Segregation and Lynching

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

September 24, 2014

Abstract

The empirical relationship between segregation and racial violence is unknown. Since racial violence (lynching) occurred in rural and urban areas, traditional measures of segregation cannot be used to estimate the relationship. Earlier analysis has used racial proportions, a poor proxy for segregation. We use a newly developed measure of residential segregation based on individual-level data (Logan and Parman 2014), which exploits complete census manuscript files to derive a measure of segregation based upon the racial similarity of next door neighbors. With this new measure, we distinguish 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 county in the United States, we estimate the relationship between racial segregation and lynching. We find that conditional on racial composition, 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 1935. The result is robust to numerous controls, functional form assumptions, and the inclusion of traditional segregation measures. 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 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

[†]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

[§]We thank Richard H. Steckel for suggestions on this project. Jacob D. Ginsberg, Terry L. Pack and Stephen Prifti provided excellent research assistance. The usual disclaimer applies.

"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

Segregation has long been viewed as a factor in explaining racial violence in the past. 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 (1992) argue that racial violence was related to black migration patterns in the early twentieth century. The number of lynchings in the United States reached their peak in 1892 (Cook, 2012). 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. 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. In particular, the role of segregation in lynching has not been empirically validated.

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.

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.

We review the existing theories of 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. We also develop a more general framework for the effect of segregation on lynchings that draws on more general concepts in social conflict. The theories have clear predictions for the relationship between segregation and racial violence. Ultimately, however, the relationship between segregation and lynching is an empirical question.

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 (2014), 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 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 individual 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 correlation of segregation and lynching for all areas where lynchings occurred.

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 concerns about possible reverse-causality in the lynching-segregation relationship. Additionaly, Logan and Parman (2014) 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 unlikely 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 dispersion, 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, and state fixed effects.

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.

In what follows, we 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 empirical results and 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 Lynching

There are four prominently proposed hypotheses of lynching in the literature which can be grouped into three spheres of community life. The theories of Labor Control and Economic Competition concern economic life, the theory of Status Competition is concerned with social life and the Power-Threat theory concerns political life. Each of these hypotheses deal with competition between blacks and whites in one sector or another. Below, we review the existing theories to incorporate racial segregation into those frameworks and also show a more general framework for social interactions.

We show that the existing theories of lynching can be extended to include segregation as a factor. Most important, we show that the theories lead to different conclusions about the nature of the segregation/lynching relationship. Theories which stress the economic factors relating to lynching, where lynching is used to maintain control over the black labor force, racial segregation itself would play a passive role– whether whites and blacks were in socially isolated would have little influence on the use of lynching for economic gain as it would be driven by market forces of labor supply and demand. Specifically, the relative size of the black labor force would matter, but not integration with whites. In social theories of lynching, however, segregation could alleviate the problems caused by status competition– when the social superiority of whites is upheld by other means (such as racial segregation) then lynching would be less likely to happen. In political theories of lynching, political concerns about African American enfranchisement could be positively or negatively related to lynching. Close interactions could lessen the belief that political gains for African Americans would come at the expense of whites. On the other hand, increased interactions could increase fears of African American political power if those interactions suggested that African American political gains would come at the expense of whites. As such, the relationship could move in either direction depending on how whites and African Americans viewed interactions over political aims.

Beyond this, we also show how segregation could be linked to more general ideas of social cohesion in a community. In the more general setting, segregation is related to racial violence either through social isolation or familiarity. If racial violence is related to isolation, more isolated environments would be more likely to experience a lynching. If familiarity led to more racial hostility, then segregation would be negatively related to lynching. We briefly review each of these theories below.

2.1 Economic Theories

The Theory of Labor Control is tied to the idea of the use of lynching as a form of social control over black workers. The theory, proposed by Beck and Tolnay, 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 is also closely linked to the Theory of Labor Control. The idea being that as whites and blacks began to compete for the same jobs (southern whites were becoming more and more economically disadvantaged as the southern economy stagnated), 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 for the remaining African Americans) 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 the measures 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 grows, 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,

however, 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.[1]

The incorporation of segregation into the economic theories of lynching leads to a passive role for segregation. 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.[2] Since economic competition is relatively silent on the effects of segregation itself the models would predict a passive role.

