power-Threat hypothesis the idea is that after widespread segregation, and disenfranchisement of black voters that lynchings would greatly decline as the dominant group no longer felt threatened by the African-American vote (Soule, 1992; Corzine et al., 1983).

A key for the political theory of lynching is that African Americans be viewed as a threat to whites. This competition for resources presumes that African American access to resources would inherently come at the detriment of whites. The role of segregation in the political theory, therefore, is nuanced as it depends on the effect of segregation on perceptions. While segregated environments may be the end result of black disenfranchisement, the premise of whites viewing African Americans as a threat presupposes that there is little interaction between the groups which would counteract such perceptions. In integrated environments, the sustained interactions between racial groups could act to obviate the need for racial violence if whites did not view African Americans as a threat *because* of their integrated environment. But the opposite could also be true. Sustained interracial interaction

could "breed fear" of black political gains if such interactions revealed that African American political power would come at the expense of white political power.

As such, the predicted effect of segregation in political theories is indeterminate. The effect of segregation could lead to more racial violence or less. The direction of the effect depends on how whites view the potential outcomes of black political advancement. Most narrative histories suggest that whites held great apprehension of black political advances irrespective of their interaction with blacks. At the same time, whether segregation mediated or enhanced any of those sentiments is unknown.

While the existing literature on lynchings has emphasized the size of the black population, the theories that have arisen from that literature all suggest a role for residential segregation. While previous researchers have not had access to measures of segregation applicable to the rural communities in which lynchings took place, the new availability of complete count census data and new techniques for measuring segregation now make it possible to investigate this relationship between segregation and lynchings. Estimates of the impact of residential segregation on the incidence of lynchings will not only shed light on the social and economic forces leading to interracial violence, they will also allow us to separately identity the effects of the size of the minority population and the extent to which that population is residentially segregated from the majority group, two very distinct concepts that previous empirical studies could not disentangle.

3 The Logan-Parman Measure of Segregation

The Logan-Parman measure is an intuitive approach to residential segregation. They assert that the location of households in adjacent units can be used to measure the degree of integration or segregation in a community, similar to Schelling's classic model of household alignment. Areas that are well integrated will have a greater likelihood of opposite race neighbors that corresponds to the underlying racial proportion of households in the area. The opposite is also true— segregated areas will have a lower likelihood of opposite race neighbors than the racial proportions would predict. The measure relies on the individual-level data available in federal census records. With the 100%

sample of the federal census available through the Minnesota Population Center's Integrated Public Use Microdata Series (IPUMS), it is possible to identify the races of next-door neighbors. Census enumerators went door to door to record households, meaning that next-door neighbors are adjacent to one another on the census manuscript page. The number of black households with white neighbors in a county can therefore be calculated by looking at the order and races of all household heads on the census manuscript pages.

The measure is based on comparing this actual number of black households in a community with white neighbors to the number expected under complete integration and under complete segregation. Formally, the measure is calculated as

$$\eta = \frac{E(\overline{x_{b,w}}) - x_{b,w}}{E(\overline{x_{b,w}}) - E(\underline{x_{b,w}})} \tag{1}$$

where $x_{b,w}$ is the actual number of black households with a white next-door neighbor, $E(\overline{x_{b,w}})$ 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).[5] 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 county.

This measure of segregation has two distinct advantages for a study of lynchings. First, lynchings were often a rural phenomenon. Traditional segregation measures such as the index of dissimilarity and index of isolation are difficult to apply to rural areas. These measures rely on comparing racial composition of individual wards in a city to the racial composition of the city as a whole. Rural communities often lack comparable geographical subunit, making it difficult to calculate or interpret these traditional measures. The next-door neighbor measure of segregation does not require geographic subunits and can therefore be easily applied and interpreted in rural settings. for a city as a whole to the racial proportions.

---

[5]See the appendix of Logan and Parman (2015) for complete details about the derivation and estimation of the measure.

The second key advantage is that this next-door neighbor measure is a better proxy for interracial interactions than traditional measures based on racial proportions. Social interaction models of segregation are inherently spatial and assume that close proximity is related to social interactions (Echenique and Fryer, 2007; Reardon et al., 2008).Two counties with the same overall black and white populations will have very different levels of interracial interactions depending on how integrated those populations are. This variation is obscured by traditional segregation measures and by the fragmentation index common to studies of ethnic conflict but picked up by the next-door neighbor measure. While the measure is based on residential location and does not directly measure interracial interactions that occur through other work or social settings, it is reasonable to assume that, all other things equal, areas in which more black households live in very close proximity to white households will have higher levels of interracial interracial interaction. A segregation measure that captures this close proximity is particularly important to test the social and political theories of lynching that hinge on white perceptions of the black community, perceptions that will be heavily influenced by personal contact with members of that community.

4   Methods and Approach

4.1   Data

4.1.1   Southern Lynching Data

Our data on lynchings come from the Historical American Lynching (HAL) Project and cover the time period from 1882 to 1930– this data is identical to the lynching data of Beck and Tolnay. This dataset is the most extensively verified set of lynchings in the literature (Cook, 2012). The lynchings in the database conform to the NAACP definition of lynchings which requires a murder to meet the following criteria to be counted as a lynching: (1) there must be evidence that someone was killed, (2) the killings must have occurred illegally, (3) three or more persons must have taken part in the killing, and (4) the killers must have claimed to be serving justice or tradition.

The HAL database contains detailed information on 2,805 lynchings including name, race and gender of the victim, the race of the mob, the stated reason for the lynching, the date of the lynching

and the county in which the lynching took place. Given our interest in the relationship between segregation and lynching, we are primarily interested in lynchings in which the victim and mob have different races. These constitute the vast majority of the recorded lynchings: 88 percent of victims were black while only 6 percent of the mobs were black. Furthermore, of the 155 black lynch mobs, only 4 targeted white victims. Therefore, a large proportion of the lynching victims were black individuals who were the victim of interracial violence. The white lynching victims were almost entirely victims of intraracial violence. These white lynching victims will help us test whether any observed relationships between residential segregation and lynching work solely through the channel of interracial violence or whether they are instead proxying for more general tendencies to for a community to resort to mob justice.

The distribution of lynchings over time is given in Figure 1. Consistent with the existing literature on lynchings, the HAL data reveal that lynchings reached their peak in the 1890s, with over 100 lynchings per year in the middle of that decade. Despite peaking in the 1890s, there remain a sizable number of lynchings each year through 1930, the end of the time period covered by the data. The geographical distribution of lynchings across the southern states is given in Figure 1. Two features of this map are worth noting. First, there is substantial variation in the number lynchings across counties within each state. This suggests that we can still exploit meaningful variation in lynchings across counties even with the inclusion of state fixed effects in our analysis. Second, lynchings are not concentrated in urban areas. Instead, they appear to be a rather rural phenomenon. This underscores the importance of using a measure of segregation that can be meaningfully applied to rural areas.

4.1.2 Segregation Data

We calculate the Logan-Parman measure of segregation for every county in the United States using the Integrated Public Use Microdata Series (IPUMS) 100 percent sample of the 1880 federal census. Consequently, we are measuring segregation prior to the start of the lynchings data and therefore looking at how segregated counties were leading up to the lynchings, not how those levels of segregation changed in response to lynchings. We also calculate the percentage of household heads who are black in each county. The geographical distribution of these two measures of the location of black households

is given in Figure 1. What is immediately obvious from Figure 1 is that residential segregation and the overall racial composition of counties are correlated but not perfectly so. Figure 2 plots the segregation index against percent black for individual counties. While there is a weak, positive relationship between the two, there is substantial heterogeneity in levels of segregation at any given level of percent black. This neighbor-based segregation index provides information on significant variation in residential sorting that is not picked up by percent black; it reveals an important dimension of the residential patterns of communities that prior studies of lynchings could not exploit.

Table 1 provides state-level summary statistics for lynchings by race, segregation, and racial composition for the South, quantifying several of the patterns visible on the maps. Lynching and segregation vary substantially across states. Georgia and Mississippi have the highest levels of lynching activity. While they also exhibit relatively high levels of segregation, it is Florida and Louisiana that stand out as the most segregated. There is also substantial variation within states, with the standard deviations in percent black and the segregation index within states consistently larger than the standard deviation of the means across states. As with the variation in lynching, the within-state variation in segregation and racial composition makes us hopeful that we can estimate a relationship between segregation and lynching even after controlling for state fixed effects.

