When the Race between Education and Technology Goes Backwards: The Postbellum Decline of White School Attendance in the Southern US

Hoyt Bleakley[†] University of Chicago

Sok Chul Hong[‡] Sogang University

April 2013

Abstract

This study examines a sharp decline of school attendance among white children in the Southern US after the Civil War. According to Census data, the school-attendance rate among whites in the Confederate states declined by almost half from 1860 to 1870, whereas that in the Northern states was approximately stable. This shock left the South approximately three decades behind its antebellum trend. We use micro data to examine a variety of hypotheses for this drop. In statistical terms, the decline is related to the postwar drop in local wealth and public-school income. Yet our analysis shows that the relationship between literacy and school attendance appears to be quite stable pre- and post-war, which suggests for only a minor role for a drop in school quality (or constraints on time in school). As supporting evidence, we show that the return to schooling, measured by the wage premium for skilled workers, declined substantially in the South after the War. Using longitudinally-linked census samples, we show that well-educated Southerners migrated out of the South among cohorts who attended schools after the War.

Keywords: School Attendance, Literacy, Return to Skill, Civil War, Southern U.S.
JEL Classification: I2, J2, J6, N3

[†] University of Chicago, Booth School of Business, and NBER, bleakley@uchicago.edu.
[‡] Sogang University, Department of Economics, shong@sogang.ac.kr.

We have benefited from comments and suggestions by William Collins, Sun Go, Chulhee Lee, Jonathan Pritchett, Donghyu Yang, and by the participants in the economic history workshop at Seoul National University, the seminars at Asian-Pacific Economic and Business History Conference, and Korean Economics Association Meetings.

I. Introduction

The rise of schooling is one of the more remarkable phenomena of the past two centuries. More education has made not only the labor force more productive but also better able to embrace new technologies. And the advance of technology has increased the demand for educated workers, thus pushing out the demand for skill roughly in parallel with the growth of supply. Most developed countries in the twentieth century experienced a "race between education and technology," and many developing countries are still experiencing that race today (Goldin and Katz 2010). In particular, primary school education has evolved toward extending public education on the basis of government investment, introducing compulsory attendance, and expanding the education for women and minority groups.

Despite this general tendency, on various occasions the transition to widespread school attendance and literacy was marked by sometimes-protracted slowdowns and even retrenchment. Moreover, the rise of school attendance did not always follow the increase of national income, as observed in England and Wales in the nineteenth century (Lindert 2009). Further, disparities by region, race, and gender are commonplace. Numerous economic researchers have invoked various demand- and supply-side factors to explain these patterns, including demand for skilled or well-educated labor, change in skilled workers' wage premium, demographic forces such as fertility rate and life expectancy, discrimination in labor market, restricting entry, cost of education, and change of tax support for basic education (Collins and Margo 2006, Lindert 2009). In many cases, the transition to universal primary schooling starts with marked increases, episodes sometimes referred to as "Lindert moments" (e.g., de Carvalho Filho and Colistete 2010 referencing the work of Peter Lindert).

The present study analyzes an episode in which Goldin and Katz' race between education and technology appeared to run in reverse and thus might be deemed an anti-Lindert moment. Specifically, we examine something rarely explored in the existing literature: the decline of school attendance among Southern whites following the Civil War (1861–1865). Comparing 1860 to 1870, we document a sharp decline in school attendance that equals approximately a third of the 1860 rate and translates into more than a year of schooling. This decline left the former Confederacy[1] with half the school-attendance rate of the Northern states in 1870, a gap

[1] We use the terms "Confederacy", "former Confederacy", "Confederate states" and variations below to refer to those

that took more than a generation to cut in half. Further, the shock left the region approximately three decades behind its antebellum trend.

In Section III, we document this decline using census data from the nineteenth century and provide greater detail on school attendance across time and space. In and of itself, this information is novel in that much of the quantitative literature on schooling patterns in the South has focused on the postbellum 19th century, and largely on blacks rather than whites, or the early 20th century (Collins 2007, Collins and Margo 2006, Bleakley 2007, Margo 1990, Sacerdote 2005). (We review the literature further in Section II.)