2.2 Social Theories

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. Basically, 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 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.

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

[2]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.

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.

2.3 Political Theories

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 any of those sentiments is unknown.

2.4 Lynching and Social Conflict

While the review of the preceding theories shows that segregation can be integrated into the existing theories of lynching, a more general framework can be proposed. Ultimately, the relationship between segregation and racial violence hinges on the relationship between social isolation and racial aggression. There are two possible effects. Segregation could increase racial violence if it is related to underlying racial hostility and animus. 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, 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.

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.

This measure does not suffer from the limitations of using political boundaries for geographical subunits and in fact does not require geographical subunits at all, making it possible to look at segregation in *any* geographical area, a key innovation of their approach to segregation. 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. Rather than asking whether an individual lives in a ward or tract with many black residents, a question that hinges on how wards or tracts are defined, they ask whether an individual lives next to a black or white neighbor, a question that can be consistently and universally applied to all households.

This approach to segregation has a number of additional advantages. First, it focuses on households as opposed to the population. The degree of residential segregation depends on the number of *households* of different types, not the number of individuals. If members of one group have larger household sizes or different household structure (for example, more likely to live in multiple generation households) there will be a difference between the population share and the household share. Household structure and size are known to vary by race historically and at present (Ruggles, 1994). Another advantage is that this measure is also an intuitive proxy for social interactions. Neighbors are quite likely to have some interactions with each other, and an increasing likelihood of opposite race neighbors implies that the average level of interactions across racial lines would be higher. Indeed, 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).

Specifically, the measure compares the observed number of black households in a area living next to a white neighbor to the predicted number given the overall racial composition of the area. They calculate the predicted number of black households with white neighbors given the number of black and white households in the area assuming that households are randomly located by race and assuming that households are completely segregated (only the households on the edge of the all black community have white neighbors). The segregation measure is then simply an estimate of how far the actual number of black households with white neighbors is between these two extremes. In essence, the measure is a counterfactual between the observed and hypothetical distribution of households in a given area.

Using the alignment of households in the census, the Logan-Parman measure identifies household race using the race of the household head. Construction of the measure begins by identifying neighbors

in manuscript census records.[3] The method requires the complete, 100% census since all households are needed. 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 next-door neighbors appear next to one another. Institutions, boarding houses and other non-households (dormitories, etc.) are excluded from the calculation. Households in apartments or other multi-family units are recorded as separate households and are retained. The analysis focuses on black households, assessing whether they have a neighbor of a different race. However, all racial groups other than black or white constituted less than 0.5% of the total population from 1870 to 1940 in census returns. As such, a black household with a neighbor of a different race is essentially the equivalent to saying he has a white neighbor.[4] Given the extremely low levels of interracial marriage in the past (fewer than 0.2% of households had opposite race spouses from 1870 to 1940) the measure assumes the race of the household head applies to all household members.

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 all observed next-door neighbors are of the same race as the household based on the race assigned at enumeration. As such, the measure of opposite race neighbors is measured at the extensive margin and is measured for each household in the manuscript census.

Summing this indicator variable across all black households for the entire county gives the number of black households with a next-door neighbor of the opposite race, x_b. The segregation measure compares this number of black households with opposite-race neighbors to the expected number under complete segregation, $E(\underline{x_b})$, and the expected number under complete integration (random assignment of neighbors), $E(\overline{x_b})$. These two values are calculated based on the total number of black households and white households in a county. $E(\underline{x_b})$ is calculated assuming that only the two households on either end of the black neighborhood have white neighbors.[5] $E(\overline{x_b})$ is calculated

[3]The full derivation of the segregation measure is given in the appendix.

[4]People with their race given as 'mulatto' are considered to be in the same category as people with their race given as 'black'.