4.2 Empirical Strategy

We take the county level estimates of segregation in 1880 and merge them with the lynching data from the Historical American Lynching (HAL) project for the number of lynchings by county from 1882-1930. Our basic specification is

$$\lambda_{i,s} = \alpha + \beta_1 \eta_{i,s} + \beta_2 \eta_{i,s}^2 + \beta_3 PctBlack_{i,s} + \beta_4 PctBlack_{i,s}^2 + \Gamma X_{i,s} + \theta_s + \epsilon_{i,s} \tag{2}$$

where λ is the number of lynchings in a county or the presence of lynching in county i in state s, η is the Logan-Parman measure of segregation, and $PctBlack$ is the proportion black in a county. Since at least Raper (1933), the relationship between the percent black of a county and lynching has been known to be non-linear. For this reason we include quadratic terms for both percent black and

the segregation measure. To control for common factors that could drive the relationship, a state fixed effect θ_s is included. $X_{i,s}$ is a vector including county characteristics which, depending upon the specification, include traditional measures of segregation (dissimilarity and isolation), measures of migration, agricultural characteristics, and counts of black elected officials. To be clear, the equation above estimates the relationship between segregation and lynching exploiting within-state variation in the segregation measure at the county level.[6]

5   Segregation and Southern Lynching

In Table 2 we regress the number of lynchings per county between 1882 and 1935 on the segregation measure and the percent of households that were black in 1880. As lynching was highly differential by region we control for state fixed effects in all specifications. Since state fixed effects are included the estimates exploit the within-state variation in segregation and lynching only. As noted earlier, a key advantage here is that lynchings in the HAL data come from years after the 1880 census used to measure segregation. To asses the sensitivity of the relationship to functional form assumptions, we estimate the relationship in four different ways.

Following the existing literature, we first estimate the relationship between lynching and segregation using count models. In both the negative binomial and Poisson specifications, we see that increasing segregation was strongly related to the number of lynchings in a county.[7] In the third set of regressions, we control for the fact that many localities had no lynchings recorded during this period. Still using state fixed effects, we estimate a probit regression where the dependent variable is an indicator for whether or not a given county experienced a lynching (since some states had no lynchings the inclusion of state fixed effects reduces the sample size). As with the count models, increases in segregation were strongly correlated with whether or not a lynching occurred in a county.

---

[6]We restrict our attention to the Census-defined South. Results are unchanged when using all counties as the Beck and Tolnay data only covers Southern lynching.

[7]Across all specifications, the coefficients on percent black squared and segregation squared are negative. Note that the tipping point at which the relationship between lynchings and percent black turns negative is typically above one, suggesting that overall all possible racial compositions, the relationship is positive. The tipping point for the segregation index is consistently above 0.9, suggesting that for the vast majority of counties, the greater segregation is associated with a higher likelihood of lynchings. The median of the segregation measure for counties with a lynching was 0.33, and the 75th percentile was 0.43.

The final set of regressions estimate the relationship using a Tobit model, which is designed to account for the fact that many counties do not experience a lynching and could be modeled as being censored. The results confirm that increases in the segregation measure were strongly related to lynchings per county. Indeed, a one standard deviation increase in the segregation measure results in an additional lynching in a county, on average. In general, the results of Table 2 suggests that segregation was strongly related to lynching at both the extensive and intensive margins in the late nineteenth and early twentieth centuries. Furthermore, in all four models the inclusion of segregation reduces the magnitude of the percent black coefficients suggesting that the unavailability of segregation data for earlier studies led to overestimates of the relationship between racial proportions and lynchings.

One interesting feature of the results is that controlling for traditional measures of segregation does not alter the results. The neighbor-based segregation coefficients remain large and significant when including the isolation index and the dissimilarity index as controls. The coefficients on these traditional segregation measures coefficients are small in magnitude and typically statistically insignificant. In some respects, the results add quantitative support for the histories of racial cleansing offered by Jaspin (2008) and Loewen (2013). In particular, Jaspin (2008) notes that general county characteristics, such as racial makeup, did not predict racial cleansing in a county, and Loewen (2013) found that towns which forbade African American inhabitants did not share observable characteristics. While boundary-based measures imply that segregation had no impact on lynching, the neighbor-based measure of segregation shows that an increasing likelihood of opposite race neighbors is correlated with less lynching activity. In light of the theories discussed earlier, the results of Table 2 are most consistent with the political theories of lynching in which whites see blacks as a threat to political power. Economic theories predicted a passive role and social theories predicted a negative relationship, both of which are rejected.

5.1   Black and White Lynching

Since most lynchings were initiated by white mobs, the race of the victim can be used to distinguish whether the lynching was interracial or intraracial. The theories of lynching outlined earlier each involve interracial conflict. As a useful check of the role of segregation in explaining lynching we

separate the sample by the race of the lynching victim. Intuitively, if segregation plays a role in lynching it would work through interracial conflict. Segregation would have explanatory power for black lynchings but would not explain white lynching. In other words, racial segregation is thought to have explanatory power in interracial violence, not intraracial violence.

In Table 3 and Table 4 we estimate the relationship between segregation and lynching by race of the victim. The results are striking with respect to segregation. Segregation is highly correlated with black lynching but has no explanatory power with white lynching.[8] A comparison of the coefficients in Table 2 and Table 3 shows that the effect of segregation on lynching is slightly higher when the victim is African American than overall and disappears when looking only at white victims. The results suggest that the measure of segregation does not work through a channel that is related to all mob violence, but rather a channel that operates through interracial violence. We take these results are suggestive evidence that the measure of segregation is related to interracial violence and that segregation itself does not appear to be a predictor of intraracial violence.

5.2 Segregation, Southern Lynching, and Contextual Factors Over Time

The results presented earlier showed the correlation of segregation with the total number of lynchings. While suggestive that segregation plays a role in lynching, the specifications above do not incorporate important demographic and economic factors that have been shown to be correlated with lynching. As such, the results presented earlier could be due to the exclusion of these omitted factors. Scholars have noted that lynchings varied over time and that the number of lynchings during a given time period could have been related to time-varying factors. While the segregation measure we use is unchanging, we investigate the degree to which the effect of segregation on lynching is robust to temporal changes in lynching.[9]

To incorporate these economic factors we follow the seminal work of Tolnay and Beck (1992b), where the authors estimate the effects of black migration and differences in black and white illiteracy and tenancy on lynchings per decade. Here, we concentrate on the relationship between black

---

[8]Due to the large number of counties with no white lynchings, the negative binomial specification is preferred.

[9]In Table A1, we find that the time-series of lynching is strong and persistent. The counties that had more lynchings in 1882-1900 continued to be the same counties in later decades, even as the total number of lynchings declined.

migration, literacy, tenancy, and lynching. The main focus of the previous literature suggests that lynching was used by whites as a means to control economic or social competition. When net out-migration occurs and blacks leave (or remain in) the area, the likelihood that there is a black lynching will therefore be diminished as previous lynchings have had their desired effect or either restricting black movement or encouraging it, depending on the direction of the effect. The racial difference in farm tenancy and illiteracy attempt to reflect a higher status of whites; a large difference would be predicted to have fewer lynchings, in line with the status competition and economic competition theories described above[10] The absolute percent of black farm tenancy is included on the basis that racial violence might have been more common in areas where blacks were landless, as they may have been more susceptible to racial violence. Percent black is predominately used in the majority of the lynching literature as an indirect measure of competition, the idea being the larger the proportion of the black population the higher the perceived threat of the minority group. Population density has commonly been used as a predictor of lynching as lynchings generally occurred in less populated areas. Finally, the absolute black population is included because the larger the black population the higher the likelihood of a black lynching.[11]

We compare specifications where black lynchings by decade (1900-1910, 1910-1920, and 1920-1930) is the dependent variable and covariates include those noted above in Table 5. Following the previous literature, we restrict ourselves to Southern states in these regressions. As noted earlier, the peak year of lynchings occurred before 1900, yet we find in Table 5 that variation in segregation was correlated with lynchings after 1900. Areas that were more segregated experienced more lynchings from 1900-1910 and 1910-1920. The inclusion of segregation intensifies the tenancy effect and lessens the effects of racial differences in illiteracy and tenancy. This suggests that the effect of tenancy is downward biased when measures of segregation are excluded from the specification, and that black-white differences in illiteracy and tenancy are upward biased when segregation is excluded from the specification. Indeed, the effect of black farm tenancy share increases by more than twenty-five percent

---

[10]Significant differences in tenancy and illiteracy suggest a lack of substitutability of white labor for black labor and therefore less intense competition between the two groups.

