The Effect of Places on Income and Educational Attainment: 1920-1940

By

James M. Bernard

Dissertation

Submitted in fulfillment of the requirements for the degree of

Doctor of Philosophy in the Department of Economics at Brown University

PROVIDENCE, RHODE ISLAND

May 2018

This dissertation by James M. Bernard is accepted in its present form
by the Department of Economics as satisfying the
dissertation requirement for the degree of Doctor of Philosophy.

Date______________________ ____________________________________
John Friedman, Advisor

Recommended to the Graduate Council

Date______________________ ____________________________________
John Friedman, Reader

Date______________________ ____________________________________
Brian Knight, Reader

Date______________________ ____________________________________
Emily Oster, Reader

Approved by the Graduate Council

Date______________________ ____________________________________
Andrew G. Campbell, Dean of the Graduate School

James M. Bernard

EDUCATION

- **Brown University, Ph.D. Candidate in Economics, expected graduation, May 2018**
- **Brown University, M.A. in Economics, 2014**
- **Swarthmore College, B.A., May 2011**

RESEARCH ASSISTANTSHIPS

- **Brown University, Research Assistant for Professor Brian Knight, 2015-present**
    - 2015-2016: I performed research relating to the efficiency loss that results when public universities offer tuition discounts to in-state but not out-of-state students. This research contributed to a working paper titled, "The Out-of-State Tuition Distortion" by Brian Knight and Nathan Schiff. I helped implement a quasi-experimental estimate of this efficiency loss by comparing the costs of college attendance and the choice of college for students living close to, but on opposite sides of, state borders.
    - 2016-present: In my current work on this project I am developing and structurally estimating a more intricate discrete choice model which incorporates a student's decision regarding which colleges to apply to, the colleges' decision to accept or reject the student, and the student's ultimate choice of college to attend. Ultimately, I will use the model to run policy simulations which investigate the effect on college applications, admissions, and attendance if colleges were to change or eliminate in-state tuition discounts.
- **Wharton School of the University of Pennsylvania, Zell Lurie Real Estate Center, Full-Time Research Assistant, 2011-2013**
    - Assisted Professors Albert Saiz, Grace Bucchianeri, Fernando Ferreira, Todd Sinai, Gilles Duranton and Jesse Handbury.

TEACHING

- **Brown University, Instructor, Intermediate Microeconomics (ECON 1110), Fall 2016**
- **Brown University, Teaching Assistant, 2014-17**
    - Taught ECON 1760 (Financial Institutions) and ECON 1110 (Intermediate Microeconomics).

OTHER

- Swarthmore College varsity baseball player, pitcher and outfielder, 2007-2011

Acknowledgements

I would like to thank the Brown University Department of Economics, and especially John Friedman, Brian Knight, and Emily Oster for their support throughout the research that led to this dissertation.

Table of Contents

The Effect of Places on Income and Educational Attainment: 1920-1940

1 Chapter 1: Intergenerational Mobility in the 1920-1940 United States

1.1 Introduction

A founding premise of the United States holds that every child should enter life with an equal level of economic opportunity. Recent research (Chetty and Hendren 2017a and 2017b, Chetty, Hendren, Kline, and Saez 2014 (CHKS), Chetty, Hendren, and Katz 2016, and Chyn 2016), however, suggests that opportunity varies substantially by the location and level of wealth into which a child is born. Children born to poor families tend to stay poor, and children born to wealthy families tend to stay wealthy. CHKS, however, documents that some parts of the modern United States exhibit substantial economic mobility, while others exhibit very little upward mobility for children born into poor households. Thus, geographic inequality in economic mobility leaves some children disadvantaged relative to others just by an accident of the place their parents chose to raise them. Yet, the United States's national myth posits that it represents a "land of opportunity." How did this myth take such strong hold if the facts do not support it? This paper mirrors the work in CHKS and Chetty and Hendren

2017a using historical data from the 1920-1940 censuses to investigate how the relationship between place and economic mobility differed then from what it is today.

The logic of this paper proceeds in three steps. First, I estimate intergenerational mobility by place (counties or commuting zones) for kids who remain in the same place between 1920 and 1930. Using these estimates, I generate predicted adult income percentiles in the national income distribution (which I call "place predictions" throughout the paper) for children born to parents in the 25th and 75th percentile of the 1920 national income distribution. I then map these estimates and compare them to their equivalent estimates from CHKS. This investigation of historical place effects yields some differences and some similarities with Chetty and Hendren's modern results. While I find that places where below-median-income children were likely to achieve higher adult incomes in the 1920s and 1930s still tend to be good places for below-median-income children, the same is not true for above-median-income children. The places where wealthy children saw greater long-term earnings in the past no longer enjoy similar great outcomes for wealthy children today.

In the second step, I estimate the correlation between these mobility estimates and place-level measures of racial composition, income inequality, and college graduation rate. My finding that the fraction of black residents correlates with less intergenerational mobility mirrors the evidence in CHKS, but unlike their research, I find a positive and significant relationship between intergenerational mobility and racial segregation for below-median-income children. I then perform Oaxaca decompositions to see whether changes in county and commuting-zone-level characteristics explain the change in predicted outcomes for 25th percentile children. I find that these changes do not explain much of the change in predictions, and that instead, changes in the relationship between these variables and upward mobility for below-median children must be a larger driver of the change in upward mobility.

Finally, I estimate the impact of an additional year of exposure to a better place, using children who relocate between 1920 and 1930. Unlike in Chetty and Hendren 2017a, the mover regressions in this paper do not reveal strong evidence that exposure time to better neighborhoods matters. These results do show a large positive relationship between moving to a better neighborhood as a child and adult income, but this effect does not vary by the child's age at the time of the move. This means that either there was no causal benefit of moving to a place with a better place prediction, or that all the causal benefit was conferred immediately upon moving to the new place.

While there is little preexisting literature exploring historical place and neighborhood effects on children, much research has investigated these questions using modern data. The Moving to Opportunity (MTO) Experiment has perhaps provided the best evidence for the effects of moving poor families from high-crime, high-poverty neighborhoods into safer, wealthier neighborhoods. Kling et al. (2007) find that adults who moved because of MTO saw improvements in mental health, but few other benefits. They find more significant but mixed results for children who moved; girls experienced better education outcomes and less risky behavior such as drug use, but boys experienced no improvement in education and an increase in risky behaviors. Kling et al. (2005) also find mixed evidence that indicates girls who moved were less likely to be arrested, yet while boys were less likely to commit violent crimes, they were also more likely to be arrested for property crime. Ludwig et al. (2013) examine longer term results from MTO, and find qualitatively similar impacts to the prior research.

In a recent paper, however, Chetty et al. (2016) examine long term impacts on children who moved through MTO at young ages, and find large positive effects. These children achieve substantially higher incomes as adults, specifically, earning 31% more in the authors' preferred treatment-on-treated specification. Their evidence corroborates the findings of Chetty and Hendren in that they suggest that it is the accumulated effect of exposure to new neighborhoods that drives neighborhood effects rather than the mere fact that a child moves. Chyn (2016), also finds large long-term effects on children who move as a result of housing demolitions, but in contrast to the two Chetty papers, these effects do not appear to vary by differences in children's ages. My paper offers further evidence on the side that the causal effect of places does not depend substantially on exposure time.

This literature, however, has little to conclude about why places matter for children. Good places may improve the lives of children through multiple channels. Possibly the most likely mechanism whereby places influence the children who grow up in them is the difference in access to good schools. Regardless of how one defines a "better" places, those "better" places likely correlate with better school districts or school catchment zones. Given the evidence indicating that good teachers confer massive long-term benefits on children (e.g. Chetty, Friedman, Rockoff 2014), stronger school districts could explain much of the benefits of good neighborhoods by themselves. Alternatively, segregation or neighbor quality could explain why some places cause good outcomes for children, but others do not. Any number of other factors could make a neighborhood a better or worse place to grow up; access to better public facilities, lower crime, cleaner air or water, or more amenities could all impact a child's prospects. Both Chetty and Hendren 2017b and Chetty, Hendren, Kline and Saez assess the (non-causal) correlation between place place predictions and neighborhood characteristics reflecting residential segregation, income inequality, school quality, social capital, and family stability. While it represents a new, important frontier to obtain causal estimates of these effects, the work in this paper also presents only correlational estimates of the relationship between neighborhood characteristics and place predictions for children.

1.2 Data

This paper uses data from the 1920, 1930, and 1940 United States censuses. It uses a panel that links 0-10-year-old males in the 1920 census to the same individuals in the 1930 and 1940 censuses. All three censuses contain information on location, race, marital status, literacy, and occupation among other demographic and economic variables. But only the 1940 census records additional useful information like income, employment status, and highest level of schooling completed.

To investigate intergenerational economic mobility, however, I need a measure of household income for children in 1920. I, therefore, impute the income of the head of household in 1920 using the 1940 census. I do this in two steps. 1) For all men in the 1940 census, I regress income on age and occupation fixed effects. 2) I then take the predicted incomes for each age-occupation cell and apply them to heads of household with the same age and occupation in the 1920 census. The imputation method employed here successfully imputes income for almost 85% of heads of household.

The biggest hurdle to analyzing long-term impacts using census data lies in linking individuals from the 1920 census to themselves in 1930 and 1940. Linking five-year-old John Smith in 1920 to 25-year-old John Smith in 1940 requires determining which of the John Smiths in 1940 is the correct match. In fact, perhaps the correct John Smith appears as "Jon Smith" or John Smit" in 1940 due to a name change or misspelling. Given that there are roughly 15 million boys between ages zero and ten in 1920, this process requires an automatable method for selecting the correct match.

I begin by assembling a list of potential 1930 matches for the boys[1] to be matched from the 1920 census and a separate list of potential matches of boys in the 1930 census to their counterparts in the 1940 census. To do this, I first match each of these boys to all men in all 48 states with the same birth state, race, age (plus or minus one year), and the same first letters of their first and last names. Ideally, the matching process would be more flexible than this. Previous efforts to match individuals in similar contexts (e.g. Feigenbam 2015) have found benefits to considering matches among people with different recorded races and with ages as far as two years apart. Similarly, it would be better not to restrict the set of potential matches to people with the same first letters of their first and last names – this method fails to match anyone who, for instance, goes by "Bill" in 1920, but goes by "Will" in either the 1930 or 1940 censuses. The matching method used here, however, employs these imperfect methods because they dramatically reduce the computing time necessary to assemble a list of potential matches.

The initial step of the matching process yields an enormous number of potential matches, so the next step is to trim this list by using first and last names to determine which names are close enough that they might be a true match. The Jaro-Winkler string similarity measure provides a gauge of whether or not two names likely belong to the same person. This measure takes the value of one if two strings are identical, and decreases towards zero as the two strings become more dissimilar. For example, "Bill" and "Billy" register a Jaro-Winkler value of 0.96 while "Bill" and "William" score 0.73. Using this measure, I remove from the list of potential matches any match in which both the first and last

[1] I focus on boys, because girls are more difficult to match across censuses. While boys names are unlikely to change over time, women were more likely to change their last names due to marriage. The linking method used here would fail to match any women (or men for that matter) who's last names changed between two iterations of the census.

name fail to achieve a similarity score of at least 0.8.

I then join together the two sets of potential matches – the list matching from 1920 to 1930 and the list matching from 1930 to 1940. This leaves a set of potential matches from 1920 to 1930 to 1940 with many children appearing more than once in the list. I then eliminate any potential matches where the match is not the "best" match for a given child from either the 1920-1930 or 1930-1940 matches. To define the "best" match between any two censuses, I first consider whether a given match weakly dominates all others on name similarity. That is, a "best" match between 1920 and 1930 occurs when it has the highest (or weakly highest) Jaro-Winkler similarity score among all potential 1930 matches for the 1920 child. If this method fails to yield a "best" match, then I call any match which weakly dominates on last name a potential match. I then restrict the sample to include only those potential 1920-1930-1940 matches where the match is "best" both between 1920 and 1930 and between 1930 and 1940.

This leaves a data set with only the best possible matches for each child. This is still not an ideal data set, as for some children there are dozens of best matches. Thus, as a final step, I eliminate all 1920 children for whom there are more than three best matches. In the end, this leaves a data set with about 3 million observations, 2 million of which are unique 1920 children, 700 thousand of which are 1920 children with two best matches (meaning 350 thousand unique children), and 228 thousand of which are 1920 children with three best matches (meaning ~75 thousand unique children). This is the sample of children that this paper analyzes.

Table 1 presents the 1920-1930-1940 match rate for each age cohort, race, and for the five largest states in 1920. It shows fairly consistent match rates across groups, but it does show lower match rates for blacks than whites, as well as a lower match rate for Texas than the other states.

It is worth noting here, that there exist more robust algorithms to match individuals between censuses. Feigenbaum (2015) proposes one such method, which relies on a manually-matched “training” data set. By carefully matching a subset of children by hand, this method derives a near-perfect match for this subset. Then, the researcher measures the impact of various rules for determining which pairs are and are not correct linked on how closely the automatically-matched data mimic the manually-matched data. This provides a framework for determining matches in a way that optimally minimizes the probability both of falsely assigning matches to pairs which are not correct matches, and of failing to assign matches to pairs who actually are correct matches. While I have not yet implemented this type of matching process, it is likely that Feigenbaum's, or another similar method, would lead to an improvement in the quality of my matched data.

In a paper looking at the effects of places on children, it is also important to define what places are relevant. In order to include all children in the analysis, an ideal place definition should cover the entire country. At the same time, an ideal place should be as small enough to capture variation by place, but also big enough to include a meaningful sample of children. In this paper, I use two main definitions of place: counties and commuting zones. For counties, I use 1920 county definitions. A small number of counties split up, or merged into several different counties between 1920 and 1930. I exclude these counties from the county level analysis, as it is impossible to tell whether a person in one of these counties moved to a new county by 1930. When considering commuting-zone-level (CZ) results, I use 1990 CZ definitions. These are places which join together one or several counties based on 1990 commuting flows, and span the entire U.S. map. For some of my analysis of movers between either counties or CZs, I exclude from the sample any movers who move to a new county or CZ, but do not move to a new state in the process. This is meant to limit the sample to movers who move across larger distances.

1.3 Empirical Strategy and Results

1.3.1 Measuring Place Effects

The first empirical goal of this paper is to characterize intergenerational mobility by place during the interwar period in the United States. To accomplish this, I begin by measuring average intergenerational mobility throughout the country. To illustrate what this means, figure 1 plots the relationship between parent and child income in two ways. The first panel plots the child's income rank in the 1940 national income distribution for his cohort, by his head of household's imputed decile rank in his cohort of the 1920 national income distribution. The second panel, on the other hand, plots the child's raw income in 1940 against his head of household's raw imputed income in 1920. Both plots reflect a strong positive relationship between parent and child income. While the relationship in terms of raw income is concave[2], the relationship in terms of percentile ranks is fairly linear, especially for children whose parents rank above the bottom decile in the income distribution.

Figures 2 and 3 display rank-rank plots like in the first panel of figure 1, for four specific commuting zones and counties respectively. Figure 2 shows results for Chicago, Philadelphia, New York, and Cleveland, while figure 3 shows results for the corresponding Cook, Philadelphia, New York, and Cuyahoga counties. While these estimations are noisier than either the national average or their equivalent plots in Chetty and Hendren 2017a, on average they display a linear relationship between parent and child income.

[2]And this concavity becomes more extreme if the graph does not trim children whose parent earns more than $1,500 imputed dollars in 1920. The concavity would also probably be more extreme if I were able to observe the actual head of household income in 1920 instead of his imputed income, which is based on occupational median income.

It is worth mentioning here the potential for measurement error in these data. Given that I only observe imputed head of household income instead of his actual income, it is possible that this could introduce some measurement error relative to the individual's true income. This measurement error could, in turn, bias the estimations in figures 1-3. This would make the slopes in figures 1-3 too flat, essentially making society look more intergenerationally mobile than it was. On the other hand, even if I could observe the true 1920 income for each head of household, it is possible that this variable could contain even more measurement error than the imputed income variable. This is because one year's income migth actually be a noisier measure of a person's average income than the average income for someone with his occupation. For this reason, the estimates in this paper may not be directly comparable to those in CHKS. Nonetheless, the results here show a linear, positive relationship between head of household and child income, and also, as I show below, reflect substantial heterogeneity in this relationship across places in the interwar United States.

