NBER WORKING PAPER SERIES

CHANGES IN BLACK-WHITE INEQUALITY: EVIDENCE FROM THE BOLL WEEVIL

Karen Clay
Ethan J. Schmick
Werner Troesken

Working Paper 27101
http://www.nber.org/papers/w27101

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2020

We would like to thank Andy Ferrara, Daniel Jones, Jessica LaVoice, Allison Shertzer, and Randy Walsh for helpful comments and suggestions. This paper also benefited from comments by seminar participants at Marquette University, the American Economic Association Annual Conference, the Liberal Arts Colleges Development Economics Conference, and the North American Regional Science Association Conference. Karen Clay acknowledges financial support from Heinz College, Carnegie Mellon University.The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Changes in Black-White Inequality: Evidence from the Boll Weevil
Karen Clay, Ethan J. Schmick, and Werner Troesken
NBER Working Paper No. 27101
May 2020, Revised June 2020
JEL No. I24,J10,J62,N32

ABSTRACT

This paper investigates the effect of a large negative agricultural shock, the boll weevil, on black-white inequality in the first half of the twentieth century. To do this we use complete count census data to generate a linked sample of fathers and their sons. We find that the boll weevil induced enormous labor market and social disruption as more than half of black and white fathers moved to other counties following the arrival of the weevil. The shock impacted black and white sons differently. We compare sons whose fathers initially resided in the same county and find that white sons born after the boll weevil had similar wages and schooling outcomes to white sons born prior to its arrival. In contrast, black sons born after the boll weevil had significantly higher wages and years of schooling, narrowing the black-white wage and schooling gaps. This decrease appears to have been driven by relative improvements in early life conditions and access to schooling both for sons of black fathers that migrated out of the South and sons of black fathers that stayed in the South.

Karen Clay
Heinz College
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
and NBER
kclay@andrew.cmu.edu

Ethan J. Schmick
Washington & Jefferson College
60 S. Lincoln St.
Washington, PA 15301
eschmick@washjeff.edu

Werner Troesken
University of Pittsburgh

1 Introduction

Black-white inequality has been remarkably persistent in recent decades, and this persistence has been a topic of much research and a focus of policymakers. Historical work has focused on the Great Migration, which helped close the black-white wage gap in the mid-twentieth century and reduced inequality (Margo, 1995; Boustan, 2009; Collins and Wanamaker, 2014; Derenoncourt, 2019). Contemporary work has focused on early environment through programs such as Moving to Opportunity, which moved families across neighborhoods within cities, and on important differences in intergenerational mobility across different locations in the United States (Chetty et al., 2014; Chetty, Hendren and Katz, 2016).

This paper uses a large negative agricultural shock that both induced large-scale migration and plausibly changed the early life environment of children born after the shock to examine how these two factors affected black-white inequality of fathers and sons in the first half of the twentieth century. The large negative agricultural shock was the boll weevil, a cotton pest that decreased cotton production and spread throughout the American South between 1892 and 1922. The boll weevil led to declines in cotton production, disruption in tenancy arrangements, increases in food production, and widespread migration (Lange, Olmstead and Rhode, 2009; Ager, Brueckner and Herz, 2017*a*). While there were many other large shocks in the first half of the twentieth century including the Mississippi Floods of 1927 and the Dust Bowl (Hornbeck and Naidu, 2014; Hornbeck, 2012), the boll weevil was a very large shock, affecting approximately 22% of the U.S. population.[1]

To examine the effect of the boll weevil, black and white fathers are observed in the 1900 or 1910 U.S. Census in the years before the boll weevil arrived in the county they resided in. These fathers are then linked to the next decadal census (1910 or 1920) after the arrival of the boll weevil. This allows us to observe whether the father migrated, where he migrated to, and any changes in his occupation. Sons observed in their fathers'

[1] 22% represents the share of the U.S. population that lived in a cotton growing county in 1900.

households in 1900, 1910, and 1920 are then linked to the 1940 Census, which allows us to observe their wage income, years of schooling, occupation, and whether they have migrated from their fathers last location.

We present three main results. First, we find the shock impacted black and white sons differently. We compare sons whose fathers initially resided in the same county and find that white sons born before and after the boll weevil had similar wage and schooling outcomes. In contrast, black sons born after the boll weevil had significantly higher wages and years of schooling, narrowing the black-white wage and schooling gaps. The black-white wage gap decreased by 12% to 22% (depending on specification) and the black-white schooling gap decreased by 6%.

Second, the relative gains for black sons born after the boll weevil were widespread. Specifically, they were not driven by the 5% of black fathers who migrated out of the South. We do find evidence that black sons born after the boll weevil to fathers who migrated out of the South experienced large additional gains wages, although the effect is not significant. Importantly, sons of black fathers that did not move out of the South still experienced a decrease in the black-white wage gap by 22 to 24%.

Third, the available evidence suggest that the gains were likely driven by early life conditions for black sons born after the boll weevil. The black population was extremely impoverished, so the boll weevil could have improved maternal, fetal, and infant health through reduction in mother's work effort both while pregnant and after birth. In addition, the variety and nutritive value of food production improved following the boll weevil (Clay, Schmick and Troesken, 2019). Both changes could have led to relative improvements in black infant health as compared to white infant health. Evidence from comparison of brothers, nutrition conditions of counties in North Carolina, and the heights of black recruits during World War II are all consistent with early life conditions having played an important role for black sons.

While the setting is historical, this work is well suited to shed light on issues of contemporary interest. It sheds light on the intergenerational effects of shocks that induce migration, particularly of disadvantaged groups. It also highlights the importance of early life environment, for the children of families that migrate and those that do not.

This paper contributes to two primary literatures. The first is the literature on black-white inequality. Scholars have discussed migration as a general mechanism through which the black-white wage gap fell (Margo, 1995; Boustan, 2009; Collins and Wanamaker, 2014; Derenoncourt, 2019). In line with this literature, our paper finds that black sons born after the boll weevil whose fathers moved out of the South saw large relative wage gains. Black sons of fathers who moved out of the South are, however, a small share of the overall sample. Our paper finds that relative wage gains for these cohorts were driven by black sons born after the boll weevil whose fathers stayed in the South. Their gains appear to be the result of improvements in early life conditions.

The second is the large literature on the effect of early life shocks on long run outcomes (Almond and Currie, 2011; Almond, Currie and Duque, 2018) and the two subliteratures that examine the role of maternal nutrition (Linnemayr and Alderman, 2011; Almond and Mazumder, 2011; Adhvaryu et al., 2016) and focus on black-white inequality (Almond, Currie and Herrmann, 2012; Bhalotra and Venkataramani, 2015; Almond and Mazumder, 2011). Our paper builds on the literatures by examining an important historical setting where there are effects on both black and white sons from the decisions of their fathers' to migrate or stay. A father's decision to migrate may change his income and the nutritional and disease environment. On the other hand, if a father chose to remain in the South the early childhood environment for his children, especially the nutrition environment, might have changed.

2 The Boll Weevil

The boll weevil appeared in Texas in 1892. From there, it progressed North and East through the cotton belt. By 1922, the boll weevil had spread throughout the entire cotton growing region of the United States. [2] Scholars have treated the arrival of the boll weevil in the cotton belt during the early 1900s as an exogenous shock that disrupted cotton production and broadly impacted the Southern economy.

The U.S. Department of Agriculture (1951), Ransom and Sutch (2001), Lange, Olmstead and Rhode (2009), and Ager, Brueckner and Herz (2017*a*) all find that the arrival of the boll weevil had large negative effects on cotton yields and production.Ager, Brueckner and Herz (2017*a*), Bloome, Feigenbaum and Muller (2017), and Ager, Brueckner and Herz (2017*b*) show that the arrival of the boll weevil had an adverse effect on tenancy, local labor markets, marriage and fertility. These adverse effects on tenancy and local labor markets induced many individuals to migrate within and across states (Lange, Olmstead and Rhode, 2009; Ager, Brueckner and Herz, 2017*a*). The arrival of the weevil had a large positive effect on the production of other crops (Ager, Brueckner and Herz, 2017*a*; Clay, Schmick and Troesken, 2019; Lange, Olmstead and Rhode, 2009). Clay, Schmick and Troesken (2019) provide evidence that the arrival of the boll weevil may have led to improved nutrition. We explore the evidence on migration and nutrition further in the next two subsections.

2.1 Migration

Using county level panel data on population and a difference in differences strategy, Lange, Olmstead and Rhode (2009) document sizeable net migration associated with the boll weevil. Counties with the highest cotton shares saw the greatest declines in population, while counties with lower cotton shares saw smaller declines. Ager, Brueckner and Herz (2017*a*) and Feigenbaum, Mazumder and Smith (2019) also find that high cotton counties experienced significant out migration with the later study linking this migration to a decrease in the violence and repression experienced by Southern blacks in cotton producing counties.

[2] See Lange, Olmstead and Rhode (2009) or Hunter and Coad (1923) for a year-by-year map of the boll weevils' progression through the cotton belt.

Collins and Wanamaker (2015) use a large linked sample of black and white men in the South for 1910-1930 to examine the early decades of the Great Migration. They find black and white migrants experienced increases in income relative to black and white non-migrants. While their focus is primarily on the Great Migration, they examine migration to more and less cotton intensive states and find that blacks and whites were less likely to move to cotton intensive states. Their work does not speak directly to the boll weevil, because in 1910 some men are observed after the boll weevil had arrived in their county, while others are observed before the boll weevil had arrived. However, their work highlights important differences in the propensity to migrate and the destinations of whites and blacks, an issue we will examine as well.

2.2 Early Life Conditions

The diet for poor whites and blacks in the South consisted primarily of salt pork, molasses, and corn. For example, Edgar (2012) argued that in South Carolina debt forced many farmers to plant cotton, which had higher expected returns than other crops. As a result, South Carolina had to import $70 to $100 million worth of food annually. For poverty-stricken tenant farmers with little ready cash, this meant that there was less to eat. The consequent increased dependence on a diet of pork, cornbread, and molasses made poor Carolinians more susceptible to disease (Edgar 1992, p. 47). For poor pregnant women, a low nutrient diet and in some cases physical labor in the fields may have adversely affected their fetuses.

Ager, Brueckner and Herz (2017*a*), Clay, Schmick and Troesken (2019), and Lange, Olmstead and Rhode (2009) provide evidence that the arrival of the boll weevil led to increases in the local production of food crops and declines in pellagra, a nutritional disease associated with niacin deficiency. Clay, Schmick and Troesken (2019) show that declines in pellagra appear to have been driven by increases in locally produced nutrient rich food. Specifically, there were increases in corn, peanut, and sweet potato acres per capita after the arrival of the boll weevil. Midwestern milled corn was degerminated, which greatly reduced its nutritional value. Increased use of local milled corn, which retained its germ, contributed to the decline in pellagra.

The early life literature provides possible channels through which the boll weevil may have affected child outcomes, which included improved nutrition and income for children whose fathers migrated and improved nutrition for children whose fathers did not migrate. Improved maternal nutrition and household income are linked to higher birth weight (Almond and Mazumder, 2011; Rossin, 2011; Almond, Hoynes and Schanzenbach, 2011; Hoynes, Miller and Simon, 2015; Lindo, 2011). Aizer and Currie (2014) (p. 856) conclude that "health at birth is an important predictor of long-term outcomes, including education, income, and disability." Improvements in early childhood nutrition, disease environment, and income are also linked to higher test scores and wages (see Almond, Currie and Duque (2018) for a review of this literature).

Importantly, for our paper, the boll weevil did have an effect on educational outcomes. Using detailed data for Georgia, a large cotton producing state, Baker (2015) finds that the boll weevil affected the enrollment in school of black children, but not white children. Black children often helped to harvest the cotton during the September to December time frame. Reductions in cotton production reduced demand for child labor, allowing more children to attend school. Baker, Blanchette and Eriksson (2020) examine the long-run impacts of the boll weevil on educational attainment. They find evidence that men who were 4-18 years old when the boll weevil arrived in a county had 0.25 years higher educational attainment than men who were 19-30 years old.

3 Data

To study the impact of the boll weevil on black-white inequality we generate a linked sample of fathers and their sons. We relate fathers' migration and sons' births to the year that the boll weevil first arrived in the county the father was initially living in. Data on the year the boll weevil first arrived in a county are taken from Lange, Olmstead and

Rhode (2009), which originally came from USDA boll weevil maps.[3] Counties invaded by the boll weevil between 1892 and 1922 are shown in Figure 1.

