NBER WORKING PAPER SERIES

LONG-RUN IMPACTS OF AGRICULTURAL SHOCKS ON EDUCATIONAL ATTAINMENT: EVIDENCE FROM THE BOLL WEEVIL

Richard B. Baker
John Blanchette
Katherine Eriksson

Working Paper 25400
http://www.nber.org/papers/w25400

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2018

We thank workshop participants at University of California Davis, the 2018 World Economic History Congress, and North American Summer Meeting of the Econometric Society for their helpful comments and suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Long-run Impacts of Agricultural Shocks on Educational Attainment: Evidence from the Boll Weevil
Richard B. Baker, John Blanchette, and Katherine Eriksson
NBER Working Paper No. 25400
December 2018
JEL No. N32,O13

ABSTRACT

The boll weevil spread across the Southern United States from 1892 to 1922 having a devastating impact on cotton cultivation. The resulting shift away from this child labor–intensive crop lowered the opportunity cost of attending school, and thus the pest increased school enrollment and attendance. We investigate the insect's long run affect on educational attainment using a sample of adults in 1940 linked back to themselves in childhood in the county in which they were likely educated. Both whites and blacks who were young (ages 4 to 9) when the boll weevil arrived saw increased educational attainment by 0.25 to 0.35 years. These findings are not driven by concurrent shocks and are not sensitive to linking method or sample selection. Our results demonstrate the potential for conflict between child labor in agriculture and educational attainment.

Richard B. Baker
School of Business
The College of New Jersey
2000 Pennington Road
Ewing, NJ 08628
rbennettbaker@gmail.com

John Blanchette
University of California, Davis
jblanchette@ucdavis.edu

Katherine Eriksson
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
kaeriksson@ucdavis.edu

1 Introduction

A substantial body of research on developing countries documents the tradeoff parents face in choosing between sending their children to work or to school by showing that child labor reduces various measures of educational attainment and achievement, including attendance, test scores, and years of schooling (see, for example, Beegle, Dehejia, and Gatti 2009; Boozer and Suri 2001; Gunnarsson, Orazem, and Sánchez 2006; Emerson, Ponczek, and Souza 2017). However, this literature largely ignores peculiar features—such as the informal employment of children on family farms and the seasonality of labor demands—of the demand for child labor in agricultural regions, which make up a majority of the developing world. Additionally, a number of papers suggest that negative agricultural shocks have a negative (Jensen 2000) or neutral (Dammert 2008) effect on educational outcomes. To fill this gap, we exploit a unique shift in agricultural production that occurred in the early twentieth-century American South to analyze the role of a child labor–intensive crop (cotton) in determining educational attainment.

This shock to cotton production was caused by the boll weevil infestation which spread from the southern tip of Texas in 1892 to affect nearly the whole Cotton Belt in 1922 (Hunter and Coad 1923). Due to reduced returns to cotton cultivation under boll weevil conditions, and at the encouragement of the state and federal agricultural agencies, farmers substituted away from cotton to alternative crops (for example, corn, peanuts, and sweet potatoes). The tasks involved in the cultivation of these alternatives were less suitable for children, as compared to picking cotton. The resulting, and plausibly exogenous, fall in the marginal product of child labor in rural areas represents a reduction in the opportunity cost of schooling. Therefore, we predict that the boll weevil increased educational attainment for children who were young when the infestation began.

To examine the boll weevil's impact on educational attainment, we match men in the 1940 census who were born in Southern states comprising the Cotton Belt to their childhood census records in 1900, 1910, and 1920.[1] As the 1940 census was the first to inquire about the years of schooling of all respondents, the result is a matched sample linking childhood location and family background information with educational attainment, as reported in adulthood. Identifying county

[1] States with territory in the Cotton Belt include: Alabama, Arkansas, Florida, Georgia, Kansas, Kentucky, Louisiana, Mississippi, Missouri, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, and Virginia. We, however, exclude Missouri, Kansas and New Mexico because they are not in the Southern United States. Moreover, the latter two were not shown to be infested by the boll weevil (Hunter and Coad 1923).

of residence during childhood allows us to calculate how old each individual was when the boll weevil arrived. We use this information to compare the educational attainment of those affected by the boll weevil at young ages with that of older cohorts, whose education should be unaffected by the insect, in a differences-in-differences framework.

The results show that white children who were 7–9 years of age when the boll weevil arrived attained 0.2314 more years of schooling, relative to those that were 19–30 when the boll weevil arrived. We find a similar result of 0.2423 years of schooling for comparatively aged black children. Overall, there is a lack of evidence to suggest that black and white children benefited differently from the boll weevil infestation with respect to educational attainment. The gains in educational attainment decrease for older children, with those aged 16–18 when the weevil arrived seeing a comparatively modest gain of one twentieth of a year of schooling.

These results show that the seasonal demand for child labor in agriculture can have substantial negative impacts on educational attainment. Child labor–intensive crops (e.g. coffee, cotton, sugarcane, tea and tobacco) are primary agricultural products of many regions of the developing world.[2] Thus, our results are suggestive of the importance of programs that encourage the production of alternative (less child labor–intensive) crops and the adoption of technologies that reduce demand for child labor in agriculture.

Additionally, this work adds to a growing literature on the broader impacts of the boll weevil. Early efforts to understand the impact of the boll weevil on the Southern economy largely limited to examinations of state-level variation (Higgs 1976; Osband 1985). Lange, Olmstead, and Rhode (2009) revived interest in the boll weevil with their analysis of the insect's effect on crop production in the South using county-level data. They show that cotton production dramatically decreased in the years following the arrival of the boll weevil with farmers shifting resources to the production of other crops. Baker (2015) finds school enrollment rates for African-American children in Georgia increased following the boll weevil's arrival, resulting from the insect's negative effect on production of child labor–intensive cotton. Ager, Brueckner, and Herz (2017) reveal the boll weevil reduced labor force participation (particularly among females), farm wages, and the number of fixed-rent

[2]The US Department of Labor (1995) report *By the Sweat and Toil of Children* finds these crops to be particularly suited to the use of child labor and details the role of children in farming them in different regions of the developing world. A recent report by the US Department of Labor (2016), known as the *List*, identifies 17 developing countries in which cotton is produced by means of child labor. A number of developing countries producing coffee, sugarcane, tea, and tobacco are also on the *List* as exploiting the labor of children.

tenant farms, with counties more reliant on cotton experiencing greater declines, illustrating the broader impact of the boll weevil on local economies. Finally, Bloome, Feigenbaum, and Muller (2017) show the boll weevil infestation reduced the proportion of farms worked by tenants, which in turn altered incentives to marry and reduced the proportion of African Americans who were wed at a young age, revealing the weevil's effect on life-altering decisions.

2 Background

2.1 Cotton, Children, and Schooling

At the beginning of the twentieth century, the South was still an agrarian economy, as agriculture employed 57 percent of the labor force in 1910. And cotton was its staple. Cotton was the single most valuable crop in 10 of the 16 Southern states at the dawn of the twentieth century. In Alabama, Arkansas, Georgia, Mississippi, South Carolina, and Texas—the states that formed the heart of the Cotton Belt—cotton comprised more than half of the value of all crops produced (US Bureau of the Census 1913).

The widespread cultivation of cotton and its long harvest season, which lasted from September to December, generated a seasonal increase in demand for labor, as harvesting cotton was not mechanized until the mid-twentieth century. Since harvesting cotton involves repetitively picking lightweight fibers from their bolls and transferring them to a sack, it is a tedious task that is performed reasonably well by anyone over the age of five. In the cotton harvest, therefore, there was a high degree of substitutability between adult and child labor. It is not surprising, then, that agriculture was by far the largest employer of children. In 1910, 34.4 percent of 10 to 15 year-olds living in the South worked, of which 86.7 percent were employed in farming. Moreover, these youth made up 17 percent of the agricultural labor force in the South (US Bureau of the Census 1924).[3] The majority of these child laborers undoubtedly worked the cotton fields.

While some adjustments were made to the school calendar in an attempt to accommodate farm-

[3] In 1910, census enumerators were given specific instructions to inquire about the occupations of women and children, as well as men. Moreover, enumerators were instructed to record children working for their parents on a farm as farm laborers. Therefore, in comparison to earlier censuses, the 1910 census is an unusually good source of data on the labor force participation of children (Moehling 1999, 82; 2004, 79). Still, these figures might understate the extent of child labor in agriculture during the fall harvest since the 1910 census recorded employment on April 15th, at the beginning of the agricultural season when the demand for child labor was comparatively low.

ing cotton (Collins and Margo 2006), the lengthy harvest period could not be avoided altogether. Baker (2015) provides anecdotal evidence of the conflict between the demand for child-labor in farming cotton and schooling in Georgia, but this conflict is not unique to that state. The superintendent of West Baton Rouge Parish, Louisiana, noted, "a falling off in attendance at several schools during the harvesting season ... due to the scarcity of labor and the need of the children in the cotton fields" (Louisiana Department of Education 1908, 48). His colleague in Calcasieu Parish expressed similar sentiments, observing, "times to pick cotton, the harvesting of rice and other crops take many out of school for a good part of the term" (Louisiana Department of Education 1902, 60-61). To accommodate the surge in attendance following the conclusion of the cotton picking season, the superintendent of Tunica County, Mississippi planned to hire an additional teacher in each African-American district (Mississippi Department of Public Education 1907, 95). Even Texas Superintendent Arthur Lefevre noted, "the average daily attendance being unusually low during 1902-3 on account of the high price and long picking season of the cotton crop of that year" (Texas Department of Education 1905, 7).

2.2 The Boll Weevil

The adult cotton boll weevil, *Anthonomus grandis*, is a small beetle, about 6 millimeters in length, and grayish in color, with a long snout and wings.[4] The westward expansion of cotton production in the United States during the nineteenth century eventually linked the Cotton Belt with the native habitat of the boll weevil, Mexico and Central America. The boll weevil first appeared in cotton fields near Brownsville, Texas, in 1892. It then spread north and east at a steady rate. By 1922, the boll weevil could be found in virtually all cotton counties in the United States, from Texas to North Carolina (Hunter and Coad 1923).

The life cycle of the boll weevil is closely intertwined with the cotton plant. Indeed, the insect lives inside the squares and bolls of the cotton plant for three of the four stages of its life cycle (egg, larvae, and pupae), and as an adult it feeds almost exclusively on the cotton plant.[5] The reason for the boll weevil's narrow appetite is that cotton is one of only a few plants (the others being wild flora with geographically small habitats) that provide the weevil with the nutrients required

[4] Lange, Olmstead, and Rhode (2009) provide a description and history of the boll weevil.

[5] A cotton square refers to a young flower bud of the cotton plant, and a cotton boll is the fiber-producing fruit.

to produce the pheromones necessary for its reproduction. This dependence on cotton causes the insect to spend its entire life in or near cotton fields (Giesen 2011).

The boll weevil's spread had disastrous effect on cotton production in the South. Damaged squares and bolls usually dropped from the plant after being fed upon and following the deposition of eggs by the boll weevil. As a result, heavily infested areas saw significant reductions in cotton output. In a county-level analysis, Lange, Olmstead, and Rhode (2009) demonstrate that within five years of the boll weevil's arrival cotton production fell by approximately 50 percent, due primarily to reduced yield but also reduced cotton acreage.

The boll weevil's negative effect on cotton yields decreased the returns to farming cotton relative to other crops. This caused farmers to substitute away from cotton in favor of more profitable alternatives, including corn, hay, potatoes, peanuts, rice, sweet potatoes, and sugar cane. These alternative crops generated less demand for child labor than cotton. Therefore, the arrival of the boll weevil, and the subsequent reduction in cotton production, had significant implications for the whole household.

3 Data

To estimate the effect of the boll weevil on years of educational attainment, as reported in the 1940 census, we need to know the year of arrival of the boll weevil in each Cotton Belt county and where men in 1940 were living as children when the insect arrived. To do this, we digitize a USDA map that tracked the boll weevil's progress and construct a linked sample of men in the 1940 census matched to themselves as children in either the 1900, 1910, or 1920 censuses.