[5]This value is a function of the probability of observing one or both of the two black households with white neighbors (a non-trivial number of households in the census do not have races given for their neighbors). Defining the number of black households with both neighbors' races observed as n_b and the total number of black households in the county as

assuming that households are randomly assigned by race: the probability of a next-door neighbor being of the opposite race is given by the fraction of the households in the county of that race.[6]

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

$$\eta = \frac{E(\overline{x_b}) - x_b}{E(\overline{x_b}) - E(\underline{x_b})} \tag{1}$$

This segregation measure increases as black residents become more segregated within an area. The measure equals zero in the case of random assignment of neighbors (no segregation) and equals one in the case of complete segregation. The measure is only defined for racially heterogeneous communities, as racially homogeneous communities are neither segregated nor integrated. The segregation measure is normalized by the population size and the percent of African Americans in the community, which allows for comparison of segregation across communities with different population sizes and racial compositions.[7]

Most important for this analysis, the measure of segregation shows that areas with large and small black populations could be segregated or integrated. Until now, segregation in rural communities could only be approximated by the percent black in a county. In Figure 1, the percent of a county that is black is a relatively poor approximation of the level of segregation in the community. At each level of percent in a county, there is significant heterogeneity in the neighbor-based measure of segregation. The correlation of the segregation measure and percent black is only 0.43. This relatively weak correlations suggest that sorting, independent of population shares, is an important dimension of segregation. It is also a dimension that traditional segregation measures cannot capture.

b_{all}, the value of $E(\underline{x_b})$ is calculated as $\frac{1}{\frac{1}{2}(n_b+1)}\left(1-\prod_{i=0}^{n_b-1}\frac{b_{all}-i-2}{b_{all}-i}\right)+2\left(1-\frac{1}{\frac{1}{2}(n_b+1)}\right)\left(1-\prod_{i=0}^{n_b-1}\frac{b_{all}-i-2}{b_{all}-i}\right)$. In the case of including households with only one observed neighbor, this equation must be modified somewhat to account for the possibility of observing one of the black households with a white neighbor but not observing the white neighbor. Full details are provided in the appendix.

[6]Following the same notation as the previous footnote and defining the total number of white households in the county as w_{all}, the value of $E(\overline{x_b})$ is calculated as $n_b\left(1-\frac{b_{all}-1}{b_{all}-1+w_{all}}\cdot\frac{b_{all}-2}{b_{all}-2+w_{all}}\right)$. As with $E(\underline{x_b})$, the equation must be modified when including households with only one observed neighbor. Details are provided in the appendix.

[7]See Logan and Parman (2014) for a comparison of this measure of segregation with traditional measures. In general, since the Logan-Parman measure does not require subdistricts or census tracts to measures segregation the measure performs better as the number and size of the black and white populations vary.

4 Methods and Approach

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, which is a compendium of lynching data recorded by Tolnay and Beck as well as other sources. To date, this is the most widely used and extensively verified lynching data in the literature (Cook, 2012). Our basic specification is

$$\lambda_i = \alpha + \beta_1 \eta_i + \beta_2 PctBlack_i + \Gamma X_i + \epsilon_i \tag{2}$$

where λ is the number of lynchings in a county or the presence of lynching in a county, η is the Logan-Parman measure of segregation, and $PctBlack$ is the proportion black in a county. To control for common factors that could drive the relationship between racial violence X includes state fixed effects, traditional measures of segregation (dissimilarity and isolation) and other controls. To be clear, the equation above estimates the relationship between segregation and lynching exploiting within-state variation.

The coefficients of interest are β_1 and β_2. As the measurement of segregation is between zero (completely integrated) and one (completely segregated) a positive coefficient on β_1 implies that higher levels of segregation are correlated with increased number of lynchings.

5 Preliminary Results

In Table 1 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. 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 five different ways.

Following the existing literature, we first estimate the relationship between lynching and segre-

gation 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. In the third column, 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 many states had no lynching 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.