3.1 Linking

Appendix Figure A.1 provides details on the linking procedure that we use to generate our sample of fathers and sons. We begin by looking for fathers in the 1900 or 1910 Censuses that had a son aged 18 or younger and who were living in a county that would be invaded by the boll weevil in the next ten years (i.e. fathers in the 1900 Census must have been living in county that would be invaded by the boll weevil between 1900 and 1910; step 1. in Figure Appendix A.1). We then link these fathers to the next decadal census (step 2. in Appendix Figure A.1). By linking these fathers we are able to observe whether they migrated out of their initial county, state, or region of residence. We then take the set of sons of successfully linked fathers (step 3 in Appendix Figure A.1) and link the sons from the 1900, 1910, or 1920 Censuses to the 1940 Census to obtain their adult outcomes (step 4 in Appendix Figure A.1). The result of this linking algorithm is a data set that allows us to observe whether a father migrated around the time the boll weevil arrived in the county he was initially living in and it provides outcomes for both fathers and their sons.

To perform all of the linking we employ the ABE algorithm, which is commonly used in economics and was developed by Abramitzky, Boustan, and Eriksson (Abramitzky, Boustan and Eriksson (2012), Abramitzky, Boustan and Eriksson (2014), Abramitzky, Boustan and Eriksson (2019)). This algorithm is similar to the algorithm used in Ferrie (1996) and Long and Ferrie (2013). We begin by adjusting first names for common nicknames and then standardize each first and surname using the NYSIIS algorithm, which transforms a name into a phonetic code. We then restrict our sample to individuals

[3]We reviewed the original USDA boll weevil map published in Hunter and Coad (1923) and found a few discrepancies between the map and the coding of the boll weevil arrival in Lange, Olmstead and Rhode (2009). The map shows that the boll weevil arrived in Cherokee County, South Carolina in 1920, but it is coded as 1921 in Lange, Olmstead and Rhode (2009). The map shows the boll weevil arriving in Iredell County and Wake County, North Carolina in 1921, but it is coded in Lange, Olmstead and Rhode (2009) as 1922. We changed the coding in these cases to align with the original map.

who are unique by NYSIIS first name, NYSIIS surname, birthplace, birth year, and race. Using these variables we search for the individual in the census we want to link them to. If we find a unique match we declare this observation to be a match. If we find multiple matches the observation is discarded. If we do not find a unique match we continue to search for individuals who match exactly on NYSIIS first name, NYSIIS surname, birthplace, and race, but we now allow birth year to differ by up to one year (e.g. if an individual in the 1910 Census reports a birth year of 1902 we will search for individuals in the 1940 Census with a birth year of 1901 and 1903). If still no unique match is found we continue to search for individuals who match exactly on NYSIIS first name, NYSIIS surname, birthplace, and race but we now allow birth year to differ by up to two years.

The results from this linking procedure are displayed in Appendix Table A.1. We begin with 565,000 black fathers and 1,010,000 white fathers with sons under the age of 19 who were observed in the 1900 or 1910 Censuses living in a county that would be invaded by the boll weevil in the next ten years. We were able to successfully link 22% of black fathers and 30% of white fathers to the next census. We then locate the sons of these successfully linked fathers and link them to the 1940 Census. We were able to successfully link 20-21% of black sons and 28-29% of white sons to the 1940 Census. The final sample has 42,000 black fathers, 55,000 black sons, 136,000 white fathers, and 195,000 white sons.

A concern with any linked sample is whether it is representative of the overall population. Appendix Table A.2 addresses this issue by comparing a number of characteristics of the linked sample to the entire set of individuals that we attempted to link for both fathers and their sons. Of note, we are significantly more likely to link fathers living in owner occupied housing and literate fathers. We are, also, significantly more likely to link sons living in owner occupied housing and urban areas. While we find significant differences between the linked sample and the sample that we attempted to link along numerous other dimensions, most of these differences are very small in magnitude (e.g. 75.9% of black fathers in the linked sample were farmers, while 75.2% of fathers in the entire sample were farmers).

3.2 Summary Statistics for Fathers and Sons

Panel A of Table 1 provides summary statistics on migration and other father characteristics by race. One striking feature is the extent to which fathers have moved in the ten year interval between censuses. Only 37% of black fathers and 46% of white fathers were still living in the county that they were observed in ten years earlier. Recall that all fathers initially lived in a county that would be invaded by the boll weevil within ten years. Blacks and whites differ by where they moved, with 27% of white father, but 38% of black fathers moving to a different county within the same state. In contrast, 24% of black fathers and 26% of white fathers moved out of the state they were first observed in. Only 5% of black fathers moved out of the South, while 8% of white fathers moved out of the South. To better understand fathers' migration decisions, Appendix Table A.3 presents a descriptive regression of the determinants of black and white fathers moving out of the county or the state they were originally observed in. Younger fathers were, generally, more likely to migrate and fathers who were initially farmers were less likely to migrate. One exception is that younger white fathers were actually less likely to move out of state than older white fathers.

Panel B of Table 1 provides descriptive statistics for the sons of these fathers by race. In 1940, only 17% of black sons and 26% of white sons lived in their fathers' initial county and 35% of black sons and 26% of white sons had moved out of their birth states. Appendix Table A.4 presents a descriptive regression of the determinants of black and white sons moving out of their father's initial county or moving out of the South. Sons were more likely to move out of the South if their father had moved states. They were less likely to move out of their father's initial county if they were white and born after the boll weevil. On the other hand, black sons born after the boll weevil were less likely to move out of the South. More years of schooling was associated with a lower propensity to move out of their father's initial county, but a higher propensity to move out of the South.

4 Empirical Strategy

Our main empirical analysis relies on a comparison of sons that were born before and after the boll weevil arrived to fathers that resided in the same county prior to the weevil's arrival. We also want to allow for the fact that the effect of being born after the boll weevil might differ by race. Accordingly, we estimate the following equation:

$$\begin{aligned} outcome_{ict} = {} & \beta_1[Born\,post\,boll\,weevil_{ct} = 1] + \beta_2[Black_i = 1] \\ & + \beta_3[Born\,post\,boll\,weevil_{ct} = 1] \times [Black_i = 1] \\ & + \theta_c + \theta_t + \theta_b + \theta_e + \epsilon_{ict} \end{aligned} \tag{1}$$

In the above equation, i indexes a son, c indexes the county that son's father was initially living in (in 1900 or 1910), and t indexes birth year.

$outcome_{ict}$ is the adult outcome of son i, whose father initially lived in county c, and who was born in year t. We use two main outcome variables: the natural log of a sons weekly wage and their years of schooling. Weekly wage is defined as an individual's yearly income in 1939 divided by the number of weeks they reported working in 1939. Census enumerators were supposed to code any individual with an annual income over \$5,000 a year as having an income of \$5,000. This practice was not universally followed as there are several individuals for whom yearly income is over \$5,000. We deal with these and other outliers in the weekly wage variable by following Acemoglu and Angrist (2000) and censor weekly wages at the 98th percentile. Weekly wages above the 98th percentile are replaced with 1.5 times the 98th percentile wage.[4]

$[Born post boll weevil_{ct} = 1]$ is a dummy variable that takes a value of one if individual i was born after the arrival of the boll weevil in their father's initial county. We measure treatment based on the son's birth year relative to when the boll weevil arrived in the father's initial county.[5] $[Black_i = 1]$ is a dummy variable that takes a value of one if

[4]For our sample, the 98th percentile of weekly wages is \$72, so incomes above this level are replaced with 1.5 times \$72, or \$108.

[5]For sons whose fathers move, we do not observe the timing of the birth and the move relative to

individual i's reported race was black.

The remaining controls are: county, birth year, census enumeration year, and birth order fixed effects. θ_c are fixed effects for father's initial county and θ_t are birth year fixed effects. θ_b is a dummy variable for individual i's location in the birth order of his family. Birth order is determined by the age of the sons who have the same father in the censuses. Thus, it does not take into account older siblings who either moved out of the house or died before the censuses were taken. θ_e are census enumeration year fixed effects. Finally, we cluster standard errors at the father's initial-county-level.

5 Black-White Inequality and the Boll Weevil

We begin by presenting results from Equation (1) where we do not include the dummy variable for being black or the interaction term. These results are displayed in Panels A and B of Table 2. Columns (1) and (2) use the log of weekly wage as the dependent variable and column (3) uses years of schooling.

In column (1) of Panel A we find that sons born after the boll weevil arrived in their fathers' initial county had about 2% higher weekly wages. Column (1) of Panel B shows that this increase was concentrated in high cotton acreage counties, which are the counties that were the most impacted by the boll weevil (high cotton acreage counties are defined as counties where the cotton acreage as a share of farm acreage is above 11%). Column (2) repeats the specification from column (1), but controls for the number years of schooling an individual completed (with a complete set of indicator variables), keeping in mind that years of schooling might be affected by the boll weevil. The coefficient in Panel A is reduced in magnitude to about 1%, while the coefficient on the interaction term in Panel B remains large and statistically significant. Finally, in column (3) we find that sons born after the arrival of the weevil attained about 0.15 years of additional schooling and it appears that at least some of this effect is concentrated in high cotton

the arrival of the boll weevil. This is particularly true for fathers who move within their original state, because we only observe a son's state of birth, not the timing of their father's move. Sons born in the year the boll weevil first arrived are coded as being treated (i.e. $[Born\,post\,boll\,weevil_{ct} = 1]$).

counties. Thus, there is broad evidence that being born after the arrival of the weevil improved both weekly wages and years of schooling.

We next examine whether the improvements in long-run outcomes differed by race and we find strong evidence that they did. Table 3 estimates Equation 1 and column (1) shows that black sons born prior to the boll weevil had weekly wages that were 0.632 log points (88%) lower than white sons born prior to the weevil.[6] White sons born after the boll weevil did not experience any significant increase in weekly wages as is shown by the coefficient on being born after the boll weevil. However, being black and being born after the boll weevil resulted in a significant increase in wages by about 8%. This reduced the black-white wage gap by approximately 12% (0.0775/0.632).

The black-white wage gap is partially attributed to differences in years of schooling. Accordingly, in column (2) we control for a son's years of schooling, keeping in mind that years of schooling are affected by the boll weevil. This reduces the black coefficient to 0.32 log points, but the coefficient on the interaction between being born after the boll weevil and black remains almost unchanged. Using these coefficient values, we find that being born after the boll weevil reduced the wage gap for blacks by 22% (0.0708/0.32). Column (3) uses years of schooling as the dependent variable. We find that blacks born prior to the boll weevil had about 3.5 fewer years of schooling than whites, but being born after the boll weevil closed this gap by about 0.2 years (6% of the gap). In conclusion, Table 3 shows that black sons differentially benefited from the arrival of the boll weevil.

Our main empirical results presented in Table 3 rely on a comparison between black and white sons whose fathers lived in the cotton growing region of the United States during the 1900 or 1910 Censuses. A natural question is how the blacks in our sample fared compared to other blacks that were not living in the cotton belt. Our analysis is not set-up to answer this question, but we can partially address the concern by noting that approximately 75% of blacks in the 1900 and 1910 censuses were living in counties that were invaded by the boll weevil. Thus, there is no reason to think that northern blacks

[6]To convert the log points into a percentage we use the formula $e^{\beta} - 1$ where β is the coefficient estimate.

are a more natural comparison group than southern whites. In fact, southern whites had many characteristics in common with southern blacks, such as a high percentage working in agriculture and a low literacy rate (see Appendix Table A.7), which were probably not shared with northern blacks.

5.1 Event studies

We next convert our main empirical results into event studies. These event studies use the following specification:

$$outcome_{ict} = \sum_{k=-6}^{6} \tau_k 1(t - h_c = k) + \xi * X_{ict} + \theta_c + \theta_t + \theta_b + \theta_e + \epsilon_{ict} \quad (2)$$

In the above equation, i indexes a son, c indexes the county that son's father was initially living in (in 1900 or 1910), and t indexes birth year. h_c is the year the boll weevil arrived in the father's initial county. Thus, $1(t - h_c = k)$ are a series of dummy variables, which take a value of one if a son was born between six years before or after the boll weevil arrived in their father's initial county. The remaining controls are: county, birth year, census enumeration year, and birth order fixed effects. We omit the indicator for being born in the year the boll weevil arrived and we group sons born more than six years before or after the boll weevil into the -6 and +6 indicators. Standard errors are clustered at the father's initial-county-level.