3.1 Mapping the Progress of the Boll Weevil

Information on the presence of the boll weevil was collected from the USDA-produced map tracking the insect's spread provided by Hunter and Coad (1923). For each year 1892–1922 the counties of the Cotton Belt were assessed as being either boll weevil free, partially infested, or fully infested. A county was considered to be partially infested in a given year if the line denoting the frontier of the boll weevil infestation in that year passed through the county leaving parts of the county visible on both sides of the line. Using this panel data, we then define three measures for each county that capture the timing of the arrival of the boll weevil: We define the *first arrival year* as the first year in which the county is coded as either partially or fully infested. The *complete infestation year* is the first year in which the county is coded as fully infested and not subsequently found to be boll weevil free or partially infested. The *year of infestation* is the ceiling of the average of the *first arrival year* and *complete infestation year*.

Throughout, we restrict our analysis to certain counties based on characteristics of the timing of the boll weevil's arrival, to increase the precision with which we date treatment. We limit our sample to 844 counties, as they were defined in 1920, of the Cotton Belt that were fully infested by the boll weevil in 1922. This excludes 104 boll weevil free and 48 partially infested, as of 1922, Cotton Belt counties, approximately half of which are in the Western High Plains and Southwestern Tablelands regions of Texas, Oklahoma, and New Mexico. Then, we exclude counties that experienced a full retreat of the boll weevil (assessed as being boll weevil free in a year subsequent to being partially or fully infested), which drops 120 counties. Finally, we exclude 38 additional counties where the absolute difference between the *first arrival year* and the *complete infestation year* is greater than four years; 30 of these are in Texas, 6 in Arkansas, and 2 in Oklahoma.[6]

3.2 Linked Census Data

We start with by extracting all men born in 13 Southern states from the 1940 Full Count census, provided by IPUMS and accessed through the NBER server. Because we need to assign exposure to the boll weevil during childhood, we link these men to themselves as children in the 1900, 1910, or 1920 censuses. We restrict our search to individuals who were between the ages of 3 and 18 in the earlier census years; this means that they were between 23 and 58 years old in 1940.[7] In the case where men are found in two census years, we keep the earlier observation.[8] We link individuals using first and last names, birth year, race, and birth state.

We link individuals using the following procedure (Abramitzky, Boustan, and Eriksson 2012) [9]:

[6]We show in Appendix Tables A.1 and A.2 that our main results are robust to dropping these restrictions, as well as imposing additional sample restrictions.

[7]When matching from 1940 backwards we do allow for the possibility that, for example, 23 year olds in 1940 could best match 1 or 2 year olds in 1920. However, we only retain in our linked sample individuals between the ages of 3 and 18, inclusive, in childhood census years.

[8]Our results are not sensitive to this restriction.

[9]See Abramitzky et al. (2018) for a description of the small differences between the matching method used in this paper and the original method described by Abramitzky, Boustan, and Eriksson (2012). Specifically, we use the abematch command provided at: https://ranabr.people.stanford.edu/matching-codes.

1. In each dataset, we remove initials, standardize nicknames, and then standardize first and last names using the NYSIIS procedure (Atack, Bateman, and Gregson 1992).

2. Starting in 1940 we match backwards to the other three years in the following way:

   (a) Restrict the 1940 dataset to individuals who are unique by standardized name, birth year, race, and birth state.

   (b) For each observation in 1940, look for an exact match on the specified matching variables in the earlier dataset. If there is an exact match, consider this observation matched. If there is more than one observation, drop the observation in 1940 and leave it unmatched.

   (c) If no match was made in the previous step and the observation was not dropped, look for a unique exact match one year off in birth year in either direction. Follow the same decision rules as (b).

   (d) If the observation is still unmatched, look for a unique exact match two years off in either direction. If the observation in 1940 fails to match, leave it unmatched.

3. Conduct the matching process described above in 2, but starting with the earlier years and matching forwards to 1940.

4. Take the intersection of the resulting backward- and forward-matched samples.

For our primary results, we aim to reduce false positive matches by further requiring individuals to be unique in each dataset within a three year age band (plus or minus one year). In particular, given that the census is not taken at the same time each year, we worry that, even with accurately reported age, calculated year of birth will be off by one year for individuals born in the months surrounding census dates. Requiring uniqueness within three year age bands eliminates the possibility of false positive matches due to this anomaly.

Our linking method results in a match rate of 27.51 percent for whites and 18.63 percent for blacks. These numbers are somewhat higher than the literature because some individuals have a chance to match twice. For example, someone who is 43 years old in 1940 could have been 13 in 1910 or 3 in 1900. As mentioned above, we keep the individual at the youngest age we find him, but having two chances to find a match mechanically increases match rates.

Robustness samples carry out a range of alternative matching procedures: First, we relax the assumption that individuals are unique within a three year age band; this maximizes the sample size at the risk of increasing false positive matches. Second, we require uniqueness by name and age within a five year band (plus or minus two years of age). Third, we exclude matches whose calculated birth years differ by more than one year across the two censuses. Fourth, we impose the five year age band uniqueness requirement and require individuals to have birth years within one year. Finally, we return to our initial restriction and match based on exact reported names instead of standardizing them with NYSIIS. Match rates are shown at the bottom of Table 7, which shows the robustness of our main results to choices in the matching procedure. The inherent tradeoff between accuracy (reducing false matches) and match rates (Bailey et al. 2017) leads to our choice of matching procedure, but our results are robust to a wide variety of methods.

We show in Table 1 that men in the matched sample come from higher socioeconomic backgrounds on average than the population at risk to match. Specifically, childhood household heads of matched individuals are more likely to be literate, own their homes, and have higher occupational income scores, relative to those of the population. Additionally, matched individuals themselves completed more years of schooling on average than those in the population. Because of this, we later weight linked individuals using inverse probability weights constructed so that the primary linked sample matches the population on the variables shown in Table 1. We find that our results are not driven by differential selection into the matched sample.

4 Estimation and Results

4.1 The Boll Weevil, Enrollment, and Attendance

As Baker (2015) shows, the boll weevil increased the enrollment rate of black children by 4 percent, and had a positive but not statistically significant effect on the white enrollment rate, in Georgia. In this section, we confirm that the boll weevil had a positive effect on enrollment using county-level data from 1900 to 1934 for six states in the South: Alabama, Arkansas, Georgia, North Carolina, South Carolina, and Tennessee.[10] This yields a sample of 347 geographically consistent counties

[10] Data for the period 1910–1934 were provided by Carruthers and Wanamaker (2017). Data for 1900–1909 were collected from annual or biennial reports of the state departments of education for those years.

which were invaded by the boll weevil between 1906 and 1922.[11]

To begin, we estimate the effect of the boll weevil on enrollment using the following specification:

$$ln(enrollment)_{ct} = \beta boll\ weevil_{ct} + \gamma ln(teachers)_{ct} + \theta_c + \theta_t + \varepsilon_{ct}, \quad (1)$$

where $boll\ weevil_{ct}$ is an indicator variable for the presence of the boll weevil in county c in year t. Additionally, we include the natural log of the number of teachers as a county-level control for school quality, county fixed effects θ_c, year fixed effects θ_t, and county-specific linear time trends $\phi_c t$.

Table 2 presents the results. Columns (1) through (3) provide estimates for whites, while columns (4) through (6) show estimates for blacks. The coefficients on the presence of the boll weevil indicator suggest the infestation had a positive and statistically significant effect on enrollment of both races. The coefficient in column (3) of 0.040 suggests the boll weevil infestation led to a 4.1 percent ($e^{0.040} - 1$) increase in school enrollment of white children. That is an increase in enrollment of 165 white children in the average county. The coefficient in column (6) of 0.060 suggests the boll weevil infestation led to a 6.2 percent ($e^{0.060} - 1$) increase in school enrollment of black children, an increase of 147 black children in the average county. While the estimated coefficients for blacks are consistently higher than those for whites, they are not statistically different from one another at the 90 percent confidence level.

To better understand the timing of the boll weevil's effect on enrollment, we replace the presence of the boll weevil indicator with 6 leads and 12 lags for the year of the insect's arrival. The empirical specification becomes

$$ln(enrollment)_{ct} = \sum_{k \geq -6, k \neq 0}^{k \leq 12} \beta_k * \mathbb{1}\{t - BW_c = k\} + \gamma ln(teachers)_{ct} + \theta_c + \theta_t + \phi_c t + \varepsilon_{ct}, \quad (2)$$

where BW_c represents the first year the boll weevil infested county c. As above, the specification also includes the natural log of the number of teachers, county fixed effects, year fixed effects, and

[11]Of the 545 counties that existed at some point during the 1900–1934 period in the six state sample, the boll weevil affected 466. A total of 62 counties are dropped due to border changes, of which four were unaffected by the boll weevil. The sample excludes 50 additional counties where the boll weevil entered and then fully retreated before reentering, because the treatment date is ambiguous. Finally, 11 counties with little to no black population are excluded for comparability.

county-specific linear time trends. The indicator for the initial year of arrival is omitted; therefore, all effects are relative to the arrival date of the boll weevil. The coefficient β_{12} represents the average effect twelve or more years after being infested (relative to the year of arrival), while β_{-6} gives the effect six or more years before infestation.

Figure 1 presents transformed coefficients of the leads and lags. Solid lines provide the main effects, while dashed lines represent 95 percent confidence intervals. Panel (a) shows school enrollment of whites remained fairly flat in the years leading up the the boll weevil's arrival, with a dip in enrollment centered around 4 years prior to contact. Relative to the arrival date of the boll weevil, white enrollment had increased by approximately 6 percent five years later, with the estimated effect remaining near 6 percent up to 12 years later. Panel (b) shows school enrollment of blacks also seems relatively flat prior to the boll weevil, with what appears to be a slight but persistent dip in enrollment in the 4 years prior to contact. Still, the coefficients on these leads are not statistically different from zero at the 95 percent confidence level. Three years after the arrival date, black enrollment had increased by approximately 6 percent compared to enrollment in the year of arrival. Moreover, these enrollment gains persisted up to 12 years after the initial boll weevil infestation.

While the boll weevil had a positive effect on enrollment, it is unclear if this effect translated to higher attendance. In Table 3, we resolve this uncertainty by showing results of specifications replacing enrollment in Equation (1) with average daily attendance. The coefficient in column (3) of 0.041 suggests the boll weevil infestation led to a 4.2 percent ($e^{0.041}-1$) increase in school attendance of white children. Meanwhile, column (7) suggests the boll weevil's arrival led to a statistically significant 5.4 percent ($e^{0.053}-1$) increase in school attendance of black children. For both races, the boll weevil's effect on attendance is remarkably similar to its effect on enrollment shown in Table 2. As with enrollment, comparable coefficients across race are not statistically different from one another even though the point estimates are consistently higher for blacks. Columns (4) and (8) show estimates of the effect on attendance conditional on enrollment. In both columns the coefficient on the boll weevil indicator is small and not statistically different from zero, suggesting new enrollees attributable to the weevil had similar attendance profiles to those that would have been enrolled regardless.

Unlike Baker (2015), we find statistically significant effects of the boll weevil on enrollment and attendance for both races when considering a broader set of states. However, it should be noted that these results are not directly comparable to those of Baker, as his dependent variable is enrollment rate while we consider enrollment without controlling for the population of school-age children.[12]

In fact, the results presented here must be interpreted with some degree of caution with respect to mechanisms, because we do not account for the school-age population. They provide suggestive evidence that children, previously not in school, enrolled after the arrival of the boll weevil, or already enrolled children stayed in school longer than they would have in the absence of the boll weevil's arrival. While we cannot rule out the possibility that the increases in enrollment are generated by boll weevil driven migration, Lange, Olmstead, and Rhode (2009) show that population swelled in advance of the boll weevil infestation, not after. So the timing of this pre-weevil population increase seems inconsistent with the post-weevil enrollment increase shown in Figure 1.[13]

Holding the school-age population constant, a persistent increase in enrollment implies an increase in average years of schooling, assuming that newly enrolled students are attending classes and completing grades. Multiplying average years of schooling of cohorts not affected by the weevil with our estimates of the weevil's effects on enrollment suggests that the infestation increased years of schooling by 0.329 (8.034×0.041) for whites and 0.287 (4.623×0.062) for blacks who were affected by the weevil prior to the school entrance age. In the next section, we use our linked-census sample to directly investigate the weevil's impact on years of schooling.