Exploiting this linearity, I measure what I call "intergenerational immobility coefficients" by commuting zone and county. Limiting the sample to children who remain in the same place from 1920 to 1930, I regress child percentile rank on parent percentile rank by place to obtain a coefficient reflecting how much a child's expected income rank improves for a one percentile increase in his parent's rank. This coefficient essentially measures intergenerational immobility as higher values reflect a more persistent relationship between parent and child income. Because, furthermore, the rank-rank relationship tends to be linear, I calculate the expected adult income rank for a child born to 25th percentile parents in each specific place by adding the place-specific intercept from this regression to 25 times the place-specific coefficient. Similarly, adding the place-specific intercept to 75 times the place-specific coefficient yields the prediction for a child born to 75th percentile parents.

Tables 3 and 4 list the ten best and worst of the 50 largest commuting zones and counties (by total 1920 population) to have been born to 25th percentile parents. They show substantial variation in predicted outcomes throughout the country. Figures 4-7 illustrate these place effects for the entire country. They display these estimates in a heat map of the 48 states. Figures 4 and 5 plot the estimates for children with 25th percentile and 75th percentile parents respectively for all commuting zones with at least 25 children in my data who remain in the zone from 1920 until 1930. Figures 6 and 7 do the same except for the county level estimates, also limiting to counties with at least 25 stayers. All four heat maps show substantial geographic heterogeneity in place predictions, with the western places tending to produce better outcomes, and southern locations tending to produce worse outcomes for both above and below-median income children. The results for children born into above-median families, however, differ in meaningful ways from the results for children born into below-median families. For instance, deep southern states like Georgia, Mississippi, and Louisiana show some of the worst outcome for below-median children, but better outcomes for above-median children.

These maps mirror the results in Chetty, Hendren, Kline, and Saez in several ways; both the historical and modern maps indicate that much of Michigan, Illinois, and eastern Minnesota have good child outcomes. Yet, the maps also differ in significant ways from their modern counterparts. The difference is especially stark in much of the West, where the census data suggests both poorer and wealthier children tend to do well, but CHKS find relatively poor outcomes in many parts of California, Nevada, and Arizona. To formalize the differences between the historical place effects and the modern, I regress each of these four place-level estimates on their equivalents from CHKS. That is, table

5 regresses the historical county-level estimates for 25th percentile children on the CHKS estimates for 25th percentile children, and does the same for 75th percentile children, as well as for commuting-zone level estimates. The results for both commuting zone and county-level estimates show that the places that were associated with good outcomes for 25th percentile children in the first half of the 20th century still tend to be associated with good outcomes for these children today.

The estimates for the 75th percentile children, however, indicate the opposite. They suggest that places that were associated with good outcomes for wealthier children tend to be associated with worse outcomes today. These differences appear to be driven especially by children in southern states such as Tennessee, Kentucky, and Missouri and western states such as California, Nevada, and Arizona. In the historical map, the southern states mentioned above are among the worst places for above-median children, and the western states are among the best. The reverse is true for both areas in the modern maps. One potential explanation would be that places where wealthy kids tended to become wealthy adults may have exhibited less intergenerational mobility overall, which may have been correlated with lower growth. This could, in turn, lead these areas to be correlated with worse outcomes in the modern world.

1.3.2 Correlates of Place Effects

Having estimated the effects of place on children, I next consider why some places have better predicted outcomes than others. While this paper does not provide a causally-estimated answer to this question, it does investigate how place effects correlate with racial segregation and 1920 labor market conditions.

Specifically, tables 6 and 7 report place-population-weighted regression results of place-level adult income rank predictions for 25th and 75th percentile children on four place-level 1920 characteristics: the fraction of residents who were black, the dissimilarity of racial segregation, the income Gini coefficient (calculated from imputed 1920 data), and the fraction of residents with a college degree (calculated from 1940 data). The dissimilarity index compares the relative proportions of blacks and whites of each enumeration district within a county or CZ to the relative proportions in the whole county or CZ[3].

Both tables 6 and 7 display qualitatively similar results. All four characteristics display strong correlation with place predictions for both below median children. All characteristics except for dissimilarity display a similar correlation with place predictions for above-median children. The coefficient on the fraction of place residents who graduated from college, however, not only suffers from omitted variable concerns, but the fact that it was calculated from the 1940 data introduces an almost-mechanical reverse causality as well. The estimate of the correlation between place predictions and the fraction of black residents in 1920, on the other hand, does not suffer from reverse causality. I find that a high fraction of black residents is associated with a decrease in predicted outcomes for below-median children and an increase in predicted outcomes for above-median children. This is consistent with the idea that more black residents correlates with less intergenerational mobility. The fact, also, that the dissimilarity index correlates positively with outcomes for below-median children, but not above-median children, is consistent with the argument in Bernard (2017), which shows that segregation is correlated with positive outcomes for African Americans (who would be especially likely to be born to below-median-income heads of household) in 1940.

Tables 8 presents Oaxaca decompositions aimed at characterizing the impact of changing local racial compositions, income segregation, and schooling, on

[3]The dissimilarity indices here are the same as those used in Bernard (2017). For more detail on how they were calculated, see the data section of that paper.

place predictions for children of 25th percentile parents. Column 1 of each table reports the decomposition for the 25 percentile county predictions using all four covariates except dissimilarity[4]. Column 2 does the same for the 75th percentile county predictons. Columns 3 and 4 repeat the same exercise for commuting zones instead; they report the average place prediction for 25th percentile children in the CHKS data and in the census data. They then estimate how much of the difference between these two numbers can be explained by changing covariates, and, conversely, how much cannot. Finally, the bottom two panels of each table break down how each of the three covariates affects the explained and unexplained variation between the historical and modern estimates of place effects for 25th percentile children.

The table shows that changing covariates do not explain changing place effects in the country as a whole. The one variable which appears to explain any portion of the change in place effects is the fraction of black residents. The results in column 1 show that the change in the fraction of black residents can explain about 10 percent of the change in county predictions for 25th percentile children, but the results in the columns 2-4 do not suggest it has any explanatory power over the change in county predictions for 75th percentile children, or commuting zone predictions for either 25th or 75th percentile children. Overall, the only conclusion to draw from this table is that the changes place-level Gini coefficients, fraction black residents, and fraction college graduates do not, alone, explain why place predictions have improved for some places and declined for others. Either the effect of these variables on place predictions must have changed between 1920 and today, or other place characteristics must have changed.

[4]CHKS do not calculate dissimilarity indices in their modern data

1.3.3 Mover Regressions

The results above suggest that there existed substantial heterogeneity in outcomes for children throughout the country. This, however, does not prove that different places causally impacted outcomes, as the differences across places could have arisen merely because similar types of people sorted together. To investigate whether, and to what degree, exposure to places causally impacted child outcomes, I follow along the lines of Chetty and Hendren 2017a by exploiting children who move with their families between 1920 and 1930. I define place quality for a child of a given age and head of household decile rank as the predicted percentile rank in adulthood for children of that age and head income decile who spend their entire childhood in that place. This leads to a measurement of the change in place quality for all children who moved between 1920 and 1930. Then, the coefficient arising from a regression of child outcomes on his change in place quality reflects the benefit of moving to a place with one percentile better outcomes.

This strategy, however, still allows the possibility that selection could upwardly bias the estimates of place effects. While movers might have less ability to sort into certain places than better-informed locals, there is still concern that, for instance, wealthier movers could see systematically different changes in place quality than poorer movers. To make the mover regressions more robust, Chetty and Hendren include family fixed effects in their estimations. Comparing siblings who moved at different ages, they recover the impact of an additional year of exposure to a new place. The identification assumption for this estimation is much weaker: families can select into different types of neighborhoods, it just cannot be that the magnitude of this selection varies by the age of the child at the move.

These results indicate that there is some negative selection on literacy and especially on income. This suggests that wealthier and better-educated families tend to move to worse neighborhoods than those in which they began. If families were more likely to move subsequent to negative income shocks, it would explain this pattern. The results further suggest that the magnitude of this negative selection is larger for families who move with older children than those who move with younger children. While this goes against the identification assumption, the magnitude of this difference in selection, and the magnitude of selection for all ages, is very small relative to the standard deviations in the data. A one standard deviation change in income or literacy correlates to roughly one tenth of a standard deviation change in the change in place quality between the old and the new place. This relationship, furthermore, is even weaker for literacy.

Figures 8 and 9 present visual evidence on the validity of this identification assumption. These present the results of two sets of regressions meant to assess the magnitude of selection by child age in 1920. They present the coefficients by age group from the following regression:

$$y_i = \alpha + \beta x_i + \epsilon_i$$

Where y_i is the z-scored change in place quality for child i who moves, and x_i is either head of household income decile or a dummy equal to 1 if the head of household is literate (these are also both z-scored). These coefficients reflect whether movers select into different changes in neighborhood quality based on their income and literacy. They are z-scored to make the interpretation of the coefficients more intuitive. A coefficient of 1 would mean that a one standard deviation increase in, for instance, head of household income leads to a 1 standard deviation increase in the change in neighborhood quality variable.

Next I present results from two regression specifications. The columns 1, 2, 4, and 5 of each table show results from this specification:

$$y_i = \alpha_f + \sum_a \alpha_a + \sum_a \alpha_a p_i + z_{opa} + z_{dpa} + \beta \Delta z_{odpa} * (11 - a) + \epsilon_i,$$

where y_i is the child's income percentile in 1940, α_f is a family fixed effect, α_a is an indicator equal to 1 if a child is of age $a \in \{0, 1, 2, \ldots, 10\}$ in 1920, z_{dpa} is the average outcome in destination d for a mover of age a and parental income decile p, and Δz_{odpa} is the change in predicted place quality for a mover of age a from orgin o to destination d, of parental income decile p. Predicted place quality is the average adult income percentile for children of a given cohort and parental income decile who stay in a given place between 1920 and 1930. $(11-a)$ gives the maximum number of years that a child could have been exposed to the new neighborhood between the ages of 0 and 10. Thus, β reflects the impact of one additional year of exposure to a one percentile better place.

The third and sixth columns of each table run a slightly different specification:

$$y_i = \sum_a \alpha_a + \sum_a \alpha_a p_i + z_{opa} + \sum_a b_a \Delta z_{odpa} \alpha_a + \epsilon_i.$$

This is just a less parametric version of the first equation, which allows exposure effects to be nonlinear by age in 1920. The coefficients b_a reflect the benefit of moving to a one percentile better place at age a. If greater exposure to better neighborhoods improves outcomes, then we would expect greater coefficients, for instance, for $a = 0$ or $a = 1$ than for $a = 9$ or $a = 10$.

These two regression specifications closely mirror the estimations in Chetty and Hendren (2017a). The only difference in the specifications here arises from the fact that the census data do not allow me to identify when a child moved. Chetty and Hendren also interact the origin and destination place predictions,

z_{opa} and z_{dpa} with age-at-move dummies, but here I cannot distinguish a child's age-at-move from his cohort. Therefore, interacting place effects with cohort would be perfectly collinear with $\Delta z_{odpa} * (11 - a)$ or $\sum_a \Delta z_{odpa}\alpha_a$, so I can only include z_{opa} and z_{dpa} without interacting them[5].

Tables 9 and 10 present the results from these estimations for CZ and county movers respectively. The first three columns of each table present results from the sample of all movers across county or commuting zone lines, while columns 4-6 limit the sample to "far movers" who also cross state lines as part of their move. Columns 2 and 5 of each table include family fixed effects. There are strong positive coefficients on every cohort-change-in-place-quality interaction in columns 3 and 6 of both tables. This unsurprisingly suggests that moving to better places is associated with better outcomes for children.

The results in this table do not, however, suggest that exposure time to better neighborhoods matters for children's long-term outcomes. In columns 3 and 6 of each graph, exposure effects would appear as greater coefficients on the change in neighborhood quality for kids who move at young ages than at older ages. This, however, is not the case. Figure 10 plots these coefficients (the coefficients from column 1 of each table) by child age in 1920 to clearly illustrate that these coefficients appear to stay roughly constant regardless of the child's 1920 age.

Columns 1-2 and 4-5 of both tables corroborate the story from columns 3 and 6; there is no consistent relationship between child outcome and the interaction of change in place quality and exposure time to the new place. In table 10, for instance, this coefficient is negative in column 1, but positive for columns 2, 4, and 5, and statistically insignificant in every column. For both county and commuting zone movers, the coefficient is positive but insignificant when

[5]The second of these equations also excludes the destination place prediction, z_{dpa}, because it is collinear with z_{opa} and $\sum_a \Delta z_{odpa}\alpha_a$. This is not a problem for Chetty and Hendren because they can differentiate birth cohorts from age groups.

restricting to the sample of movers moving across state lines. Thus, it is possible that exposure matters more for those individuals making a larger move, while it matters less for those making smaller moves. Given the large standard errors here, however, the conclusion here is that there is no evidence that exposure to places with better place-level predictions mattered for children in the interwar United States.

1.4 Conclusion

This paper replicates much of CHKS and Chetty and Hendren's work for the United States in the 1920s and 1930s. While it finds many qualitative similaries in the geographic variation in intergenerational mobility between the past and present, it also finds meaningful differences between the two time periods. Many of the places that saw good average outcomes for below-median-income children in 1920-1940 still see good outcomes for these children, but the same is not true for places which saw good outcomes for above-median children. Like the modern findings, these data suggest that the fraction of black residents in a place negatively correlates with intergenerational mobility, but unlike in the modern world, segregation was associated with better outcomes for poor children.

Finally, unlike in Chetty and Hendren 2017a, the results here fail to produce strong evidence that the effect of moving to a better place varies by the length of exposure to the new place. Thus, I can conclude from the investigation of movers in section 3.3, that either place had a smaller causal effect on children's long-term outcomes in the 1920s and 1930s than it does today, or that the full benefit of moving to a better place were immediately conferred upon moving to the new place in the 1920s. If the second explanation is true, it might be explained by changes in the importance of human capital in the labor market. Perhaps the main benefit of growing up in a better place today derives from greater human capital formation, so it is best to spend as much of one's childhood as possible in a place with access to good schools. It is possible, though, that education was less necessary in the more labor-intensive, industrial economy of the 1920s and 1930s. In this case, much of the benefits of better places would come from better access to jobs or higher wages in the labor market. Then, it would make little difference whether a child moved to a better place at the age of two or nine - either way, he could enjoy the full labor market benefits of his new location. In fact, if this were true, then it might have been better to move to a better place at age nine than age two, if moving at age nine meant a child would have been less likely to move out of the new, better, place before he entered the labor market.

In the historical research, as in the research that focuses on the modern world, it is unclear what specific aspects of places make them better places to grow up in than others. What is clear from the research here, and from the research in Bernard (2017), is that not all of what we know about good and bad places in the modern world was necessarily true in the historical United States. Both the place characteristics that are associated with good place outcomes and the influence on children of exposure time to different places, appear to have changed since the 1920s and 1930s. In both the historical and modern data, a focus of future research should be isolating exactly what it is that makes some places better for children, and why. If we can learn what it is about some places that allows poor children economic opportunities that other places do not, perhaps we can develop policies which help turn all places into lands of opportunity.

1.5 References

Bayer, Patrick, Steve Ross, and Giorgio Topa. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," Journal of Political Economy, 116.6 (2008): 1150-96.

Bernard, James. "Segregation, the Great Migration, and African American Outcomes." (2017).

Chetty, Raj, John Friedman, and Jonah Rockoff. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates" , American Economic Review 104.9 (2014): 2593–2632.

Chetty, Raj, Nathaniel Hendren. The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. NBER Working Paper No. 23001 (2017).

Chetty, Raj, Nathaniel Hendren. The Impacts of Neighborhoods on Intergenerational Mobility II: County Level Estimates. NBER Working Paper No. 23021 (2017).

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," American Economic Review 106.4 (2016).

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States", Quarterly Journal of Economics 129.4 (2014): 1553-1623.

Chyn, Eric. "Moved to Opportunity: The Long-Run Effect of Public Housing Demolition on Labor Market Outcomes of Children." Working, (2016).