The coefficients, τ_k, when weekly wages are the dependent variables are plotted in Panel A of Figure 2. There is no noticeable trend in weekly wages for sons who were born one to five years prior to the boll weevil's arrival. Sons born six or more years prior to the boll weevil's arrival did have significantly lower wages, but recall that this is a catch-all category that includes sons born six to 28 years prior to the weevil's arrival.[7]

[7]It is possible for a son in our sample to be born 28 years prior the boll weevil's arrival. For example, if the son was 18 years old in the 1900 Census and the weevil did not arrive in their father's initial county until 1910 than that individual would have been born 28 years prior the weevil's arrival.

We do find evidence that sons born after the arrival of the boll weevil had higher weekly wages and in some cases this is significant. Panel B of Figure 2 plots a similar graph using years of schooling as the dependent variable. There is, again, no noticeable pre-trend for sons who were born one to five years prior to the boll weevil's arrival. Again, sons born six or more years before the boll weevil did have significantly lower amounts of schooling. The increase in years of schooling appears to initially occur for sons born in the year the weevil arrived and the upward trend continues after that. This is in contrast to weekly wages, where the initial increase appeared to occur for sons born one to two years after the weevil's arrival.

In Appendix Figure A.2 we repeat the analysis in Figure 2, but now interact the event time indicators with an indicator if the son was black. This figure, therefore, shows how the black-white wage and schooling gap evolved relative to the boll weevil's arrival. In Panel A it appears that the black-white wage gap was fairly constant prior the weevil arriving. Black sons born after the arrival of the boll weevil saw increases in their wages relative to white sons so the black-white wage gap decreased. However, these estimates are noisy and not significant. In Panel B, we find that the black-white schooling gap was increasing prior to the boll weevil (i.e. blacks had decreasing years of schooling relative to whites) and this trend reversed after the weevil arrived. Again, the estimates are noisy.

These event study figures demonstrate that there were not noticeable pre-trends in either schooling or weekly wages prior to the boll weevil's arrival and they both show increases after the weevil arrived.

5.2 Sensitivity of Results to Linking Procedure

We next explore the sensitivity of our main results to different linking procedures. One major concern with our empirical results is that the effects we are identifying are simply the result of false matches. For our results to be driven by false matches the rate of false positives would have to be systematically related to both the race of the child and the year that they were born relative to the boll weevil's arrival. While we believe that this in unlikely we, nevertheless, repeat our linking procedure using two different algorithms that should reduce the number of false positive matches.

First, we link both fathers and sons that match exactly on first name (not phonetically cleaned), last name (not phonetically cleaned), birthplace, birth year, and race. We re-estimate our baseline results (Table 3) with this sample and display the results in Appendix Table A.5. The results using only individuals that match exactly on name and birth year are almost identical to our main results in magnitude, although the significance is slightly reduced. The second method links only fathers and sons whose NYSIIS cleaned first and last names are unique within a five-year age band in both the initial census and the 1940 Census. We, again, re-estimate Table 3 using only unique names and ages and display the results in Appendix Table A.6. The results using this sample are, again, similar to our main results, although the magnitude of the wage effect is slightly reduced and the magnitude of the schooling effect is slightly larger. We conclude that our results are likely not being driven by false matches.

6 Migration Decisions and Early Life Conditions

6.1 Fathers' Migration Decisions

In this section we explore the migration decisions of fathers around the time the boll weevil arrived in the county they were living in, the impact of these migration decisions on their sons, and the change in early life conditions that occurred even for families that did not migrate.

First, we examine the locations that fathers moved to. Panel A of Table 1 provides a rough idea of where fathers are moving, but we are more precise about the exact locations in Figure 3. Panel A shows the net change in black fathers in our sample residing in each county and Panel B displays the same for white fathers.[8] Fathers migrated to many

[8] To construct these maps we calculated the total number of black or white fathers from our sample that moved into a county between the first time they were observed (either in 1900 or 1910) and the next time they were observed (either in 1910 or 1920). We then subtracted the total number of fathers

locations, including Southern cities, Northern cities, and to the West. The most common locations for white fathers were counties a short distance to the north of the Cotton Belt. The most common locations for black fathers were more diffuse, although some also moved north of the Cotton Belt. Figure 4 shows the migration patterns of black and white sons in our sample.[9] Again, urban areas, both within and outside of the South, and the West appear to be popular locations to migration to.

Fathers that decided to migrate were, generally, positively selected in that they had higher socioeconomic statuses and, in some cases, were more likely to be literate. Appendix Table A.7 shows a number of father characteristics broken down by migrant status for each race. Fathers that migrated had significantly higher average initial LIDO scores, were less likely to be a farmer and were less likely to own their home. LIDO scores are lasso-adjusted industry, demographic, and occupation (LIDO) score from Saavedra and Twinam (2020). These scores calculate occupational standing in industry-occupation-state-age-race cells and we use these scores to proxy for income since income is not available in censuses prior to 1940.[10] Fathers that migrated within the same state or within the South generally had lower literacy rates than fathers that did not migrant. However, fathers that migrated out of the South had much higher literacy rates than fathers that did not migrate. There is not a distinct pattern between the age of migrants and non-migrants in our sample. For example, black fathers that migrated within the same state were younger, but black fathers that migrated to a different state within the South were older than non-migrants. Finally, it appears that when whites moved states within the South or out of the South they moved at least 100 miles further than blacks.

Did the fathers that moved experience better or worse outcomes in these new locations? We conclude this section by showing how the occupational status of fathers that

from our sample that moved out of that county to get a measure of net migration.

[9]To construct these maps we calculated the total number of black or white sons from our sample that moved into a county between the first time they were observed (in 1900, 1910 or 1920) and 1940. We then subtracted the total number of sons that moved out of that county from our sample to get a measure of net migration.

[10]LIDO scores are similar to the commonly used occupational income score, but they produce results that are closer to those from earnings regressions. We prefer LIDO scores to OCC scores because they compute occupational standing in industry-occupation-state-age-race cells as opposed to just occupation cells.

migrated changed. In Table 4 the dependent variable in column (1) is the change in fathers' occupational income score (OCC score) from the first time they are observed (in 1900 or 1910) to the next time they are observed (in 1910 or 1920).[11] The dependent variable in the second column is the change in fathers' LIDO scores. Black fathers that did not migrate, actually saw a slight decreases in their OCC and LIDO scores compared to white fathers that did not migrate. There is conflicting evidence on whether fathers that moved out of the South saw increases in their occupational status. Fathers that moved out of the South saw decreases in their OCC scores, but increases in their LIDO scores. Fathers that migrated from a rural area to an urban area saw large, positive, increases in both their OCC and LIDO score. Finally, black fathers that migrated out of the South saw differential increases in both their OCC and LIDO scores. We conclude that black fathers that migrated out of the South generally experienced improved economic status.

6.2 The Impact of Fathers' Migration on Sons

We next show that both black and white sons benefited from their fathers moving, but that neither race differentially benefited from these moves. To do this, we estimate Equation 1 but now include a triple interaction term for being born after the boll weevil, being black, and having a father that migrated. The results are displayed in Table 5. The dependent variable in all columns is the log of sons' weekly wages in 1940 and all columns control for dummy variables for the number of years of schooling a son attained. In column (1) we look at fathers that moved within the same state, column (2) examines fathers that moved to a different state within the South, and column (3) looks at the fathers that moved out of the South. All columns control for dummy variables for the other types of moves (e.g. column (1) controls for dummy variables, which are not reported, for a father moving to a different state within the South and moving out of the South).

Similar to Table 3, we find a large black-white wage gap for black sons born prior to

[11] OCC score is defined by IPUMS (Ruggles et al. (2020)) as the median income received by persons employed in a particular occupation in the 1950 Census.

the boll weevil (the coefficient on the black indicator variable) and we find no evidence that white sons born after the boll weevil had higher weekly wages (the coefficient on the post boll weevil indicator variable). Fathers moving, regardless of type of move, increased all sons' weekly wages with the biggest effect coming for sons whose father moved out of the South (coefficient on the father moved indicator variable). It appears that sons born after the arrival of the boll weevil differentially benefited from these moves, especially if the father moved within the same state or out of the South (the interaction term between fathers moving and born post boll weevil). However, black sons born after the boll weevil to a father that migrated (the triple interaction term) did not differentially benefit from the migration. Importantly, black sons born after the boll weevil whose father did not migrate still saw a significant decrease in the black-white wage gap, of a similar magnitude to what was found in Table 3. This can be seen through the interaction between being black and being born after the boll weevil. In particular, we find that sons of black fathers that did not move still experienced a decrease in the black-white wage gap by 22% to 24%.[12]

We repeat these specification, but now use sons' locations relative to the fathers' initial county in Appendix Table A.8. Importantly, in all specifications, black sons that did not migrate out of their fathers' initial counties still experienced a 30% reduction in the black-white wage gap. In addition, there is no evidence that black sons born after the boll weevil that moved out of their father's initial county differentially benefited from these moves. To conclude this section, while we find strong evidence that sons benefited from their fathers' migrating, we find no evidence that black sons differentially benefited from these moves. Thus, black sons born after the boll weevil whose fathers did not migrate still saw significant increases in their wages relative to white sons. We next turn to potential improvements in early childhood environment that might have occurred within the fathers' initial counties after the arrival of the boll weevil.

[12]We have run a specification that includes triple interactions for all three types of moves (within state, within South, and out of South) in the same regression and we find similar results. In particular, none of the triple interactions are significant and the coefficient on the interaction between black and born post boll weevil is 0.0778 and significant at the 1% level. The black-white wage gap for sons of black fathers that did not migrate closed by 24% in this specification.

6.3 Early Life Conditions

We conclude our empirical analysis by offering evidence that early life conditions were significantly better for sons born after the arrival of weevil, regardless of whether their father migrated, and black sons differentially benefited from these improved early life conditions. Clay, Schmick and Troesken (2019) argue that Southern nutrition significantly improved after the arrival of the boll weevil because Southern farmers switched from growing cotton, which had no nutritional value, to growing nutritionally rich foods such as corn, peanuts, and sweet potatoes. They also provide evidence that prior the arrival of the boll weevil many southerners were consuming imported food that was not nutritional rich. As a result of the boll weevil, the pellagra death rate, a disease caused by nutritionally deficiencies, significantly decreased. An improved nutrition environment could significantly impact young children and children still in utero.[13]

We begin by exploring the relationship between the arrival of the boll weevil and nutrition. To do this we would, ideally, use the pellagra death rate to measure the nutrition environment for all counties in the South. However, this data was not systemically reported until after the boll weevil arrived in most of the South. One exception is North Carolina, which reported county-level counts of pellagra deaths prior the the boll weevil's invasion of the state in 1919.

In the first two columns of Table 6 we explore the relationship between the arrival of the boll weevil and the pellagra death rate controlling for county and year fixed effects. Column (1) shows that counties in North Carolina saw decreases in the pellagra death rate (i.e. improvements in the nutrition environment) after the arrival of the boll weevil. This effect is concentrated in counties that were heavily black in the 1910 Census (column (2)). Share black was standardized so that it has a mean of zero and a standard deviation of one. Thus, counties with the average share black saw a decrease in the pellagra death rate by 16%, while every standard deviation increase in the share black resulted in an

[13]There is strong evidence that the long-run effects of nutrition on health and socioeconomic status accrue mainly during the first three years of life. See, for example, Hoddinott et al. (2008) and Schroeder et al. (1995)

additional 15% decrease in the pellagra death rate. Note that the main effect for share black is not included because it is absorbed by the county fixed effects.

Next, we demonstrate that men born after the arrival of the boll weevil were significantly taller than men born prior to its arrival, and that these results are driven by black men. To perform this analysis we use the *U.S. World War II Army Enlistment Records, 1938-1946* from the National Archives and Records Administration. These records contain enlistment information for over 8.5 million individuals who served in World War II. Importantly, the records contain information on the state and county an individual was residing in when they enlisted as well as the individual's height (in inches). We restrict these records to men who lived in the same state they were born in, were drafted, were born after 1915 and before 1925 (there are very few records for individuals born after 1925), had a valid height and weight and were living in a state that had at least one county that was invaded by the boll weevil after 1914.[14] We restrict to individuals born before 1925 because those born after 1925 might still have been growing when they enlisted. After making these restrictions we make the assumption that individuals are living in the same county they were born in when they enlisted. We recognize that, in light of the information presented in this paper, this is likely not an accurate assumption. However, any movements made within the state of birth (since we are restricting to men who were living in their state of birth) will only introduce measurement error and bias our coefficient estimates towards zero.