4.2 Boll Weevil and Long-Run Schooling Outcomes

In our linked-census data we follow individuals over time, but observe their years of schooling only in the 1940 census manuscripts, when they are adults. Therefore, we examine the boll weevil's impact on educational attainment by comparing adjacent birth cohorts. Our main specification is a differences-in-differences approach where we compare those who were children (ages 4 through 18),

[12]The difference between Baker's (2015) results and those presented here could be explained by increasing school-age population coincident with the onset of the boll weevil infestation, with the increase in white population being relatively greater than that of blacks.

[13]If a boll weevil induced increase in fertility explained the uptick in enrollment, we would expect enrollment to increase with a lag of five or more years. Furthermore, Bloome, Feigenbaum, and Muller (2017) find a decline in the share of African Americans who married at young ages in response to the boll weevil, which would likely have a negative effect on fertility rates.

and still eligible for public schooling, when the boll weevil arrived to young adults (ages 18 through 30), who likely completed their schooling prior to the infestation. This approach is represented by the following regression equation:

$$y_{ica} = \sum_{k=0}^{4} \beta_k * \mathbb{1}\{3(5-k)+1 < BW_c - a \leq 3(6-k)\} + \gamma X_i + \theta_c + \theta_a + \varepsilon_{ica}, \tag{3}$$

where y_{ica} is an educational outcome of interest for individual i, who was born in year a and observed in county c in childhood (between the ages of 3 and 18, inclusive). Boll weevil exposure occurs at age $BW_c - a$, where BW_c represents the *year of infestation* of county c.[14] Therefore, the indicator function ($\mathbb{1}\{3(5-k)+1 \leq BW_c - a \leq 3(6-k)\}$) returns a 1 if the individual was aged $3(5-k)+1$ to $3(6-k)$ in the *year of infestation*. The θ_c and θ_a are county and birth-cohort fixed effects, respectively, while X_i is a vector of individual-level family background controls as measured in the childhood census year. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. The β_k's are the coefficients of interest, representing the average effect of the boll weevil on individuals in exposure cohort k. Because the boll weevil arrived at different times in different counties we are able to separately identify the full set of cohort and county fixed effects as well as the age of exposure treatment effects for those exposed at age 18 or under.

The first two columns of Table 4 show the results of regressions following Equation (3) with years of schooling as the dependent variable. Children who were first exposed to the boll weevil between the ages of 4 and 6, inclusive, saw the greatest gains in years of schooling as a result. White children exposed at ages 4–6 attained 0.2609 more years of schooling on average, as compared with those exposed at ages 19–30 (shown in column [1]). White children exposed at 7–9 years of age experienced similar relative gains of 0.2314 years of schooling. Both results are statistically significant at the 99.9 percent confidence level. Those exposed to the boll weevil at older ages saw smaller increases in educational attainment: exposure at 10–12 0.1459 years and 13–15 0.1017 years of schooling. If the opportunity cost of schooling is increasing in years of schooling and work experience, and the timing of the transition from schooling to working maximizes expected lifetime

[14] Throughout the remainder of the text, we refer to $BW_c - a$ as age at exposure. Additionally, we define exposure cohorts as groups of individuals that were first exposed to the boll weevil at the same age.

earnings, then a modest negative shock to the opportunity cost of schooling will induce those in school to stay in school longer but will affect few who have already left school for work. Thus, not surprisingly, those exposed to the boll weevil between the ages of 16 and 18 saw an increase in educational attainment of only 0.0423 years on average, relative to those exposed at ages 19–30, and this result is statistically different from zero only at the 90 percent confidence level. Column (2) shows remarkably similar gains in educational attainment for blacks resulting from the boll weevil infestation. The only coefficient that differs notably, but not statistically, from those of whites is that for the youngest age group. Black children exposed at ages 4–6 attained 0.3581 more years of schooling on average, compared to those exposed at ages 19–30.

Columns (3) and (4) of Table 4 instead consider the boll weevil's impact on eighth grade completion. Again, children exposed to the boll weevil at younger ages were more responsive in their likelihood of completing at least eight years of schooling. As shown in column (3) White children exposed at ages 7–9 were more likely to complete the eighth grade by 2.55 percentage points, compared with those exposed at ages 19–30. Black children saw a similar increase, of 2.47 percentage points, in the likelihood of graduating grammar school due to boll weevil exposure at ages 7–9, as indicated in column (4). Differences in point estimates are not statistically significant across race and are generally quite small.

In interpreting the β_k's as average treatment effects, we are implicitly making the assumption that individual's between the ages of 19 and 30 when the boll weevil arrives do not alter their years of schooling or school completion rates in response. Indeed, there are structural reasons to expect 19 to 30 year olds to be little affected by the weevil; children generally attend school, while adults generally work. When the outcome of interest is eighth grade completion, this seems like a quite reasonable assumption as the typical age for finishing eighth grade is 14. While age-for-grade statistics were not commonly published by Southern states during the early twentieth century, a report on the 1923 school year in Tennessee reveals that approximately 1.3 percent of whites and 3.3 percent of blacks ages 19–21 were enrolled in grades one through eight (Tennessee Department of Education 1924). Thus, the boll weevil is unlikely to have had a discernible impact on eighth grade completion for exposure cohorts over the age of 19. However, with years of schooling as the outcome, the validity of this assumption is not as clear. The 1920 census shows that 4.3 (2.2) percent of white (black) Southern-born men ages 19–30 were still attending school.[15] If the boll weevil had a positive (negative) effect on the educational attainment of individuals in older exposure cohorts, then the coefficients shown in Table 4 would understate (overstate) the true effects. Therefore, we next consider the possibility of pre-trends (or the validity of 19–30 year old exposure cohorts as a comparison group).

It is common for studies taking an empirical approach of the form specified in Equation (3) to also show results of a regression equation with both leads and lags on treatment, to allay concerns regarding the possibility of preexisting trends, like the following:

$$y_{ica} = \sum_{k\geq -4, k\neq -1}^{k\leq 4} \beta_k * \mathbb{1}\{3(5-k)+1 < BW_c - a \leq 3(6-k)\} + \gamma X_i + \theta_c + \theta_a + \varepsilon_{ica}. \tag{4}$$

Such results are often shown graphically. However, Borusyak and Jaravel (2017) point out the results of such a regression are underidentified due to the absence of an untreated control group. In our case, it is not possible to fully disentangle the birth cohort effects θ_a from exposure cohort effects β_k in the presence of county fixed effects. The simplest solution to this underidentification problem would be to drop the county fixed effects. However, county fixed effects are necessary for identification since the timing of the boll weevil's arrival is not independent of time invariant county characteristics (for example, longitude, latitude, and climate).

Instead, we follow the recommendation of Borusyak and Jaravel (2017), imposing the restrictions that $\hat{\beta}_{-4} = \hat{\beta}_{-1} = 0$. In practice, this means dropping two indicators for older exposure cohorts: those exposed at 19–21 and 28–30. Then, an F-test on the joint significance of the remaining pre-trends provides a test for the existence of non-linear pre-trends.[16] Given the nonlinearity in attendance rates between 19–21 year olds and older age groups, we would expect to see evidence of non-linear pre-trends if those exposed at 19–21 were modifying their schooling behavior. The results of these restricted regressions are shown in Table 5. The fact that the coefficients on the 4–6 to 16–18 age of exposure indicators are little affected, compared to those presented in Table 4, is encouraging. The coefficients on the indicators for exposure to the boll weevil at ages 22-24 and

[15]Calculated using the USA Full Count sample for 1920 (Ruggles et al. 2018). We exclude those born in Kentucky, Oklahoma, Virginia, West Virginia, Maryland, Delaware, and the District of Columbia, to better compare with our sample of boll weevil–affected counties.

[16]Additionally, we show in Appendix Table A.3 that our main results are robust to excluding the 19–21 exposure cohort from the omitted category and top coding years of schooling at 12.

25–27 are small in magnitude and not statistically different from zero. Moreover, the F-statistics on their joint significance are below 2.31, the critical value of F for the $\alpha = 0.1$ significance level, so we cannot reject the null hypothesis that the coefficients on the remaining pre-trends jointly equal zero. This provides reasonable reassurance against the presence of non-linear pre-trends, as would be likely if older exposure-cohorts modified their schooling behavior due to, or in anticipation of, the boll weevil's arrival.[17]

We also investigate whether the boll weevil's effect on years of schooling differed according to the intensity of cotton production in one's childhood county of residence. We use the ratio of acres in cotton to improved farm acreage in 1889 as a measure of cotton intensity prior to the boll weevil's arrival.[18] Figure 2 reveals, for our linked-sample, the county-level distribution of the share of improved acreage planted in cotton in 1889. The median county in our sample grew cotton on 30.4 percent of improved acreage, while the 95th percentile was 50.6 percent.

We modify Equation (3) by interacting each exposure cohort bin with a 4th-order polynomial in cotton intensity. Figure 3 reveals the predicted impact of exposure to the boll weevil at ages 4–6 (panel [a]), 7–9 (panel [b]), 10–12 (panel [c]), and 13–15 (panel [d]) on years of schooling as it varies by cotton intensity in childhood county of residence.[19] Figure 3 shows substantial positive effects on educational attainment even in counties with little cotton production in 1889. The boll weevil's effect on years of schooling increases further as intensity of cotton production rises to approximately 15 percent. The youngest exposure cohort (4–6 year olds in panel [a]) experienced a 0.15 year increase in schooling as cotton acreage went from 0 to 15 percent, that is a 50 percent increase in the effect relative to those in counties with no cotton acreage. Then, the impact of the weevil declines from its peak as cotton intensity increases from 15 to 50 percent of farm acreage. Panel

[17]We can look at educational outcomes because years of schooling and completion of grammar school are largely determined before the boll weevil arrives for those exposed at 19 years of age and older, making them an arguably valid comparison group for the estimation of average treatment effects. However, labor market outcomes of those in the 19–30 year old exposure cohort were likely affected by the boll weevil, as were those of younger exposure cohorts. Therefore, we lack a valid comparison group for quantifying the boll weevil's effect on occupational score, wages, employment status, et cetera using this data.

[18]This approach is similar to Lange, Olmstead, and Rhode (2009), but they use total farm acres as the denominator. The US Census Office (1890) clarified for enumerators of the farm schedules that "land once plowed is *improved* unless afterward abandoned for cultivation" and "rocky, hill, and mountain pastures are *not improved*" (36). However, such pasture land, and otherwise uncultivated land, was to be included in total farm acreage. Therefore, using improved acreage allows us to better capture the share of cultivated land devoted to cotton as cotton intensity and improves cross-county comparability.

[19]Comparable graphs of the boll weevil's impact by cotton intensity on eighth grade completion are shown in Appendix Figure A.1.

(a) shows the decline for the youngest exposure cohort mirrors the rise, falling by approximately 0.15 years of schooling. However, the 7–9 (panel [b]) and 10–12 (panel [c]) exposure cohorts saw larger declines relative to the preceding rise.

The predicted effects by cotton intensity suggest that children in counties where cotton production was marginally optimal benefited the most from the weevil with respect to educational attainment, while children in counties that were heavily dependent on cotton agriculture benefited to a lesser extent. At first glance, this might seem counterintuitive. However, these estimates reflect the overall impact of the boll weevil, and do not control for the insect's impact on household income. Empirical and anecdotal evidence suggests that the boll weevil did greater damage where cotton was more intensely grown and substitution away from cotton, in terms of acreage, was greatest where the crop was marginally optimal (Giesen 2011; Lange, Olmstead, and Rhode 2009). Combined with the effects shown in Table 3, this suggests education gains were most pronounced in locations where substitution away from cotton was greatest and income effects were limited. While counties where the boll weevil had large negative effects on income, and yet substitution away from cotton was relatively limited, saw more modest gains in educational attainment.