[11]Our specification uses Tolnay and Beck (1992b) as a guide, but with modifications due to data restrictions. Our results include the Net Migration of blacks rather than the Net Out-Migration of blacks and does not account for mortality.

in all specifications.[12] For illiteracy differences by race, the effect is reduced by nearly ten percent across all specifications, and ceases to be statistically significant in the 1910-1920 specification once segregation is included.

The results in Table 5 imply that segregation not only explains some of the effects attributed to black population shares, but to other factors as well. The size of the black labor force in tenant farming was at slightly greater risk of experiencing a lynching in the county once segregation is included. At the same time, inclusion of segregation lessens the effect of educational differences by race. Given the results of Table 5, it appears that the channels through which segregation impacted the number of lynchings were likely nuanced and merit additional analysis. At a minimum, the results support the contention that segregation is a missing component in the empirical analysis of lynching over time, and also shows that the inclusion of important economic and demographic factors does not alter the relationship between racial segregation and lynching.

5.3 Segregation, Southern Lynching, and Executions

One key issue in the existing lynching literature is the degree to which lynchings were a function of the lack of rule of law. For example, in social theories of lynching, racial violence is used to enforce a racial hierarchy. This supposes that other means of enforcement are not readily available or were less likely to be used. The previous literature has analyzed the relationship between lynching and legal executions as a novel test of this hypothesis. In particular, researchers have looked to see if black lynching had a positive or negative elasticity with black executions. For example, Tolnay et al. (1992) test the notion that lynching was related to the practice of "rough justice" as opposed to racial animus. The previous literature has found that black lynchings were not a substitute for black executions, suggesting that lynchings were not a replacement for rule of law and therefore were likely to serve an alternative purpose.

It could certainly be the case that segregation could be related to environments where legal executions were more or less likely to occur, and therefore the effect we estimate between segregation and lynching could be spuriously driven by this factor. In Table 6 we estimate the reduced form

---

[12]Neither result is statistically significant in 1920-1930.

relationship between all black executions by decade from 1910 to 1930 using state fixed effects and a number of determinants used in previous literature, predominately Tolnay and Beck (1992b); Tolnay et al. (1992). Our models differ slightly from that of the original study as different time periods are used. We find that in various specifications that although segregation is positively related to black executions, it does not change the empirical findings of earlier studies which found that lynchings and executions were not statistically related to each other. While more segregated environments did experience more black executions, the inclusion of segregation does not alter the lack of a relationship between black executions and black lynchings.

6 Robustness Checks

Is the relationship shown between segregation and lynching causal? Ideally, we would like to exploit a source of exogenous variation in Southern segregation. Such a measure would be used as an instrument, allowing us to estimate the local average treatment effect of residential segregation on lynching. Unfortunately, nearly any factor that would be related to racial segregation in the South would not satisfy the exclusion restriction required for an instrumental variable. If there was an antebellum or Reconstruction-era factor that was related to segregation, it would likely be directly related to the social, economic, and/or political factors related to lynching and vary significantly at the county level within southern states. For example, if antebellum agricultural productivity or the extent of slaveholding were related to racial segregation, they are likely related to the economic conditions which persisted thereafter and would partially explain lynching. It is also important to note that our specifications include state fixed effects. That is, we are only exploiting within-state variation in segregation to estimate these effects. Any potential common factor must also vary at this level to drive the relationship between segregation and lynching. Since finding an instrument which would satisfy the exclusion restriction for Southern areas is highly unlikely, we take the approach of determining whether the most likely factors that would be related to segregation would alter the effects of segregation and lynching.

Specifically, since the measure of segregation comes from 1880 and lynchings are measured from

1882 onward, we are less concerned about reverse-causality per se than a common factor related to both segregation and lynching which would give rise to a spurious relationship. By considering these most likely factors directly, we further extend the empirical literature on segregation and racial violence while at the same time establishing that the estimated relationship is not driven by a common predetermined factor, such that the effect we estimate is unlikely to be driven by omitted variable bias. Since the previous analysis of contextual factors in Table 5 included the measures of economic racial and socioeconomic differences established in the literature and the results in Table 6 considered social factors established in the literature, we focus on political measures and antebellum (historical) factors that could, presumably, drive the relationship we estimate between segregation and lynching.[13]

6.1   Segregation and Antebellum Southern Development

One factor to consider is the possibility that the segregation measure is simply a proxy for antebellum developments which left a long and persistent effect that was correlated with segregation and lynching. It is true that the areas which experienced more lynchings at any one point in time were also more likely to experience more lynchings during other time periods.[14] As such, the effect of segregation on lynching could be due to predetermined factors which are related to segregation, but for which segregation is simply a noisy proxy. Naturally, antebellum factors could drive both segregation and racial violence, and the failure to control for antebellum factors would lead to an overestimate of the relationship between segregation and lynching at best, and a spurious relationship at worst.

We address this issue by including several factors which could be related to antebellum development, segregation, and lynching in the same specification. Our goal is two-fold. First, to see whether the antebellum factors are correlated with lynching and, second, to see if their inclusion alters the relationship between segregation and lynching. We consider several proxies for antebellum development– the percent of the black population that was free in 1860, the average cash value of farms in a county in 1860, the output of one of three cash crops (cotton, sugar, and tobacco) in local

---

[13]A related point is that if all potential instruments are exhausted, then the only conceivable factor which would drive segregation and its relationship with lynching would be *unmeasurable*. While we do not make that assertion, our aim is to show that the most likely variables that could be thought of as instruments for segregation (1) do not satisfy the exclusion restriction and (2) do not alter the estimated relationship between segregation and lynching.

[14]See Table A1 for the time series of lynching.

agricultural output in 1860 and the fraction of all slaveholders in a county with more than 50 slaves. The first factor relates to the likelihood of greater acceptability of black autonomy in locations with larger shares of free blacks. The second factor is related to antebellum wealth. The third factor is related to agricultural development– to include the entire South we include the three most prominent crops since tobacco- and sugar-growing locations also experienced significant numbers of lynchings. The last proxy is a measure of the fraction of farms in the county that were large plantations, but rather than looking over all households, only estimates it as a fraction of the slaveholding class.[15]

In Table 7, we show the results. Some of the antebellum factors do have a strong relationship with lynching. For example, sugar and tobacco production in 1860 are positively correlated with lynching, but cotton production is not. The cash value of farms is also positively correlated with lynching, underscoring the partially economic roots of interracial violence. The percent of the black population which was free is not well correlated with lynching. Most important, the inclusion of these antebellum factors does not alter the relationship between segregation and lynching. It is still the case that counties which were more segregated experienced more lynchings, even controlling for antebellum factors.

6.2   Segregation and Black Political Involvement

While some scholars have asserted that segregation could reflect the rule of law, others believe that political involvement of African Americans would play a role in either causing or preventing lynchings. If African Americans had significant political power, it may be difficult to participate in extra-legal executions as blacks would, presumably have more access to legal redress. On the other hand, if lynchings were retaliatory, areas which saw black political gains during Reconstruction could have been places where racial intimidation was more likely to occur. If either of these political channels was related to segregation, it is possible that the correlation between segregation and lynching would be biased upwards by this omitted political factor. Similarly, the results thus far are most consistent with the political theory of the relationship between segregation and lynching. As such, the inclusion of a measure of black political involvement, which would be the root of white resistance to black

---

[15]Due to the high correlation of percent black with measures of the slave population, they are excluded from the specifications.

political gains, may eliminate the correlation of segregation and lynching.

To assess this possibility we measure black political participation in a novel way– the number of African American officeholders during Reconstruction at the county level. This measure captures black political involvement but is not contaminated with the effects of black disenfranchisement, which occurred during the same period as racial violence. For example, using contemporaneous voting outcomes would not ensure that one was capturing black political involvement if previous political success was related to later racial violence and disenfranchisement. This measure also allows us to asses robustness of the effect of political participation as the officials can be disaggregated by type of office (federal, state, or local) and branch of government (executive, legislative, or judicial). This count of African American office holders is one of the few quantitative measures of black political involvement and success during Reconstruction.[16]

In Table 8 we present the results. Although in each instance the point estimate for the number of officials of each type is negative, the effect is not statistically significant. More important, the inclusion of the number of black officials does not alter the relationship between segregation and lynching. While Table 8 shows the results for all officials, state office holders, local office holders, and judicial officials, no other type of official (e.g., executive or legislative) was correlated with black lynching nor altered the relationship between segregation and lynching. Similarly, while Table 8 shows the results for black lynchings, the same pattern is seen for lynchings over all and for white lynchings. Thus, the results are inconsistent with the hypothesis that lynchings were greater in places where blacks were more likely to hold office during Reconstruction, and they are also inconsistent with the hypothesis that black political power during Reconstruction left a lasting effect on racial violence. Most important, the results are inconsistent with the argument that the segregation measure captures an omitted political factor which explains the segregation measure's correlation with lynching.