Collins, William J. and Marianne H. Wanamaker. "Selection and Economic Gains in the Great Migration of African Americans: New Evidence from Linked Census Data." American Economic Journal: Applied Economics 6.1 (2014): 220-52.

Currie, Janet and Aaron Yelowitz. "Are Public Housing Projects Good for Kids?" Journal of Public Economics 75.1 (2000): 99-124.

Dahl, Gordon B. "Mobility and the return to education: Testing a Roy model with multiple markets." Econometrica 70.6 (2002): 2367-2420.

Feigenbaum, James J. "Automated Census Record Linking." (2015).

Galster, George C. "The mechanism (s) of neighbourhood effects: Theory, evidence, and policy implications." Neighbourhood effects research: New perspectives. Springer Netherlands, (2012): 23-56.

Gibbons, Stephen, Olmo Silva, and Felix Weinhardt. "Everybody needs good neighbours? Evidence from students' outcomes in England." The Economic Journal 123.571 (2013): 831-874.

Jacob, B.A. "Public housing, housing vouchers, and student achievement: Evidence from public housing demolitions in Chicago" American Economic Review, 94 (2004): 233–258

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. "Experimental analysis of neighborhood effects." Econometrica 75.1 (2007): 83-119.

Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz. "Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment." The Quarterly Journal of Economics (2005): 87-130.

Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu. "Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity." American Economic Review, 103(3) (2013): 226-31.

Logan, J. R., Spielman, S., Xu, H., & Klein, P. N. "Identifying and bounding ethnic neighborhoods." Urban geography, 32.3 (2011), 334-359.

Minnesota Population Center and Ancestry.com. IPUMS Restricted Complete Count Data: Version 1.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2013.

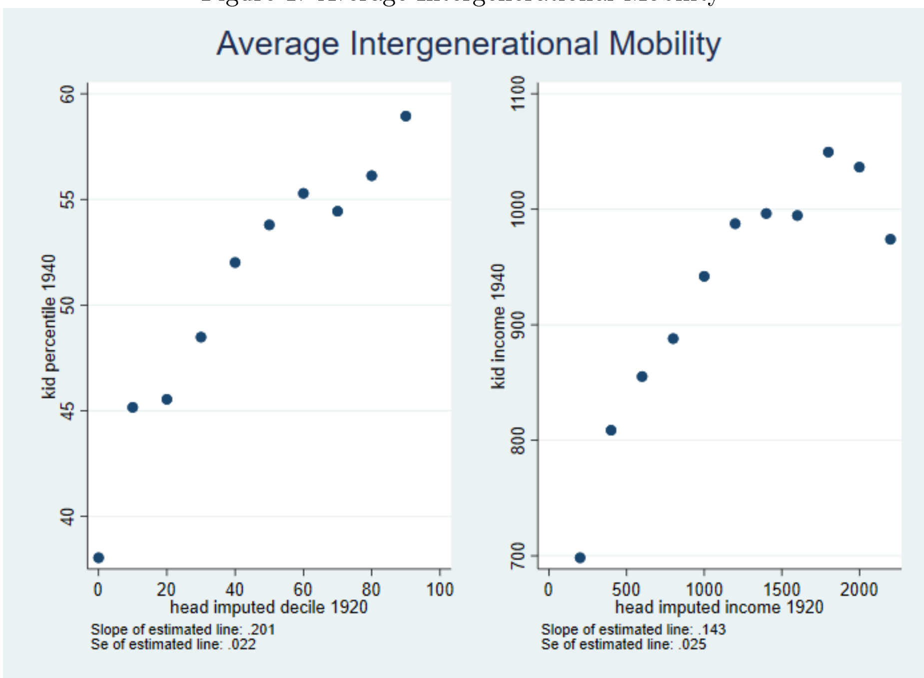
Oreopoulos, Philip. “The long-run consequences of living in a poor neighborhood.” The quarterly journal of economics (2003): 1533-1575.

Reardon, Sean F. and Glenn Firebaugh. “Measures of Multigroup Segregation." Sociological Methodology, 32.1 (2002): 33{67.

Wodtke, Geoffrey T, David J Harding, and Felix Elwert. “Neighborhood Effects in Temporal Perspective: The Impact of Long-term Exposure to Concentrated Disadvantage on High School Graduation.” American Sociological Review, 76.5 (2011): 713–736.

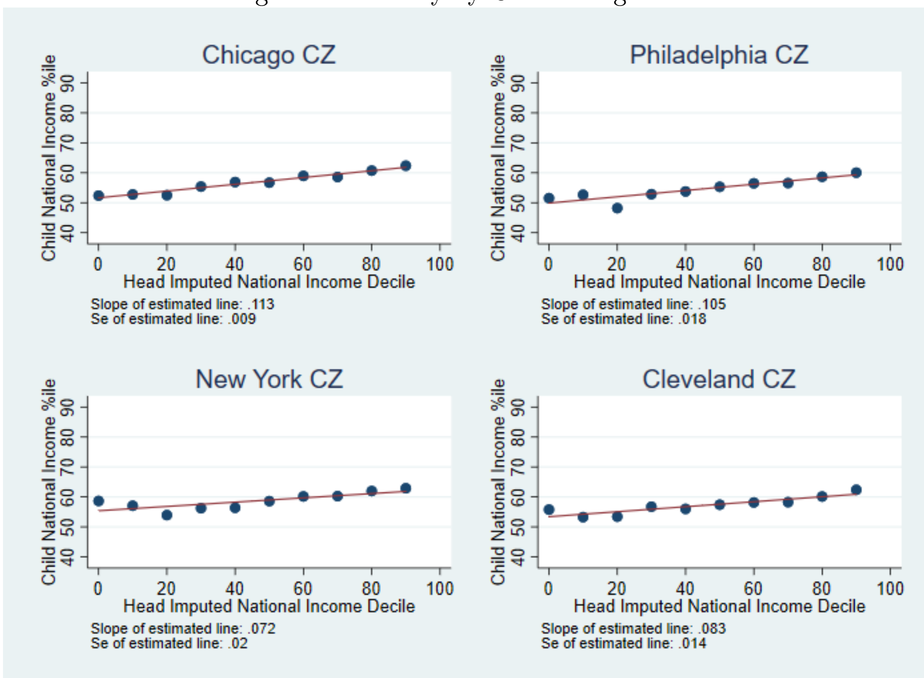
1.6 Tables and Figures

Figure 1: Average Intergenerational Mobility



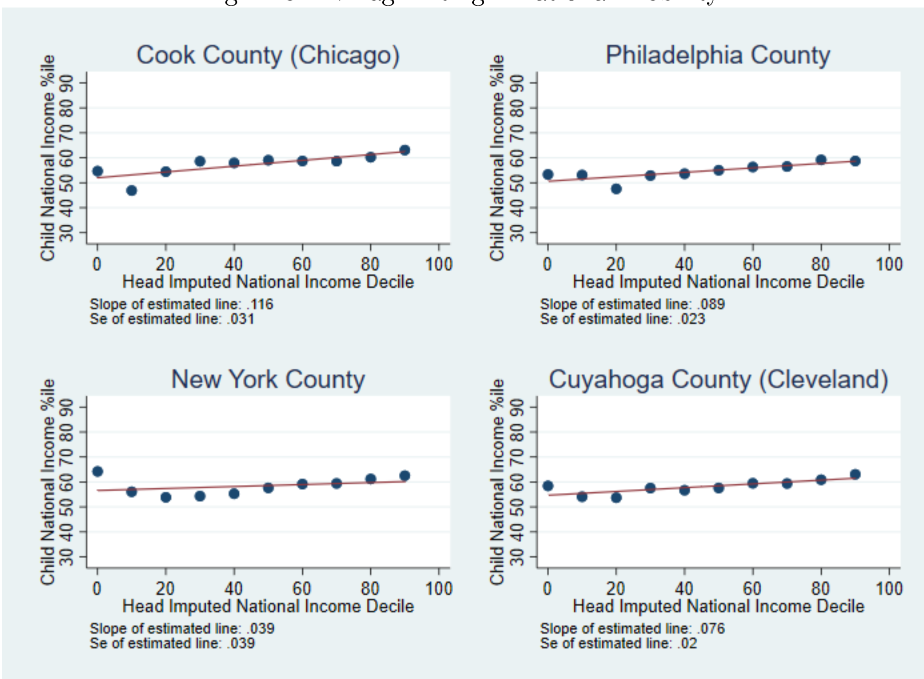
Note: The first figure plots the child's percentile rank in the 1940 national income distribution by his head of household's imputed decile rank in the 1920 national income distribution. The second figure plots average child raw income in 1940 by the head of household's binned raw imputed income in 1920.

Figure 2: Mobility by Commuting Zone



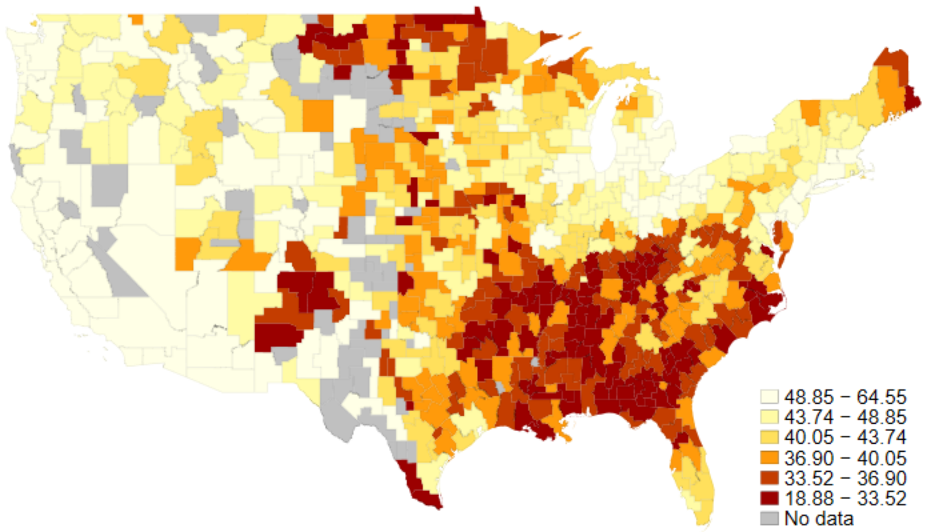
Note: These plots show average child percentile rank in the 1940 national income distribution by binned head of household decile rank in the 1920 national income distribution for four of the most populous commuting zones in 1920.

Figure 3: Average Intergenerational Mobility



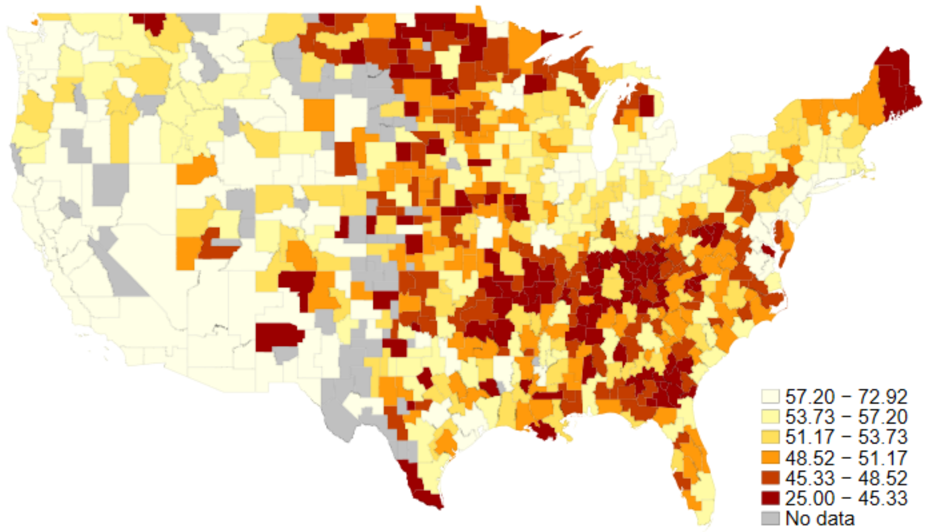
Note: Note: These plots show average child percentile rank in the 1940 national income distribution by binned head-of-household decile rank in the 1920 national income distribution for four of the most populous counties in 1920.

Figure 4: Predicted Income Rank for Children Born into the 25th Percentile (By CZ)



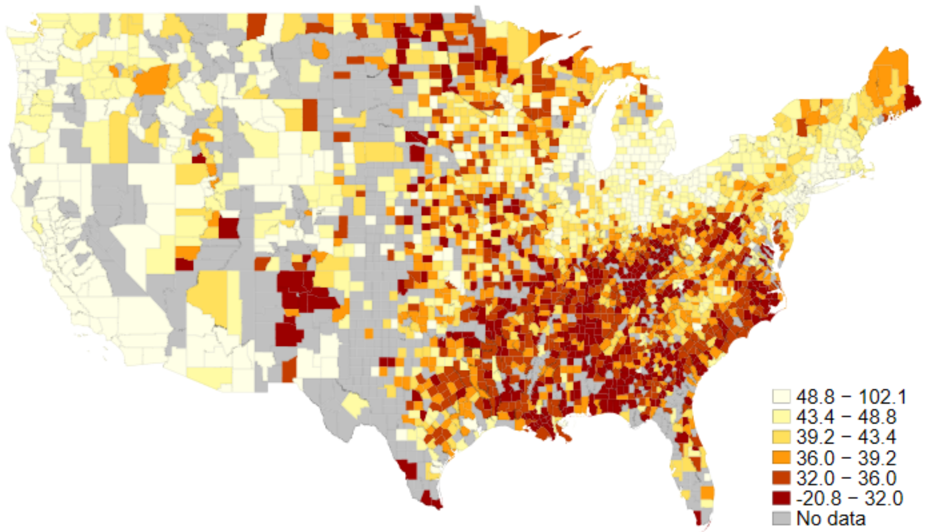
This is a heat map reflecting the predicted adult income percentile rank by CZ for children born into a 25th percentile household who remain in one CZ from 1920 to 1930. This prediction is based on a regression of child percentile rank on head of household imputed income decile by place. Section 3.1 discusses the estimation strategy at greater length.

Figure 5: Predicted Income Rank for Children Born into the 75th Percentile (By CZ)



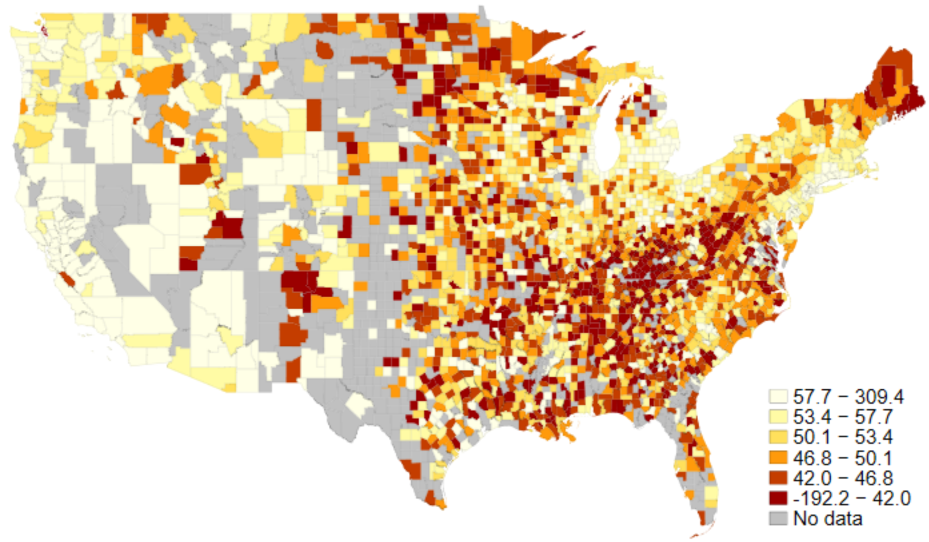
This is a heat map reflecting the predicted adult income percentile rank by CZ for children born into a 75th percentile household who remain in one CZ from 1920 to 1930. This prediction is based on a regression of child percentile rank on head of household imputed income decile by place. Section 3.1 discusses the estimation strategy at greater length.