5 Robustness of Main Results

5.1 Concurrent Shocks to Education

A possible threat to validity regards shocks to educational resources and requirements contemporaneous with the weevil infestation. Events that overlapped in timing with the boll weevil's spread and significantly affected educational attainment include the passage of compulsory school attendance laws (Lleras-Muney 2002) and the construction of Rosenwald schools (Aaronson and Mazumder 2011). If the passage of compulsory schooling laws or Rosenwald school construction were correlated with the onset of the boll weevil infestation, this could result in spurious correlation between boll weevil exposure and student outcomes.

Therefore, we assign to individuals in our sample a measure of Rosenwald school exposure, based on childhood county of residence and year of birth, and include this and its interaction with an indicator for rural status as controls. Additionally, we calculate the number of years of schooling required, by state and birth cohort, from the ages for which compulsory schooling laws required attendance and the effective dates of those laws, adding this measure as a control.[20] Table 6 shows the results of our main specifications modified to add controls for compulsory schooling and Rosenwald school exposure. The coefficients presented are comparable in magnitude, if not slightly larger, and statistical significance to the baseline results shown in Table 4, ameliorating concern that omitted variable bias is driving our results.

5.2 Robustness to Matching

Table 7 presents results of specifications, following Equation (3) with years of schooling as the dependent variable, utilizing samples generated by several different matching procedures. Columns (1) through (6) show results for whites, while columns (7) through (12) display results for blacks. The first columns, (1) and (7), reproduce results utilizing our baseline sample, those shown in Table 4, for ease of comparison. Columns (2) and (8) relax match restrictions slightly by allowing matches between individuals that are not unique within three-year age bands. The methodology in this sample is, therefore, identical to (Abramitzky, Boustan, and Eriksson 2012). This method results in a match rate of 36.25 (27.05) percent for the white (black) sample, significantly higher than the 27.51 (18.63) percent match rate of the baseline. Relaxing restrictions yields point estimates that are smaller in magnitude by 15 to 29 percent relative to the baseline for the youngest three exposure cohorts, but still statistically significant. This is likely due to attenuating noise introduced by a higher level of false positive matches.

Matching restrictions are increased relative to the baseline, to reduce the likelihood of false positive matches, in columns (3) and (9), which require individuals to be unique within five-year age bands. The match rate falls to 23.61 (15.12) percent for the white (black) sample, yet the coefficients are not meaningfully changed compared with the baseline. Columns (4) and (10) show results from a matched sample requiring that individuals be unique within three-year age bands and have absolute differences in year of birth of less than or equal to one year. The latter restriction is also aimed at reducing the frequency of false positive matches, as those with two year discrepancies in year of birth are more likely to be incorrect. Again, the coefficients are remarkably similar to the baseline. Columns (5) and (11) additionally restrict matches to individuals that are unique within

[20] For children under the entrance age when compulsory schooling laws took effect, this is simply the minimum age at which exit was legally allowed less the entrance age.

five-year age bands, and present results that are nearly identical. Columns (6) and (12) return to the three-year age band restriction, but require matches on exact names rather than standardized versions of names. Results are little changed for the black sample here, but point estimates are reduced in magnitude for whites by 20 to 35 percent relative to the baseline.[21]

5.3 Robustness to Weighting

Recall that the baseline sample is not representative of the population, as shown in Table 1. In Table 8, we examine whether the results are robust to weighting the baseline sample to be representative of the population along the characteristics presented in Table 1. Panel (a) presents specifications analogous to the baseline shown in Table 4, but uses inverse probability weights calculated based on family background characteristics provided by childhood census data. Instead, panel (b) uses inverse probability weights estimated from 1940 census characteristics. The results presented do not meaningfully differ from the baseline results. Therefore, the results cannot be explained by observable differences between the population and the matched sample.

6 Conclusion

We add to an expanding literature exploring the boll weevil's impact beyond its direct effect on cotton production. The spread of the infestation through the South, whose agrarian economy was heavily dependent on cotton cultivation, provided an exogenous shock to agricultural productivity, particularly for women and children. Our findings reveal gains in educational attainment for those at young ages when the boll weevil arrived. Whites exposed at ages 4–6 gained 0.2609 years of schooling on average, while comparatively aged blacks gained 0.3581 years. Those white and black children ages 7–9 also saw significant gains of 0.2314 and 0.2423 years of schooling, respectively. Slightly older exposure cohorts also saw gains but these declined as they approached 18.

The magnitude of these estimates can be compared to the findings of several studies of contemporaneous shocks to schooling. Our finding that exposure to the boll weevil at age 7–9 increased educational attainment by nearly a quarter of a year, is comparable to the imposition of compulsory schooling and child labor laws requiring school entrance at age 7 and allowing work permits

[21] Appendix Table A.4 shows the baseline results with eighth grade completion as the dependent variable to be similarly robust.

beginning at age 12.[22] It is important to note, however, that passage of such a law does not imply compliance with the law. Compulsory attendance laws were not well enforced in many locales and often provided myriad exceptions. Another comparison, for black children in particular, is instructive: Aaronson and Mazumder (2011) estimate that going from no (Rosenwald exposure of 0) to full (Rosenwald exposure of 1) coverage of black 7 through 17 year olds by Rosenwald teachers led to a gain in educational attainment of 1.186 years. Therefore, our result is roughly equivalent to having enough Rosenwald teachers to teach 20 percent of black children in one's childhood county during ages 7 through 17, or Rosenwald exposure of 0.20 where the mean Rosenwald exposure in 1930 was 0.27. Finally, Baker (2018) shows that a one percent increase in school resources for the first three years of schooling increased educational attainment by 0.0334 years for white children in early–twentieth century Georgia. So white children would experience approximately the same gains in years of schooling from a 7 percent increase in school financial resources. Therefore, our estimates represent economically meaningful gains on the order of a significant funding boost, which seems quite reasonable for an event that so dramatically changed agricultural production in the region.

As the boll weevil itself only directly affects cotton production, any impact of the boll weevil on student outcomes must run through its devastation of the cotton crop. Contemporary observations and empirical evidence has demonstrated that Southern farmers shifted away from cotton production after infestation (Lange, Olmstead, and Rhode 2009; Giesen 2011). This shift from a child labor–intensive crop to alternatives that generated less demand for child labor provides one potential mechanism: a fall in the opportunity cost of schooling led to increased enrollment and attendance. Over time this accumulated into higher levels of average educational attainment. A second explanation is that the boll weevil made farming a less attractive occupation, causing children at the margin to shift their occupational aspirations and their preparations accordingly. Where fieldwork might have provided suitable training for the farm profession, schooling made available a wider set of occupations.

Both of these mechanisms likely played a role in increasing educational attainment following the weevil's arrival. If the first was at play, then our results are suggestive of the benefits of programs

[22] Lleras-Muney (2002) shows that each additional year between the school entrance age and work permit age increased years of schooling by 0.05 years. This makes our result roughly equivalent to a five year gap between the school entrance age and permit age.

encouraging farmers to switch cultivation to less child labor–intensive crops and to adopt child labor–reducing technologies, which would decrease the opportunity cost of schooling in rural areas. However, if the second mechanism was also at play, then our results might overstate the potential gains from such programs, as such programs generally have neutral to positive effects on the returns to farming. Getting farmers to switch crops would be achieved by compensating them for loses incurred due to switching, and mechanization likely has a positive effect on the productivity of adult labor. Still, it should be noted that our estimates are net of the income effect of the boll weevil, suggesting that our results could in fact understate the benefits of these programs.

A third mechanism is suggested by the work of Clay, Schmick, and Troesken (2017), who claim that the diversification of crops following the boll weevil infestation reduced the incidence of a pellagra, a disease caused by having insufficient niacin, by increasing the availability of locally grown vegetables that had higher niacin content than imported foods. If those susceptible to the disease were randomly selected from the population or the disease was not commonly fatal, then we would expect increased health might have some positive effect on educational attainment on average for those surviving to adulthood. However, pellagra disproportionately affected the poor, because cheap shelf-stable foods had lower niacin content prior to enrichment. The Louisiana Board of Health noted in their 1928/29 report, well after the boll weevil infestation began: "As this disease [pellagra] rises and falls with the economic situation, there seems little we can do to prevent its prevalence in localities where crops fail or employment is not remunerative" (1930, 10). That same report shows the case fatality rate for pellagra over the 1926–29 period to be 47 percent in Louisiana. Given the positive relationship between family income and educational attainment (see, for example, Blanden and Gregg 2004; Taubman 1989), it is unclear whether this channel strengthens or attenuates our estimates.[23]

There is still much work to be done to understand the broader effects of the boll weevil on the Southern economy. Whether the spread of the infestation just prior to an unprecedented wave of migration out of the region represents a causal or coincidental relationship remains unexplored. Additionally, tracing the insects impact on occupational choice would be instructive.

[23] Indeed, studies of the income neutral health interventions, which disproportionately affected the poor, of iron fortification (Niemesh 2015) and hookworm eradication (Bleakley 2007) fail to find statistically significant effects on years of schooling when accounting for mean reversion. Rather they find these interventions had substantial educational benefits at the intensive margin.

References

Aaronson, Daniel, and Bhashkar Mazumder. 2011. "The Impact of Rosenwald Schools on Black Achievement." *Journal of Political Economy* 119 (5): 821–88.

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, Katherine Eriksson, James Feigenbaum, and Santiago Perez. 2018. "Automated Linking of Historical Data." Working paper.

Ager, Philipp, Markus Brueckner, and Benedikt Herz. 2017. "The Boll Weevil Plague and its Effect on the Southern Agricultural Sector, 1899–1929." *Explorations in Economic History* 65: 94–105.

Atack, Jeremy, Fred Bateman, and Mary Eschelbach Gregson. 1992."'Matchmaker, Matchmaker, Make Me a Match': A General Personal Computer-Based Matching Program for Historical Research." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 25 (2): 53–65.

Bailey, Martha, Connor Cole, Morgan Henderson, and Catherine Massey. 2017. "How Well Do Automated Linking Methods Perform in Historical Data? Evidence from New U.S. Ground Truth." *NBER Working Paper Series*, No. 24019.

Baker, Richard B. 2015. "From the Field to the Classroom: The Boll Weevil's Impact on Education in Rural Georgia." *Journal of Economic History* 75 (4): 1128–60.

Baker, Richard B. 2018. "School Resources and Labor Market Outcomes: Evidence from Early Twentieth-Century Georgia." Working paper.

Beegle, Kathleen, Rajeev Dehejia, and Roberta Gatti. 2009. "Why Should We Care about Child Labor? The Education, Labor Market, and Health Consequences of Child Labor." *The Journal of Human Resources* 44 (4): 871-89.

Blanden, Jo, and Paul Gregg. 2004. "Family Income and Educational Attainment: A Review of Approaches and Evidence for Britain." *Oxford Review of Economic Policy* 20 (2): 245—63.

Bleakley, Hoyt. 2007. "Disease and Development: Evidence from Hookworm Eradication in the American South." *The Quarterly Journal of Economics* 122 (1): 73–117.

Bloome, Deirdre, James Feigenbaum, and Christopher Muller. 2017. "Tenancy, Marriage, and the Boll Weevil Infestation, 1892–1930." *Demography* 54 (3): 1029–49.

Boozer, Michael A., and Tavneet K. Suri. 2001. "Child Labor and Schooling Decisions in Ghana." Unpublished manuscript, September. Yale University, New Haven, CT.

Borusyak, Kirill, and Xavier Jaravel. 2017. "Revisiting Event Study Designs, with an Application to the Estimation of the Marginal Propensity to Consume." Working paper.

Carruthers, Celeste K., and Marianne H. Wanamaker. 2017. "Separate and Unequal in the Labor Market: Human Capital and the Jim Crow Wage Gap." *Journal of Labor Economics* 35 (3): 655–96.

Clay, Karen, Ethan Schmick, and Werner Troesken. 2017. "The Rise and Fall of Pellagra in the American South." *NBER Working Paper Series*, No. 23730.

Collins, William J., and Robert A. Margo. 2006. "Historical Perspectives on Racial Differences in Schooling in the United States." In *Handbook of the Economics of Education*, vol. 1, edited by Eric A. Hanushek and Finis Welch, 107-54. Elsevier.

Dammert, Ana C. 2008. "Child Labor and Schooling Response to Changes in Coca Production in Rural Peru." *Journal of Development Economics* 86 (1): 164–80.