---

[16]Our count comes from *Freedom's Lawmakers*, the most comprehensive measure of black officeholders during Reconstruction.

7 Discussion

In this paper we have used a new measure of segregation from the complete 1880 census which used the simple criteria of the race of a neighbor to investigate the relationship between segregation and Southern lynching. This interracial violence has had lingering effects on levels of trust and the development of the American South. Incorporating segregation into the analysis of lynching offers an opportunity to understand how interracial contact influences conflict. The Logan-Parman measure of segregation based on the races of next-door neighbors offers a much stronger proxy for interracial interactions than the racial proportions prior studies were restricted to. We showed that the relationship between this neighbor-based segregation index and Southern lynching was quite strong and robust, with greater segregation associated with higher levels of lynchings conditional on the overall racial composition of a county.

The results are consistent with the political theory, where whites fear disenfranchisement and a more general theory of social conflict, where segregation reflects underlying animosity between races. The results are inconsistent with status competition and with economic theories of Southern lynching. Rather than being a substitute for racial violence, segregation appears to be a complementary factor. Also, segregation does not play a passive role in lynching, as predicted by economic theories of lynching. The results are most consistent with the hypothesis that segregation was correlated with social conflict to the extent that areas which were more segregated were also areas which experienced more racial violence and/or where the externalities of lynching are greater in more racially segregated environments.

We also considered how the inclusion of segregation alters the existing empirical literature on lynchings. We found that the inclusion of segregation altered some of the relationship between lynchings and black/white differences in economic factors, lessening their impact. We found that the inclusion of segregation had no impact on the relationship between lynchings and executions. Most important, we found that it is unlikely that the relationship we estimate between segregation and lynching is driven by omitted factors. We found that including black political gains during Reconstruction and antebellum economic and demographic factors did not alter the relationship between segregation and lynching. Overall, we found that segregation has a large and robust relationship with racial violence

in the American past.

At a minimum, this project shows that segregation is an important part of the Southern lynching story in the United States that should be investigated. The empirical relationship between segregation and racial violence shows that the effects of segregation are not confined to urban communities but also have a strong relationship with those in rural areas. As we noted earlier, more than three quarters of the population lived in rural areas in 1880. Given the recent calls to expand the scope of lynching data to incorporate the national trend in lynching over this time period (Cook, 2012), future work can extend this analysis beyond the South and allow us to investigate regional differences in the relationships described here. Understanding the relationship between segregation and racial violence helps us understand the dynamics of segregation in rural communities in the twentieth century.

References

Alesina, A. and La Ferrara, E. (2002). Who trusts others? *Journal of Public Economics*, 85(2):207–234.

Alesina, A. and La Ferrara, E. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43:762–800.

Bailey, A. K., Tolnay, S. E., Beck, E., and Laird, J. D. (2011). Targeting lynch victims social marginality or status transgressions? *American Sociological Review*, 76(3):412–436.

Beck, E. and Tolnay, S. E. (1990). The killing fields of the deep south the market for cotton and the lynching of blacks, 1882-1930. *American Sociological Review*, 55(4):526–539.

Beck, E. and Tolnay, S. E. (1992). A season for violence the lynching of blacks and labor demand in the agricultural production cycle in the american south. *International Review of Social History*, 37(1):1–24.

Blalock, H. M. (1967). *Toward a theory of minority-group relations*. Wiley.

Cassar, A., Grosjean, P., and Whitt, S. (2013). Legacies of violence: trust and market development. *Journal of Economic Growth*, 18(3):285–318.

Coale, A. and Rives, N. (1973). A statistical reconstruction of the black population of the united states, 1880-1970: Estimates of true numbers by age and sex, birth rates, and total fertility. *Population Index*, 39(1):3–36.

Collier, P. and Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers*, 56(4):563–595.

Collier, P. and Rohner, D. (2008). Democracy, development, and conflict. *Journal of the European Economic Association*, 6(2-3):531–540.

Cook, L. (2012). Converging to a national lynching database recent developments and the way forward. *Historical Methods*, 45(2):55–63.

Corzine, J., Creech, J., and Corzine, L. (1983). Black concentration and lynchings in the south testing blalocks power-threat hypothesis. *Social Forces*, 61(1):774–796.

Costa, D. L. and Kahn, M. E. (2003). Understanding the american decline in social capital. *Kyklos*, 56:17–46.

DeFina, R. and Hannon, L. (2011). The legacy of black lynching and contemporary segregation in the south. *The Review of Black Political Economy*, 38(2):165–181.

Eblen, J. (1974). New estimates of vital rates of united states black population during the nineteenth century. *Demography*, 11(2):301–319.

Echenique, F. and Fryer, R. G. (2007). A measure of segregation based on social interactions. *The Quarterly Journal of Economics*, 122(2):441–485.

Fearon, J. D. and Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97(01):75–90.

Hovland, C. I. and Sears, R. R. (1940). Minor studies of aggression correlations of economic indices with lynchings. *Journal of Psychology*, 9(1):301–310.

Jaspin, E. (2008). *Buried in the bitter waters: The hidden history of racial cleansing in America.* Basic Books.

Kantrowitz, S. (2012). *More Than Freedom: Fighting for Black Citizenship in a White Republic, 1829-1889.* Penguin.

Loewen, J. W. (2013). *Sundown towns: A hidden dimension of American racism.* The New Press.

Logan, T. D. and Parman, J. M. (2015). The national rise in residential segregation. Technical report, National Bureau of Economic Research.

National Association for the Advancement of Colored People (1919). *Thirty years of lynching in the United States, 1889-1918.* NAACP, New York.

Nunn, N. (2008). The long-term effects of africa's slave trades. *The Quarterly Journal of Economics*, pages 139–176.

Nunn, N. and Wantchekon, L. (2011). The slave trade and the origins of mistrust in africa. *American Economic Review*, 101:3221–3252.

Ottaviano, G. I. and Peri, G. (2006). The economic value of cultural diversity: evidence from us cities. *Journal of Economic Geography*, 6(1):9–44.

Preston, S., Elo, I., Foster, A., and Fu, H. (1998). Reconstructing the size of the african american population by age and sex, 1930-1990. *Demography*, 35(1):1–21.

Price, G., Darity Jr, W. A., and Headen, A. (2008). Does the stigma of slavery explain the maltreatment of blacks by whites? the case of lynchings. *Journal of Socioeconomics*, 37(1):167–193.

Raper, A. (1933). *The Tragedy of Lynching*. University of North Carolina Press.

Reardon, S. F., Matthews, S. A., OSullivan, D., Lee, B. A., Firebaugh, G., Farrell, C. R., and Bischoff, K. (2008). The geographic scale of metropolitan racial segregation. *Demography*, 45(3):489–514.

Roediger, D. R. and Esch, E. D. (2012). *The Production of Difference: Race and the Management of Labor in U.S. History*. Oxford University Press.

Rohner, D., Thoenig, M., and Zilibotti, F. (2013). Seeds of distrust: Conflict in uganda. *Journal of Economic Growth*, 18(3):217–252.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of mathematical sociology*, 1(2):143–186.

Soule, S. (1992). Populism and black lynching in georgia, 18901900. *Social Forces*, 71(1):431–449.

Tolnay, S. E. and Beck, E. (1992a). Racial violence and black migration in the american south 1910 to 1930. *American Sociological Review*, 57(1):103–16.

Tolnay, S. E. and Beck, E. (1992b). Racial violence and black migration in the american south, 1910 to 1930. *American Sociological Review*, 57(1):103–116.

Tolnay, S. E. and Beck, E. (1995). *A Festival of Violence: An Analysis of Southern Lynchings*. University of Illinois Press.

Tolnay, S. E., Beck, E., and Massey, J. L. (1992). Black competition and white vengeance legal execution of blacks as social control in the cotton south, 1890 to 1929. *Social Science Quarterly*, 73(3):627–644.

Voigtländer, N. and Voth, H.-J. (2012). Persecution perpetuated: the medieval origins of anti-semitic violence in nazi germany. *Quarterly Journal of Economics*, 127(3):1339–1392.

Williams, D. T. (1968). Amid the gathering multitude: The story of lynching in america, a classified listing. *Unpublished manuscript. Tuskegee University, Tuskegee, Alabama*.