The fourth column of Table 1 is an OLS specification that restricts attention only to those counties that experienced a lynching. There, the question is whether the number of lynchings is related to segregation, conditional on being a location where lynching occurred. The results show that there is a strong relationship between the number of lynchings and segregation even for the locations where lynchings occurred. Multiple lynchings were more likely in highly segregated environments. Indeed, a one standard deviation change in the segregation measure increases the number of lynchings per county by 0.10 standard deviations, a sizeable effect.[8] The fifth column estimates 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 1 suggests that segregation was strongly related to lynching at both the extensive and intensive margins in the late nineteenth and early twentieth centuries.

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. One interesting feature of the results in Table 1 is that the traditional measures of segregation are not well correlated with the number or presence of lynchings in a county. While boundary-based measures imply that segregation had no impact on lynching, the neighbor-based measure of segregation shows

[8]The result is robust to non-linear models which account for the well-established non-linear relationship between the percent African American in a county and lynching.

that an increasing likelihood of opposite race neighbors is correlated with less lynching activity.

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 2 and Table 3 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. A comparison of the coefficients in Table 1 and Table 2 shows that the effect of segregation on lynching is slightly higher when the victim is African American than overall. 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.

6 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 lynching. If neighbors are less much likely to be of a different race than random assignment would predict then that location is more segregated than another that is closer to random assignment. We showed that the relationship between segregation and lynching was quite strong and robust.

The results are consistent with the political theory, where whites fear disenfranchisement and social

conflict, where segregation reflects underlying animosity between races. The results are inconsistent with status competition and with economic theories of lynching. Rather than being a substitute for racial violence, segregation appears to be a complementary factor. Future specifications and the inclusion of additional data will allow us to further refine this relationship.

In particular, matching the data used here to a more extensive list of county controls will allow to test particular features of the existing theories of lynching. For example, the segregation measure can be interacted with changes in prices of agricultural productivity, tenant farming, political election outcomes, and other measures that have been used in the literature. This will allow us to extend the exiting analysis to include a fuller range of controls and characteristics that have been shown to be related to lynching.

At a minimum, this project shows that segregation is an important part of the 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. Understanding the relationship between segregation and racial violence helps us understand the dynamics of segregation in rural communities in the twentieth century. Better knowledge of segregation's past will give us the tools to outline its full impact in the past and, most important, the present.

References

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. (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.

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.

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.

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.

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. (2014). The national rise in segregation. Technical report, The Ohio State University.

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.

Ruggles, S. (1994). The origins of african-american family structure. *American Sociological Review*, 59(13):136–151.

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. (1992). 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. (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.