Emerson, Patrick M., Vladimir Ponczek, and André Portela Souza. 2017. "Child Labor and Learning." *Economic Development and Cultural Change* 65 (2): 265–96.

Giesen, James C. 2011. *Boll Weevil Blues: Cotton, Myth, and Power in the American South.* Chicago: University of Chicago Press.

Gunnarsson, Victoria, Peter F. Orazem, and Mario A. Sánchez. 2006. "Child Labor and School Achievement in Latin America." *The World Bank Economic Review* 20 (1): 31-54.

Haines, Michael R., and Inter-university Consortium for Political and Social Research (ICPSR). 2010. Historical, Demographic, Economic, and Social Data: The United States, 1790–2002 [computer file]. ICPSR02896-v3. Ann Arbor, MI: ICPSR [distributor]. http://doi.org/10.3886/ICPSR02896. (2010-05-21).

Higgs, Robert. 1976. "The Boll Weevil, the Cotton Economy, and Black Migration 1910–1930." *Agricultural History* 50 (2): 334–50.

Hunter, W. D., and B. R. Coad. 1923. *The Boll Weevil Problem.* Washington, DC: GPO.

Jensen, Robert. 2000. "Agricultural Volatility and Investments in Children." *The American Economic Review* 90 (2): 399-404.

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.

Lleras-Muney, Adriana. 2002. "Were Compulsory Attendance and Child Labor Laws Effective? An Analysis from 1915 to 1939." *Journal of Law and Economics* 45 (2): 401–35.

Louisiana Department of Education. 1902. *Biennial Report of the State Superintendent of Education of Public Education to the General Assembly.* Baton Rouge, LA: The Advocate.

Louisiana Department of Education. 1908. *Proceedings of the Conference of Parish Superintendents of Public Education.* Baton Rouge, LA: Daily State Publishing.

Louisiana Board of Health. 1930. *Biennial Report of the Louisiana State Board of Health to the Legislature of the State of Louisiana, 1928–1929.* New Orleans, LA: State Printer.

Mississippi Department of Public Education. 1907. *Biennial Report and Recommendations of the State Superintendent of Public Education to the Legislature of Mississippi.* Nashville, TN: Brandon Printing.

Moehling, Carolyn M. 1999. "State Child Labor Laws and the Decline of Child Labor." *Explorations in Economic History* 36 (1): 72-106.

Moehling, Carolyn M. 2004. “Family Structure, School Attendance, and Child Labor in the American South in 1900 and 1910.” *Explorations in Economic History* 41 (1): 73-100.

Niemesh, Gregory T. 2015. “Ironing Out Deficiencies: Evidence from the United States on the Economic Effects of Iron Deficiency.” *The Journal of Human Resources* 50 (4): 910–58.

Osband, Kent. 1985. “The Boll Weevil Versus ‘King Cotton’.” *Journal of Economic History* 45 (3): 627–43.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2018. *IPUMS USA: Version 8.0* [dataset]. Minneapolis, MN: IPUMS. https://doi.org/10.18128/D010.V8.0

Taubman, Paul. 1989. “Role of Parental Income in Educational Attainment.” *The American Economic Review* 79 (2): 57–61.

Tennessee Department of Education. 1924. *Annual Report of the Department of Education*. Nashville, TN: Ambrose-Nashville.

Texas Department of Education. 1905. *Biennial Report of the State Superintendent of Public Instruction*. Austin, TX: Gammel-Statesman Publishing.

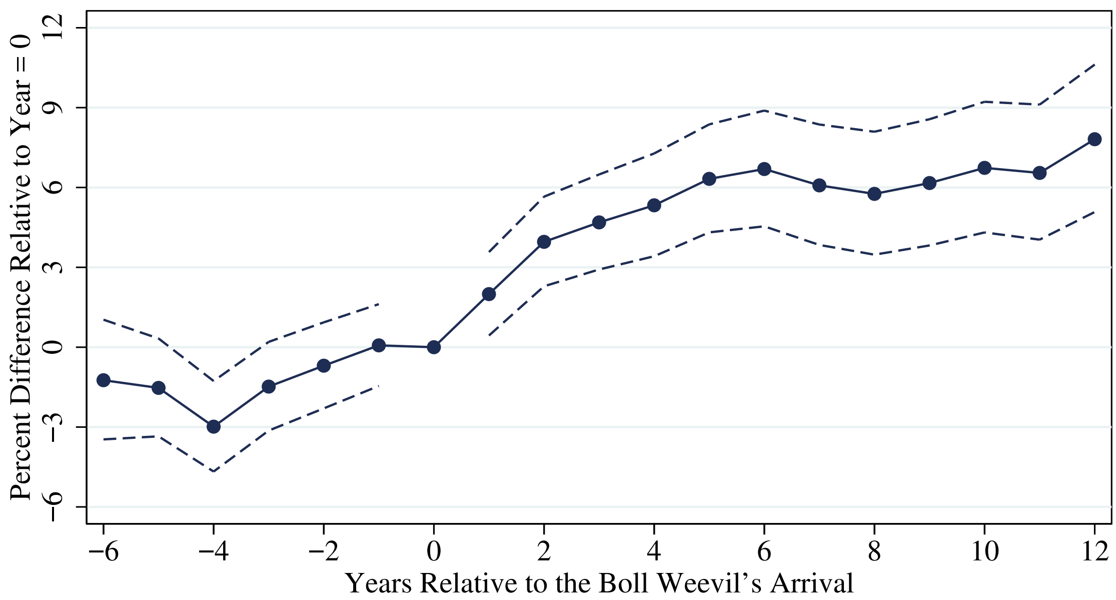
US Bureau of the Census. 1913. *Thirteenth Census of the United States: 1910*. 11 vols. Washington, DC: GPO.

US Bureau of the Census. 1924. *Children in Gainful Occupations at the Fourteenth Census of the United States*. Washington, DC: GPO.

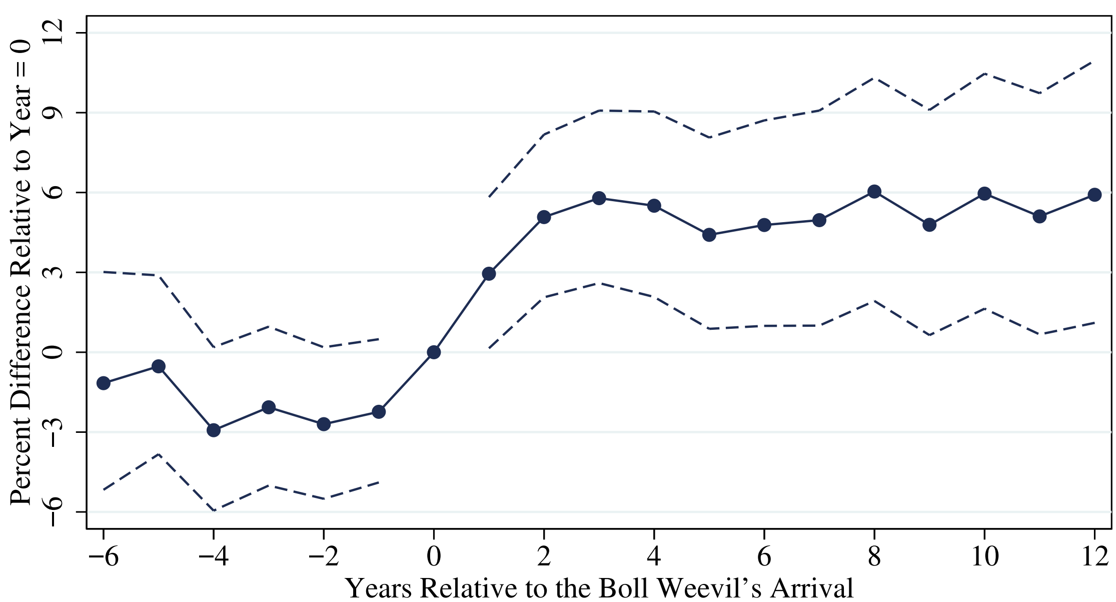
US Census Office. 1890. *Eleventh Census of the United States: Instructions to Enumerators*. Washington, DC: GPO.

US Department of Labor. 1995. *By the Sweat and Toil of Children Volume II: The Use of Child Labor in U.S. Agricultural Imports and Forced and Bonded Child Labor*. Washington, DC.

US Department of Labor. 2016. *List of Goods Produced by Child Labor or Forced Labor*. Washington, DC.



(a) White school enrollment



(b) Black school enrollment

Figure 1: School Enrollment Relative to the Boll Weevil's Arrival

Note: The y-axis shows the percent change in enrollment relative to the boll weevil's year of arrival. The dashed lines indicate 95 percent confidence intervals.
Source: See the text.

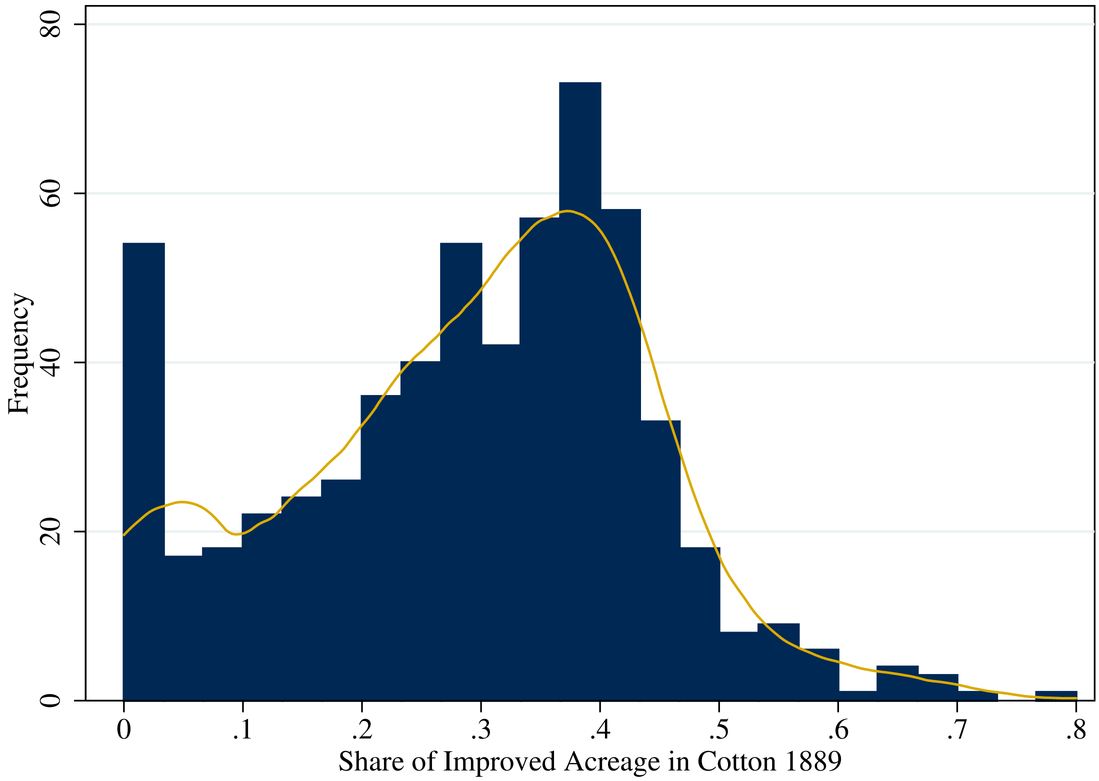
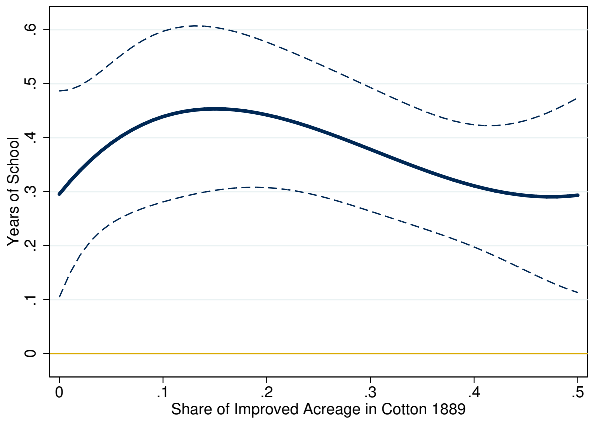
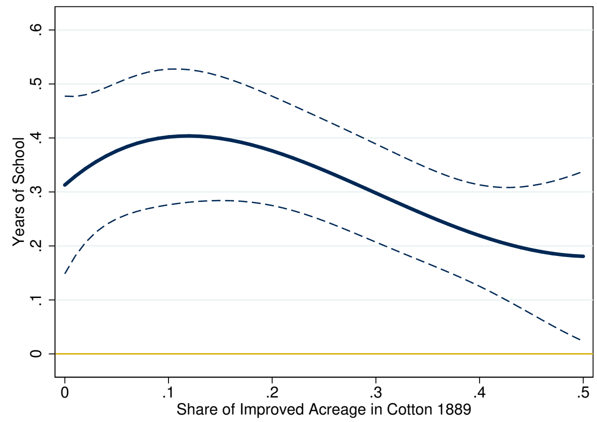