Figure 6: Predicted Income Rank for Children Born into the 25th Percentile (By County)



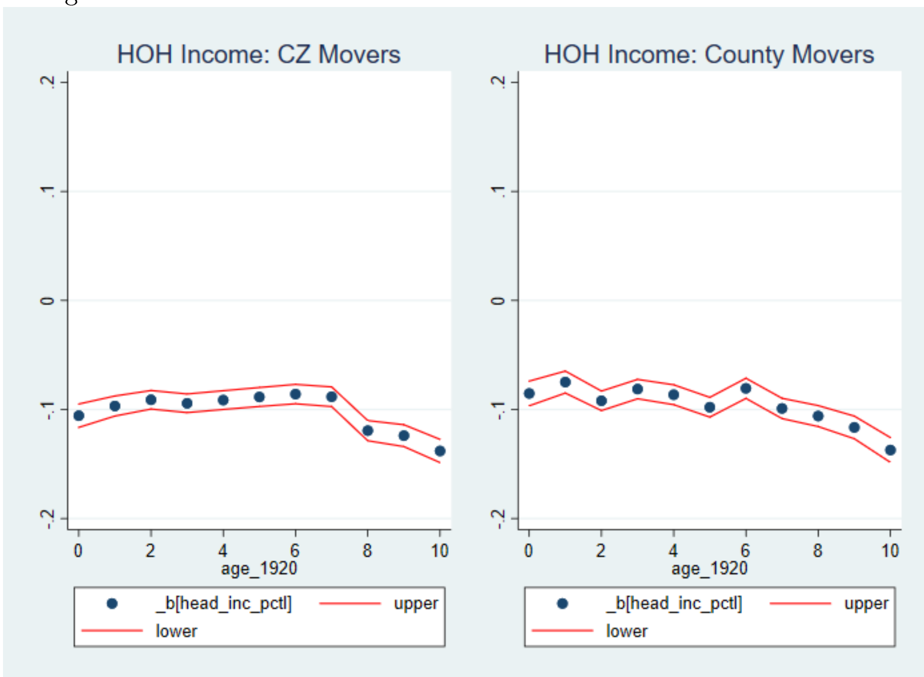
This is a heat map reflecting the predicted adult income percentile rank by county for children born into a 25th percentile household who remain in one county from 1920 to 1930. This prediction is based on a regression of child percentile rank on head of household imputed income decile by place. Section 3.1 discusses the estimation strategy at greater length.

Figure 7: Predicted Income Rank for Children Born into the 75th Percentile (By County)



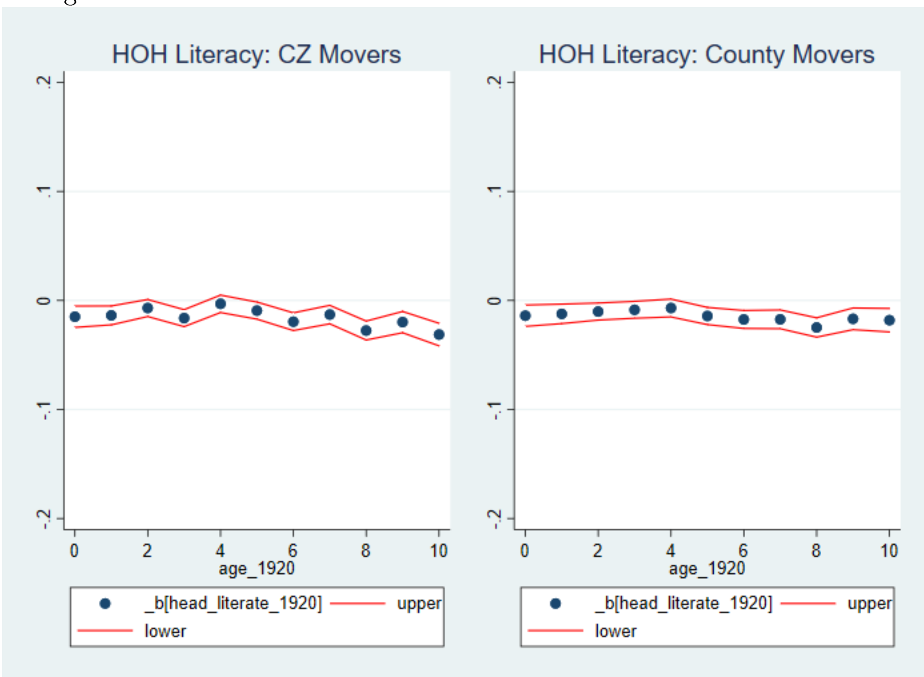
This is a heat map reflecting the predicted adult income percentile rank by county for children born into a 75th percentile household who remain in one county from 1920 to 1930. This prediction is based on a regression of child percentile rank on head of household imputed income decile by place. Section 3.1 discusses the estimation strategy at greater length.

Figure 8: Evidence on Selection: Relationship between Income and Place Change



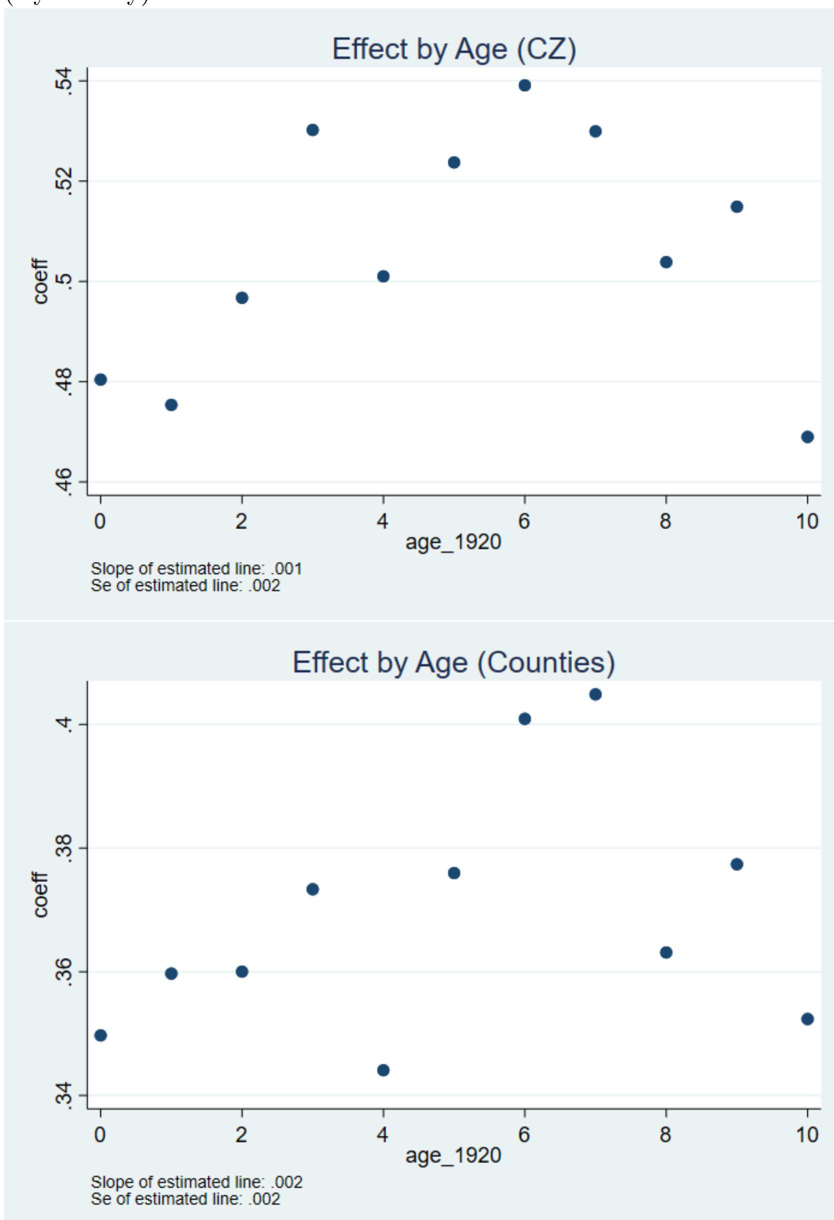
These figures plot coefficients from a regression of z-scored change in place quality on z-scored head of household income percentile by age group for children who move between 1920 and 1930. The first figure focuses on children who move from one CZ to another, and the second figure focuses on those who move from one county to another. The "upper" and "lower" lines plot 95% confidence bounds on the coefficents.

Figure 9: Evidence on Selection: Relationship between Literacy and Place Change



These figures plot coefficients from a regression of z-scored change in place quality on z-scored head of household literacy by age group for children who move between 1920 and 1930. The first figure focuses on children who move from one CZ to another, and the second figure focuses on those who move from one county to another. The "upper" and "lower" lines plot 95% confidence bounds on the coefficents.

Figure 10: Predicted Income Rank for Children Born into the 75th Percentile (By County)



These two figures plot the coefficients in column 1 of tables 9 and 10 respectively. They give the effect of moving to a one percentile better place at each age from 0-10.

Table 1: Summary Stats on Match Quality

		Match Rate
age in 1920		
	0	0.27
	1	0.28
	2	0.28
	3	0.27
	4	0.27
	5	0.27
	6	0.26
	7	0.26
	8	0.25
	9	0.25
	10	0.21
5 Largest States:		
	New York	0.26
	Pennsylvania	0.24
	Illinois	0.24
	Ohio	0.26
	Texas	0.19
Race:		
	White	0.26
	Black	0.22

Table reports the percentage of children in each category in the 1920 census for whom my algorithm successfully assigns a match in 1930 and 1940.

Table 2: Summary Stats of Key Variables

	Stayers	SD	County Movers	SD	CZ Movers	SD
Number	2,141,497	.	1,040,171	.	774,030	.
Head Income	998.2891	661.0147	974.8042	671.1149	957.1839	681.5349
Head literate	.9263179	.2612529	.9121558	.2830683	.9133987	.2812501
Head Inc Pctl	45.98042	26.61772	44.78532	26.75104	43.74719	27.29146
Child Income	903.712	677.1102	882.859	682.6046	861.8028	682.0823
Child Inc Pctl	49.8267	28.86287	48.64476	29.043	47.27984	29.10921
Chg Ct Qual	.	.	1.547966	10.77327	.	.
Chg CZ Qual	.	.	.	.	1.337346	8.834328

Summary stats presented for stayers and movers separately. Stayers are children who remained in the same county between 1920 and 1930. County movers are those who moved to a new county between 1920 and 1930. CZ movers are the subset of county movers who moved to a county in a new commuting zone between 1920 and 1930.

Head income and income percentile are calculated using a measure of income imputed by occupation and age using the 100 percent sample of the 1940 census. Child income percentile is calculated using the 100 percent sample of the 1940 census. The change in place quality variables measure the change in the predicted outcome between origin and destination places for a child born to a head of household in a given income percentile.

Table 3: Best and Worst Commuting Zones

Place	State	Prediction for 25th percentile child	Prediction for 75th percentile child
Best CZs:			
Oakland-Fremont-Hayward	California	65.35	66.5
Los Angeles	California	61.25	64.14
Flint	Michigan	59.67	62.09
New York	New York	56.67	58.76
New York	New York	56.28	58.95
Oak Harbor	Washington	56.27	59.34
Bridgeport-Stamford	Connecticut	56.12	58.11
Cleveland	Ohio	55.75	57.87
Chicago	Illinois	55.53	58.16
Denver	Colorado	55.22	58.47
Worst CZs:			
Nashville	Tennessee	37.45	52.86
Atlanta	Georgia	39.09	55.01
Memphis	Tennessee	39.14	59.31
New Orleans	Louisiana	40.33	51.44
Birmingham	Alabama	40.66	55.2
Dallas	Texas	41.19	55.7
Duluth	Minnesota	41.56	47.74
Bloomsburg-Berwick	Pennsylvania	42.56	51.54
Louisville	Kentucky	43.25	51.54
Bedford	Virginia	44.02	45.47

The top panel presents the 10 of the 50 largest 1920 commuting zones with the highest predicted income percentile for children born to 25th percentile parents. The bottom panel presents the 10 of the 50 largest with the worst predictions for these children.

Table 4: Best and Worst Counties

Place	State	Prediction for 25th percentile child	Prediction for 75th percentile child
Best counties: San Francisco	California	67.16	67.16
Alameda	California	63.67	65.71
Los Angeles	California	61.9	64.53
Bronx	New York	60.73	60.7
Wayne	Michigan	60.32	62.23
Queens	New York	60.02	61.21
King	Washington	58.22	60.2
Hartford	Connecticut	58.22	60.2
Essex	New Jersey	58.22	60.28
Monroe	New York	57.79	58.97
Worst counties:			
Lackawanna	Pennsylvania	41.04	47.59
Orleans	Louisiana	42.06	51.94
Jefferson	Alabama	42.89	56.04
Luzerne	Pennsylvania	43.45	45.49
Jefferson	Kentucky	46.35	52.42
Bristol	Massachusetts	46.35	52.42
Westmoreland	Pennsylvania	47.31	50.93
Onondaga	New York	50.3	52.42
Providence	Rhode Island	50.33	52.36
Allegheny	Pennsylvania	50.97	54.35

The top panel presents the 10 of the 50 largest 1920 counties with the highest predicted income percentile for children born to 25th percentile parents. The bottom panel presents the 10 of the 50 largest with the worst predictions for these children.

Table 5: Regressions of 1920-1940 estimates on CHKS Estimates

	(1) p25 cty	(2) p75 cty	(3) p25 cz	(4) p75 cz
CHKS p25 ct	0.325***			
	(0.0335)			
CHKS p75 ct		-0.365***		
		(0.0522)		
CHKS p25 cz			0.203***	
			(0.0526)	
CHKS p75 cz				-0.362***
				(0.0639)
_cons	27.11***	71.24***	32.27***	73.08***
	(1.461)	(3.134)	(2.331)	(3.863)
N	2362	2362	664	664

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The left-hand side variables are the 25th and 75th percentile place predictions for 1920 counties and commuting zones. The right-hand side variables are the equivalent predictions from CHKS.

Table 6: Relationship between CZ Predictions and Place Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	p25	p25	p25	p25	p25	p75	p75	p75
frac black	-19.10***				-4.405**			
	(1.526)				(1.500)			
dissim		7.132**				-0.0359		
		(2.200)				(1.975)		
frac coll			583.4***				467.9***	
			(30.63)				(28.64)	
gini				-116.1***				-63.98***
				(3.759)				(4.584)
N	675	675	675	675	675	675	675	675

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

This table presents the regressions of the 25th and 75th percentile place predictions for 1920 commuting zones on four 1920 CZ characteristics. The right-hand side variables are the fraction of the CZ that was black in 1920, the 1920 CZ dissimilarity index of racial segregation, the 1940 fraction of the CZ population that had at least four years of college, and the 1920 Gini coefficient computed using imputed income.

Table 7: Relationship between County Predictions and Place Characteristics

	(1) p25	(2) p25	(3) p25	(4) p75	(5) p75	(6) p75	(7) p75	(8) p75
frac black	-18.13***				11.32***			
	(0.847)				(1.062)			
dissim		11.65***				-3.583***		
		(0.752)				(0.924)		
frac coll			536.5***				437.7***	
			(17.12)				(22.16)	
gini				-112.8***				-39.70***
				(3.155)				(4.527)
N	2362	2362	2362	2362	2362	2362	2362	2362

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

This table presents the regressions of the 25th and 75th percentile place predictions for 1920 counties on four 1920 county characteristics. The right-hand side variables are the fraction of the county that was black in 1920, the 1920 county dissimilarity index of racial segregation, the 1940 fraction of the county population that had at least four years of college, and the 1920 Gini coefficient computed using imputed income.