Needless to say, the Civil War figures as the central explanation in understanding the decline in school attendance, although this event bundles together various possible impacts on the human-capital decision, including but not limited to the direct effects of war and the institutional changes that came with the failure of Secession. As a first pass, we examine in Section IV various candidate mechanisms by testing how much of the postwar decline in schooling can be explained by various observable factors at the household, county, and state levels. Among these factors, the postwar devastation of household/local wealth and public school income primarily based on local property taxes almost completely explains the estimated coefficient for the postwar decline of school attendance in the former Confederacy. The Southern counties, which were more prosperous than those in the North prior to the Civil War, experienced a sharp decline in per-capita wealth by 68% in 1860-1870, whereas it increased by 11% in the Northern counties. Our analysis suggests that if two regions had experienced similar changes in these local income variables, the decline in school attendance would have been completely attenuated. However, we find that these economic conditions are considerably related to the change of school attendance, but do not account for the postwar change of school-quantity or quality measures in the former Confederacy, such as number of schools per capita and number of teachers per school. This suggests that local wealth variables are more associated with change in demand for schooling rather than supply-side shocks. Moreover, the results for other types of local conditions provide some novel implications. Increasing black schooling as a legacy of

---

states that attempted to secede from the Union during the Civil War. We recognize that the Confederacy did not exist in 1860 and was defunct by 1870, so these terms are inaccurate descriptions for all of the census years. We use them nonetheless to save space.

emancipation and Reconstruction might not crowd out white children's school attendance in the former Confederate states. Infectious disease environment in mid-19th-century America, measured by malaria and hookworm risk, was obviously a detriment to schooling, but it was not the major cause of the postbellum decline in school attendance among Southern white children.

In the second half of the study, we adopt a more model-based approach to decomposing the decline in schooling into supply vs. demand factors. In Section V, we use a standard model of the time-in-school decision and discuss the implications of various shocks that might provoke students to leave school at an earlier age. On the one hand, a decline in the tax base shrinks the government's ability to provide schools, which could constrain students from attending as much school as is optimal and/or depress the quality of the instruction received per unit time in school. Considering downward-sloping marginal benefit (*MB*) curve and upward-sloping marginal cost (*MC*) curve, the cases above have contrasting implications for marginal benefits (additional benefits that accrue from the additional time in school). In the "constrained" case, school attainment is being rationed down from the initial equilibrium level, which pushes up the *MB* of schooling. Thereby, the gross *MB* and net *MB* (i.e. *MB* minus *MC*) will go up. In contrast, a decline in school quality leaves the net *MB* unchanged in equilibrium, but reduces the gross *MB* of schooling by shifting both curves downward. On the other hand, the labor-market return to schooling may have declined postwar. This would also lead to less schooling in the long run. The net *MB* of schooling would remain unchanged as people adjust their optimal schooling decision, and the gross *MB* would not change much because this can increase the opportunity cost of schooling. In summary, each hypothesis has different results in terms of net and gross *MB* of schooling. In Section V, we take these implications of the model by using literacy and occupational income score as proxies of gross marginal benefits.

First, Southern students in 1880 became literate at the same rate, per unit time in school, as they did in 1860, but the later cohorts simply spent less time in school. We measure the path to literacy as a function of time in school by treating the cross-section of school ages in 1860 and 1880 as a pseudo-panel. In other words, we compare how quickly both literacy and (imputed) years of schooling rise with age in those two years, and interpret the relationship between the two as the marginal benefit to literacy from time in school. If school quality had declined, we would expect this relationship to attenuate. In contrast, the data show essentially the same relationship before and after the war between literacy and time in school. Our analysis suggests

that Southern students in 1880 were on track to achieve the 1860 level of literacy if had they just spent the same amount of time in school. The stability of this output per unit input is inconsistent with the hypothesis that school quality declined.

Second, we examine several data sets constructed by longitudinally linking observations in the Census manuscripts in the years 1850–80. This allows us to examine the adult outcomes of individuals depending on whether they attended school in their teens, and differentially by region and by antebellum status. Again, we do not find evidence supporting the hypotheses of constrained schooling or diminished school quality. Being in school predicts literacy and occupational status similarly for cohorts of school age in the South before or after the War. Instead, we find that, at the individual level, being in school predicts leaving the South by 1880 for the cohorts of school age in 1870, but it does not predict outmigration from the region for those of school age prior to the War. This 'brain drain' suggests a lower skill premium in the South during the postwar period, although it had less effect on older cohorts that had already made their location-specific investments.