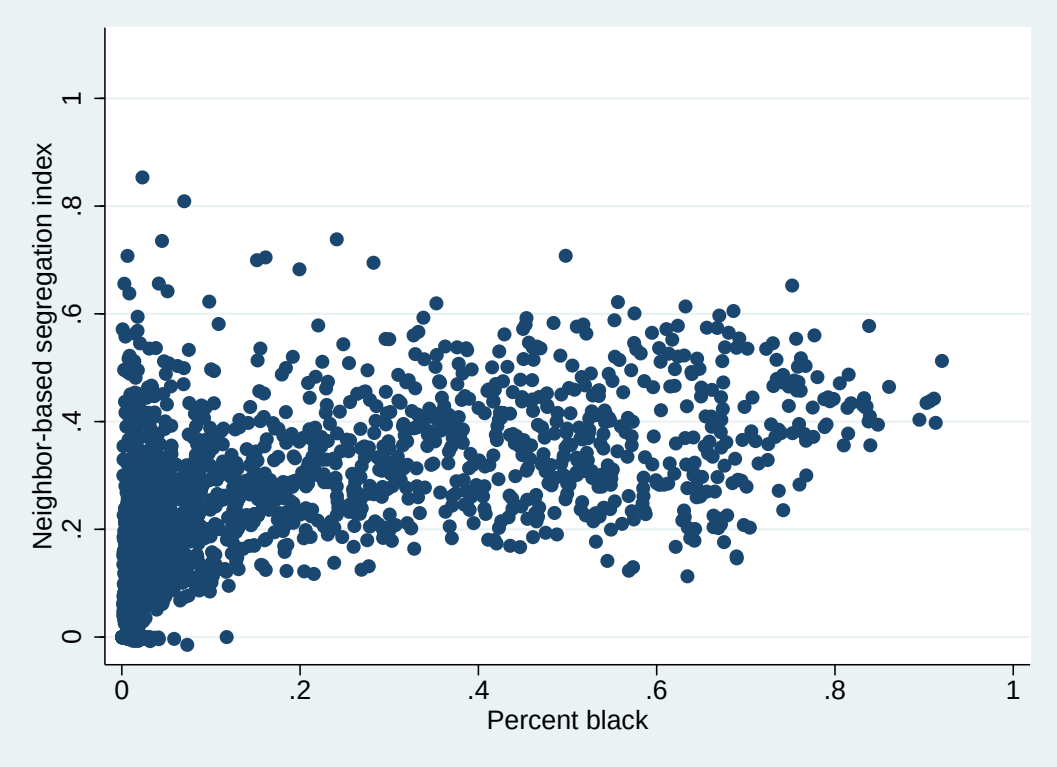

Figure 1: Percentage of population African American and the Logan-Parman measure of segregation, 1880.

Table 1: The Correlation of Segregation with Number of Lynchings per County

Segregation and Lynching

Method	Negative Binomial	Poisson	Probit	OLS	Tobit
Dependent Variable	Number of Lynchings	Number of Lynchings	Lynching in a County	Lynchings (Lynchings>0)	Number of Lynchings
Segregation Index	1.917***	1.464***	0.544***	3.188*	5.965***
	[0.398]	[0.208]	[0.154]	[1.698]	[1.660]
Percent Black	1.348***	1.252***	0.220**	5.264***	5.801***
	[0.216]	[0.105]	[0.102]	[0.906]	[0.961]
Isolation Index	-0.0455	0.405	-0.257	0.333	-1.980
	[0.820]	[0.455]	[0.243]	[3.721]	[3.123]
Dissimilarity Index	-1.511***	-1.362***	-0.206	-2.396	-3.067
	[0.518]	[0.289]	[0.176]	[2.238]	[2.065]
Constant	0.650***	0.765***		2.328***	0.961
	[0.202]	[0.108]		[0.861]	[0.874]
State Fixed Effects	X	X	X	X	X
Observations	2,100	2,100	783	597	2,100
R-squared				0.243	

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

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

Segregation and Black Lynching

Method	Negative Binomial	Poisson	Probit	OLS	Tobit
Dependent Variable	Number of Lynchings	Number of Lynchings	Lynching in a County	Lynchings (Lynchings>0)	Number of Lynchings
Segregation Index	2.330***	1.644***	0.736***	3.334*	7.259***
	[0.443]	[0.224]	[0.185]	[1.822]	[1.711]
Percent Black	1.686***	1.450***	0.452***	5.070***	6.626***
	[0.232]	[0.112]	[0.117]	[0.899]	[0.947]
Isolation Index	-0.212	0.612	-0.354	0.0610	-2.717
	[0.954]	[0.529]	[0.312]	[4.351]	[3.417]
Dissimilarity Index	-2.111***	-1.850***	-0.515**	-2.525	-5.117**
	[0.591]	[0.327]	[0.212]	[2.436]	[2.154]
Constant	0.461**	0.623***		1.993**	-5.117**
	[0.219]	[0.117]		[0.873]	0.388
State Fixed Effects	X	X	X	X	X
Observations	2,100	2,100	783	540	2,100
R-squared				0.232	