Columns (3) and (4) of Table 6 shows results from regressions when height (in inches) is the dependent variable. Both columns control for fixed effects for: the enlistee's county of residence, birth year, and year of enlistment. Column (3) shows that enlistees born after the arrival of the weevil were about 0.07 inches taller than enlistees born prior to its arrival. Column (4) shows that this effect is almost entirely concentrated in black enlistees.

[14] To serve in WWII an individual had to be between 5 and 6.5 feet tall and weigh over 105 pounds. Thus, a valid height is between 60 and 78 inches and a valid weight is 105 pounds and above. States that had at least one county invaded by the boll weevil after 1914 are: Alabama, Arkansas, Florida, Georgia, Missouri, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, and Virginia.

If sons born after the arrival of the boll weevil grew up in a better nutrition environment then we would expect to see differences in long-run outcomes between brothers in the same household. We find evidence supportive of this idea in Table 7, which re-estimates the results from Table 3 on the set of brothers in our sample (individuals that have the same father) where at least one brother was born prior to the arrival of the boll weevil and at least one brother was born after its arrival. While none of the coefficients are significant, the magnitudes are economically meaningful and suggest an improvement in early life conditions. In particular, black brothers born after the boll weevil had wages that were about 3.5% higher and attained 0.22 years of additional schooling compared to their siblings born prior to the weevil.[15]

In this section, we presented evidence that demonstrated that individuals born after the arrival of the boll weevil experienced plausibly better early childhood nutrition environments and that black children disproportionately benefited from this improved nutrition. We showed that the pellagra death rate (a disease caused by nutritional deficiencies) significantly decreased after the arrival of the boll weevil and these decreases were concentrated in counties with high African American shares of the population. We also showed the black WWII enlistees born after the boll weevil were significantly taller. When comparing brothers within the same household we found that black brothers born after the boll weevil had higher wages and years of schooling than their brothers that were born before its arrival.

7 Conclusion

This paper studied the effects of a large exogenous and sustained negative agricultural shock, the boll weevil, on black-white inequality in the first half of the twentieth century. To do this we use complete count census data to generate a linked sample of fathers and their sons and compare sons whose fathers initially resided in the same county. We find

[15]These percentages are the linear combination of coefficients from the born post boll weevil term and the interaction term in columns (1) and (3).

the shock impacted black and white sons differently. Although white sons born before and after the boll weevil had similar wages and schooling outcomes, black sons born after the boll weevil had significantly higher wages and years of schooling. The relative gains of black sons born after the boll weevil narrowed the black-white wage gap by 12% without controls for schooling and by 22% with controls for schooling. The relative gains for black sons born after the boll weevil were widespread. Specifically, they were not driven by the 5% of black fathers who migrated out of the South. Sons of black fathers that did not move out of the South still saw decreases in the black-white wage gap of up to 24%. The available evidence suggests that the gains were likely driven by improvements in early life conditions, especially early life nutrition, for black sons born after the weevil's arrival.

This paper sheds new light on two issues of importance – the drivers of changes in the black-white wage gap and the effects of a large negative agricultural shock on a highly impoverished group. Most of the literature on the black-white wage gap focuses on migration as a source of change. We find evidence of this as well, but also find large effects of the boll weevil as a driver of change for black sons born after its arrival to families that did not migrate. The boll weevil had negative short run effects on cotton production and led to large scale migration. Strikingly, it appears that boll weevil improved early life conditions for black sons born after its arrival. The likely mechanism is improved nutrition and possibly less strenuous working conditions for pregnant women. Thus agricultural disasters can lead to relative improvements for impoverished populations in the next generation.

References

Abramitzky, Ran, Leah Boustan, and Katherine Eriksson. 2019. "To the new world and back again: Return migrants in the age of mass migration." *ILR Review*, 72(2): 300–322.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2012. "Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration." *American Economic Review*, 102(5): 1832–56.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A nation of immigrants: Assimilation and economic outcomes in the age of mass migration." *Journal of Political Economy*, 122(3): 467–506.

Acemoglu, Daron, and Joshua Angrist. 2000. "How large are human-capital externalities? Evidence from compulsory schooling laws." *NBER macroeconomics annual*, 15: 9–59.

Adhvaryu, Achyuta, Steven Bednar, Teresa Molina, Quynh Nguyen, and Anant Nyshadham. 2016. "When It Rains It Pours: The Long-run Economic Impacts of Salt Iodization in the United States." *Review of Economics and Statistics*, 1–45.

Ager, Philipp, Markus Brueckner, and Benedikt Herz. 2017*a*. "The boll weevil plague and its effect on the southern agricultural sector, 1889–1929." *Explorations in Economic History*, 65: 94–105.

Ager, Philipp, Markus Brueckner, and Benedikt Herz. 2017*b*. "Structural Change and the Fertility Transition in the American South."

Aizer, Anna, and Janet Currie. 2014. "The intergenerational transmission of inequality: maternal disadvantage and health at birth." *Science*, 344(6186): 856–861.

Almond, Douglas, and Bhashkar Mazumder. 2011. "Health capital and the prenatal environment: the effect of Ramadan observance during pregnancy." *American Economic Journal: Applied Economics*, 3(4): 56–85.

Almond, Douglas, and Janet Currie. 2011. "Killing me softly: The fetal origins hypothesis." *Journal of economic perspectives*, 25(3): 153–72.

Almond, Douglas, Hilary W Hoynes, and Diane Whitmore Schanzenbach. 2011. "Inside the war on poverty: The impact of food stamps on birth outcomes." *The review of economics and statistics*, 93(2): 387–403.

Almond, Douglas, Janet Currie, and Mariesa Herrmann. 2012. "From infant to mother: Early disease environment and future maternal health." *Labour Economics*, 19(4): 475–483.

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood circumstances and adult outcomes: Act II." *Journal of Economic Literature*, 56(4): 1360–1446.

Baker, Richard B. 2015. "From the field to the classroom: the Boll Weevil's impact on education in Rural Georgia." *The Journal of Economic History*, 75(4): 1128–1160.

Baker, Richard B, John Blanchette, and Katherine Eriksson. 2020. "Long-run impacts of agricultural shocks on educational attainment: Evidence from the boll weevil." *The Journal of Economic History*, 80(1): 136–174.

Bhalotra, Sonia R, and Atheendar Venkataramani. 2015. "Shadows of the captain of the men of death: Early life health interventions, human capital investments, and institutions." *Human Capital Investments, and Institutions (August 8, 2015).*

Bloome, Deirdre, James Feigenbaum, and Christopher Muller. 2017. "Tenancy, Marriage, and the Boll Weevil Infestation, 1892–1930." *Demography*, 54(3): 1029–1049.

Boustan, Leah Platt. 2009. "Competition in the promised land: Black migration and racial wage convergence in the North, 1940–1970." *The Journal of Economic History*, 69(3): 755–782.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz. 2016. "The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment." *American Economic Review*, 106(4): 855–902.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the land of opportunity? The geography of intergenerational mobility in the United States." *The Quarterly Journal of Economics*, 129(4): 1553–1623.

Clay, Karen, Ethan Schmick, and Werner Troesken. 2019. "The rise and fall of pellagra in the American South." *The Journal of Economic History*, 79(1): 32–62.

Collins, William J, and Marianne H Wanamaker. 2014. "Selection and economic gains in the great migration of African Americans: new evidence from linked census data." *American Economic Journal: Applied Economics*, 6(1): 220–52.

Collins, William J, and Marianne H Wanamaker. 2015. "The great migration in black and white: New evidence on the selection and sorting of southern migrants." *The journal of economic history*, 75(4): 947–992.

Derenoncourt, Ellora. 2019. "Can you move to opportunity? Evidence from the Great Migration."

Edgar, Walter. 2012. *South Carolina in the Modern Age.* Univ of South Carolina Press.

Feigenbaum, James, Soumyajit Mazumder, and Cory Smith. 2019. "When Coercive Economies Fail: The Political Economy of the US South After the Boll Weevil."

Ferrie, Joseph P. 1996. "A new sample of males linked from the public use microdata sample of the 1850 US federal census of population to the 1860 US federal census manuscript schedules." *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 29(4): 141–156.

Hoddinott, John, John A Maluccio, Jere R Behrman, Rafael Flores, and Reynaldo Martorell. 2008. "Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults." *The lancet*, 371(9610): 411–416.

Hornbeck, Richard. 2012. "The enduring impact of the American Dust Bowl: Short- and long-run adjustments to environmental catastrophe." *American Economic Review*, 102(4): 1477–1507.

Hornbeck, Richard, and Suresh Naidu. 2014. "When the levee breaks: black migration and economic development in the American South." *American Economic Review*, 104(3): 963–90.

Hoynes, Hilary, Doug Miller, and David Simon. 2015. "Income, the earned income tax credit, and infant health." *American Economic Journal: Economic Policy*, 7(1): 172–211.

Hunter, Walter David, and Bert Raymond Coad. 1923. *The boll-weevil problem.* US Dept. of Agriculture.

Lange, Fabian, Alan L Olmstead, and Paul W Rhode. 2009. "The impact of the boll weevil, 1892–1932." *The Journal of Economic History*, 69(3): 685–718.

Lindo, Jason M. 2011. "Parental job loss and infant health." *Journal of health economics*, 30(5): 869–879.

Linnemayr, Sebastian, and Harold Alderman. 2011. "Almost random: Evaluating a large-scale randomized nutrition program in the presence of crossover." *Journal of Development Economics*, 96(1): 106–114.

Long, Jason, and Joseph Ferrie. 2013. "Intergenerational occupational mobility in Great Britain and the United States since 1850." *American Economic Review*, 103(4): 1109–37.

Margo, Robert A. 1995. "Explaining black-white wage convergence, 1940–1950." *ILR Review*, 48(3): 470–481.

Ransom, Roger L, and Richard Sutch. 2001. *One kind of freedom: The economic consequences of emancipation.* Cambridge University Press.

Rossin, Maya. 2011. "The effects of maternity leave on children's birth and infant health outcomes in the United States." *Journal of health Economics*, 30(2): 221–239.

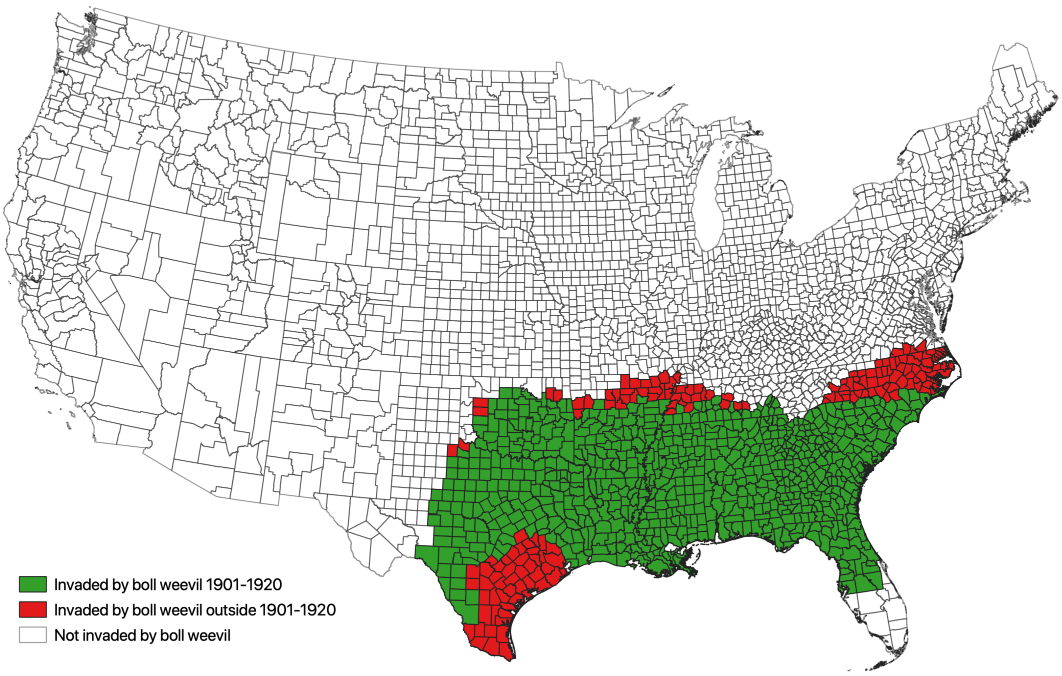
Saavedra, Martin, and Tate Twinam. 2020. "A machine learning approach to improving occupational income scores." *Explorations in Economic History*, 75: 101304.