Table 8: Oaxaca Decompositions

	(1) Cty: 25th	(2) Cty: 75th	(3) CZ: 25th	(4) CZ: 75th
overall				
group_1	42.33***	59.08***	43.09***	59.68***
	(0.134)	(0.102)	(0.211)	(0.154)
group_2	40.18***	49.44***	41.15***	51.21***
	(0.178)	(0.256)	(0.296)	(0.264)
difference	2.152***	9.644***	1.936***	8.473***
	(0.223)	(0.276)	(0.364)	(0.305)
explained	-1.106***	-1.225***	-2.169***	-2.024***
	(0.176)	(0.201)	(0.319)	(0.229)
unexplained	3.258***	10.87***	4.104***	10.50***
	(0.207)	(0.263)	(0.341)	(0.322)
explained				
gini	-1.205***	-1.052***	-2.183***	-1.832***
	(0.134)	(0.157)	(0.272)	(0.230)
frac black	0.259***	-0.0219	0.113	-0.0722
	(0.0652)	(0.0536)	(0.0690)	(0.0467)
frac coll	-0.160***	-0.151***	-0.0984	-0.120**
	(0.0323)	(0.0306)	(0.0526)	(0.0450)
unexplained				
gini	20.46***	13.81***	32.96***	19.89***
	(1.732)	(3.829)	(2.606)	(3.015)
frac black	-0.839***	-0.742*	-2.104***	-1.387***
	(0.134)	(0.348)	(0.208)	(0.229)
frac coll	-14.59***	-15.05***	-12.44***	-12.37***
	(0.731)	(0.995)	(0.915)	(1.031)
N	3754	3754	1256	1256

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
This table presents Oaxaca decompositions of the change in place predictions using three of the variables from tables 6 and 7.

Table 9: CZ Mover Regressions

	(1) Spec 1	(2) Spec 1 w/ family FE	(3) Spec 2	(4) Spec 1	(5) Spec 1 w/ family FE	(6) Spec 2
chg x exp	-0.00123	0.000324		0.00130	0.000891	
	(0.00146)	(0.00259)		(0.00227)	(0.00523)	
0 x change			0.483***			0.479***
			(0.0173)			(0.0309)
1 x change			0.479***			0.416***
			(0.0160)			(0.0267)
2 x change			0.504***			0.440***
			(0.0155)			(0.0245)
3 x change			0.532***			0.481***
			(0.0147)			(0.0231)
4 x change			0.503***			0.438***
			(0.0147)			(0.0223)
5 x change			0.524***			0.462***
			(0.0140)			(0.0215)
6 x change			0.541***			0.493***
			(0.0141)			(0.0213)
7 x change			0.535***			0.504***
			(0.0142)			(0.0213)
8 x change			0.504***			0.442***
			(0.0142)			(0.0211)
9 x change			0.519***			0.469***
			(0.0140)			(0.0205)
10 x change			0.472***			0.383***
			(0.0159)			(0.0231)
N	508956	508956	508956	188748	188748	188748

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
All columns include age-head-income-decile fixed effects and place predictions for both the origin and destination places for an individual of a given head of household income decile. "chg qual" is the change in the predicted outcome for a given child between his origin and destination CZ. "change x exp" is 11 minus the child's 1920 age interacted with that change. Columns 3 and 6 instead interact age dummies with the change in place prediction. Columns 1-3 consider all children who moved to a new CZ between 1920 and 1930. Columns 4-6 consider the subset of those children who moved to a new CZ in a new state.

Table 10: County Mover Regressions

	(1) Spec 1	(2) Spec 1 w/ family FE	(3) Spec 2	(4) Spec 1	(5) Spec 1 w/ family FE	(6) Spec 2
chg x exp	-0.00184	-0.00182		0.00207	0.00403	
	(0.00123)	(0.00231)		(0.00214)	(0.00552)	
0 x change			0.362***			0.354***
			(0.0143)			(0.0285)
1 x change			0.376***			0.332***
			(0.0127)			(0.0240)
2 x change			0.377***			0.320***
			(0.0127)			(0.0229)
3 x change			0.389***			0.344***
			(0.0120)			(0.0213)
4 x change			0.357***			0.303***
			(0.0118)			(0.0208)
5 x change			0.387***			0.320***
			(0.0116)			(0.0201)
6 x change			0.412***			0.363***
			(0.0120)			(0.0203)
7 x change			0.419***			0.345***
			(0.0122)			(0.0208)
8 x change			0.375***			0.320***
			(0.0121)			(0.0204)
9 x change			0.389***			0.335***
			(0.0123)			(0.0202)
10 x change			0.365***			0.280***
			(0.0144)			(0.0232)
N	491690	491690	491690	140518	140518	140518

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All columns include age-head-income-decile fixed effects and place predictions for both the origin and destination places for an individual of a given head of household income decile. "chg qual" is the change in the predicted outcome for a given child between his origin and destination counties. "change x exp" is 11 minus the child's 1920 age interacted with that change. Columns 3 and 6 instead interact age dummies with the change in place prediction. Columns 1-3 consider all children who moved to a new county between 1920 and 1930. Columns 4-6 consider the subset of those children who moved to a new county in a new state.

2 Chapter 2: Segregation, the Great Migration, and African American Outcomes: 1920-1990

2.1 Introduction

In the modern United States, cities with high levels of residential segregation by race have seen dramatically worse outcomes for African Americans than less segregated cities. African Americans experience increased poverty rates, greater unemployment, and lower educational attainment in segregated cities, while whites do not suffer in the same way. Research (e.g. Ananat (2011), Cutler and Glaeser (1997)) suggests, furthermore, that this effect on blacks is causal. According to their research, segregation makes black outcomes worse, due most likely, to some combination of two factors: 1) segregation induces selective migration patterns in which higher-educated, higher-earning African Americans tend to locate in less segregated cities, and 2) segregation imposes negative externalities on African Americans.

It is not theoretically obvious, however, that segregation should have a harmful effect on minority communities. Segregation may reduce human capital formation among blacks by cutting them off from more affluent, better-educated peer groups and the higher-quality schools they attend (Card and Rothstein (2007), Ananat), or it may make it harder for African Americans to effect political change (Ananat and Washington (2009)). On the other hand, segregation may allow minority communities to build better informal labor markets (Bayer, Ross, and Topa (2006)) or, alternatively, it may foster infant industries by protecting black businesses from white competition. It is possible, in addition, that segregation might have differential impacts on different groups of African Americans. If educational segregation exists alongside racial segregation, segregation may actually increase average neigborhood human capital for low-educated blacks who now live near some higher-educated African Americans, instead of a mixture of low-educated blacks and whites. This, of course, would lower average peer human capital for high-education blacks.

The uncertainty of segregation's theoretical effect on black outcomes, furthermore, mirrors the empirical ambiguity in segregation's effect over time. While research in the modern U.S. has consistently found severe negative consequences of segregation for African Americans, the data show that this has not always necessarily been the case. As this paper documents, African Americans in 1940 tended to achieve better outcomes in segregated cities than unsegregated cities. This stands in contrast to both OLS and instrumental variables estimates of the effect of segregation on modern outcomes. Thus, the data raise the possibility that segregation may once have been beneficial, or at least not harmful, to African Americans.

This paper seeks to explain why African Americans saw better outcomes in segregated cities in 1940, despite evidence that suggests segregated cities cause worse outcomes for their African American populations in the modern world. There are three major classes of explanation to consider: 1) Omitted place-level characteristics: good labor market conditions in industrial Midwestern and Northeastern cities may have masked the negative effect of segregation in 1940 blacks, while the relative decline of these cities exposed it in the latter 20th century. 2) Actual change in the causal effect of segregation on African American individuals: it may be that segregation fostered African American economic networks or improved black educational outcomes in the first half of the 1900s, only to reverse later in the century. 3) Omitted individual-level characteristics: selective migration periods may have led higher socioeconomic status blacks to move disproportionately to segregated cities (or their ancestors may have moved to more segregated cities), thus improving the observed relationship between segregation and outcomes.

After describing the data, this paper begins by using census data to illustrate that the OLS relationship between segregation and black outcomes is positive for 1940, whereas it is negative for 1990. It then investigates the impact of omitted place-level characteristics on this regression. While including state fixed effects, industrial composition variables, and other place-level characteristics, reduces the observed coefficient, it, nonetheless, remains positive and statistically significant. While this is not conclusive proof that the relationship between segregation and black outcomes would remain positive even if the econometrician could control for all omitted place-level characteristics, these results represent a more thorough investigation of these relationships than, to my knowledge, has been published elsewhere. That the coefficient remains statistically and economically significant even with a large set of controls does, however, prove that there is a major difference between the 1940 and 1990 relationships between segregation and outcomes, whether or not this difference represents a change in the causal effect of segregation.

I then argue that selective migration did not inflate the relationship between segregation and black outcomes. Using a panel of African American males linked across the 1920, 1930, and 1940 censuses, I show that migration, at least between 1920 and 1940, actually decreased the black human capital stock in segregated cities. This should have reduced rather than increased the observed relationship between segregation and outcomes. In fact, the migration pattern, in which less-educated blacks relocated to segregated cities during the great migration, may, however, explain the long-term decline in the relationship between segregation and black outcomes. If the average human capital of African Americans' neighbors decreased over time in segregated cities, then the neighborhood externalities they experienced might have been positive in 1940, but negative in

1990.

2.2 Data

This paper uses three main sources of U.S. census data to investigate the effect of segregation from 1920-1990, with a focus on 1940. It also makes use of data assembled by Cutler, Glaeser, and Vigdor (1999) and Ananat (2011) to support the analyses of the census data. I describe each data set in more detail throughout this section.

2.2.1 IPUMS Census Data 1920-1990

The first data source I use is publicly available and downloadable IPUMS U.S. census samples for 1920-1990. These consist of 1% samples for 1920, 1940, 1950, and 1970, and 5% samples for 1930, 1960, 1980, and 1990. The censuses from 1940 and after include measures of income from wages and years of education completed. They also offer geographic detail to at least the MSA level. While the 1920 and 1930 censuses also contain geographic data, they do not record any measures of income or education. For some analysis, I therefore impute income for individuals in these two censuses using the approach described below.

To investigate the effect of segregation on income before 1940, I impute wages for individuals in the 1920 and 1930 censuses. I did this using 1940 wages by occupation. Specifically, I computed the median wage by occupation in the publicly available 1940 IPUMS census sample (using 1940 dollars), and then assigned this to every individual of that same occupation in the 1920 or 1930 census. Similarly, for the sake of comparability to 1920 and 1930, I also calculated income for the censuses from 1940 and later, using the same method.

2.2.2 1920-1940 Census-Linked Males

To more fully investigate the relationship between segregation and individual outcomes, I use the full population of adult African American males in the 100% 1940 census sample.

2.2.3 1940 100% Census Sample

For the final main data set, I construct a panel linking boys in the 1920 census to themselves in the 1930 and 1940 censuses to isolate the effect of selective migration decisions on this relationship in 1940. To do this, I used the 100% samples of the 1920, 1930, and 1940 censuses, and a linking algorithm that matches individuals between censuses based on individuals' birth state, birth year, race, and first and last names. The panel includes only males between the ages of zero and ten (in 1920; they were 20-30 years old in 1940) throughout the 48 states. Appendix A provides more information on the exact details of the linking method. All three censuses contain information on location, race, marital status, literacy, and occupation among other demographic and economic variables. Yet, only the 1940 census records additional useful information like income, employment status, and years of schooling completed.

For the analysis, I also wanted to assemble a measure of the head of household income for these boys in 1920. However, given that the 1920 census does not report income, I use the 1940 census to impute head of household income like I did for the publicly available IPUMS samples in 1920. The only difference between the imputed wages in these data and those for the publicly available 1920 samples is that for this data set, I calculate median income by occupation using the 100% 1940 census sample, rather than the publicly available 1% sample.

2.2.4 Supplementary Data Sources

Segregation The main measure of segregation employed here is the dissimilarity index, given below:

$$\frac{1}{2}\sum_{d=1}^{D}\left|\frac{black_d}{black_c} - \frac{nonblack_d}{nonblack_c}\right|. \tag{1}$$

Here, $black_c$ measures the number of African Americans living in a given city or county, and $black_d$ measures the number of African Americans living in neighborhood d within place c. $nonblack_c$ and $nonblack_d$ represent the same, except for non-blacks living in the same city or county and neighborhood. This paper makes use of dissimilarity indices computed in two ways. The first set of dissimilarity indices comes from Cutler, Glaeser, and Vigdor (1999), who use 1990 MSAs as the city definition and census tracts as the neighborhood definition. For the second set of dissimilarity indices, I used the 1920-1940 censuses to calculate indices using counties as the city definition and enumeration districts[6] as the neighborhood defintion.

The dissimilarity index measures how dissimilar the racial composition of individual neighborhoods is from the racial composition of the city as a whole. In a city with two neighborhoods, if half of all African Americans and half of all non-blacks lived in each of the neighborhoods (i.e. perfect integration), then the dissimilarity index would take a value of zero. Conversely, if one neighborhood were entirely black, and one entirely white, then this city would be perfectly segregated, and its dissimlarity index would take a value of one.

Cutler, Glaeser, and Vigdor show that U.S. segregation, as measured by a dissimilarity index, rose over the course of the 20th century until about 1970, before falling slightly by 1980 and 1990.

[6] Enumeration districts measure the area that a single census enumerator was able to cover, and are smaller than census tracts.

Railroad Density Instrument In appendix B, to assess the impact of segregation on African Americans in the 1980 and 1990 censuses, I use an instrument developed by Ananat. The Railroad Density Instrument (RDI) measures the degree to which railroad tracks divide MSAs into distinct neighborhoods. This instrument takes a value of zero if railroads do not divide the city into more than one section, increases as railroads divide the city into more distinct sections, and approaches a value of one as the number of sections delineated by railroad tracks approaches infinity. Ananat makes the argument that RDI is a valid instrument because 1) when railroads divide cities into more sections, it fosters segregation by providing natural neighborhood boundaries, and 2) conditional on the total amount of railroad tracks in a city, the layout of these tracks is randomly assigned across cities.

It is worth pointing out that this instrument would seem to be a perfect way to derive a causal estimate of the effect of segregation on black outcomes in 1940. Unfortunately, however, RDI does not have sufficient explanatory power over segregation in 1940 to qualify as a relevant instrument. That a strong statistical relationship between RDI and segregation did not develop until the late 20th century, is consistent with Ananat's explanation that it is, in fact, the layout of railroad tracks combined with migration patterns over time that fosters segregation by filtering new arrivals to a city into their respective black and white neighborhoods.

2.3 Explaining the 1940 Segregation Coefficient

In a simple model of the 1940 effect of segregation, individual i's outcome in place j, y_{ij}, is a function segregation, $seg_{i,1940}$, a vector of his individual and household characteristics, Z_{1i}, and a vector of place characteristics, Z_{2j}:

$$y_{ij} = f\left[\left(seg_{j,1940}\right), Z_{1i}, Z_{2j}\right]. \tag{2}$$

Assuming additive separability yields the following regression equation:

$$y_{ij} = \alpha + \beta\left(seg_{j,1940}\right) + \gamma_1 Z_{1i} + \gamma_2 Z_{2j} + \epsilon_{ij}. \tag{3}$$

In this analysis, the two outcome variables of interest are log wages and total years of schooling, and segregation is measured by a dissimilarity index.

The biggest concern in identifying the causal effect of segregation on outcomes using this framework is that $seg_{j,1940}$ is likely correlated with Z_{1i} and Z_{2j}. Specifically, there are two main omitted variables that are especially likely to bias the coefficient on segregation. First, place characteristics might have driven both segregation and wages. For instance, segregation was particularly high in many Midwestern, Rust Belt cities, where there also may have been many good industrial jobs for Afrian Americans without an advanced education. This may lead us to observe a strong positive relationship between segregation and outcomes, even if it was, in actuality, the industrial composition of segregated cities causing the good outcomes. Similarly, it is possible that variation in racial tolerance and attitudes across places could bias the OLS relationship if it influences both residential segregation and black outcomes. The second class of potential confounding variables concerns selective migration; if people who choose to live in more segregated cities tend to have positive characteristics, then this would make segregated cities look better for African Americans, even if it had not had a causal impact. There could, of course, also be other omitted variables that influence both segregation and outcomes, such as varying preferences over interracial contact and integration, or city geographic characteristics.

This section seeks first to establish that the 1990 relationship between segregation and African American outcomes was materially different from the 1940 relationship. It then shows that the 1940 OLS relationship between segregation and black outcomes remains positive and statistically significant, even when I include control variables meant to proxy for omitted place-level characteristics. Finally, it demonstrates that migration to segregated cities between 1920 and 1940 was actually negatively selected, biasing the OLS estimates down from the true causal estimate of segregation's impact on black outcomes.