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

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

Segregation and White Lynching

Method	Negative Binomial	Poisson	Probit	OLS	Tobit
Dependent Variable	Number of Lynchings	Number of Lynchings	Lynching in a County	Lynchings (Lynchings>0)	Number of Lynchings
Segregation Index	1.146	0.828	0.0729	1.426	1.087
	[0.792]	[0.530]	[0.156]	[1.032]	[1.251]
Percent Black	-0.401	-0.126	-0.138	0.232	-0.893
	[0.473]	[0.331]	[0.0956]	[0.594]	[0.763]
Isolation Index	-0.320	-0.0374	-0.140	0.984	-0.738
	[1.417]	[0.798]	[0.265]	[2.018]	[2.121]
Dissimilarity Index	0.473	0.635	0.0903	0.124	0.878
	[0.952]	[0.611]	[0.185]	[1.245]	[1.475]
Constant	-1.287***	-1.340***	0.0903	1.022*	-2.107***
	[0.426]	[0.298]		[0.555]	[0.688]
State Fixed Effects	X	X	X	X	X
Observations	2,100	2,100	783	197	2,100
R-squared				0.103	

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.[9] Two versions of this indicator variable are constructed, one in

[9]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

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 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

mulatto would be considered to be of the same race as his neighbors.

function of these probabilities by expressing it as:

$$p(\text{white neighbor}) = 1 - \left(\frac{b_{all} - 1}{b_{all} - 1 + w_{all}}\right)\left(\frac{b_{all} - 2}{b_{all} - 2 + w_{all}}\right) \tag{3}$$

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:

$$E(\overline{x_b}) = p(\text{white neighbor}) \cdot n_b \tag{4}$$

$$E(\overline{x_b}) = n_b\left(1 - \left(\frac{b_{all} - 1}{b_{all} - 1 + w_{all}}\right)\left(\frac{b_{all} - 2}{b_{all} - 2 + w_{all}}\right)\right) \tag{5}$$

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:

$$E(\overline{x_b}) = p(\text{white neighbor}|B = 1) \cdot n_{b,B=1} + p(\text{white neighbor}|B = 0) \cdot n_{b,B=0} \tag{6}$$

$$E(\overline{x_b}) = n_{b,B=1}\left(1 - \left(\frac{b_{all} - 1}{b_{all} - 1 + w_{all}}\right)\left(\frac{b_{all} - 2}{b_{all} - 2 + w_{all}}\right)\right) + n_{b,B=0}\left(1 - \frac{b_{all} - 1}{b_{all} - 1 + w_{all}}\right) \tag{7}$$

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:

$$E(\underline{x_b}) = p(\text{observe one of the two in } n_b \text{ draws}) \cdot 1 + p(\text{observe both in } n_b \text{ draws}) \cdot 2 \tag{8}$$

$$E(\underline{x_b}) = \frac{1}{\frac{1}{2}(n_b+1)} \left(1 - \prod_{i=0}^{n_b-1} \frac{b_{all}-i-2}{b_{all}-i}\right) + 2\left(1 - \frac{1}{\frac{1}{2}(n_b+1)}\right)\left(1 - \prod_{i=0}^{n_b-1} \frac{b_{all}-i-2}{b_{all}-i}\right) \tag{9}$$

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:

$$\begin{aligned} E(\underline{x_b}) &= \left(\frac{n_{b,B=1}}{n_b} + \frac{n_{b,B=0}}{n_b} \cdot \frac{1}{2}\right) && (10)\\ &\quad \cdot \left[\frac{1}{\frac{1}{2}(n_b+1)} \left(1 - \prod_{i=0}^{n_b-1} \frac{b_{all}-i-2}{b_{all}-i}\right) + \right. && (11)\\ &\quad \left. 2\left(1 - \frac{1}{\frac{1}{2}(n_b+1)}\right)\left(1 - \prod_{i=0}^{n_b-1} \frac{b_{all}-i-2}{b_{all}-i}\right)\right] && (12) \end{aligned}$$

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:

$$\eta = \frac{E(\overline{x_b}) - x_b}{E(\overline{x_b}) - E(\underline{x_b})} \tag{13}$$

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.[10][11]

[10]Note that it is possible for this measure to be less than zero if the particular 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.

[11]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 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.