Schroeder, Dirk G, Reynaldo Martorell, Juan A Rivera, Marie T Ruel, and Jean-Pierre Habicht. 1995. "Age differences in the impact of nutritional supplementation on growth." *The Journal of nutrition*, 125(suppl_4): 1051S–1059S.

U.S. Department of Agriculture, Bureau of Agricultural Economics. 1951. *Statistics on cotton and related data. Statistical Bulletin No. 99.* GPO.

Figures and Tables

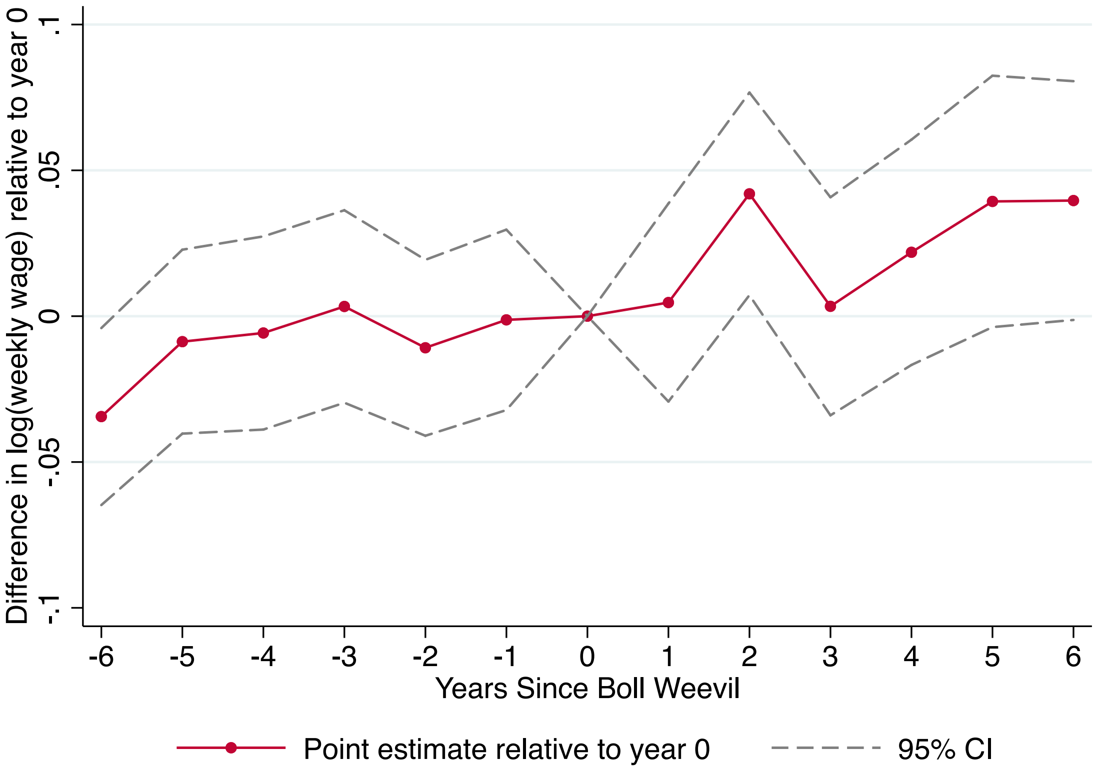
Figure 1: Counties invaded by boll weevil



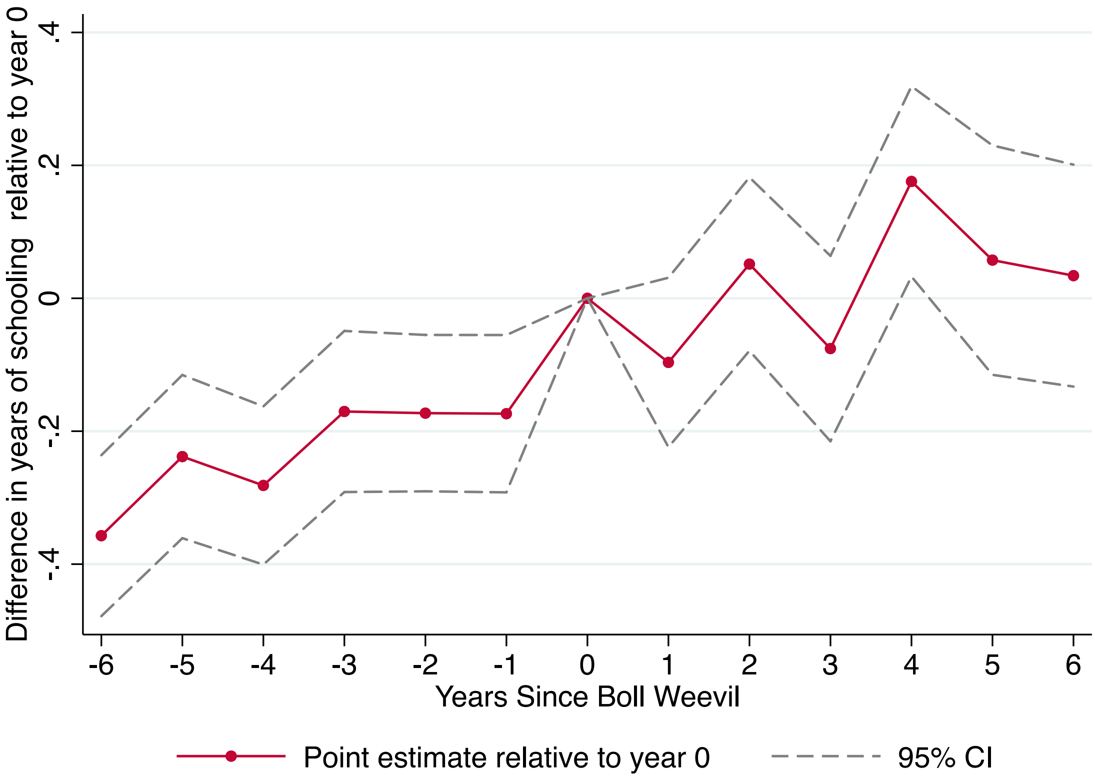
Notes: This map displays counties that were invaded by the boll weevil.

Figure 2: Event studies

Panel A: Weekly wage



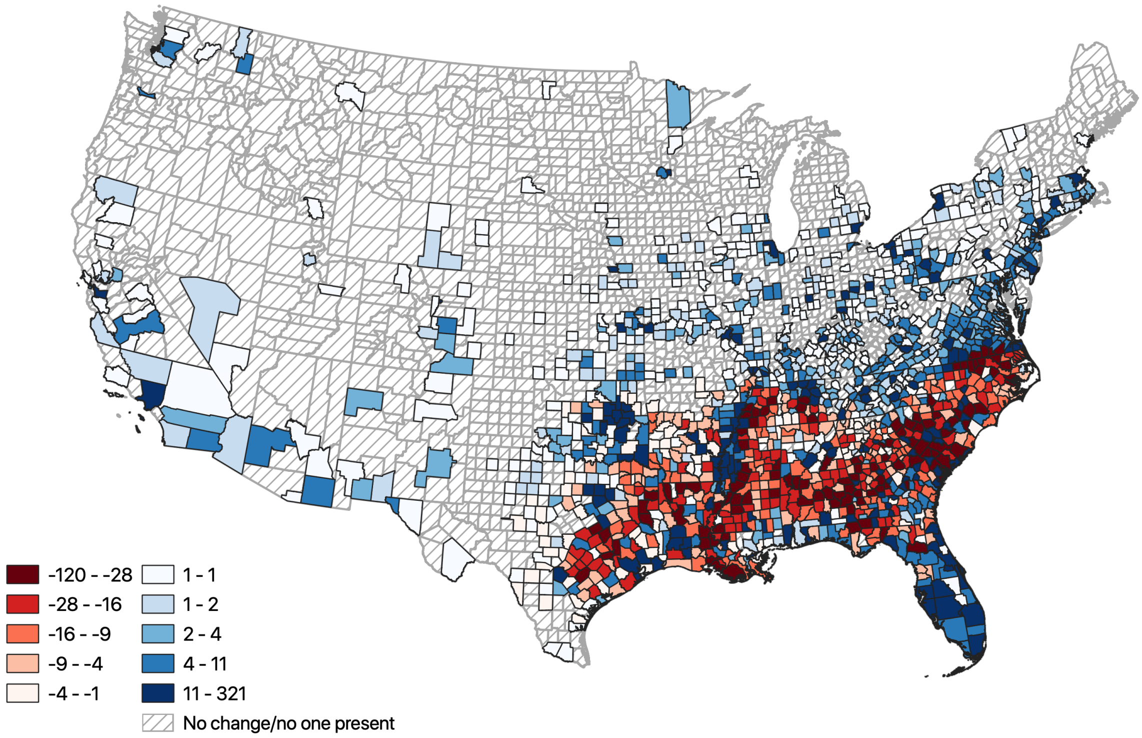
Panel B: Years of schooling



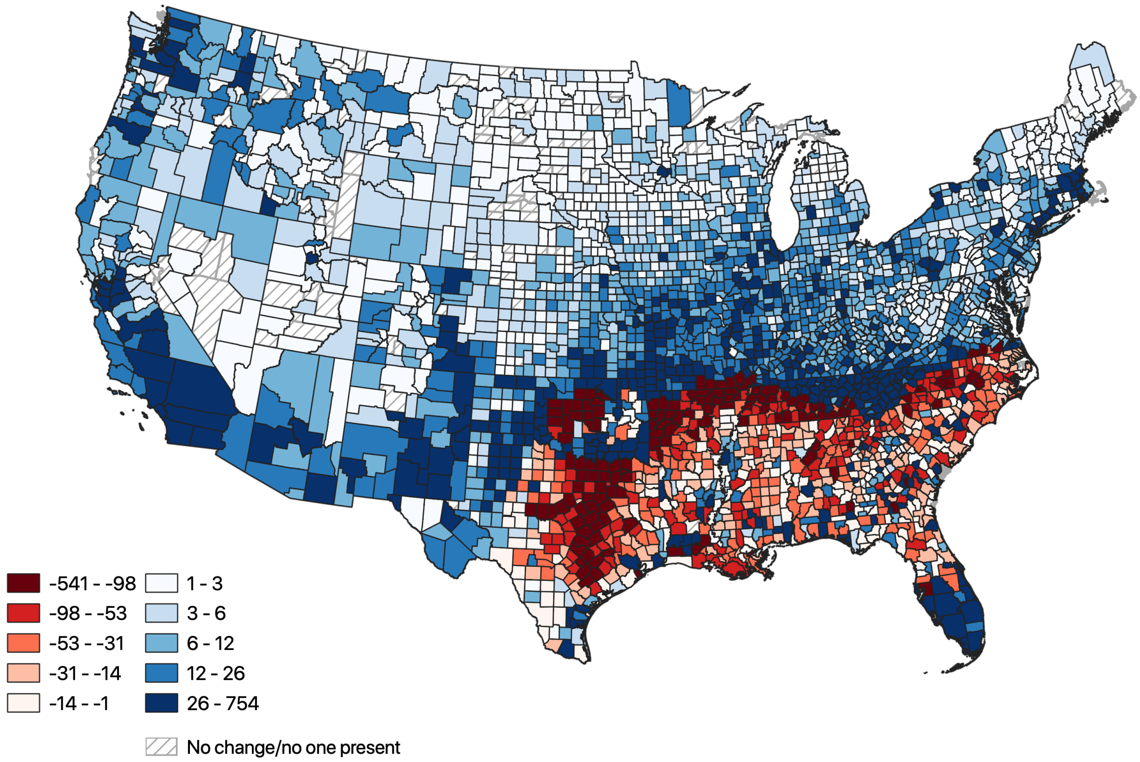
Notes: These figures show estimates of the coefficients τ_k from Equation 2 in the text. The coefficients for -6 and 6 years since the boll weevil are estimated using sons born 6 or more years before/after the boll weevil.