Figure 2: Distribution of Sample Counties by Intensity of Cotton Production

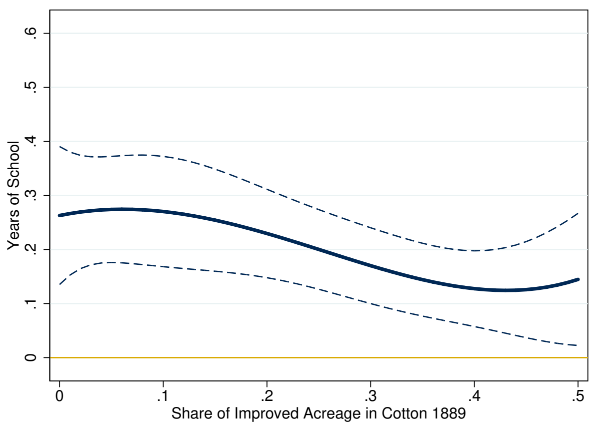
Notes: Cotton intensity is defined as the ratio of acres planted in cotton to improved farm acreage as measured in 1889. The solid line provides the kernel density estimate of the distribution.
Source: Haines and ICPSR (2010).



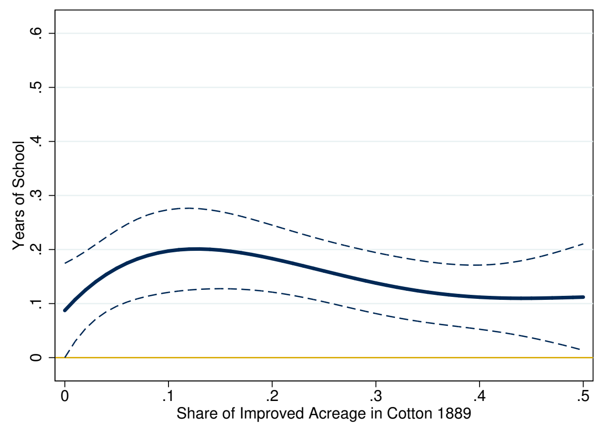
(a) Children exposed at 4–6



(b) Children exposed at 7–9



(c) Children exposed at 10–12



(d) Children exposed at 13–15

Figure 3: Impact of the Boll Weevil on Years of Schooling by Intensity of Cotton Farming in 1889

Note: The y-axis shows the difference in years of schooling relative to individuals exposed to the boll weevil between the ages of 19 and 30, inclusive. The dashed lines indicate 95 percent confidence intervals.
Source: See the text.

Table 1: Comparing the Matched Sample to the Full Population

	White				Black			
	(1) Matched Sample		(2) Difference from Pop.		(3) Matched Sample		(4) Difference from Pop.	
Panel a: Childhood Characteristics								
age	9.3954	[4.4440]	-0.1390***	(0.0069)	9.4542	[4.4115]	-0.0782***	(0.0104)
farm status	0.6428	[0.4792]	0.0060***	(0.0007)	0.6701	[0.4702]	0.0394***	(0.0011)
urban status	0.1414	[0.3484]	0.0120***	(0.0007)	0.1050	[0.3065]	-0.0041***	(0.0006)
year of BW arrival	1915.8720	[5.6512]	0.3920***	(0.0089)	1915.4190	[5.3168]	0.4290***	(0.0125)
Household head:								
literate	0.8799	[0.3251]	0.0251***	(0.0005)	0.5387	[0.4985]	0.0550***	(0.0012)
homeowner	0.5722	[0.4948]	0.0373***	(0.0008)	0.2672	[0.4425]	0.0390***	(0.0010)
mortgage free	0.4563	[0.4981]	0.0300***	(0.0008)	0.1890	[0.3915]	0.0264***	(0.0009)
occupation score	17.2518	[11.0228]	0.4920***	(0.0171)	13.9362	[5.7530]	0.2050***	(0.0137)
Census year:								
1900	0.3736	[0.4837]	-0.0623***	(0.0008)	0.3898	[0.4877]	-0.0740***	(0.0011)
1910	0.4468	[0.4972]	0.0322***	(0.0008)	0.4706	[0.4991]	0.0253***	(0.0011)
1920	0.1796	[0.3839]	0.0300***	(0.0005)	0.1396	[0.3466]	0.0487***	(0.0008)
State of residence:								
Alabama	0.1552	[0.3621]	-0.0069***	(0.0005)	0.1516	[0.3586]	-0.0083***	(0.0008)
Arkansas	0.0413	[0.1990]	-0.0010***	(0.0003)	0.0354	[0.1849]	0.0015***	(0.0004)
Florida	0.0378	[0.1908]	0.0054***	(0.0003)	0.0408	[0.1978]	0.0043***	(0.0005)
Georgia	0.1530	[0.3599]	-0.0099***	(0.0005)	0.1501	[0.3572]	-0.0255***	(0.0008)
Kentucky	0.0016	[0.0402]	-0.0001**	(0.0001)	0.0007	[0.0267]	0.0000	(0.0001)
Louisiana	0.0555	[0.2290]	0.0013***	(0.0003)	0.0545	[0.2270]	-0.0091***	(0.0005)
Mississippi	0.0867	[0.2813]	-0.0003	(0.0004)	0.1362	[0.3431]	-0.0139***	(0.0008)
North Carolina	0.1941	[0.3955]	0.0126***	(0.0006)	0.1939	[0.3954]	0.0505***	(0.0009)
Oklahoma	0.0039	[0.0626]	0.0008***	(0.0001)	0.0006	[0.0255]	0.0005	(0.0001)
South Carolina	0.1110	[0.3141]	0.0096***	(0.0004)	0.1423	[0.3494]	-0.0152***	(0.0008)
Tennessee	0.0910	[0.2876]	0.0028***	(0.0004)	0.0514	[0.2208]	0.0119***	(0.0005)
Texas	0.0689	[0.2533]	-0.0144***	(0.0004)	0.0423	[0.2013]	0.0033***	(0.0004)
Observations	512,165		2,602,100		197,214		2,157,831	
Panel b: Adulthood Characteristics								
age	37.5143	[9.3294]	-0.6620***	(0.0076)	38.0608	[9.2987]	0.0587***	(0.0131)
years of schooling	8.5055	[3.5877]	0.3220***	(0.0029)	5.4350	[3.3490]	0.2480***	(0.0047)
Observations	2,504,982		6,974,037		675,030		2,535,635	

Notes: The matched sample includes male children between the ages of 3 and 18 in 1900, 1910, and 1920, matched to 1940 adult educational outcomes in 1940. Columns (1) and (3) report means for the matched sample with standard deviations in brackets. Coefficients in columns (2) and (4) are from regressions of individual characteristics of interest on an indicator for being in the matched sample. Thus, these columns show the difference between the matched sample and population with standard errors in parentheses.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Table 2: Estimates of the Boll Weevil's Effect on ln(Enrollment)

	White			Black		
	(1)	(2)	(3)	(4)	(5)	(6)
boll weevil	0.063***	0.029**	0.040***	0.079***	0.045***	0.060***
	(0.010)	(0.009)	(0.007)	(0.017)	(0.013)	(0.012)
ln(teachers)?	NO	YES	YES	NO	YES	YES
Time trends?	NO	NO	YES	NO	NO	YES
Observations	8,242	8,104	8,104	8,208	8,046	8,046
Number of counties	347	347	347	347	347	347
R-squared	0.9556	0.9710	0.9853	0.9563	0.9726	0.9800

Notes: The dependent variable is the natural log of enrollment. Standard errors adjusted for clustering by county are in parentheses. All regressions include year and county fixed effects. Columns (3) and (6) also include county-specific linear time trends.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Table 3: Estimates of the Boll Weevil's Effect on ln(Attendance)

	White				Black			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
boll weevil	0.061***	0.026*	0.041***	0.004	0.072***	0.037*	0.053***	0.001
	(0.014)	(0.012)	(0.011)	(0.008)	(0.022)	(0.017)	(0.015)	(0.010)
ln(teachers)?	NO	YES	YES	YES	NO	YES	YES	YES
Time trends?	NO	NO	YES	YES	NO	NO	YES	YES
ln(enrollment)?	NO	NO	NO	YES	NO	NO	NO	YES
Observations	8,234	8,096	8,096	8,088	8,187	8,026	8,026	8,016
Number of counties	347	347	347	347	347	347	347	347
R-squared	0.9348	0.9513	0.9693	0.9817	0.9389	0.9562	0.9670	0.9808

Notes: The dependent variable is the natural log of attendance. Standard errors adjusted for clustering by county are in parentheses. All regressions include year and county fixed effects. Columns (3), (4), (7), and (8) also include county-specific linear time trends. Additionally, columns (4) and (8) add the natural log of enrollment as a control.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Table 4: Estimates of the Boll Weevil's Effect on Long-Run Educational Outcomes

	Years of Schooling		Completed 8th Grade	
	(1) White	(2) Black	(3) White	(4) Black
Age exposed:				
4–6	0.2609***	0.3581***	0.0176*	0.0325***
	(0.0605)	(0.0724)	(0.0086)	(0.0093)
7–9	0.2314***	0.2423***	0.0255***	0.0247***
	(0.0494)	(0.0601)	(0.0069)	(0.0074)
10–12	0.1459***	0.1512**	0.0172***	0.0126*
	(0.0374)	(0.0467)	(0.0052)	(0.0057)
13–15	0.1017***	0.1407***	0.0140***	0.0115*
	(0.0283)	(0.0390)	(0.0040)	(0.0047)
16–18	0.0423+	0.0612*	0.0054+	0.0035
	(0.0222)	(0.0304)	(0.0031)	(0.0036)
Observations	432,235	170,923	432,235	170,923
R-squared	0.1630	0.0909	0.1172	0.0567
Dependent variable:				
Mean	8.3481	4.9775	0.5728	0.1981
Std. Dev.	3.6760	3.2238	0.4947	0.3986

Notes: The dependent variables are given in the column headings. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Table 5: The Boll Weevil's Effect on Long-Run Educational Outcomes with Pre-Trends

	Years of Schooling		Completed 8th Grade	
	(1) White	(2) Black	(3) White	(4) Black
Age exposed:				
4–6	0.2620***	0.3654***	0.0163+	0.0369***
	(0.0631)	(0.0742)	(0.0089)	(0.0097)
7–9	0.2312***	0.2494***	0.0243***	0.0283***
	(0.0509)	(0.0622)	(0.0070)	(0.0078)
10–12	0.1435***	0.1588**	0.0159**	0.0155**
	(0.0387)	(0.0483)	(0.0053)	(0.0060)
13–15	0.0973**	0.1486***	0.0127**	0.0135**
	(0.0297)	(0.0404)	(0.0042)	(0.0050)
16–18	0.0358	0.0692*	0.0042	0.0047
	(0.0235)	(0.0314)	(0.0032)	(0.0038)
22–24	-0.0225	0.0234	-0.0036	0.0020
	(0.0193)	(0.0274)	(0.0027)	(0.0034)
25–27	-0.0251	0.0097	-0.0014	-0.0057
	(0.0232)	(0.0308)	(0.0033)	(0.0037)
Observations	432,235	170,923	432,235	170,923
R-squared	0.1630	0.0909	0.1172	0.0567
F-stat on pre-trend	1.0039	0.3639	0.9042	1.8891