Table 1: Lynchings and Segregation in the South

	Number of lynchings		County-level percent black		County-level segregation index	
	Black victims	White victims	Mean	Standard devation	Mean	Standard devation
Alabama	273	24	0.37	0.26	0.36	0.13
Arkansas	184	48	0.24	0.24	0.34	0.16
Florida	224	19	0.32	0.23	0.45	0.13
Georgia	435	21	0.42	0.22	0.31	0.13
Kentucky	128	43	0.13	0.11	0.22	0.12
Louisiana	304	53	0.52	0.24	0.42	0.12
Mississippi	509	22	0.50	0.22	0.37	0.09
North Carolina	82	16	0.32	0.19	0.25	0.11
South Carolina	148	6	0.58	0.16	0.33	0.11
Tennessee	175	37	0.17	0.15	0.22	0.10

All lynching data are taken from the Project HAL database. Percent black and segregation numbers are from the authors' calculations based on the 100 percent sample of the 1880 federal census.
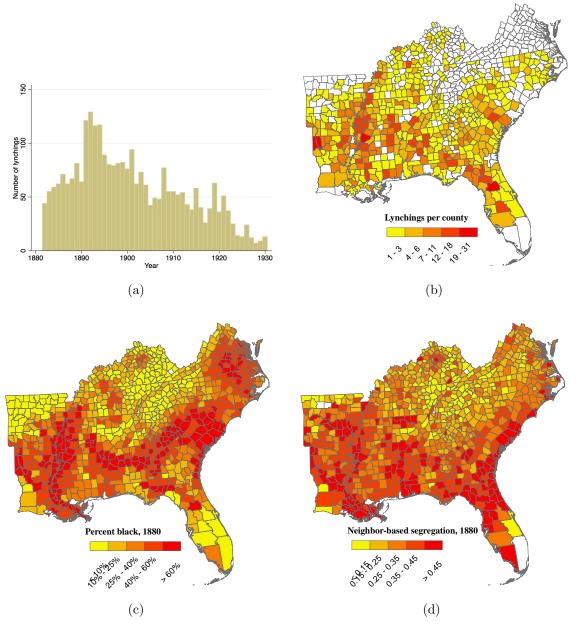
Figure 1: Segregation, Lynching, and Racial Population Shares. (a) Number of lynchings by year, 1882-1930; (b) Lynchings per county, 1882-1930; (c) Percent black by county in 1880; and, (d) Segregation by county in 1880. Source: Project HAL Data and Logan and Parman (2015).
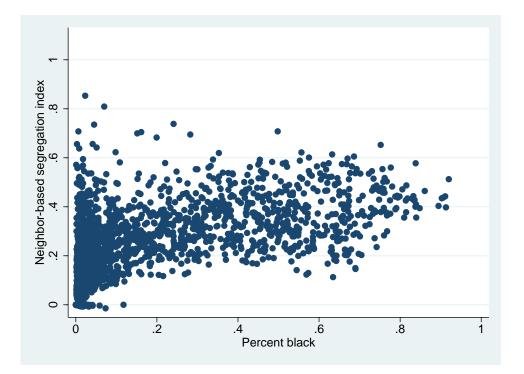
Figure 2: Neighbor-based segregation and percent black by county, 1880. Source: Logan and Parman (2015).

Table 2: The Correlation of Segregation with Number of Lynchings per County

Method	Negative Binomial			Poisson			Probit			Tobit		
Dependent Variable	Number of Lynchings			Number of Lynchings			Number of Lynchings			Number of Lynchings		
Percent Black	5.390***	3.970***	3.475***	4.775***	3.534***	3.101***	1.448***	0.936***	0.883***	17.22***	12.22***	10.90***
	[0.591]	[0.686]	[0.726]	[0.344]	[0.403]	[0.429]	[0.235]	[0.288]	[0.314]	[2.563]	[3.094]	[3.352]
Percent Black^2	-4.276***	-3.068***	-2.545***	-3.519***	-2.512***	-2.082***	-1.347***	-0.875***	-0.823**	-11.79***	-7.421**	-6.117*
	[0.677]	[0.737]	[0.777]	[0.366]	[0.401]	[0.430]	[0.296]	[0.332]	[0.355]	[3.045]	[3.394]	[3.632]
Segregation Index	4.637***	4.483***		4.360***	4.274***		1.370***	1.307***		13.40**	12.49**	
	[1.256]	[1.282]		[0.865]	[0.869]		[0.423]	[0.431]		[5.201]	[5.317]	
Segregation Index^2	-5.115***	-4.474***		-4.760***	-4.253***		-1.725***	-1.560**		-14.67**	-11.96	
	[1.656]	[1.727]		[1.096]	[1.112]		[0.612]	[0.637]		[7.141]	[7.465]	
Isolation Index		-0.0363			0.0785			-0.178			-1.663	
		[0.857]			[0.519]			[0.249]			[3.280]	
Dissimilarity Index		-0.944*			-0.789**			-0.0207			-1.618	
		[0.542]			[0.329]			[0.186]			[2.215]	
Constant	0.240	-0.0962	-0.419*	0.307***	-0.321**	-0.0602				0.0805	-1.624*	-0.794
	[0.157]	[0.268]	[0.235]	[0.0927]	[0.158]	[0.174]				[0.690]	[0.942]	[1.116]
State Fixed Effects	X	X	X	X	X	X	X	X	X	X	X	X
Observations	1,174	1,111	1,111	1,174	1,111	1,111	786	783	783	1,174	1,111	1,111

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Table 3: The Correlation of Segregation with Number of Black Lynchings per County

Method	Negative Binomial			Poisson			Probit			Tobit		
Dependent Variable	Number of Black Lynchings			Number of Black Lynchings			Number of Black Lynchings			Number of Black Lynchings		
Percent Black	7.223***	5.646***	5.074***	6.127***	4.824***	4.322***	2.583***	2.035***	1.870***	23.66***	18.27***	16.40***
	[0.661]	[0.754]	[0.804]	[0.393]	[0.452]	[0.485]	[0.281]	[0.332]	[0.365]	[2.652]	[3.129]	[3.405]
Percent Black^2	-5.948***	-4.609***	-3.988***	-4.701***	-3.647***	-3.150***	-2.411***	-1.898***	-1.743***	-18.09***	-13.33***	-11.51***
	[0.743]	[0.801]	[0.852]	[0.409]	[0.445]	[0.481]	[0.344]	[0.379]	[0.408]	[3.090]	[3.397]	[3.653]
Segregation Index	5.600***		5.461***	5.073***		4.918***	1.800***		1.728***	16.63***		15.57***
	[1.449]		[1.483]	[1.016]		[1.019]	[0.538]		[0.559]	[5.707]		[5.854]
Segregation Index^2	-6.368***		-5.432***	-5.691***		-4.893***	-2.418***		-2.108***	-19.55**		-15.55*
	[1.880]		[1.946]	[1.278]		[1.290]	[0.750]		[0.790]	[7.695]		[8.002]
Isolation Index		-0.678			-0.220			-0.352			-3.648	
		[1.045]			[0.645]			[0.342]			[3.737]	
Dissimilarity Index			-1.144*			-0.943**			-0.152			-2.263
			[0.635]			[0.390]			[0.231]			[2.364]
Constant	-0.225	-1.024***	-0.618**	-0.105	-0.844***	-0.512**				-1.550**	-3.649***	-2.461**
	[0.171]	[0.272]	[0.305]	[0.105]	[0.187]	[0.203]				[0.696]	[1.025]	[1.183]
State Fixed Effects	X	X	X	X	X	X	X	X	X	X	X	X
Observations	1,175	1,112	1,112	1,175	1,112	1,112	786	783	783	1,175	1,112	1,112

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Table 4: The Correlation of Segregation with Number of White Lynchings per County

Method	Negative Binomial			Poisson			Probit			Tobit		
Dependent Variable	Number of White Lynchings			Number of White Lynchings			Number of White Lynchings			Number of White Lynchings		
Percent Black	-0.540	-1.869	-1.805	-0.628	-2.003*	-1.681	-0.114	-0.192	-0.195	-1.194	-2.590	-2.471
	[1.234]	[1.501]	[1.582]	[0.876]	[1.062]	[1.143]	[0.224]	[0.275]	[0.297]	[1.991]	[2.455]	[2.638]
Percent Black^2	0.643	1.718	1.675	0.952	2.079*	1.781	0.0369	0.118	0.122	0.919	2.161	2.054
	[1.462]	[1.631]	[1.703]	[1.024]	[1.134]	[1.207]	[0.272]	[0.304]	[0.324]	[2.401]	[2.704]	[2.873]
Segregation Index		2.446	2.871		3.121*	3.435*		0.538	0.570		4.806	5.354
		[2.445]	[2.525]		[1.793]	[1.802]		[0.437]	[0.448]		[3.910]	[4.016]
Segregation Index^2		-1.582	-2.508		-2.748	-3.629		-0.837	-0.904		-5.988	-7.192
		[3.434]	[3.706]		[2.387]	[2.499]		[0.624]	[0.662]		[5.546]	[5.909]
Isolation Index			0.889			0.686			0.0790			1.131
			[1.683]			[0.848]			[0.259]			[2.357]
Dissimilarity Index			-0.0988			0.237			-0.0216			-0.0692
			[1.115]			[0.706]			[0.187]			[1.682]
Constant	-0.934***	-1.324***	-1.411**	-0.973***	-1.410***	-1.603***				-1.828***	-2.433***	-2.569***
	[0.359]	[0.470]	[0.570]	[0.254]	[0.349]	[0.415]				[0.577]	[0.751]	[0.903]
State Fixed Effects	X	X	X	X	X	X	X	X	X	X	X	X
Observations	1,175	1,112	1,112	1,175	1,112	1,112	786	783	783	1,175	1,112	1,112