Figure 3: Father Mobility by Race

Panel A: Black Fathers



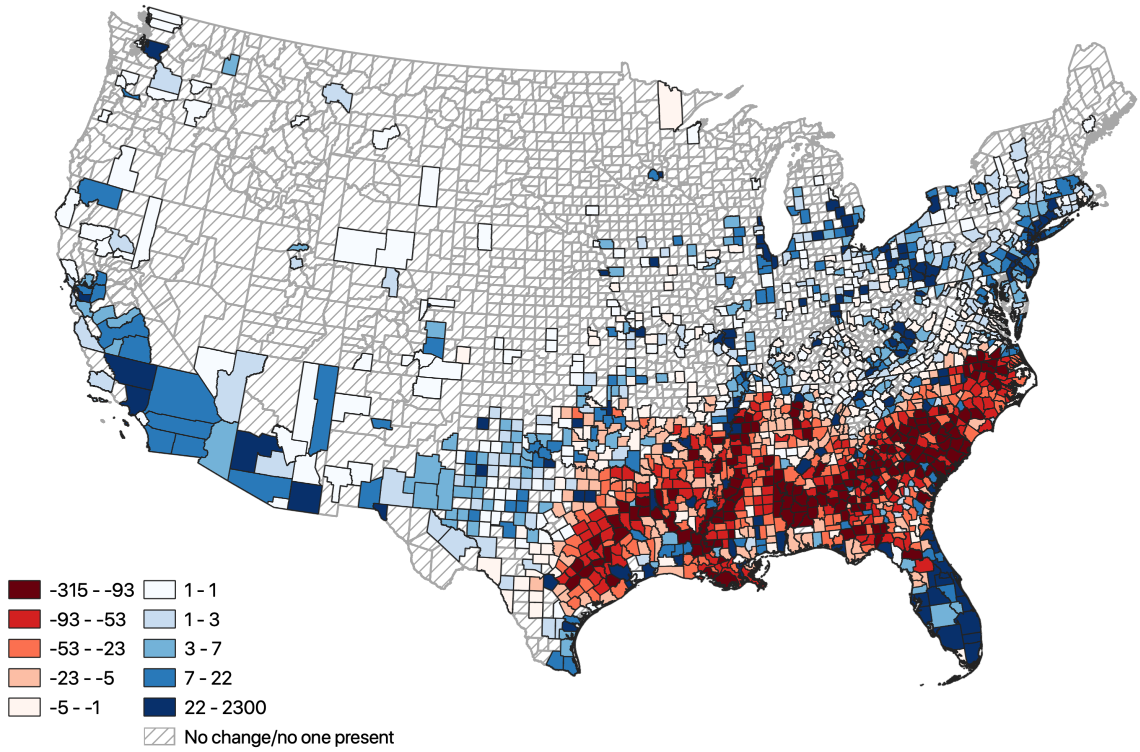
Panel B: White Fathers



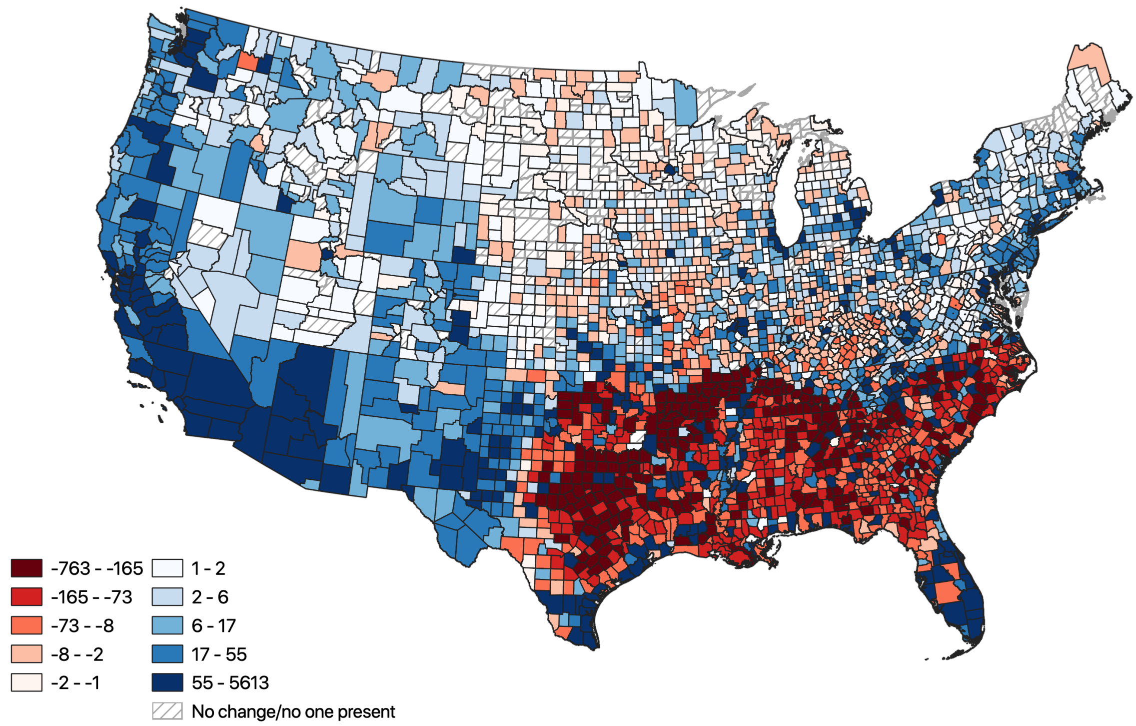
Notes: These maps shows the change in the number of black and white fathers from our sample that are living in a county. To construct these maps we calculated the total number of black or white fathers from our sample that moved into a county between the first time they were observed (either in 1900 or 1910) and the next time they were observed (either in 1910 or 1920). We then subtracted the total number of fathers from our sample that moved out of the county to get a measure of net migration. Counties shaded in blue saw an increase in the number of fathers (in-migration), while counties shaded in red saw a decrease (out-migration). The values of the bins were chosen so that an equal number of counties would be in each bin. Because we require that a father must be living in a county invaded by the boll weevil when we first observe him (either in 1900 or 1910) counties not invaded by the boll weevil can only experience an in-migration of fathers.

Figure 4: Son Mobility by Race

Panel A: Black Sons



Panel B: White Sons



Notes: These maps shows the change in the number of black and white sons from our sample that are living in a county. To construct these maps we calculated the total number of black or white sons from our sample that moved into a county between the first time they were observed (in 1900, 1910 or 1920) and 1940. We then subtracted the total number of sons that moved out of the county from our sample to get a measure of net migration. Counties shaded in blue saw an increase in the number of sons (in-migration), while counties shaded in red saw a decrease (out-migration). The values of the bins were chosen so that an equal number of counties would be in each bin.

Table 1: Summary statistics

	Black (1)	White (2)
Panel A: Fathers		
Percent living in same county	0.37	0.46
Percent living in same state but different county	0.38	0.27
Percent living in different state in South	0.19	0.18
Percent moving out of South	0.05	0.08
Percent moving from rural to urban	0.15	0.12
Father's Initial LIDO score	6.36	14.52
Change in Father's LIDO score (2nd Census - 1st Census)	0.90	0.08
Father is a farmer	0.78	0.67
Father's age	37.46	38.69
Observations	41862	136357
Panel B: Sons		
Living in father's initial county	0.17	0.26
Moved out of father's initial county; same state	0.45	0.42
Moved out of father's initial state; stayed in South	0.19	0.20
Moved out of father's initial state; moved out of South	0.20	0.12
Moved out of birth state	0.35	0.26
Weekly wage	12.72	24.29
Years of schooling	4.97	8.46
Percent born post BW	0.09	0.11
Age in 1940	37.70	37.62
Observations	54871	194762

Table 2: The boll weevil and long-run outcomes

	Log(weekly wage)		Schooling
	(1)	(2)	(3)
Panel A			
Born post BW	0.0230**	0.00729	0.148***
	(0.00926)	(0.00813)	(0.0376)
Years of schooling		X	
Observations	154664	154664	249633
Groups	656	656	656
R-sq.	0.0801	0.265	0.0683
Panel B			
Born post BW	0.00192	-0.00976	0.100*
	(0.0124)	(0.0111)	(0.0531)
Born post BW * High cotton	0.0349**	0.0282**	0.0783
	(0.0144)	(0.0132)	(0.0574)
Years of schooling		X	
Observations	154664	154664	249633
Groups	656	656	656
R-sq.	0.0802	0.265	0.0683

Notes: Panel A of this table displays estimates for Equation 1 in the text, but excludes the black indicator variable and the interaction term. Panel B provides estimates for Equation 1, but includes an interaction if a county had a high cotton acreage as a share of total farm acreage (defined as having a cotton share of total farm acreage above 11%). All columns control for: father's initial county fixed effects, birth order dummies, birth year dummies, and census enumeration year dummies. Birth order is determined by the age of the sons who have the same father in the censuses. Thus, it does not take into account older siblings who either moved out of the house or died before the censuses were taken. Standard errors are clustered at the fathers-initial-county level.

$* = p < 0.10$

$** = p < 0.05$

$*** = p < 0.01$

Table 3: Differences in long-run outcomes by race

	Log(weekly wage)		Schooling
	(1)	(2)	(3)
Black	-0.632***	-0.320***	-3.542***
	(0.00950)	(0.0106)	(0.0440)
Born post BW	0.00172	-0.00753	0.0639*
	(0.00916)	(0.00826)	(0.0367)
Born post BW * Black	0.0775***	0.0708***	0.202***
	(0.0180)	(0.0170)	(0.0678)
Observations	154664	154664	249633
Groups	656	656	656
R-sq.	0.153	0.280	0.174
Years of Schooling		X	

Notes: This table displays estimates for Equation 1 in the text. All columns control for: father's initial county fixed effects, birth order dummies, birth year dummies, and census enumeration year dummies. Birth order is determined by the age of the sons who have the same father in the censuses. Thus, it does not take into account older siblings who either moved out of the house or died before the censuses were taken. Standard errors are clustered at the fathers-initial-county level.
$* = p < 0.10$
$** = p < 0.05$
$*** = p < 0.01$

Table 4: Changes in father SES from migrating out of boll weevil counties

	Change in OCC score (1)	Change in LIDO score (2)
Black	-0.0203	-0.149*
	(0.0865)	(0.0843)
Father moves out of South	-0.838***	3.266***
	(0.202)	(0.201)
Black * Father moves out of South	1.851***	3.059***
	(0.325)	(0.312)
Father moves to urban from rural	10.04***	9.559***
	(0.166)	(0.177)
Black * Father moves urban	-3.720***	-1.586***
	(0.217)	(0.218)
Observations	155414	136677
R-sq.	0.0936	0.166
Mean of initial LIDO score	17.69	12.28

Notes: This table displays a descriptive regression where the dependent variable is the change in fathers' OCC or LIDO scores from the first time they are observed to the second time they are observed. All columns control for: father's initial county fixed effects, age dummies, and census enumeration year dummies. Standard errors are clustered at the fathers-initial-county level.
$* = p < 0.10$
$** = p < 0.05$
$*** = p < 0.01$

Table 5: Differences in the returns to father migration by race

	Log(weekly wage)		
†Father moved:	Within state	Within South	Out of South
	(1)	(2)	(3)
Black	-0.330***	-0.315***	-0.318***
	(0.0119)	(0.0106)	(0.0106)
Born post BW	-0.0104	-0.00325	-0.00656
	(0.00890)	(0.00910)	(0.00850)
Born post BW * Black	0.0814***	0.0684***	0.0693***
	(0.0213)	(0.0190)	(0.0172)
Father moved†	0.0221***	0.0431***	0.124***
	(0.00754)	(0.00773)	(0.00976)
Black * Father moved†	0.0270**	-0.0278**	-0.0453**
	(0.0110)	(0.0132)	(0.0185)
Born post BW * Father moved†	0.0231*	-0.000379	0.0511**
	(0.0128)	(0.0188)	(0.0256)
Born post BW * Black * Father moved†	-0.0315	0.00827	0.0774
	(0.0343)	(0.0406)	(0.0738)
Observations	154664	154664	154664
Groups	656	656	656
R-sq.	0.282	0.282	0.282

Notes: This table displays estimates for Equation 1 in the text, but includes a triple interaction for a father migrating. All columns control for: father's initial county fixed effects, birth order dummies, birth year dummies, census enumeration year dummies, dummies for the number of years of schooling, and a full set of dummy variables for other possible moves (e.g. column (1) controls for dummy variables if sons fathers moved states within the South or moved out of the South). Birth order is determined by the age of the sons who have the same father in the censuses. Thus, it does not take into account older siblings who either moved out of the house or died before the censuses were taken. Standard errors are clustered at the fathers-initial-county level.