Notes: The dependent variables are given in the column headings. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Table 6: Robustness of the Boll Weevil's Effect on Long-Run Educational Outcomes to Concurrent Shocks

	Years of Schooling		Completed 8th Grade	
	(1) White	(2) Black	(3) White	(4) Black
Age exposed:				
4–6	0.2975***	0.3867***	0.0217*	0.0362***
	(0.0619)	(0.0729)	(0.0088)	(0.0094)
7–9	0.2569***	0.2589***	0.0284***	0.0267***
	(0.0496)	(0.0602)	(0.0070)	(0.0074)
10–12	0.1690***	0.1671***	0.0197***	0.0146*
	(0.0381)	(0.0477)	(0.0052)	(0.0060)
13–15	0.1248***	0.1544***	0.0166***	0.0131**
	(0.0289)	(0.0400)	(0.0041)	(0.0049)
16–18	0.0546*	0.0677*	0.0068*	0.0043
	(0.0229)	(0.0305)	(0.0031)	(0.0036)
Rosenwald exposure?	YES	YES	YES	YES
Compulsory schooling laws?	YES	YES	YES	YES
Observations	432,151	170,911	432,151	170,911
R-squared	0.1630	0.0911	0.1172	0.0569

Notes: The dependent variables are given in the column headings. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. Additionally, specifications include a measure of Rosenwald school exposure, its interaction with an indicator for rural status, and a measure of exposure to compulsory school attendance laws as controls for contemporaneous shocks to education.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Table 7: Robustness of the Boll Weevil's Effect on Years of Schooling to Matching

	White					
	(1)	(2)	(3)	(4)	(5)	(6)
Age exposed:						
4–6	0.2609***	0.2054***	0.2414***	0.2704***	0.2464***	0.2050**
	(0.0605)	(0.0538)	(0.0664)	(0.0640)	(0.0703)	(0.0646)
7–9	0.2314***	0.1957***	0.2244***	0.2325***	0.2258***	0.1603**
	(0.0494)	(0.0444)	(0.0523)	(0.0523)	(0.0554)	(0.0524)
10–12	0.1459***	0.1146***	0.1304**	0.1398***	0.1247**	0.1165**
	(0.0374)	(0.0340)	(0.0409)	(0.0402)	(0.0443)	(0.0417)
13–15	0.1017***	0.0922***	0.0912**	0.0936**	0.0837*	0.0664*
	(0.0283)	(0.0251)	(0.0312)	(0.0304)	(0.0336)	(0.0312)
16–18	0.0423+	0.0281	0.0392	0.0395+	0.0347	0.0276
	(0.0222)	(0.0204)	(0.0243)	(0.0237)	(0.0263)	(0.0253)
Match requirements:						
Unique within 3 years?	YES	NO	YES	YES	YES	YES
Unique within 5 years?	NO	NO	YES	NO	YES	NO
Difference in birth year ≤ 1?	NO	NO	NO	YES	YES	NO
Match on exact names?	NO	NO	NO	NO	NO	YES
Match Rate	27.51	36.25	23.61	24.85	21.24	24.87
Observations	432,235	570,388	370,005	384,553	327,345	361,012
R-squared	0.1630	0.1375	0.1717	0.1699	0.1790	0.1658

	Black					
	(7)	(8)	(9)	(10)	(11)	(12)
Age exposed:						
4–6	0.3581***	0.2539***	0.4050***	0.3416***	0.3687***	0.3631***
	(0.0724)	(0.0627)	(0.0798)	(0.0821)	(0.0936)	(0.0828)
7–9	0.2423***	0.1914***	0.2412***	0.2897***	0.2702***	0.2331**
	(0.0601)	(0.0512)	(0.0659)	(0.0697)	(0.0800)	(0.0705)
10–12	0.1512**	0.1260**	0.1691***	0.1813***	0.1799**	0.1626**
	(0.0467)	(0.0398)	(0.0509)	(0.0544)	(0.0611)	(0.0544)
13–15	0.1407***	0.0990**	0.1475***	0.1370**	0.1262*	0.1544***
	(0.0390)	(0.0338)	(0.0430)	(0.0434)	(0.0498)	(0.0438)
16–18	0.0612*	0.0268	0.0736*	0.0819*	0.0933*	0.0663+
	(0.0304)	(0.0258)	(0.0343)	(0.0336)	(0.0393)	(0.0365)
Match requirements:						
Unique within 3 years?	YES	NO	YES	YES	YES	YES
Unique within 5 years?	NO	NO	YES	NO	YES	NO
Difference in birth year ≤ 1?	NO	NO	NO	YES	YES	NO
Match on exact names?	NO	NO	NO	NO	NO	YES
Match Rate	18.63	27.05	15.12	14.49	11.46	15.48
Observations	170,923	254,122	136,188	129,877	100,179	133,916
R-squared	0.0909	0.0706	0.1054	0.0975	0.1148	0.0952

Notes: The dependent variable is years of schooling. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. Columns (1) and (7) repeat the baseline results shown in Table 4 for ease of comparison.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

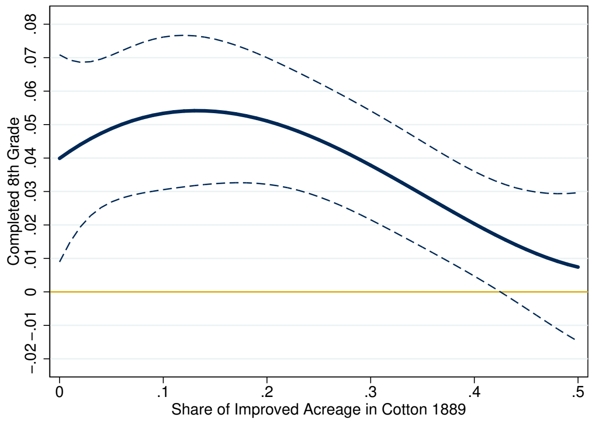
Table 8: Robustness of the Boll Weevil's Effect on Long-Run Educational Outcomes to Weighting

	Years of Schooling		Completed 8th Grade	
	(1) White	(2) Black	(3) White	(4) Black
Panel a: Weighting on Childhood Characteristics				
Age exposed:				
4–6	0.2764***	0.3568***	0.0214*	0.0319***
	(0.0628)	(0.0728)	(0.0090)	(0.0096)
7–9	0.2414***	0.2430***	0.0269***	0.0243**
	(0.0516)	(0.0600)	(0.0071)	(0.0074)
10–12	0.1522***	0.1521**	0.0180***	0.0121*
	(0.0392)	(0.0462)	(0.0053)	(0.0058)
13–15	0.1071***	0.1459***	0.0141***	0.0122*
	(0.0298)	(0.0390)	(0.0042)	(0.0048)
16–18	0.0462*	0.0594+	0.0054+	0.0034
	(0.0234)	(0.0308)	(0.0032)	(0.0036)
Observations	431,541	170,923	431,541	170,923
R-squared	0.1607	0.0829	0.1165	0.0510
Panel b: Weighting on Adulthood Characteristics				
Age exposed:				
4–6	0.2241***	0.3485***	0.0171+	0.0312***
	(0.0639)	(0.0747)	(0.0087)	(0.0089)
7–9	0.2244***	0.2328***	0.0265***	0.0224**
	(0.0530)	(0.0620)	(0.0071)	(0.0071)
10–12	0.1367**	0.1484**	0.0168**	0.0114*
	(0.0414)	(0.0481)	(0.0054)	(0.0055)
13–15	0.1018**	0.1401***	0.0140***	0.0109*
	(0.0309)	(0.0394)	(0.0042)	(0.0044)
16–18	0.0335	0.0612*	0.0047	0.0033
	(0.0245)	(0.0311)	(0.0032)	(0.0034)
Observations	429,267	169,460	429,267	169,460
R-squared	0.1678	0.0929	0.1202	0.0548

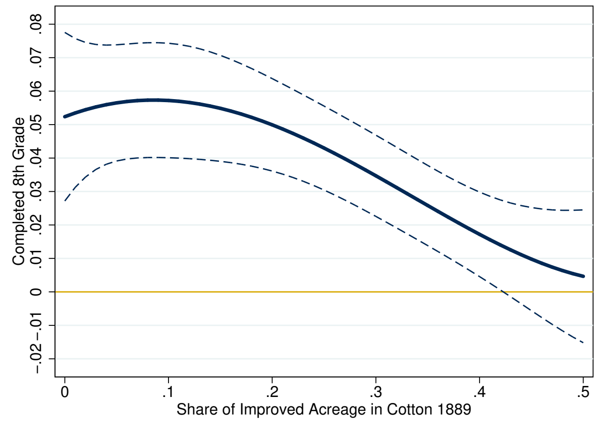
Notes: The dependent variables are given in the column headings. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. The specifications in panel (a) use an inverse proportional weighting method to weight our matched sample to be reflective of the population with respect to observable childhood characteristics: census year; state of residence; race; age; household head's occupational score, homeownership, mortgage status, and literacy; household farm status; and year of infestation of county of residence. The specifications in panel (b) use an inverse proportional weighting method to weight our matched sample to be reflective of the population with respect to observable adulthood characteristics: state of residence; age; and years of schooling.

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

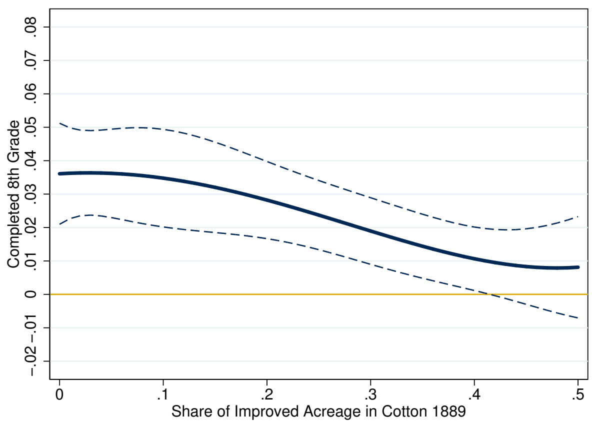
A Appendix Figures and Tables



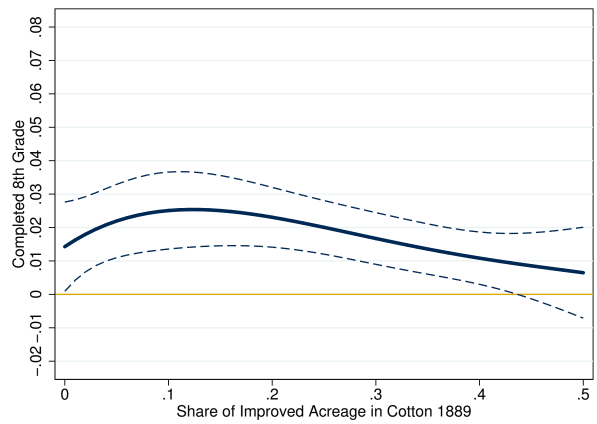
(a) Children exposed at 4–6



(b) Children exposed at 7–9



(c) Children exposed at 10–12



(d) Children exposed at 13–15

Figure A.1: Impact of the Boll Weevil on Eighth Grade Completion by Intensity of Cotton Farming in 1889

Note: The y-axis shows the difference in likelihood of completing eighth grade relative to individuals exposed to the boll weevil between the ages of 19 and 30, inclusive. The dashed lines indicate 95 percent confidence intervals.
Source: See the text.