2.3.1 Segregation over Time

I begin by presenting baseline figures showing the relationship between segregation and wages and education for whites and blacks over time. These figures both use the sample of 20-60-year-old males in the publicly available IPUMS census waves from 1920-1990. Each plot shows the coefficient from regressions of the outcome variable (log wages for figure 1, and years of schooling for figure 2) on an MSA-level dissimilarity index from that year (as calculated by Cutler, Glaeser, and Vigdor (1999)). Each regression also controls for median wages in the metropolitan area. For each figure, I also extend the results back to 1920 and 1930 using imputed income and schooling.

These figures show that the relationship between segregation and outcomes was positive for African Americans in 1940 and before, but worsened over time. Whites, on the other hand, saw fairly little change in the relationship between segregation and outcomes. The standard errors clustered at the MSA level, however, are large, especially before 1970, making it impossible to statistically rule out equality between the 1940 and 1990 coefficients (although the 1990 coefficient in figure 1 is statistically-significantly negative).

Next, to remove some of the omitted variable bias from figures 1 and 2, figures 3 and 4 report the results of similar regressions for African Americans with a number of additional controls. Specifically, they report coefficients and 95%

confidence intervals from regressions of log income on segregation, controlling for MSA proportion black, population, MSA white employment shares in the top ten largest industries nationally, white median wage, white median income, individual age, and state fixed effects. Like figures 1 and 2, figures 3 and 4 show that the relationship between segregation was lower in 1990 than 1940, but the relationship is not as striking as in figure 1, and the standard errors again make it impossible to reject the equality of the two coefficients. It should also be noted that some of the variables used as controls here may be the causal result of segregation, in which case the coefficients could be biased away from the true causal relationship between segregation and the outcome variables. Nonetheless, the figures show that, even with variables to proxy for place-level omitted variables, the relationship between segregation and outcomes is lower for 1990 than 1940.

2.3.2 Place-Level Omitted Variables

Given the evidence on the relationship between segregation and outcomes over time, I then investigate whether the 1940 coefficient reflects a statistically significant and positive relationship between segregation and wages. Tables 1-3 present four similar sets of regressions designed to delve deeper into the 1940 relationship. Table 1 reports the results of regressions of individual log income on county-level dissimilarity, while table 2 reports the results of regressions of years of school on dissimilarity. Tables 1 and 2 use the whole sample of age 20-59 African American men in the 100% sample of the 1940 census. Table 3, on the other hand, uses the census linked panel of African American males, and reports the results of regressions of individual log income on county-level dissimilarity in the top panel and regressions of years of school on dissimilarity in the second panel. The first column of each table includes no other controls, but columns 2-5 successively add in more controls, including the share of the county's white population in the top 20 largest national industries, state fixed effects, county proportion and number of African Americans, county median white log income and years of school, and individual age.

Tables 1-3 show strongly significant and positive relationships between segregation and income and education. This relationship holds up despite the addition of a large set of industrial control variables as well as demographic, individidual, and geographic controls. The strong positive relationships here stand in stark contrast to the negative relationships on the equivalent regressions in the publicly available 1990 data.

Nonetheless, by all measures in tables 1-3, the relationships between segregation and income and education dramatically decreases with more controls. This suggests that omitted variables play a large role throughout all the specifications here, however, the coefficient on segregation remains positive and statistically significant. The question, therefore, is whether, even when including the controls in column 5 of the tables, omitted factors are strong enough to make the coefficient positive in spite of the true causal relationship between segregation and income being either nonexistent or negative. Appendix B presents a bounding exercise from Oster (2016) which suggests that the remaining omitted variables would only have to be about 1-5% as influential as the controls in column 5 of these tables. It is not obvious, however, whether the omitted characteristics are likely to be more than 5 % as important as the included variables. The variables in column 5 were selected because they were the best proxies for the likely sources of omitted variable bias, but I cannot rule out the possibility that the remaining omitted characteristics are influential enough to produce positive and significant estimates of the relationship between segregation and outcomes, even if the true causal effect of segregation on African Americans in 1940 was negative. I simply argue that the evidence in these tables represents a more thorough OLS investigation of the relationship between segregation and black outcomes in 1940, and it offers suggestive evidence that the causal effect of segregation may have been nonnegative.

2.3.3 Selective Migration

The controls in tables 1-3 mostly proxy for place-level omitted variables. Selective migration patterns to and away from segregated places could, however, also bias the OLS relationship between segregation and black outcomes. If African Americans with better education, or otherwise more favorable unobservable characteristics, were more likely to choose to locate in segregated cities, then we would see positive coefficients in tables 1-3 even without any causal impact of segregation.

To address this concern, I performed another analysis meant to elucidate whether the average socioeconomic stock increased or decreased in segregated cities between 1920 and 1940. In column 1 of the top panel of table 4, I regress the 1940 dissimilarity index of an individual's 1940 location on his head of household's 1920 imputed income. The second column is the same, except I use the 1940 dissimilarity of the child's 1920 location, a proxy for the level of dissimilarity that the individual would have experienced in 1940 if no one had moved between 1920 and 1940. That the coefficient in the second column is significantly higher than that in the first column, indicates that selective migration between 1920 and 1940 actually lowered the average socioeconomic stock for black men in segregated cities. To address this concern, the third column regresses head of household's imputed income on the 1920 dissimilarity in the the child's 1920 county. This specification does not allow for migration in any capacity to influence the estimation. That this coefficient is also higher than that in column 1 proves that if no black men had migrated between 1920 and 1940, the human capital stock among black men in segregated cities would have been higher than it actually was in 1940. The second panel of table 8 repeats the same exercise, except instead of using head of household imputed income as the outcome variable, it uses an indicator variable equal to 1 if the child's head of household was literate in 1920. The results here, while not as statistically precise as the top panel, support the conclusion that migration patterns decreased the human capital stock of segregated cities between 1920 and 1940. Head of household literacy and imputed income do not, of course, measure the entirety of individual-level unobservable characteristics. Perhaps harder-working, more diligent young African Americans were more willing to migrate in search of better jobs, in which case selective migration would not have biased down the coefficients on segregation by as much as table 4 suggests it would have. Yet, by the two important measures here, selective migration between 1920 and 1940 meaningully decreased the relationship between segregation and black socioeconomic status.

Having shown in table 4 that migration reduced the human capital among the parents of young black men living in segregated cities in 1940, I turn to the young black men themselves to confirm that this is reflected in their own educational outcomes. Table 5 compares the relationship between segregation and outcomes for 20-30-year-old black males who moved to a different state as children between 1920 and 1940 and black males who did not move between 1920 and 1940. Using the panel data set, I define movers as anybody who lived in a different state in 1940 than in 1920. I then regress years of school on 1940 dissimilarity for blacks who moved (column 1) and those who did not move (column 2). These results show that movers saw less disparity in education between segregated cities and unsegregated cities than those who did not move. Columns 3 and 4 repeats the exercise for income, regressing log income on segregation for movers and nonmovers. This analysis also shows that movers saw a weaker relationship between income and segregation than did nonmovers, but this disparity is smaller, both in absolute terms and relative to the standard deviations of years of school and log income in the sample[7], than the disparity for education.

Thus, new migrants to segregated cities exhibited lower human capital than the African Americans already living in segregated cities, but they saw a smaller, not-statistically-significant difference in their wages relative to those already living in segregated cities. Perhaps education was not essential for high wages in 1940, or the robust industrial economy of segregated cities masked the costs of lower education for African Americans. Yet, for whatever reason, blacks who moved to segregated cities enjoyed the same strong relationship between segregation and their wages, without showing quite as strong a relationship between segrgation and education as the African Americans already living there.

The results in table 5 are more suggestive than those in table 4; it is quite possible that nonmovers show a stronger relationship between segregation and outcomes because nonmovers had greater exposure to beneficial causal influence of segregation. Yet, the findings in table 5 are at least consistent with the story in table 4 - they suggest that migration flows weakened the 1940 relationship between human capital and segregation. Taking the results in table 5 in conjunction with the results in table 4, the message is clear: the national environment which saw large-scale African American migration out of the unsegregated South into the segregated cities of the North reduced the human capital stock in segregated cities. Migration patterns during this time period, therefore, cannot explain the puzzle as to why segregation is correlated with good black outcomes in 1940. In fact, this presents convincing evidence that at least between 1920-1940, selective migration made the relationship between

[7]The difference between the coefficients in the first two colums, 1.35, is about 39% of the 3.5 standard deviation in years of school for black men in the sample, whereas the difference between the last two colums, 0.11, is about 13% of the the 0.88 standard deviation in income.

segregation and outcomes look worse than it truly was for African Americans.

That does not mean, of course, that migration patterns *prior* to 1920 did not influence the relationship between segregation and black outcomes. It is possible that the legacy of slavery, for instance, sent higher-skilled workers out of the South to more segregated Northern cities. Given the large scale of the Great Migration, however, the influence of preexisting selective location patterns would have to have been enormous to outweigh the strongly negatively selected migration patterns between 1920 and 1940.

2.4 Discussion and Conclusion

My analysis shows that the observed relationship between residential segregation and black outcomes worsened between 1940 and 1990. This finding is robust to the inclusion of place-level covariates, and does not stem from selective migration, at least during the 1920s and 1930s period of the Great Migration. Selective migration, in fact, reduced the observed relationship between segregation and outcomes in 1940.

There are many theories that might explain why segregation was better for African Americans in 1940 than 1990. For instance, segregation in 1940 may have fostered informal labor markets among African Americans that helped blacks find better-paying jobs. Additionally, segregation could have protected black-owned businesses from the competition of white-owned businesses. It is conceivable, then, that the benefits of these informal labor markets may have eroded over the 20th century as the manufacturing industry in segregated cities evolved, or that the Civil Rights Movement might have harmed black-owned businesses, as the end of economic segregation exposed African American entrepreneurs to new competition. It is also likely that the omitted place characteristics which made segregated cities look so appealing to African Americans in

1940 would have changed over time. For instance, Berman et al. (1994) show evidence of a shift in demand from unskilled to skilled labor in the manufacturing industry in the 1980s, which would have hit lower-educated African Americans in segregated cities especially hard. If the unobservable place characteristics changed in a way that made segregated cities less favorable environments for African Americans, then the relationship between segregation and black outcomes would have declined, even without any change in either selection on unobservable individual-level characteristics or the causal effect of segregation.

Yet, the evidence on selective migration presented here suggests one particular explanation for why segregation would have been more beneficial (or at least less harmful) to African Americans in 1940 than in 1990. The black human capital stock was higher in segregated cities in 1940 than in less segregated cities. Thus, segregation would have led African Americans to live near other African Americans with above average educational and economic achievement. The migration patterns between 1920 and 1940, however, represented an enormous shock to the segregated cities of the North. Segregated cities saw an influx of African Americans who, while they did not necessarily possess less human capital than the average African American in 1940, possessed less human capital on than the average African American in the well-educated segregated cities. When migration lowered the human capital stock of segregated cities, it left African Americans segregated into neighborhoods with less-educated neighbors. This reduction in human capital in predominantly black neighborhoods might have made those neighborhoods less desirable to whites and better-educated blacks, thus giving rise to a cycle where lower human capital increased segregation, and, in turn, segregation lowered human capital. This is consistent with the history of suburbanization and "white flight," which saw segregation increase across American cities until 1970.

Given the evidence that where children grow up dramatically influences their adult outcomes (e.g. Chetty, Hendren, and Katz (2016) and Chetty and Hendren (2015)), that neighbors influence each other's labor market outcomes (e.g. Bayer, Ross, and Topa), and that low human capital in one generation persists into future generations (e.g. Borjas (1992), Borjas (1995), Black, Devereux, and Salvanes (2005), and Curie and Moretti (2003)), it is very likely that the Great Migration's negative effect on the African American human capital stock in segregated cities would have lowered the human capital stock in these cities for succeeding generations as well. The realignment of black human capital during the Great Migration, therefore, supports Ananat's argument that the negative effect of segregation in 1990 results from selective migration and negative human capital externalities. Specifically, it was migration to segregated cities during the Great Migration which caused later generations of African Americans in segregated cities to be born to less-educated parents on average, and to grow up around less-educated neighbors. During the Great Migration, there were so many movers, and the preexisting level of black human capital in segregated cities was so high, that large-scale migration could not help but lower the average human capital in segregated cities. An unfortunate byproduct, therefore, of the Great Migration was that the economic opportunities which originally drew African Americans to segregated cities inadvertantly planted the seeds which eventually turned these once-vibrant cities into places where African Americans suffer, instead, from a lack of economic opportunity.

2.5 References

Ananat, Elizabeth Oltmans. "The Wrong Side(s) of the Tracks: The Causal Effects of Racial Segregation on Urban Poverty and Inequality." American Economic Journal: Applied Economics, 3.2 (2011): 34-66.

Bayer, Patrick, Steve Ross, and Giorgio Topa. “Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes,” Journal of Political Economy, 116.6 (2008): 1150-96.

Berman, Eli, John Bound, and Zvi Griliches.”"Changes in the demand for skilled labor within US manufacturing: evidence from the annual survey of manufactures.” The Quarterly Journal of Economics 109.2 (1994): 367-397.

Bernard, James. “Intergenerational Mobility in the 1920-1940 United States.” (2017).

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. “Why the apple doesn’t fall far: Understanding intergenerational transmission of human capital.” American economic review 95.1 (2005): 437-449.

Borjas, George J. “Ethnic capital and intergenerational mobility.” The Quarterly Journal of Economics 107.1 (1992): 123-150.

Borjas, George J. “Ethnicity, Neighborhoods, and Human-Capital Externalities.” The American Economic Review 85.3 (1995): 365.

Card, David, and Jesse Rothstein. “Racial Segregation and the Black-White Test Score Gap.” Journal of Public Economics, 91(11–12) (2007): 2158–84.

Chetty, Raj, Nathaniel Hendren. The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. NBER Working Paper No. 23001 (2017).

Chetty, Raj, Nathaniel Hendren. The Impacts of Neighborhoods on Intergenerational Mobility II: County Level Estimates. NBER Working Paper No. 23021 (2017).

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. “The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment,” American Economic Review (2016) 106 (4).

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. “Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States", Quarterly Journal of Economics 129.4 (2014): 1553-1623.

Chyn, Eric. "Moved to Opportunity: The Long-Run Effect of Public Housing Demolition on Labor Market Outcomes of Children." Working, 2016.

Collins, William J. and Marianne H. Wanamaker. "Selection and Economic Gains in the Great Migration of African Americans: New Evidence from Linked Census Data." American Economic Journal: Applied Economics 6.1 (2014): 220-52.

Currie, Janet, and Enrico Moretti. "Mother's education and the intergenerational transmission of human capital: Evidence from college openings." The Quarterly Journal of Economics 118.4 (2003): 1495-1532.

Currie, Janet and Aaron Yelowitz. "Are Public Housing Projects Good for Kids?" Journal of Public Economics 75.1 (2000): 99-124.

Cutler, David M., and Edward L. Glaeser. "Are Ghettoes Good or Bad?" Quarterly Journal of Economics, 112.3 (1997): 827–72.

Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor. "The rise and decline of the American ghetto." Journal of political economy 107.3 (1999): 455-506.

Dahl, Gordon B. "Mobility and the return to education: Testing a Roy model with multiple markets." Econometrica 70.6 (2002): 2367-2420.

Feigenbaum, James J. "Automated Census Record Linking." (2015).

Galster, George C. "The mechanism (s) of neighbourhood effects: Theory, evidence, and policy implications." Neighbourhood effects research: New perspectives. Springer Netherlands, (2012): 23-56.

Gibbons, Stephen, Olmo Silva, and Felix Weinhardt. "Everybody needs good neighbours? Evidence from students' outcomes in England." The Economic Journal 123.571 (2013): 831-874.

Jacob, B.A. "Public housing, housing vouchers, and student achievement: Evidence from public housing demolitions in Chicago" American Economic Review, 94 (2004), pp. 233–258

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. "Experimental analysis of neighborhood effects." Econometrica 75.1 (2007): 83-119.

Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz. "Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment." The Quarterly Journal of Economics (2005): 87-130.

Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu. "Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity." American Economic Review, 103.3 (2013): 226-31.

Logan, J. R., Spielman, S., Xu, H., & Klein, P. N. "Identifying and bounding ethnic neighborhoods." Urban geography, 32(3) (2011): 334-359.

Massey, Douglas S., and Nancy A. Denton. American Apartheid: Segregation and the Making of the Underclass. (1993) Cambridge, MA: Harvard University Press.

Minnesota Population Center and Ancestry.com. IPUMS Restricted Complete Count Data: Version 1.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2013.

Oreopoulos, Philip. "The long-run consequences of living in a poor neighborhood." The quarterly journal of economics (2003): 1533-1575.

Oster, Emily. "Unobservable Selection and Coefficient Stability: Theory and Validation." Journal of Business Economics and Statistics, forthcoming.

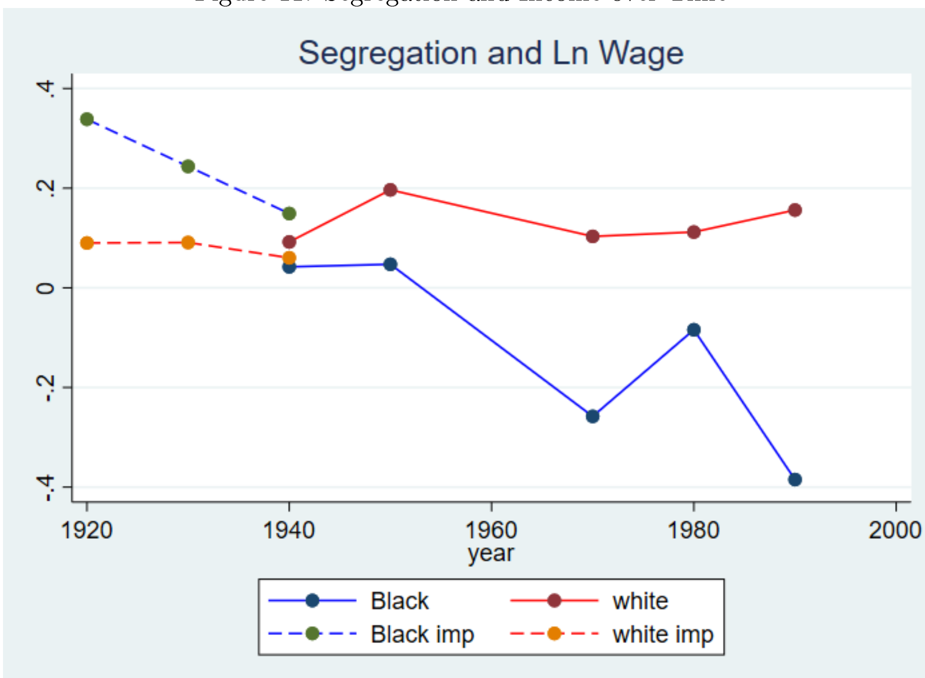
Reardon, Sean F. and Glenn Firebaugh. "Measures of Multigroup Segregation." Sociological Methodology, 32.1 (2002): 33-67.

Vigdor, Jacob L. "The pursuit of opportunity: Explaining selective black migration." Journal of Urban Economics 51.3 (2002): 391-417.

Wodtke, Geoffrey T, David J Harding, and Felix Elwert. "Neighborhood Effects in Temporal Perspective: The Impact of Long-term Exposure to Concentrated Disadvantage on High School Graduation." American Sociological Review, 76.5 (2011): 713–736.

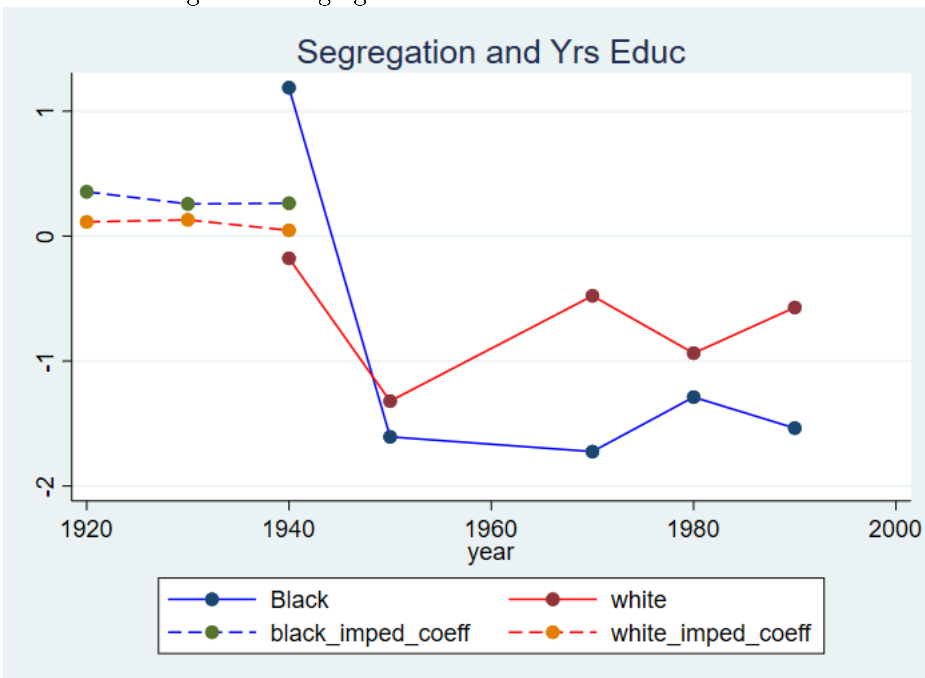
2.6 Tables and Figures for Chapter 2

Figure 11: Segregation and Income over Time



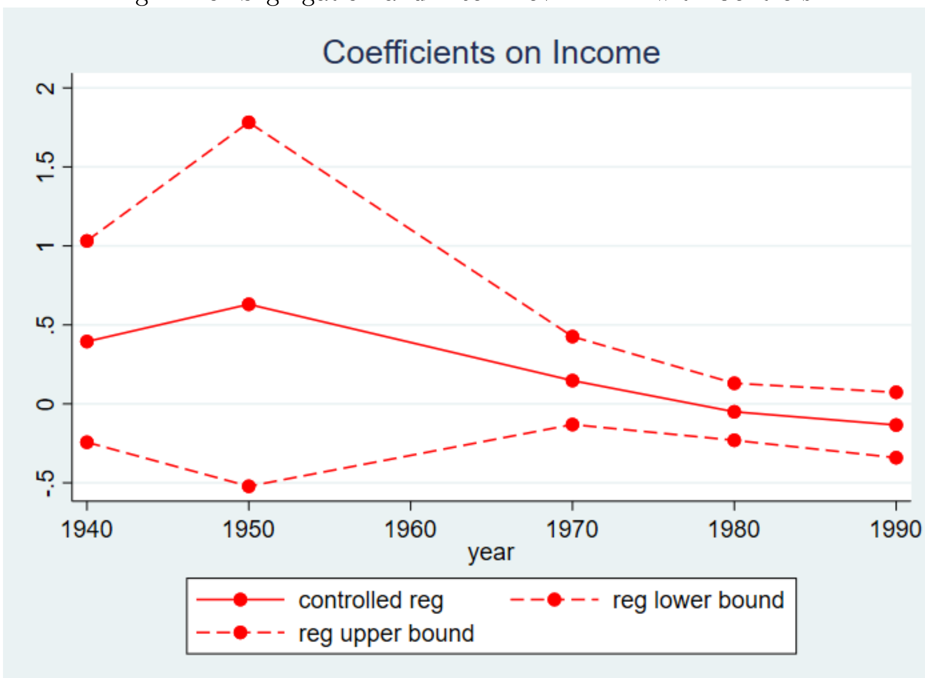
Wage data from IPUMS census sample for each year 1940-1990. The solid lines report coefficients from regressions of log income on segregation (also controlling for median MSA wage). The dashed lines report coefficients from regressions of imputed log income (imputed by occupation using 1940 log income) on segregation. The segregation measure is a dissimilarity index as calculated by Cutler, Vigdor, and Glaeser. Sample limited to males in MSAs age 20-59 reporting nonzero income.

Figure 12: Segregation and Years School over Time



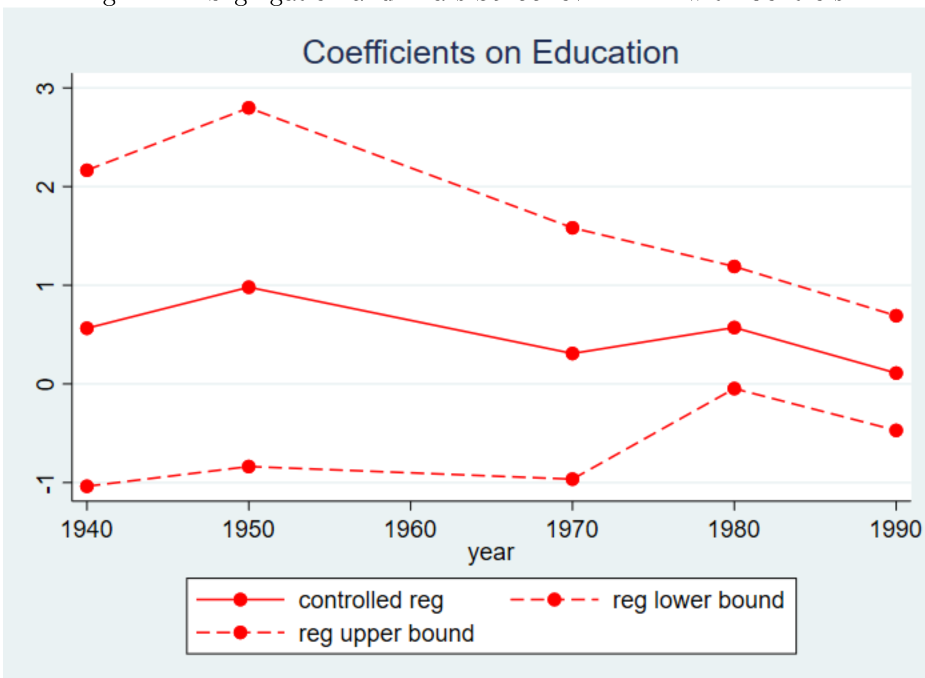
Schooling data from IPUMS census sample for each year. The solid lines report coeffficients from regressions of years of school on segregation (also controlling for median MSA wage). The dashed lines report coefficients from regressions of imputed years schooling (imputed by occupation using 1940 years schooling) on segregation. The segregation measure is dissimilarity index as calculated by Cutler, Vigdor, and Glaeser. Sample limited to males in MSAs age 20-59 .

Figure 13: Segregation and Income over Time with Controls



Wage data from IPUMS census sample for each year 1940-1990. The solid line reports coefficients from regressions of log income on segregation controlling for the MSA number and proportion of blacks, MSA population, MSA white employment shares in the top ten largest industries nationally, MSA median white wage and median years of school, individual age, and state fixed effects. The dashed lines report the 95% confidence interval. The segregation measure is dissimilarity index as calculated by Cutler, Vigdor, and Glaeser. Sample limited to black males in MSAs age 20-59 reporting nonzero income.

Figure 14: Segregation and Years School over Time with Controls



Data on years of schooling from IPUMS census sample for each year 1940-1990. The solid line reports coefficients from regressions of years of school on segregation controlling for the MSA number and proportion of blacks, MSA population, MSA white employment shares in the top ten largest industries nationally, MSA white median wage and median years of school, individual age, and state fixed effects. The dashed lines report the 95% confidence interval. The segregation measure is dissimilarity index as calculated by Cutler, Vigdor, and Glaeser. Sample limited to black males in MSAs age 20-59.

Table 11: Regressions of Log Wages on Dissim with Controls: Full Sample

	(1)	(2)	(3)	(4)	(5)
	Log Inc	Log Inc	Log Inc	Log Inc	Log Inc
dism	1.605***	0.654***	0.656***	0.448***	0.394***
	(0.0482)	(0.0572)	(0.0575)	(0.0558)	(0.0542)
ct pop (1000s)				0.000288***	0.000290***
				(0.0000707)	(0.0000659)
ct black prop				-0.852***	-0.548***
				(0.0763)	(0.0988)
age 1940					0.00739***
					(0.000250)
medwage					0.0250***
					(0.00622)
Industry Shares	No	Yes	Yes	Yes	Yes
N	2008740	2008740	2008740	2008740	2008740
R^2	0.162	0.205	0.218	0.223	0.232
State FE			Yes	Yes	Yes

Sample consists of all black males in my 1920-1940 census-linked sample. Industry shares refers to the county share of white workers in each of the 20 largest national industries in 1940. Medwage and medschool refer to the median log wages and years of school for white 20-60-year-old men in the 100% census sample living in each county.

Table 12: Regressions of Log Wages on Dissim with Controls: Full Sample

	(1) Years School	(2) Years School	(3) Years School	(4) Years School	(5) Years School
dism	5.128***	2.963***	1.686***	1.323***	1.097***
	(0.225)	(0.230)	(0.162)	(0.196)	(0.181)
ct pop (1000s)				0.000340*	0.000402*
				(0.000170)	(0.000169)
ct black prop				-1.541***	-0.180
				(0.241)	(0.331)
age 1940					-0.0698***
					(0.00151)
medwage					-0.0682**
					(0.0213)
medschool					0.152***
					(0.0224)
Industry Shares	No	Yes	Yes	Yes	Yes
N	2465036	2465036	2465036	2465036	2465036
R^2	0.106	0.125	0.149	0.150	0.196
State FE			Yes	Yes	Yes

Sample consists of all black males in my 1920-1940 census-linked sample. Industry shares refers to the county share of white workers in each of the 20 largest national industries in 1940. Medwage and medschool refer to the median log wages and years of school for white 20-30-year-old men in the panel data set living in each county.

Table 13: Regressions of Years School on Dissim with Controls: Linked Sample

	(1) Log Inc	(2) Log Inc	(3) Log Inc	(4) Log Inc	(5) Log Inc
dism	1.402***	1.029***	0.925***	0.629***	0.508***
	(0.0479)	(0.0496)	(0.0549)	(0.0547)	(0.0544)
N	139361	139361	139361	139361	139361
R^2	0.134	0.156	0.173	0.179	0.214
	Years School	Years School	Years School	Years School	Years School
dism	6.961***	5.859***	3.477***	2.530***	2.273***
	(0.249)	(0.320)	(0.242)	(0.234)	(0.219)
N	179445	179445	179445	179445	179445
R^2	0.190	0.207	0.263	0.269	0.280
ct pop(1000s)				Yes	Yes
ct black prop				Yes	Yes
age 1940					Yes
medwage					Yes
medschool					Yes
Industry Shares		Yes	Yes	Yes	Yes
State FE			Yes	Yes	Yes

Sample consists of all black males in my 1920-1940 census-linked sample. Industry shares refers to the county share of white workers in each of the 20 largest national industries in 1940. Medwage and medschool refer to the median log wages and years of school for white 20-30-year-old men in the panel data set living in each county. The top panel presents regressions of log income on 1940 dissimilarity, and the second panel presents regressions of years of school on 1940 dissimilarity. The bottom panel indicates the controls included in each column.

Table 14: Regressions of Dissim on HoH Characteristics

	(1) 1940 Dissim	(2) 1940 Dissim of 1920 County	(3) 1920 Dissim of 1920 County
head inc pctl	0.00277***	0.00449***	0.00384***
	(0.00018)	(0.00027)	(0.00028)
N	161,291	151,015	161,291
head literate	0.0597***	0.0826***	0.0789***
	(0.0057)	(0.0082)	(0.009)
N	177,810	166,897	177,810

County-level-clustered standard errors in parentheses.
Sample includes all linked African American males ages 20-30. Dissimilarity is calculated by comparing the racial makeup of enumeration districts to the racial makeup of the county. The outcome variable in column 1 is 1940 dissimilarity, the outcome variable in column 2 is the 1940 dissimilarity of the county in which a child lived in 1920, and the outcome variable in column 3 is the 1920 dissimilarity of a child's 1920 county. The first panel presents the results of regressions of the dissimilarity measures on imputed head of household income, and the second panel presents results of regressions of the dissimilarity measures on an indicator for whether or not the head of household was literate in 1920.