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

35

Table 5: Segregation, Black Lynching, and Contextual Factors by Decade, 1900-1930

	1900-1910	1900-1910	1910-1920	1910-1920	1920-1930	1920-1930
Percent Black	-2.934***	-3.785***	6.346***	4.145***	7.875***	6.557***
	[0.928]	[0.963]	[1.031]	[1.128]	[1.600]	[1.774]
Percent Black^2	1.347	1.940**	-5.368***	-3.745***	-7.384***	-6.645***
	[0.933]	[0.955]	[1.094]	[1.139]	[1.756]	[1.844]
Segregation Index		4.349**		7.977***		0.354
		[1.778]		[2.673]		[3.485]
Segregation Index^2		-3.709*		-7.539**		2.590
		[2.159]		[3.389]		[4.221]
Net Migration of Blacks	4.33e-05***	4.10e-05***	-1.70e-05	-1.63e-05	2.96e-05	1.54e-05
	[1.59e-05]	[1.56e-05]	[1.72e-05]	[1.70e-05]	[3.29e-05]	[3.21e-05]
Percent Black Farm Tenancy Among Black Farmers	0.695***	0.898***	0.876**	1.260***	0.282	0.589
	[0.231]	[0.237]	[0.379]	[0.394]	[0.518]	[0.502]
Population Density	-0.00951***	-0.00779***	-0.000406	-0.000408	-0.00853**	-0.00663**
	[0.00248]	[0.00235]	[0.000281]	[0.000284]	[0.00344]	[0.00317]
Difference Between Black - White Illiteracy	3.265***	2.982***	2.236***	1.634	0.425	0.0749
	[0.346]	[0.351]	[0.866]	[1.013]	[1.646]	[1.662]
Difference Between Black - White Tenanacy	-0.602*	-0.489	-0.155	-0.156	0.717	0.664
	[0.327]	[0.327]	[0.544]	[0.542]	[0.649]	[0.646]
Black Population Size	2.74e-05***	2.33e-05***	2.32e-05***	2.05e-05***	2.65e-05**	2.41e-05**
	[7.26e-06]	[7.17e-06]	[4.45e-06]	[4.53e-06]	[1.08e-05]	[1.05e-05]
Constant	0.430	-0.545	-3.305***	-4.645***	-3.107***	-3.340***
	[0.269]	[0.413]	[0.287]	[0.506]	[0.485]	[0.679]
Observations	772	772	772	772	778	778

Standard errors in brackets   *** p<0.01, ** p<0.05, * p<0.1

All specifications estimated with Poisson regression. Dependent variable in each regression: Number of Black Lynchings by decade

Net Migration of Blacks defined as net change in black population over decade.

Percent Black Farm Tenancy Among Black Farmers, Population Density, Difference between Black -

White Illiteracy, Difference between Black - White Tenancy, and Black Population Size defined at end of decade.

Table 6: Segregation, Black Executions, Black Lynching, and Contextual Factors by Decade, 1910-1930

	Black Executions from 1911-1920				Black Executions from 1921-1930			
	Poisson		Negative Binomial		Poisson		Negative Binomial	
Percent Black	5.571***	4.593***	5.152***	4.181***	5.624***	4.370***	5.367***	4.244***
	[1.068]	[1.182]	[1.250]	[1.384]	[1.015]	[1.132]	[1.145]	[1.266]
Percent Black^2	-4.497***	-3.589***	-4.352***	-3.376**	-4.234***	-3.162**	-4.341***	-3.349**
	[1.197]	[1.274]	[1.415]	[1.508]	[1.213]	[1.287]	[1.379]	[1.450]
Black Population	2.80e-05***	2.84e-05***	3.87e-05***	4.02e-05***	2.42e-05***	2.36e-05***	3.45e-05***	3.34e-05***
	[3.13e-06]	[3.29e-06]	[6.73e-06]	[6.80e-06]	[1.56e-06]	[1.59e-06]	[5.25e-06]	[5.09e-06]
Difference Between Black - White Tenanacy	0.0232	0.0191	0.111	0.0299	-0.473	-0.466	-0.529	-0.518
	[0.391]	[0.404]	[0.506]	[0.536]	[0.378]	[0.378]	[0.454]	[0.455]
Difference Between Black - White Illiteracy	-1.075	-1.139	-0.986	-1.095	-1.834	-1.738	-1.440	-1.376
	[1.273]	[1.296]	[1.519]	[1.578]	[1.412]	[1.432]	[1.557]	[1.582]
Black Lynchings During Time Period	-0.00779	-0.00782	-0.0200	-0.0203	0.133	0.140	0.0670	0.0770
	[0.0467]	[0.0469]	[0.0625]	[0.0628]	[0.0894]	[0.0886]	[0.114]	[0.113]
White Executions During Time Period	0.187*	0.153	0.301*	0.272	0.313***	0.300***	0.301***	0.289***
	[0.112]	[0.117]	[0.174]	[0.175]	[0.0415]	[0.0423]	[0.0821]	[0.0815]
Segregation Index		4.512*		6.720**		4.509**		4.398*
		[2.315]		[2.882]		[2.129]		[2.452]
Segregation Index^2		-5.758*		-9.690**		-4.870*		-5.031
		[3.031]		[3.957]		[2.674]		[3.300]
Constant	-2.307***	-2.894***	-2.328***	-3.169***	-2.187***	-2.783***	-2.276***	-2.824***
	[0.460]	[0.562]	[0.493]	[0.639]	[0.377]	[0.477]	[0.409]	[0.519]
State Fixed Effects	X	X	X	X	X	X	X	X
Observations	1,073	1,073	1,073	1,073	1,103	1,103	1,103	1,103

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Table 7: Segregation, Lynching, and Antebellum Factors

Dependent Variable	Total Number of Lynchings					
Method	Negative Binomial			Poisson		
Segregation Measure	4.870***	4.231***	3.763***	5.151***	4.550***	4.090***
	[1.281]	[1.286]	[1.283]	[0.896]	[0.898]	[0.893]
Segregation Measure Squared	-4.457***	-3.682**	-2.870*	-4.675***	-3.955***	-3.133***
	[1.725]	[1.727]	[1.727]	[1.151]	[1.152]	[1.138]
Percent Free Blacks, 1860	-1.148	-0.733	-0.397	-1.454**	-1.077*	-0.724
	[0.834]	[0.840]	[0.844]	[0.641]	[0.640]	[0.620]
Percent Large Plantations, 1860	1.034	-0.416	-0.168	0.706*	-0.635	-0.525
	[0.813]	[0.900]	[0.945]	[0.410]	[0.494]	[0.527]
Cash value of farms, 1860		5.71e-08***	5.51e-08***		5.20e-08***	5.71e-08***
		[1.70e-08]	[2.06e-08]		[9.73e-09]	[1.20e-08]
Tobacco output, 1860			1.24e-06***			1.10e-06***
			[4.12e-07]			[2.38e-07]
Cane sugar output, 1860			-6.32e-07***			-6.48e-07***
			[2.34e-07]			[1.50e-07]
Cotton output, 1860			1.15e-08			-1.42e-08
			[8.58e-08]			[4.56e-08]
Constant	0.410	0.376	0.408	0.367**	0.336*	0.366*
	[0.267]	[0.265]	[0.266]	[0.187]	[0.186]	[0.188]
State Fixed Effects	X	X	X	X	X	X
Observations	520	519	519	520	519	519

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Table 8: Segregation, Black Lynching, and Black Elected Officials