Table A.1: Robustness of the Boll Weevil's Effect on Years of Schooling to Sample Selection

	White					
	(1)	(2)	(3)	(4)	(5)	(6)
Age exposed:						
4–6	0.2609***	0.3071***	0.2817***	0.2984**	0.2337***	0.2641***
	(0.0605)	(0.0543)	(0.0579)	(0.1067)	(0.0646)	(0.0619)
7–9	0.2314***	0.2495***	0.2293***	0.2178*	0.2047***	0.2249***
	(0.0494)	(0.0439)	(0.0473)	(0.0882)	(0.0526)	(0.0510)
10–12	0.1459***	0.1581***	0.1481***	0.1133	0.1229**	0.1326***
	(0.0374)	(0.0335)	(0.0359)	(0.0723)	(0.0405)	(0.0384)
13–15	0.1017***	0.0893***	0.0963***	0.0646	0.0756*	0.0931**
	(0.0283)	(0.0255)	(0.0275)	(0.0521)	(0.0307)	(0.0297)
16–18	0.0423+	0.0279	0.0324	-0.0179	0.0337	0.0437+
	(0.0222)	(0.0195)	(0.0215)	(0.0394)	(0.0242)	(0.0234)
County-level sample requirements:						
Full infestation prior to 1923?	YES	YES	YES	YES	YES	YES
No full retreat of boll weevil?	YES	NO	YES	YES	YES	YES
Full infestation in $\leq$ 4 years?	YES	NO	NO	YES	YES	YES
Infestation prior to 1915?	NO	NO	NO	YES	NO	NO
No border changes?	NO	NO	NO	NO	YES	NO
1889 acres in cotton $\geq$ 100?	NO	NO	NO	NO	NO	YES
Observations	432,235	554,808	454,308	145,269	367,047	408,250
R-squared	0.1630	0.1590	0.1614	0.1604	0.1591	0.1629
	Black					
	(7)	(8)	(9)	(10)	(11)	(12)
Age exposed:						
4–6	0.3581***	0.3559***	0.3654***	0.4763**	0.3683***	0.3655***
	(0.0724)	(0.0694)	(0.0725)	(0.1470)	(0.0782)	(0.0730)
7–9	0.2423***	0.2422***	0.2489***	0.3708**	0.2350***	0.2397***
	(0.0601)	(0.0579)	(0.0603)	(0.1171)	(0.0646)	(0.0611)
10–12	0.1512**	0.1582***	0.1524**	0.1965*	0.1636**	0.1429**
	(0.0467)	(0.0448)	(0.0467)	(0.0878)	(0.0498)	(0.0473)
13–15	0.1407***	0.1287***	0.1377***	0.1702*	0.1567***	0.1446***
	(0.0390)	(0.0368)	(0.0387)	(0.0739)	(0.0420)	(0.0399)
16–18	0.0612*	0.0525+	0.0596*	0.0555	0.0771*	0.0646*
	(0.0304)	(0.0287)	(0.0302)	(0.0493)	(0.0318)	(0.0311)
County-level sample requirements:						
Full infestation prior to 1923?	YES	YES	YES	YES	YES	YES
No full retreat of boll weevil?	YES	NO	YES	YES	YES	YES
Full infestation in $\leq$ 4 years?	YES	NO	NO	YES	YES	YES
Infestation prior to 1915?	NO	NO	NO	YES	NO	NO
No border changes?	NO	NO	NO	NO	YES	NO
1889 acres in cotton $\geq$ 100?	NO	NO	NO	NO	NO	YES
Observations	170,923	187,167	173,022	65,409	143,631	164,609
R-squared	0.0909	0.0978	0.0911	0.0811	0.0914	0.0901

Notes: The dependent variable is years of schooling. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. Columns (1) and (7) repeat the baseline results shown in Table 4 for ease of comparison.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Table A.2: Robustness of the Boll Weevil's Effect on Eighth Grade Completion to Sample Selection

	White					
	(1)	(2)	(3)	(4)	(5)	(6)
Age exposed:						
4–6	0.0176*	0.0241**	0.0197*	0.0334*	0.0141	0.0202*
	(0.0086)	(0.0079)	(0.0083)	(0.0149)	(0.0091)	(0.0089)
7–9	0.0255***	0.0251***	0.0237***	0.0286*	0.0230**	0.0267***
	(0.0069)	(0.0062)	(0.0067)	(0.0120)	(0.0073)	(0.0071)
10–12	0.0172***	0.0144**	0.0157**	0.0142	0.0146**	0.0173**
	(0.0052)	(0.0046)	(0.0050)	(0.0098)	(0.0055)	(0.0054)
13–15	0.0140***	0.0106**	0.0121**	0.0106	0.0105*	0.0139**
	(0.0040)	(0.0036)	(0.0039)	(0.0072)	(0.0042)	(0.0043)
16–18	0.0054+	0.0020	0.0030	-0.0032	0.0037	0.0065*
	(0.0031)	(0.0027)	(0.0030)	(0.0053)	(0.0032)	(0.0032)
County-level sample requirements:						
Full infestation prior to 1923?	YES	YES	YES	YES	YES	YES
No full retreat of boll weevil?	YES	NO	YES	YES	YES	YES
Full infestation in $\leq$ 4 years?	YES	NO	NO	YES	YES	YES
Infestation prior to 1915?	NO	NO	NO	YES	NO	NO
No border changes?	NO	NO	NO	NO	YES	NO
1889 acres in cotton $\geq$ 100?	NO	NO	NO	NO	NO	YES
Observations	432,235	554,808	454,308	145,269	367,047	408,250
R-squared	0.1172	0.1125	0.1150	0.1138	0.1147	0.1168
	Black					
	(7)	(8)	(9)	(10)	(11)	(12)
Age exposed:						
4–6	0.0325***	0.0281**	0.0329***	0.0558***	0.0325**	0.0355***
	(0.0093)	(0.0089)	(0.0092)	(0.0164)	(0.0101)	(0.0092)
7–9	0.0247***	0.0230**	0.0252***	0.0455***	0.0235**	0.0243**
	(0.0074)	(0.0072)	(0.0074)	(0.0122)	(0.0080)	(0.0074)
10–12	0.0126*	0.0142*	0.0126*	0.0163+	0.0133*	0.0120*
	(0.0057)	(0.0055)	(0.0057)	(0.0096)	(0.0061)	(0.0058)
13–15	0.0115*	0.0118**	0.0118*	0.0139+	0.0133**	0.0119*
	(0.0047)	(0.0045)	(0.0046)	(0.0080)	(0.0050)	(0.0048)
16–18	0.0035	0.0031	0.0036	0.0004	0.0051	0.0035
	(0.0036)	(0.0034)	(0.0036)	(0.0055)	(0.0037)	(0.0037)
County-level sample requirements:						
Full infestation prior to 1923?	YES	YES	YES	YES	YES	YES
No full retreat of boll weevil?	YES	NO	YES	YES	YES	YES
Full infestation in $\leq$ 4 years?	YES	NO	NO	YES	YES	YES
Infestation prior to 1915?	NO	NO	NO	YES	NO	NO
No border changes?	NO	NO	NO	NO	YES	NO
1889 acres in cotton $\geq$ 100?	NO	NO	NO	NO	NO	YES
Observations	170,923	187,167	173,022	65,409	143,631	164,609
R-squared	0.0567	0.0635	0.0569	0.0461	0.0563	0.0555

Notes: The dependent variable is an indicator for reporting completion of eight or more years of schooling in the 1940 census. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. Columns (1) and (7) repeat the baseline results shown in Table 4 for ease of comparison. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Table A.3: Robustness of Estimates of the Boll Weevil's Effect on Long-Run Educational Outcomes

	Years of Schooling				Completed 8th Grade	
	(1) White	(2) Black	(3) White	(4) Black	(5) White	(6) Black
Age exposed:						
4–6	0.2992***	0.2985**	0.2095***	0.3474***	0.0237*	0.0341*
	(0.0795)	(0.1062)	(0.0514)	(0.0682)	(0.0112)	(0.0132)
7–9	0.2652***	0.1899*	0.2105***	0.2351***	0.0310**	0.0261*
	(0.0697)	(0.0910)	(0.0428)	(0.0565)	(0.0096)	(0.0109)
10–12	0.1746**	0.1069	0.1335***	0.1535***	0.0218**	0.0138
	(0.0561)	(0.0747)	(0.0324)	(0.0437)	(0.0077)	(0.0092)
13–15	0.1253**	0.1047+	0.0966***	0.1403***	0.0178**	0.0124+
	(0.0436)	(0.0609)	(0.0247)	(0.0364)	(0.0060)	(0.0073)
16–18	0.0605+	0.0333	0.0369+	0.0585*	0.0084+	0.0043
	(0.0346)	(0.0490)	(0.0191)	(0.0286)	(0.0047)	(0.0058)
Excludes 19–21 from omitted category?	YES	YES	NO	NO	YES	YES
Top codes years of schooling at 12?	NO	NO	YES	YES	NO	NO
Observations	432,235	170,923	432,235	170,923	432,235	170,923
R-squared	0.1630	0.0909	0.1598	0.0915	0.1172	0.0567
Adjusted r-squared	0.1617	0.0874	0.1585	0.0880	0.1158	0.0531

Notes: Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.

Table A.4: Robustness of the Boll Weevil's Effect on Eighth Grade Completion to Matching

	White					
	(1)	(2)	(3)	(4)	(5)	(6)
Age exposed:						
4–6	0.0176*	0.0187*	0.0154+	0.0182*	0.0163+	0.0093
	(0.0086)	(0.0076)	(0.0092)	(0.0090)	(0.0097)	(0.0094)
7–9	0.0255***	0.0239***	0.0249***	0.0251***	0.0248**	0.0164*
	(0.0069)	(0.0060)	(0.0075)	(0.0072)	(0.0079)	(0.0073)
10–12	0.0172***	0.0151**	0.0146**	0.0159**	0.0137*	0.0134*
	(0.0052)	(0.0046)	(0.0056)	(0.0055)	(0.0060)	(0.0058)
13–15	0.0140***	0.0147***	0.0127**	0.0128**	0.0120**	0.0090*
	(0.0040)	(0.0037)	(0.0043)	(0.0043)	(0.0046)	(0.0044)
16–18	0.0054+	0.0062*	0.0053	0.0048	0.0048	0.0033
	(0.0031)	(0.0027)	(0.0033)	(0.0033)	(0.0035)	(0.0033)
Match requirements:						
Unique within 3 years?	YES	NO	YES	YES	YES	YES
Unique within 5 years?	NO	NO	YES	NO	YES	NO
Difference in birth year ≤ 1?	NO	NO	NO	YES	YES	NO
Match on exact names?	NO	NO	NO	NO	NO	YES
Observations	432,235	570,388	370,005	384,553	327,345	361,012
R-squared	0.1172	0.0997	0.1232	0.1214	0.1277	0.1176
	Black					
	(7)	(8)	(9)	(10)	(11)	(12)
Age exposed:						
4–6	0.0325***	0.0214**	0.0319**	0.0306**	0.0308*	0.0321**
	(0.0093)	(0.0080)	(0.0106)	(0.0109)	(0.0123)	(0.0109)
7–9	0.0247***	0.0191**	0.0197*	0.0307***	0.0271**	0.0264**
	(0.0074)	(0.0062)	(0.0085)	(0.0087)	(0.0101)	(0.0090)
10–12	0.0126*	0.0108*	0.0111+	0.0150*	0.0143+	0.0160*
	(0.0057)	(0.0050)	(0.0066)	(0.0069)	(0.0080)	(0.0068)
13–15	0.0115*	0.0076+	0.0097+	0.0107+	0.0080	0.0149**
	(0.0047)	(0.0041)	(0.0054)	(0.0055)	(0.0065)	(0.0054)
16–18	0.0035	0.0010	0.0035	0.0057	0.0067	0.0057
	(0.0036)	(0.0030)	(0.0041)	(0.0042)	(0.0049)	(0.0043)
Match requirements:						
Unique within 3 years?	YES	NO	YES	YES	YES	YES
Unique within 5 years?	NO	NO	YES	NO	YES	NO
Difference in birth year ≤ 1?	NO	NO	NO	YES	YES	NO
Match on exact names?	NO	NO	NO	NO	NO	YES
Observations	170,923	254,122	136,188	129,877	100,179	133,916
R-squared	0.0567	0.0412	0.0672	0.0623	0.0755	0.0618

Notes: The dependent variable is an indicator for reporting completion of eight or more years of schooling in the 1940 census. Standard errors adjusted for clustering by childhood county of residence are in parentheses. All specifications include year of birth fixed effects, childhood county of residence fixed effects, and controls for family background. Family background controls include childhood household head's occupational score, homeownership status, and literacy, as well as indicators for urban location and farm residence. Columns (1) and (7) repeat the baseline results shown in Table 4 for ease of comparison.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.10$.