$* = p < 0.10$

$** = p < 0.05$

$*** = p < 0.01$

Table 6: The boll weevil's effect pellagra in North Carolina and the height of WWII enlistees

	Log(pellagra death rate)		Height (inches)	
	(1)	(2)	(3)	(4)
Post BW	-0.231***	-0.161**	0.0683***	0.0337
	(0.0662)	(0.0685)	(0.0191)	(0.0214)
Black				-0.616***
				(0.0347)
Post BW * Share black (or Black)		-0.146**		0.137***
		(0.0641)		(0.0381)
Observations	882	882	398747	398747
R-sq.	0.535	0.540	0.0293	0.0357

Notes: Columns (1) and (2) use county-level data for the 98 counties in North Carolina. In particular, the log of the pellagra death rate is the dependent variable. We measure percent black in a county in 1910. Percent black is standardized to have a mean of zero and a standard deviation of one. All columns control for: county fixed effects and year fixed effects. Accordingly, the main effect for percent black is absorbed by the county fixed effects since there is no time variation. Standard errors are clustered at the county. Columns (3) and (4) use individual level data for WWII enlistees. In particular, we use the *U.S. World War II Army Enlistment Records, 1938-1946* from the National Archives and Records Administration. All columns control for fixed effects for the enlistee's county of residence, birth year, and year of enlistment. See the text for more details on the data. Standard errors are clustered at the county-of-residence at time of enlistment level.
$* = p < 0.10$
$** = p < 0.05$
$*** = p < 0.01$

Table 7: Brother wage and schooling comparisons

	Log(weekly wages)		Schooling
	(1)	(2)	(3)
Born post BW	0.0138	0.00841	0.111
	(0.0363)	(0.0339)	(0.110)
Born post BW * Black	0.0220	0.0106	0.113
	(0.0477)	(0.0472)	(0.150)
Observations	13760	13760	27667
Groups	5963	5963	11129
R-sq.	0.605	0.642	0.599
Years of Schooling		X	

Notes: This table displays estimates for Equation 1 in the text, but uses father fixed effect as opposed to father's initial county fixed effects. All columns control for: father fixed effects, birth order dummies, birth year dummies, and census enumeration year dummies. Birth order is determined by the age of the sons who have the same father in the censuses. Thus, it does not take into account older siblings who either moved out of the house or died before the censuses were taken. Standard errors are clustered at the fathers-initial-county level.

$* = p < 0.10$

$** = p < 0.05$

$*** = p < 0.01$

A Appendix

Table A.1: Linking results by race

	1900-1910 Blacks (1)	1900-1910 Whites (2)	1910-1920 Blacks (3)	1910-1920 Whites (4)
Fathers with at least one son under 19	181,795	362,624	383,984	647,485
Linked fathers	40,441	109,020	82,890	191,619
(Match rate)	22%	30%	22%	30%
Sons of linked fathers	97,962	256,181	173,476	422,714
Linked sons of linked fathers*	19,188	71,367	35,683	123,395
(Match Rate)	20%	28%	21%	29%
Number of fathers	14,887	50,740	26,975	85,617
Linked brothers	893	7,611	3,487	15,676

∗ To be included in this sample a son has to be from a family with under 10 sons. The match rate we report for linked sons of linked fathers is after we discard matches that are clearly incorrect. An incorrect match is defined as a family where the father is initially observed in the cotton belt (in 1900 or 1910), then observed outside the cotton belt (in 1910 or 1920), but is observed with a son who was born in a state outside the cotton belt prior to the father being observed in the cotton belt. For example, an incorrect match would be a father observed in North Carolina in 1910 and Pennsylvania in 1920, but in 1920 he is observed with a son who was born in 1908 in Pennsylvania.

Table A.2: Comparison of Linked and Base Samples

	Black Males		White Males	
	Linked	Attempted to Link	Linked	Attempted to Link
Panel A: Fathers				
Age	37.751***	39.526	39.542***	40.262
Literacy	0.596***	0.554	0.911***	0.893
In owner occupied housing†	0.230***	0.220	0.523***	0.503
In urban area	0.114**	0.117	0.162	0.162
Farmer	0.759***	0.752	0.642***	0.636
N	123331	565779	300639	1010109
Panel B: Sons				
Age	8.187***	8.394	8.553***	8.696
Literacy	0.280***	0.291	0.464***	0.474
In owner occupied housing†	0.251***	0.244	0.580***	0.552
In urban area	0.100***	0.091	0.149***	0.140
Father in same county	0.372**	0.377	0.450***	0.437
Father in same state, different county	0.389*	0.385	0.278***	0.289
Father in different state, in South	0.194	0.195	0.186***	0.196
Father moved out of South	0.045	0.043	0.086***	0.079
Father moved rural to urban	0.144	0.144	0.115***	0.117
N	54871	271438	194762	726822

Notes: The stars report significance from a test of equality of means. All tests were conducted relative to the base sample that we attempted to link.

†: There are a few observations for which owner occupied housing is not available. In Panel A there are 123,282 linked black father, 565,524 black sons that we attempted to link, 300,136 linked white fathers, and 1,008,275 white sons that we attempted to link that have owner occupied housing status available. In Panel B there are 54,846 linked black sons, 271,274 black sons that we attempted to link, 194,446 linked white sons, and 725,647 white sons that we attempted to link that have owner occupied housing status available.

$* = p < 0.10$

$** = p < 0.05$

$*** = p < 0.01$

Table A.3: Determinants of Fathers' Migration

	Black fathers		White fathers	
	Moved out of county	Moved out of state	Moved out of county	Moved out of state
	(1)	(2)	(3)	(4)
Father initially farmer	-0.103***	-0.0243***	-0.0649***	-0.0631***
	(0.00959)	(0.00405)	(0.0167)	(0.00596)
Father initial age under 20	0.147***	0.0158	0.147***	-0.0352***
	(0.0355)	(0.0182)	(0.0293)	(0.0131)
Father age 20-29	0.0752***	0.0243***	0.0767***	-0.0328***
	(0.00621)	(0.00281)	(0.00433)	(0.00266)
Father age 30-39	0.0398***	0.0158***	0.0426***	-0.0146***
	(0.00591)	(0.00264)	(0.00351)	(0.00215)
Constant	0.674***	0.0549***	0.552***	0.138***
	(0.00960)	(0.00431)	(0.0164)	(0.00764)
Observations	41862	41862	136357	136357
R-sq.	0.0124	0.00468	0.00757	0.0133

Notes: In columns (1) and (3) the dependent variable is an indicator that takes a value of one if they father moved out of their initial county and a zero otherwise. In columns (2) and (4) the dependent variable is an indicator that takes a value of one if the father moved out of the South and a zero otherwise. The omitted age category is fathers aged 40 and over.

$* = p < 0.10$

$** = p < 0.05$

$*** = p < 0.01$

Table A.4: Determinants of Sons Moving out of Father's Original State

	Black sons		White sons	
	Moved out of father's initial county	Moved out of South	Moved out of father's initial county	Moved out of South
	(1)	(2)	(3)	(4)
Father moved states	-0.0888***	0.0426***	-0.172***	0.148***
	(0.00404)	(0.00428)	(0.00463)	(0.00702)
Born post BW	0.0388***	-0.0525***	-0.0367***	0.0227***
	(0.00929)	(0.00662)	(0.00624)	(0.00515)
Years of schooling	-0.00466***	0.0321***	-0.00576***	0.00753***
	(0.00102)	(0.000708)	(0.000931)	(0.000249)
Age 20-29 in 1940	-0.0296**	-0.0165**	-0.0138	-0.00346
	(0.0124)	(0.00793)	(0.0262)	(0.00812)
Age 30-39	-0.0520***	0.0393***	-0.0428**	0.0222***
	(0.00931)	(0.00708)	(0.0212)	(0.00541)
Age 40-49	-0.0382***	0.0561***	-0.00203	0.0150***
	(0.00740)	(0.00623)	(0.0184)	(0.00395)
Constant	0.245***	0.00553	0.380***	0.00360
	(0.00936)	(0.00643)	(0.0278)	(0.00472)
Observations	54871	54871	194762	194762
R-sq.	0.0157	0.0777	0.0360	0.0506

Notes: In columns (1) and (3) the dependent variable is an indicator that takes a value of one if a son moved our of their father's initial county and a zero otherwise. In columns (2) and (4) the dependent variable is an indicator that takes a value of one if the son moved out of the South and a zero otherwise. The omitted age category is sons aged 50 and over. $* = p < 0.10$
$** = p < 0.05$
$* * * = p < 0.01$

Table A.5: Differences in long-run outcomes by race using linking with exact name and age

	Log(weekly wage)		Schooling
	(1)	(2)	(3)
Black	-0.680***	-0.346***	-3.806***
	(0.0165)	(0.0178)	(0.0645)
Born post BW	0.00789	-0.0116	0.141**
	(0.0194)	(0.0185)	(0.0699)
Born post BW * Black	0.0791*	0.0733*	0.232
	(0.0461)	(0.0441)	(0.162)
Observations	33410	33410	53920
Groups	656	656	656
R-sq.	0.149	0.279	0.159
Years of Schooling		X	

Notes: This table displays estimates for Equation 1 in the text. All columns control for: father's initial county fixed effects, birth order dummies, birth year dummies, and census enumeration year dummies. Birth order is determined by the age of the sons who have the same father in the censuses. Thus, it does not take into account older siblings who either moved out of the house or died before the censuses were taken. Standard errors are clustered at the fathers-initial-county level.
$* = p < 0.10$
$** = p < 0.05$
$* * * = p < 0.01$

Table A.6: Differences in long-run outcomes by race using unique names

	Log(weekly wage)		Schooling
	(1)	(2)	(3)
Black	-0.651***	-0.336***	-3.608***
	(0.0122)	(0.0135)	(0.0548)
Born post BW	0.0174	0.00272	0.0861*
	(0.0127)	(0.0119)	(0.0502)
Born post BW * Black	0.0444	0.0461*	0.301***
	(0.0283)	(0.0267)	(0.0990)
Observations	70216	70216	115714
Groups	655	655	655
R-sq.	0.158	0.284	0.165
Years of Schooling		X	

Notes: This table displays estimates for Equation 1 in the text. All columns control for: father's initial county fixed effects, birth order dummies, birth year dummies, and census enumeration year dummies. Birth order is determined by the age of the sons who have the same father in the censuses. Thus, it does not take into account older siblings who either moved out of the house or died before the censuses were taken. Standard errors are clustered at the fathers-initial-county level.