Table 15: Regressions of Dissimilarity on Individual Characteristics

	(1) years_school movers	(2) years_school nonmovers	(3) ln_inc movers	(4) ln_inc nonmovers
dissim_ct_40	6.321***	7.675***	1.213***	1.324***
	(0.369)	(0.297)	(0.0859)	(0.0448)
N	49326	130119	41160	98201

County-level-clustered standard errors in parentheses.
Sample includes all linked African American males ages 20-30 in 1940. Dissimilarity is calculated by comparing the racial makeup of enumeration districts to the racial makeup of the county. Movers are defined as those who lived in a different state in 1940 than 1920, and all others are defined as nonmovers.

2.7 Appendix

2.7.1 Appendix A: Linking Across Censuses

Here I describe the process I used to match individuals across the 1920, 1930, and 1940 censuses to construct the panel data set used in this paper. This section is adapted from Bernard (2017), which uses the same data set.

The biggest hurdle to analyzing long-term impacts using census data lies in linking individuals from the 1920 census to themselves in 1930 and 1940. Linking five-year-old John Smith in 1920 to 25-year-old John Smith in 1940 requires determining which of the John Smiths in 1940 is the correct match. In fact, perhaps the correct John Smith appears as "Jon Smith" or "John Smit" in 1940. Given that there were roughly 15 million males between ages zero and ten in 1920, this process requires an automatable method for selecting the correct match.

I begin by assembling a list of potential 1930 matches for the males to be matched from the 1920 census and a separate list of potential matches of those males in the 1930 census to their counterparts in the 1940 census. To do this, I first match each of these individuals to all males in all 48 states with the same birth state, race, age (plus or minus one year), and the same first letters of their first and last names. Ideally, the matching process would be more flexible than this. Previous efforts to match individuals in similar contexts (i.e. Feigenbam 2015) have found benefits to considering matches among people with different recorded races and with ages as far as two years apart. Similarly, it would be better not to restrict the set of potential matches to people with the same first letters of their first and last names – this method fails to match anyone who, for instance, goes by "Bill" in 1920, but then refers to himseld as "Will" in either the 1930 or 1940 censuses. The matching method used here, however, employs these imperfect methods because they dramatically reduce the computing time necessary to assemble a list of potential matches.

The initial step of the matching process yields an enormous number of potential matches, so the next step is to trim this list by using first and last names to determine which names are close enough that they produce a true match. The Jaro-Winkler string similarity measure provides a gauge of whether or not two names likely belong to the same person. This measure takes the value of one if two strings are identical, and decreases towards zero as the two strings become more dissimilar. For example, "Bill" and "Billy" register a Jaro-Winkler value of 0.96, while "Bill" and "William" score 0.73. Using this measure, I remove from the list of potential matches any match in which both the first and last name fail to achieve a similarity score of at least 0.8.

I then join together the two sets of potential matches – the list matching from 1920 to 1930 and the list matching from 1930 to 1940. This leaves a set of potential matches from 1920 to 1930 to 1940 with many children appearing more than once in the list. I then eliminate any potential matches where the match is not the "best" match for a given child from either the 1920-1930 or 1930-1940 matches. To define the "best" match between any two censuses, I first consider whether a given match weakly dominates all others on name similarity. In other words, a "best" match between 1920 and 1930 occurs when it has the highest (or weakly highest) Jaro-Winkler similarity score among all potential 1930 matches for the 1920 child. I then restrict the sample to include only those potential 1920-1930-1940 matches where the match is "best", both between 1920 and 1930 and between 1930 and 1940.

This leaves a data set with only the best possible matches for each child. This is still not an ideal data set, as for some children there are dozens of best matches. Thus, as a final step, I eliminate all 1920 children for whom there are more than three best matches. This leaves a sample with mostly unique individuals, but some duplicated matches. This is the linked sample of children that this paper analyzes.

It is worth noting here, that there exist more robust algorithms to match individuals between censuses. Feigenbaum (2015) proposes one such method, which relies on a manually-matched "training" data set. By carefully matching a subset of children by hand, this method derives a near-perfect match for this subset. The researcher then measures the impact of various rules for determining which pairs are and are not correctly linked on how closely the automatically-matched data mimic the manually-matched data. This provides a framework for determining matches in a way that optimally minimizes both the probability of falsely assigning matches to pairs which are not correct matches, and of failing to assign matches to pairs who are in fact correct matches. Due to time constraints, I have not implemented this type of matching process, but it is likely that Feigenbaum's or another similar method would lead to an improvement in the quality of my matched data. As is, there is likely substantial measurement error arising from falsely assigned matches.

2.7.2 Appendix B: Bounding Exercise

I conduct a bounding analysis described in Oster (2016) which identifies how correlated omitted and observable characteristics would have to be, so as to imply that the true causal relationship between segregation and the 1940 outcomes in tables 1-3 is zero or negative. Revisiting equation (3), the individual and place-level covariates, Z_{1i} and Z_{2j}, are both made up of observable and unobservable components. Let the observable components be Z^o_{1i} and Z^o_{2j}, and let the unobservable components be Z^u_{1i} and Z^u_{2j}. Then rewriting the observable components as $\omega^o_{ij} = Z^o_{1i} + Z^o_{2j}$, and the unobservable components, $W_{2ij} = \gamma_1 Z^u_{1i} + \gamma_2 Z^u_{2j}$, I can rewrite equation (3) to match Oster's equation 1:

$$y_{ij} = \alpha + \beta\left(seg_{j,1940}\right) + \Psi\omega_{ij}^{o} + W_{2ij} + \epsilon_{ij}. \tag{4}$$

Oster assumes that the selection on observables must be proportional to the selection on unobservables, and defines δ as the coefficient measuring the proportionality between selection on unobservable and observable factors (other than the treatment variable, dissimilarity in this case), where $\delta\frac{Cov\left(\Psi\omega_{ij}^{o}, seg_{j,1940}\right)}{Var\left(\Psi\omega_{ij}^{o}\right)} = \frac{Cov(W_{2ij}, seg_{j,1940})}{Var(W_{2ij})}$. As δ increases, it means that greater selection on observable characteristics would imply a greater degree of selection on unobservable characteristics as well, and a greater probability that omitted variables meaningfully obscure the true β^*. $\delta = 1$ would imply that there is exactly as much selection on unobservable as observable characteristics. Given assumed values of both δ and the maximum R-squared that the regression would take on if all unobservable characteristics were included in the regression[8], it is possible to estimate the true treatment effect, β^*, from observing the residuals of a regression of $seg_{j,1940}$ on ω_{ij}^{o}, as well as the coefficients and R-squared of two OLS regressions - one without any control variables, and one with all observable control variables.

Tables 6 and 7 explore how substantial the selection on unobservables would have to be to change the interpretation of the segregation coefficients from tables 1-3. Table 6 performs this exercise for the panel data set, while table 7 considers the full 1940 census sample of black men. These tables report at least how large δ would have to be under different assumptions on the maximum theoretical R-squared[9] value to make the true β^* effect of segregation equal to 0. For the

[8]If all values in equation (4) were observable and accurate, then the maximum R-squared would be 1 - the right-hand-side variables could perfectly predict the outcome. If, however, there is measurement error in y_{ij}, then the maximum R-squared would be less than 1.

[9]I provide estimations under a range of R^2 assumptions, because the potential for measurement error in census wage and schooling variables means that even if I could observe all unobservable characteristics in equation 4, the R^2 value from a regression with these variables

panel data set, depending on the assumptions on R^2 and whether log income or years of school is used as the outcome variable, I find that δ would have to be anywhere from 0.03 to 0.055 to render the true β^* estimate nonpositive. For the 100% sample, I find lower values of δ ranging from 0.01 to 0.024. These tables also show what the β^* estimate would be under various assumptions on δ and R^2.

Table 16: Implied β^* under Assumptions on δ and R^2: Panel Data

	Income			
	R^2=0.7	R^2=0.8	R^2=0.9	R^2=1
$\beta^* = 0$ if $\delta =$	0.048	0.04	0.034	0.03
if $\delta = 0.005$	.502	.49	.477	.464
if $\delta = 0.01$	.44	.415	.39	.365
if $\delta = 0.015$	.379	.342	.305	.269
	Years of School			
	R^2=0.7	R^2=0.8	R^2=0.9	R^2=1
$\beta^* = 0$ if $\delta =$	0.055	0.044	0.037	0.032
if $\delta = 0.005$	2.041	1.987	1.933	1.879
if $\delta = 0.01$	1.815	1.708	1.604	1.5
if $\delta = 0.015$	1.593	1.438	1.286	1.137

This table reports implied treatment effect of segregation, β^*, as well as the proportionality coefficient δ under different assumptions on R^2 and δ using Oster's proportional selection on observables assumption. The top row of each panel reports the value of δ that would make β^* 0 for different assumptions on the maximum R^2. The next three rows of each panel report the β^* implied by different assumptions on δ and R^2. The top panel calculates δ and β^* estimates for regressions where log income is the outcome variable, and the bottom panel calculates for regressions where years of school is the outcome variable. These estimates use the census-linked panel of African American men.

Thus, tables 6 and 7 imply that β^* would be negative if the remaining selection on unobservables, even after controlling for all the unobservable characteristics in column 5 of tables 1-3, were both proportional to the selection on the observables in column 5 of the tables, and were more than 1-5% as important as the observables controlled for in column 5. This, in turn, raises the question: what is a likely true value of δ? The best way to answer this would be to find an instrument for segregation in 1940, and then use it to calculate

might still be less than 1.

Table 17: Implied β^* under Assumptions on δ and R^2: 100 % 1940 Data

	Income			
	R^2=0.7	R^2=0.8	R^2=0.9	R^2=1
$\beta^* = 0$ if $\delta =$	0.024	0.019	0.017	0.014
if $\delta = 0.005$	.303	.284	.266	.247
if $\delta = 0.01$	.216	.181	.146	.111
if $\delta = 0.015$	.134	.083	.033	-.015
	Years of School			
	R^2=0.7	R^2=0.8	R^2=0.9	R^2=1
$\beta^* = 0$ if $\delta =$	0.015	0.013	0.011	0.01
if $\delta = 0.005$	.736	.665	.594	.524
if $\delta = 0.01$	.381	.242	.104	-.031
if $\delta = 0.015$	.033	-.169	-.366	-.56

This table reports implied treatment effect of segregation, β^*, as well as the proportionality coefficient δ under different assumptions on R^2 and δ using Oster's proportional selection on observables assumption. The top row of each panel reports the value of δ that would make β^* 0 for different assumptions on the maximum R^2. The next three rows of each panel report the β^* implied by different assumptions on δ and R^2. The top panel calculates δ and β^* estimates for regressions where log income is the outcome variable, and the bottom panel calculates for regressions where years of school is the outcome variable. These estimates use the full 1940 census sample of African American men.

β^* directly. I could then check what values of δ would make the estimated coefficients in tables 1-4 be consistent with the true β^* . Unfortunately, I have no valid instrument for 1940 segregation.

There do, however, exist some potential instruments for segregation in later years, making it possibe to estimate a δ for later years by comparing OLS estimates of the effect of segregation to IV results. While it is reasonable to think that δ might have changed between 1940 and 1990, finding an estimate of δ for 1980 and 1990 would at least offer an estimate of a reasonable range that δ could take on. I therefore turn to Ananat's railroad density index (RDI) to instrument for segregation in 1980 and 1990, the only two years for which RDI has any meaningful statistical relationship with dissimilarity indices (although, the instrument is weak in my data even for these years). Specifically, the first stage regression for the top panel of table 7 generates a predicted MSA-level dissimilarity index by regressing dissimilarity index on RDI and total track length.

Then, the second stage regresses individual log wages on predicted MSA-level dissimilarity index. Although the instrument here has first-stage F-statistics lower than common weak instrument standards, in the absence of another convincing instrument, I take these estimates as the best available. I find that the two-stage least squares coefficients on dissimilarity are 0.24 and -0.38 for 1980 and 1990 respectively, but that neither of these estimates is statistically significantly different from 0. In the second panel of the table, I perform the same exercise with years of school as the outcome variable.

Then, assuming the instrument is relevant and satisfies the exclusion restriction (which Ananat builds a case that it does in 1990), I consider these to be the true estimates of the causal effect of segregation on log black wages in those years. With these estimates of β^* in hand, I back out what degree of proportional selection on unobservable characteristics must be true in order to reconcile these coefficients with the OLS coefficients on dissimilarity. To to do this, for 1980 and 1990 I ran a controlled regression of log wages on dissimilarity and control variables replicating the estimations in column 5 of tables 1-3 using the 1980 and 1990 census data. Comparing these results to the uncontrolled regression of log wages on segregation implied that δ in 1980 would have to have been anywhere from -0.0022 to 0.017 if the maximum R^2 was one or anywhere from -0.0028 to 0.0216 if the maximum R^2 was 0.8. The equivalent values for 1990 are 0.0028 to 0.0095 and 0.0036 to 0.012.

These estimates suggest a range of deltas which preserve the possibility that the 1940 causal effect of segregation is positive. Moreover, if this range of deltas were to apply to the 1940 data, it would be very unlikely that the true causal effect was negative.

Table 18: Delta Estimates: 1980-1990

	1980	1990
Ln(wage)	.2377	-.3506
se	.5141	.3126
delta ($R^2 = 1$)	-.0022	.0028
delta ($R^2 = 0.8$)	-.0028	.0036
Yrs Sch	-4.7558	-2.29
se	1.7091	.891
delta ($R^2 = 1$)	.0156	.0095
delta ($R^2 = 0.8$)	.0196	.0119
1st stg F-stat	7.1243	8.486

The four panels each present the results of three two-stage least squares regressions of log income (panel 1) or years schooling (panel 2) on dissimilarity indexes. The samples consist of African Americans in 1980 and 1990 using publicly available IPUMS data. Each estimation uses Ananat's RDI instrument, conditional on the total length of railroad tracks in the MSA, to predict dissimilarity. "Delta ($R^2 = 1$)" and "delta ($R^2 = 0.8$)" report the deltas that would be required for the OLS estimates from those years to be consistent with the the IV estimates under the assumptions of the maximum R-squared being 1 and 0.8 respectively when comparing the uncontrolled regressions to regressions controlling for white employment shares in the top 20 industries, white median wage and median schooling, MSA proportion black and population, state fixed effects, and individual age.

2.7.3 Appendix C: Summary Statistics

Table 19: Summary Statistics

	Black			White		
	Wage	School	Dism	Wage	School	Dism
1940	6.315	7.063	.751	7.002	9.639	.777
sd	.77	3.24	.1	.8	3.29	.09
N	11566	13910	11290	143928	176785	139448
1950	7.328	8.011	.78	7.79	10.486	.788
sd	.81	3.45	.09	.82	3.39	.08
N	5551	6393	21733	53476	64257	188493
1970	8.405	10.062	.803	8.848	11.925	.797
sd	.87	3.26	.08	.83	3.07	.08
N	91261	105353	62426	838884	940098	511900
1980	9.053	11.592	.694	9.484	12.848	.682
sd	1.04	2.86	.12	.92	2.88	.12
N	188740	237073	238513	1536631	1744210	1700705
1990	9.534	12.288	.659	10.015	13.302	.633
sd	1.09	2.41	.11	.97	2.57	.12
N	179785	224798	229066	1532933	1731987	1721815
1940 Panel	5.83	7.008	.568	6.567	10.522	.73
sd	.85	3.58	.22	.84	2.86	.19
N	139361	179445	181400	2319709	2813281	2847652
1940 100%	5.953	5.788	.577	6.749	9.105	.712
sd	.88	3.51	.22	.92	3.44	.2
N	2008740	2465036	2504403	2.13e+07	2.55e+07	2.60e+07

This table reports mean, standard deviation, and number of observations by race of log income, years of school, and dissimilarity. The top panel presents these statistics for 1940, 1950, and 1970-1990 using the publicly available IPUMS samples and the dissimilarity measures assembled by Cutler, Glaeser, and Vigdor. The second panel presents statistics from the census-linked 1920-1940 panel data set, using 1940 dissimilarity measures calculated from the 100% census sample. The final panel reports statistics assembled using the entire 100% census sample.