	Number of Black Lynchings								
Percent Black	5.553***	5.479***	4.941***	5.464***	4.934***	5.507***	4.944***	5.469***	4.922***
	[0.753]	[0.756]	[0.805]	[0.757]	[0.806]	[0.754]	[0.804]	[0.755]	[0.804]
Percent Black^2	-4.486***	-4.291***	-3.751***	-4.252***	-3.726***	-4.359***	-3.767***	-4.308***	-3.731***
	[0.802]	[0.822]	[0.869]	[0.827]	[0.872]	[0.809]	[0.858]	[0.813]	[0.861]
Segregation Index	5.614***	5.602***	5.462***	5.593***	5.453***	5.620***	5.478***	5.630***	5.483***
	[1.446]	[1.445]	[1.478]	[1.445]	[1.478]	[1.445]	[1.478]	[1.444]	[1.477]
Segregation Index^2	-6.394***	-6.319***	-5.412***	-6.327***	-5.413***	-6.331***	-5.415***	-6.415***	-5.476***
	[1.876]	[1.876]	[1.941]	[1.875]	[1.941]	[1.876]	[1.941]	[1.874]	[1.939]
Total Number of Officials		-0.0114	-0.00684						
		[0.0105]	[0.0106]						
State Officials				-0.0206	-0.0129				
				[0.0177]	[0.0178]				
Local Officials						-0.0294	-0.0217		
						[0.0259]	[0.0256]		
Judicial Officials								-0.0591	-0.0456
								[0.0474]	[0.0470]
Isolation Index			-0.681		-0.690		-0.660		-0.693
			[1.040]		[1.040]		[1.040]		[1.040]
Dissimilarity Index			-1.127*		-1.119*		-1.142*		-1.121*
			[0.633]		[0.634]		[0.633]		[0.633]
Constant	-18.05	-18.24	-17.58	-17.39	-17.47	-17.96	-18.07	-18.23	-17.58
	[708.1]	[778.9]	[629.2]	[510.6]	[596.3]	[677.4]	[804.4]	[776.0]	[631.0]
State Fixed Effects	X	X	X	X	X	X	X	X	X
Observations	1,109	1,109	1,109	1,109	1,109	1,109	1,109	1,109	1,109

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

A Deriving the Segregation Measure

Construction of the measure begins by identifying neighbors in the census. The complete set of household heads in the census is sorted by reel number, microfilm sequence number, page number and line number. This orders the household heads by the order in which they appear on the original census manuscript pages, meaning that adjacent households appear next to one another. There are two different methods for identifying each household head's next-door neighbors. The first is to simply define the next-door neighbors as the household head appearing before the individual on the census manuscript page and the household head appearing after the individual on the census manuscript page. An individual that is either the first or last household head on a particular census page will only have one next door neighbor identified using this method.

To allow for the next door neighbor appearing on either the previous or next census page and to account for the possibility that two different streets are covered on the same census manuscript page, an alternative method for identifying neighbors is also used that relies on street name rather than census manuscript page. In this alternative measure next-door neighbors are now identified by looking at the observations directly before and after the household head in question and declaring them next-door neighbors *if and only if* the street name matches the street name of the individual of interest (and the street name must be given, two blank street names are not considered a match). This approach has the advantage of finding the last household head on the previous page if an individual is the first household head on his census manuscript page or the first household head on the next page if the individual was the last household head on a manuscript page. However, the number of observations is reduced substantially relative to the first method because many individuals have no street name given. Few roads had names in historical census records. This is particularly true in rural areas.

Once next door neighbors are identified, an indicator variable is constructed that equals one if the individual has a next door neighbor of a different race and zero if both next-door neighbors are of the same race as the household head.[1] Two versions of this indicator variable are constructed, one in which all observations are used and one in which only those observations for which both next-door neighbors are observed are used. This latter version reduces the sample size but, for the remaining individuals, gives a more accurate measure of the percentage of individuals with a neighbor of a different race.

Formally, we begin with the following:

- b_{all}: the total number of black household heads in the area

- $n_{b,B=1}$: the number of black household heads in the area with two observed neighbors

- $n_{b,B=0}$: the number of black household heads in the area with one observed neighbor

- x_b: the number of black household heads in the area with a neighbor of a different race

The equivalent variables for the set of white household heads are similarly defined. These components, by themselves, can be used to derive new measures of social interaction between races. For

---

[1]Based on the race assigned at enumeration. This is similar to the *racesing* coding of race constructed by IPUMS. One key feature of *racesing* for our purposes is places people with their race given as 'mulatto' in the same category as people with their race given as 'black'. So a black individual living next to two neighbors listed on the census as mulatto would be considered to be of the same race as his neighbors.

example, using the measures above one can calculate the share of households with an opposite race neighbor.

Given these measures, the basic measure of segregation is calculated as the distance the area is between the two extremes of complete segregation and the case where neighbor's race is entirely independent of an individual's own race. There are a total of four versions of the segregation measure. Each of these measures corresponds to one of the two different methods of defining next-door neighbors (whether the specific street of residence is identified on the census manuscript form) and whether all individuals with a neighbor present are included or only those individuals with both neighbors identified are used.

In the case of random neighbors, the number of black residents with at least one white neighbor will be a function of the fraction of black households relative to all households. In particular, the probability that any given neighbor of a black household will be black will be $\frac{b_{all}-1}{(b_{all}-1)+w_{all}}$. The probability that the second neighbor will be black if the first neighbor is black will then be $\frac{b_{all}-2}{b_{all}-2+w_{all}}$. The probability that a black household head will have at least one white neighbor can be written as a function of these probabilities by expressing it as:

$$p(\text{white neighbor}) = 1 - \left(\frac{b_{all}-1}{b_{all}-1+w_{all}} \right) \left(\frac{b_{all}-2}{b_{all}-2+w_{all}} \right) \tag{1}$$

where the second term comes from the assumption that the races of adjacent neighbors are uncorrelated, a reasonable assumption given that we are considering randomly located neighbors. The expected value of x_b under random assignment of neighbors would then be:

$$E(\overline{x_b}) = p(\text{white neighbor}) \cdot n_b \tag{2}$$

$$E(\overline{x_b}) = n_b \left(1 - \left(\frac{b_{all}-1}{b_{all}-1+w_{all}} \right) \left(\frac{b_{all}-2}{b_{all}-2+w_{all}} \right) \right) \tag{3}$$

The calculation of this upper bound on x_b must be modified slightly when including household heads for which only one neighbor is observed. In this case, the expected number of black household heads with a white neighbor under random assignment of neighbors will be composed of two different terms, the first corresponding to those household heads with both neighbors observed and the second corresponding to those household heads with only one neighbor observed. Letting B be an indicator variable equal to one if both neighbors are observed and equal to zero if only one neighbor is observed, the expected total number of black household heads with a white neighbor is then:

$$E(\overline{x_b}) = p(\text{white neighbor}|B=1) \cdot n_{b,B=1} + p(\text{white neighbor}|B=0) \cdot n_{b,B=0} \tag{4}$$

$$E(\overline{x_b}) = n_{b,B=1} \left(1 - \left(\frac{b_{all}-1}{b_{all}-1+w_{all}} \right) \left(\frac{b_{all}-2}{b_{all}-2+w_{all}} \right) \right) + n_{b,B=0} \left(1 - \frac{b_{all}-1}{b_{all}-1+w_{all}} \right) \tag{5}$$

Under complete segregation, the number of black individuals living next to white neighbors would simply be two, the two individuals on either end of the neighborhood of black residents, giving a lower bound for the value of x_b. However, it is necessary to account for observing only a fraction of the household heads. The expected observed number of black household heads living next to a white neighbor when sampling from an area with only two such residents will be:

$$E(\underline{x_b}) = p(\text{observe one of the two in } n_b \text{ draws}) \cdot 1 + p(\text{observe both in } n_b \text{ draws}) \cdot 2 \tag{6}$$

$$E(\underline{x_b}) = \frac{1}{\frac{1}{2}(n_b+1)}\left(1-\prod_{i=0}^{n_b-1}\frac{b_{all}-i-2}{b_{all}-i}\right) + 2\left(1-\frac{1}{\frac{1}{2}(n_b+1)}\right)\left(1-\prod_{i=0}^{n_b-1}\frac{b_{all}-i-2}{b_{all}-i}\right) \qquad (7)$$

The product in the expression above gives the probability of selecting neither of the two black household heads with white neighbors in n_b successive draws from the b_{all} black household heads. Thus one minus this product is the probability of drawing either one or both of the two household heads with white neighbors. Note that the product notation is used above because it makes it easier to see how the probability is being derived. In practice, the product reduces to $\frac{(b_{all}-n_b)(b_{all}-n_b-1)}{b_{all}(b_{all}-1)}$. The ratio $\frac{1}{\frac{1}{2}(n_b+1)}$ gives the fraction of these cases that correspond to drawing just one of the two household heads with white neighbors. This comes from noting that with n_b draws, that there are n_b ways to draw one of the two household heads while there are $\sum_{i=1}^{n_b-1}(n_b-i)$ or $n_b(n_b-1) - \frac{(n_b-1)n_b}{2}$ ways to draw both of the household heads.