$* = p < 0.10$

$** = p < 0.05$

$*** = p < 0.01$

Table A.7: Selection Into Migration Based on Pre-migration Characteristics

	Black fathers		White fathers	
	Mean	N	Mean	N
Fathers' Initial LIDO Score				
Non-migrant	6.10	14579	14.15	58194
Within state migrant	6.28***	14652	14.21	33988
Within South migrant	6.87***	7290	14.68***	22744
Out of South migrant	7.11***	1799	17.37***	10074
Total	6.36	38320	14.52	125000
Father Initially a Farmer				
Non-migrant	0.833	15644	0.701	62825
Within state migrant	0.785***	16063	0.680***	37034
Within South migrant	0.713***	8162	0.649***	25001
Out of South migrant	0.693***	1993	0.504***	11497
Total	0.785	41862	0.669	136357
Fathers' Homeownership Rates				
Non-migrant	0.285	15643	0.629	62784
Within state migrant	0.184***	16062	0.451***	36999
Within South migrant	0.215***	8160	0.464***	24949
Out of South migrant	0.202***	1992	0.490***	11464
Total	0.229	41857	0.539	136196
Fathers' Literacy Rates				
Non-migrant	0.606	15644	0.913	62825
Within state migrant	0.566***	16063	0.899***	37034
Within South migrant	0.556***	8162	0.899***	25001
Out of South migrant	0.646***	1993	0.930***	11497
Total	0.583	41862	0.908	136357
Fathers' Initial Age				
Non-migrant	38.44	15644	39.38	62825
Within state migrant	36.12***	16063	37.03***	37034
Within South migrant	38.78**	8162	38.82***	25001
Out of South migrant	35.24***	1993	40.06***	11497
Total	37.46	41862	38.69	136357
Distance of Move (Miles)				
Within state migrant	96.11	15885	101.70	36338
Within South migrant	306.90	8101	381.76	24711
Out of South migrant	666.49	1980	838.76	11427
Total	205.37	25966	313.40	72476
N	25966		72476	

Notes: The stars report significance from a test of equality of means. All tests were conducted relative to the non-migrant mean. We do not conduct the tests for the "Total" rows.
$* = p < 0.10$
$** = p < 0.05$
$*** = p < 0.01$

Table A.8: Differences in the returns to son migration by race

	Log(weekly wage)			
†Son location (in 1940):	Father's initial county	Father's initial state; different county	Different state in South	Different state not in South
	(1)	(2)	(3)	(4)
Black	-0.394***	-0.352***	-0.413***	-0.472***
	(0.00856)	(0.0103)	(0.00939)	(0.0100)
Born post BW	-0.0149*	-0.0508***	0.00276	-0.0167*
	(0.00881)	(0.00988)	(0.00892)	(0.00878)
Born post BW * Black	0.112***	0.109***	0.113***	0.148***
	(0.0163)	(0.0193)	(0.0194)	(0.0187)
Son location (in 1940)†	-0.562***	-0.363***	-0.211***	0.136***
	(0.0305)	(0.01000)	(0.00961)	(0.0103)
Black * Son location†	-0.116***	-0.142***	0.0126	0.275***
	(0.0214)	(0.0122)	(0.0125)	(0.0136)
Born post BW * Son location†	-0.00693	0.0792***	-0.0782***	-0.0227
	(0.0242)	(0.0128)	(0.0149)	(0.0181)
Black * Born post BW * Son location†	0.0749	0.0253	0.0268	-0.0730**
	(0.0470)	(0.0274)	(0.0357)	(0.0356)
Observations	154664	154664	154664	154664
Groups	656	656	656	656
R-sq.	0.323	0.324	0.323	0.325

Notes: This table displays estimates for Equation 1 in the text, but includes a triple interaction for a son's location relative to their father's initial county. All columns control for: father's initial county fixed effects, birth order dummies, birth year dummies, census enumeration year dummies, and dummies for the number of years of schooling. Birth order is determined by the age of the sons who have the same father in the censuses. Thus, it does not take into account older siblings who either moved out of the house or died before the censuses were taken. Standard errors are clustered at the fathers-initial-county level. All columns control for a set of dummy variables for other possible locations relative to fathers' initial county. In columns (1)-(3) the omitted dummy is for sons that moved out of the South. In column (4) the omitted dummy is sons that migrated to a different state within the South.

$* = p < 0.10$

$** = p < 0.05$

$*** = p < 0.01$

Figure A.1: Multi-step linking procedure

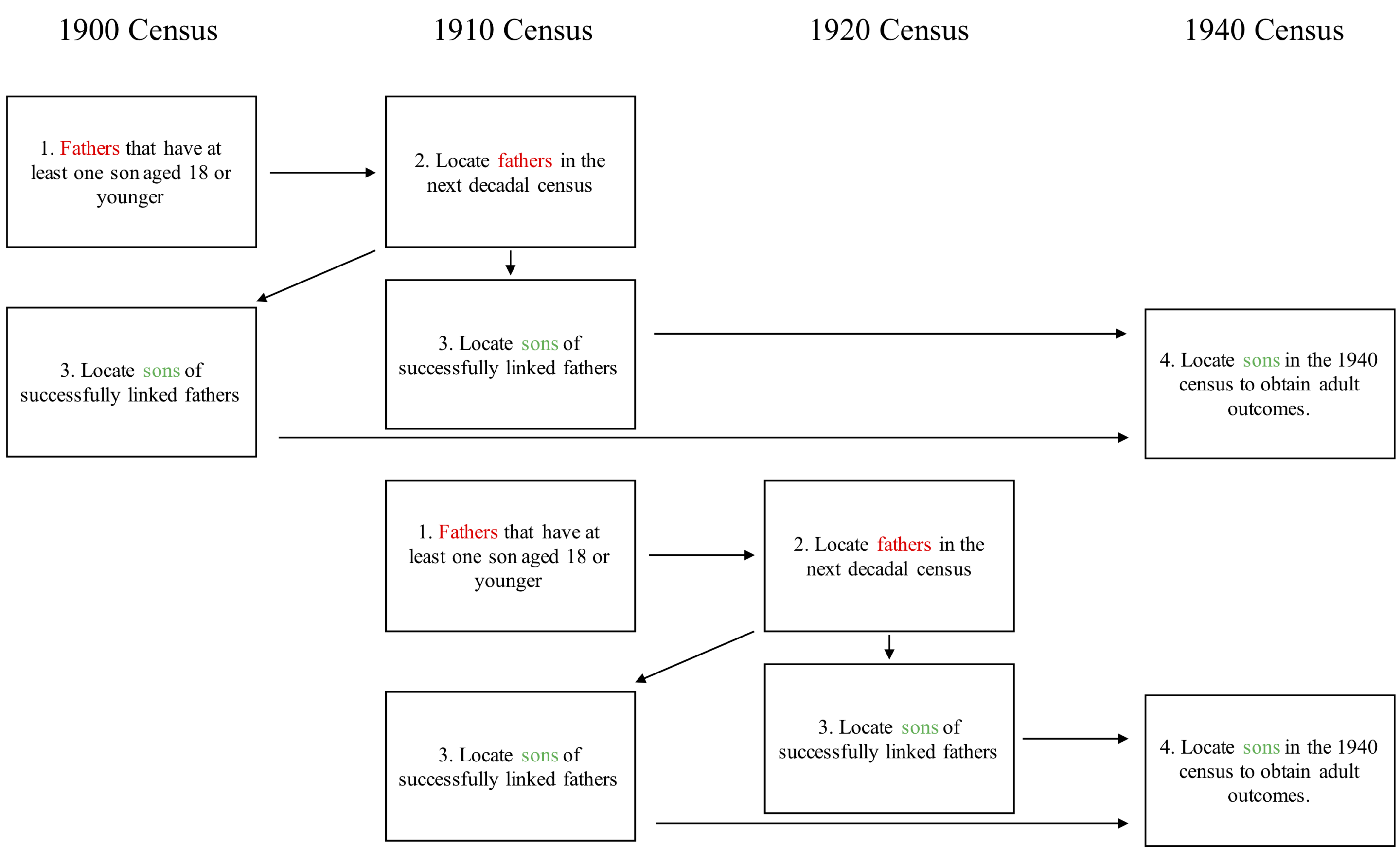
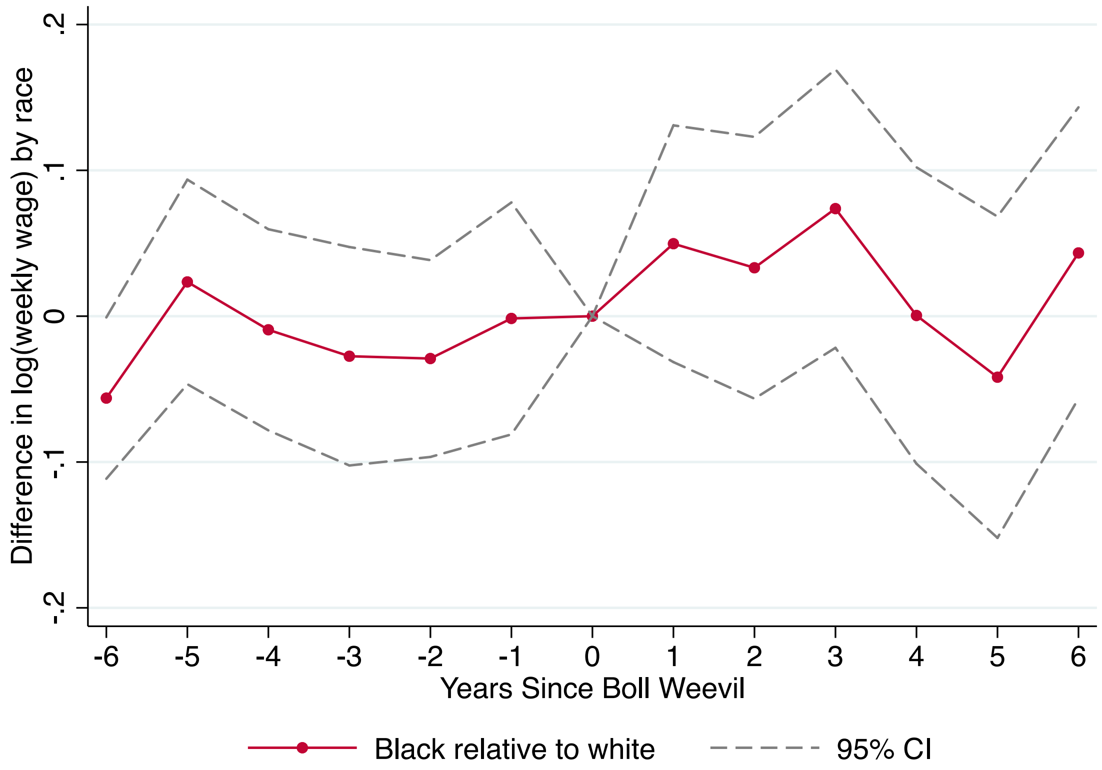
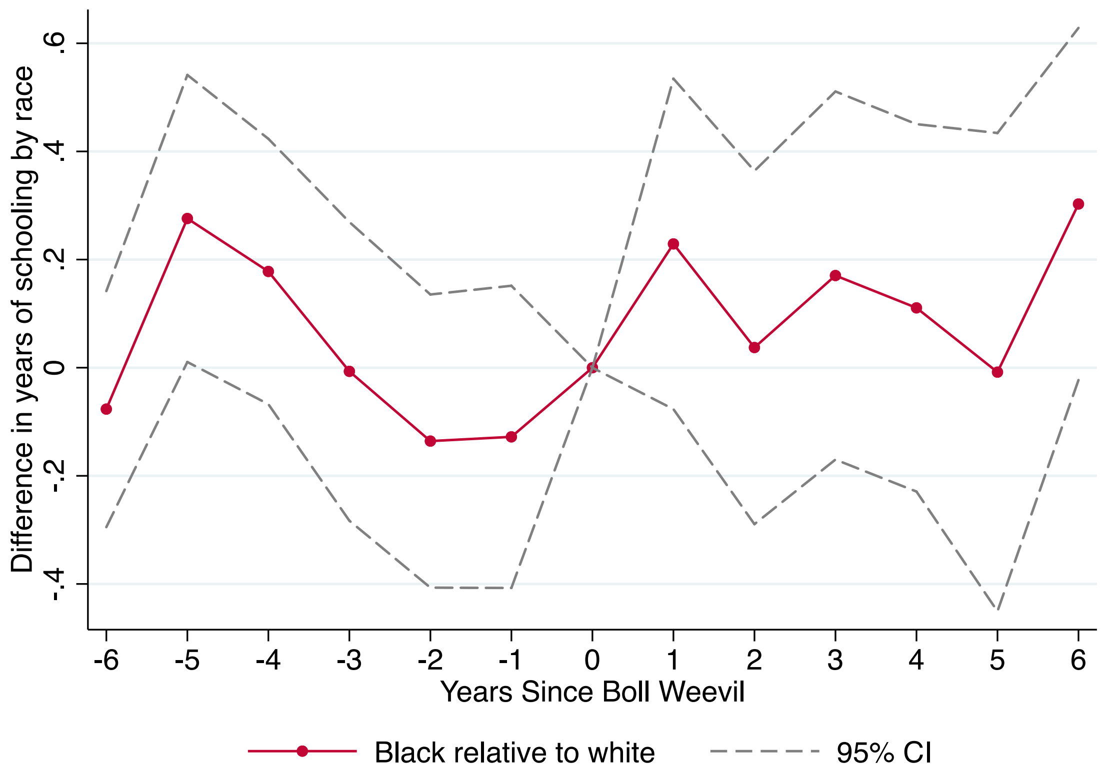


Figure A.2: Event studies

Panel A: Weekly wage



Panel B: Years of schooling



Notes: These figures show estimates of the coefficients τ_k from Equation 2 in the text, but we interact the dummy variables for year of birth relative to the boll weevil with another dummy variable if an individual is black. The coefficients for -6 and 6 years since the boll weevil are estimated using sons born 6 or more years before/after the boll weevil.