Finally, in the case where household heads with only one observed neighbor are included, it is necessary to account for the probability that a black household head with a white neighbor will be drawn but that white neighbor is not the observed neighbor. The expected value of x_b accounting for the probability that the white neighbor is unobserved for a household head with only one observed neighbor is:

$$E(\underline{x_b}) = \left(\frac{n_{b,B=1}}{n_b} + \frac{n_{b,B=0}}{n_b}\cdot\frac{1}{2}\right) \qquad (8)$$

$$\cdot\left[\frac{1}{\frac{1}{2}(n_b+1)}\left(1-\prod_{i=0}^{n_b-1}\frac{b_{all}-i-2}{b_{all}-i}\right)+ \right. \qquad (9)$$

$$\left. 2\left(1-\frac{1}{\frac{1}{2}(n_b+1)}\right)\left(1-\prod_{i=0}^{n_b-1}\frac{b_{all}-i-2}{b_{all}-i}\right)\right] \qquad (10)$$

In this equation, the fraction of black household heads with only one observed neighbor, $\frac{n_{b,B=0}}{n_b}$, has its expected value of x_b reduced by an additional factor of $\frac{1}{2}$ to account for the fact that if one of these individuals is one of the two black household heads living next to a white neighbor there is only a 50 percent chance that the white neighbor is the observed neighbor.

The degree of segregation in an area, η, can then be defined as the distance between these two extremes, measured from the case of no segregation:

$$\eta = \frac{E(\overline{x_b}) - x_b}{E(\overline{x_b}) - E(\underline{x_b})} \qquad (11)$$

This segregation measure increases as black residents become more segregated within an area, equaling zero in the case of random assignment of neighbors (no segregation) and equalling one in the case of complete segregation.[2] Note that it is possible for this measure to be less than zero if the particular

---

[2]Given the evidence that population counts of the size of the African American community in census returns is biased, we are concerned about the problem of missing African Americans (Coale and Rives, 1973; Eblen, 1974; Preston et al., 1998). While it would appear that under-reporting of African Americans would be a concern, it would only bias estimates of the segregation measure if the missing African American households had white neighbors. To see how, note that $E(\underline{x_b})$ is invariant to the number of black and white households as it estimates the minimum number of households who would have opposite race neighbors, which itself is not a function of the size of either group. Since the measure of segregation is the ratio of the two differences ($E(\overline{x_b})$ - x_b and $E(\overline{x_b})$ - $E(\underline{x_b})$), only if the estimate of x_b is biased downward would missing black households have a material effect on the estimate of segregation. Given

sample of household heads is actually more integrated than random assignment of neighbors. For example, suppose every other household head on the manuscript pages were black in an area that is 50 percent black. With random assignment of neighbors we would expect to observe at least some black household heads having black neighbors. In this case, x_b would be larger than $E(\overline{x_b})$ making η negative. The measure can also exceed one in the rare cases where only zero or one black household heads with a white neighbor are observed. In these cases x_b may actually be smaller than $E(\underline{x_b})$. We do not observe this for counties with more than ten black households.

In communities with large numbers of both black and white households, $E(\overline{x_b})$ will be substantially larger than $E(\underline{x_b})$, allowing us to approximate the above equation as

$$\eta \approx 1 - \frac{x_b}{E(\overline{x_b})} \tag{12}$$

This form of the segregation index is similar to measures of segregation based on evenness that take the following form

$$\rho = 1 - \frac{\frac{1}{N}\sum_{i=1}^{m} N_i D_i}{\frac{1}{N}\sum_{i=1}^{m} E(N_i D_i|\text{no segregation})} \tag{13}$$

where N is the total population of interest across all m subunits of the larger area, N_i is is the population of interest in subunit i, and D_i is the relevant measure of diversity in that subunit. If we define the subunit to be as small as possible, namely an individual black household, then N becomes the number of black households and N_i simply becomes one, reducing the above expression to

$$\rho = 1 - \frac{\frac{1}{n_b}\sum_{i=1}^{n_b} D_i}{\frac{1}{n_b}\sum_{i=1}^{m} E(D_i|\text{no segregation})} \tag{14}$$

where n_b is the same number of black households used in the derivation of our measure above. In the context of our measure, the diversity measure D_i is equal to one if a black household has a white neighbor and zero if it does not. Noting that x_b is the number of black households for which D_i is equal to one and $E(\overline{x_b})$ is the expected number of black households for which D_i is equal to one under no segregation, this generic expression for a measure of evenness becomes

$$\rho = 1 - \frac{\frac{1}{n_b} x_b}{\frac{1}{n_b} E(\overline{x_b})} = 1 - \frac{x_b}{E(\overline{x_b})} \tag{15}$$

which is identical to Equation 12, the approximation of our measure in the case of large numbers of black and white households. Thus our measure can be thought of as a household-level measure of evenness when the number of black and white households is large.

It is important to note that the assumption required for this reformulation of the measure, that $E(\overline{x_b})$ is substantially larger than $E(\underline{x_b})$, is a strong assumption. When applying our measure to smaller geographical areas such as individual enumeration districts and to areas with very small numbers of black households, $E(\overline{x_b})$ will not be orders of magnitude larger than $E(\underline{x_b})$. One key advantage of our measure as defined in Equation 11 relative to traditional measures is that it can still be applied to small black populations and small total populations while maintaining a consistent interpretation for the values of zero and one. As the simulations in the paper demonstrate, traditional

---

the reality of census enumeration, it is unlikely that enumerators deliberately skipped African American households in integrated communities as opposed to skipping entire groups of black households.

measures of segregation such as dissimilarity and isolation do not reliably converge to zero in the case of complete integration or one in the case of complete segregation when populations are small while our measure does.

Table A1: Time Series of Southern Lynchings

Method	Negative Binomial			Poisson		
All Lynchings	1900-1909	1910-1919	1920-1930	1900-1909	1910-1919	1920-1930
Lynchings from 1910-1919			0.0632			0.0593
			[0.0533]			[0.0451]
Lynchings from 1900-1909		0.154***	0.112***		0.132***	0.104***
		[0.0352]	[0.0382]		[0.0210]	[0.0307]
Lynchings from 1882-1899	0.0725***	0.0803***	0.0461*	0.0703***	0.0808***	0.0489**
	[0.0214]	[0.0229]	[0.0269]	[0.0123]	[0.0150]	[0.0230]
Constant	-0.186	-0.629***	-1.851***	-0.180	-0.580***	-1.850***
	[0.201]	[0.212]	[0.311]	[0.135]	[0.152]	[0.290]
Black Lynchings	1900-1909	1910-1919	1920-1930	1900-1909	1910-1919	1920-1930
Lynchings from 1910-1919			0.0595			0.0523
			[0.0585]			[0.0492]
Lynchings from 1900-1909		0.170***	0.142***		0.146***	0.134***
		[0.0369]	[0.0410]		[0.0218]	[0.0324]
Lynchings from 1882-1899	0.0745***	0.0971***	0.0464	0.0691***	0.0926***	0.0482*
	[0.0251]	[0.0258]	[0.0309]	[0.0140]	[0.0164]	[0.0256]
Constant	-0.168	-0.686***	-2.229***	-0.145	-0.627***	-2.227***
	[0.207]	[0.218]	[0.365]	[0.134]	[0.154]	[0.345]
White Lynchings	1900-1909	1910-1919	1920-1930	1900-1909	1910-1919	1920-1930
Lynchings from 1910-1919			0.610			0.721
			[0.767]			[0.669]
Lynchings from 1900-1909		0.586	0.706		0.602	0.658
		[0.791]	[0.670]		[0.733]	[0.574]
Lynchings from 1882-1899	-0.00565	-0.240	0.240	-0.00282	-0.191	0.214*
	[0.176]	[0.407]	[0.152]	[0.158]	[0.371]	[0.115]
Constant	-17.97	-4.024***	-2.873***	-17.95	-4.031***	-2.815***
	[1,038]	[1.015]	[0.557]	[1,029]	[1.002]	[0.508]
State Fixed Effects	X	X	X	X	X	X
Observations	599	599	599	599	599	599

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1