Finally, we present evidence that the skill premium plummeted during and after the Civil War by comparing regional wages series for engineers and (unskilled) laborers constructed by Coelho and Shepherd (1976). The skill wage premium, which was measured by the relative wage, was cut in half in the South during 1860-1868. This primarily resulted from a substantial decline in average *real* wage among the highly skilled. This evidence, combined with the differential outmigration by more educated Southern workers, suggests that the main mechanism for the decline in school attendance was that the labor-market return to skill dropped after the War. These results are consistent with a supply elasticity for skill (more precisely, time in school) of slightly more than one.

II. Literature Review

Emphasizing the educational advance in the twentieth century and its considerable contribution to economic growth, Goldin and Katz (2010) refer to the twentieth century as the Human Capital Century. According to their estimate, advances in education across the twentieth century account for almost 15 percent of the labor productivity change in the United States. The American educational attainment was high for most of the twentieth century by the standards of other nations.

As Goldin and Katz (2010) point out, the key features of U.S. educational institutions and human capital accumulation had largely taken shape in the nineteenth century. Most of all, the expansion of publicly-supported formal schooling, which emerged in the period before American Civil War, determined U.S. development of educational system in the twentieth century leading the world in schoolings. However, such success was not continuous. U.S. educational evolution was closely tied up with the social, political, demographic, and industrial forces that shaped the nineteenth century (Cubberley 1919). Those factors occasionally held back and even retrenched the widespread public schooling and educational attainment. In the following, we briefly review existing historical literature on how American education has changed and various factors of the long-term change.

Although it has been argued that American public education began with the establishment of the district of school of Colonial New England (Cubberley 1919), public school in the period was different from that in the nineteenth century in both theory and practice. So, many educational historians have argued that American public school did not come into being until the passage of the state public school laws, which were mostly passed after the American Revolution (Drake 1955). Following the state school laws and the subsequent establishment of the State Board of Education, some states began to make grants to the local schools, and later created state school funds. Besides the public school funds, early public schools were funded from various sources, such as local voluntary supports for town schools, tuitions paid by parents (known as rate bills), and local property taxes. But toward the mid-nineteenth century, local property taxes had been the major source as people's attitude gets toward tax-supported schools (Cubberley 1919). Accordingly, local autonomy and political voices were key factors that determined the size of tax support and consequently the local status of public education (Dabney 1936). Regarding these factors, Go and Lindert (2010) estimate that counties with higher shares of free men who actually voted for president had significantly higher school attendance rates among white children in 1850 than otherwise, and conclude that the regional disparity in the distribution of political voice and voting power led to uneven early rise of American public schooling between Northern and Southern regions.

Numerous historical studies have well documented that the uneven distribution of public schooling between Northern and Southern regions had persisted until the early twentieth century (Cubberley 1919, Drake 1955, Dabney 1936, Knight 1922, Cremin 1965). They pointed out

different economic and political interests of property owners as its key factor. Many of them considered the Civil War as one of the major obstacles that had delayed the development of public schooling in the South. It is worthy to note that the effect of the Civil War on school attendance or other educational variables has been rarely quantified in existing studies, although descriptive evidence has been much provided. Summarizing the descriptive argument, postwar economic desolation in the former Confederacy made it difficult to meet adequately its educational and other social needs. In addition, the social conservatism of the South, which had grown strong under the influence of slavery, hindered the development of public schooling; the public school and mass education was scorned as one of the fruits of reconstruction among the old landed aristocracy in the South (Knight 1929).

Much of the quantitative literature on historical schooling patterns has focused on the gap of educational attainment between whites and blacks, and its long-term trend. Collins and Margo (2006) examined this trend on the basis of the IPUMS data. They found that although the school attendance and literacy rates among black populations were very low in the immediate aftermath of the Civil War, the rates substantially increased and converged to the levels of white populations over several decades after the War. From economic perspectives, they suggested two factors as the explanations of the convergence. First, the marginal return to schooling at the beginning of exposure to formal schooling due to emancipation exceeded its marginal cost, and so the parents of black children would have desired to send them to school. Second, the marginal return to schooling, measured by wage premium for educated blacks, was substantial in the early postbellum period. Relatedly, Sacerdote (2005) compared outcomes such as literacy, school attendance, and adult occupation between former-slave families and free-black families. He finds that it took roughly two generations for the outcomes of two types of black descendants to converge, and suggests that the convergence was facilitated by intermarriage among slave and free families.

But Collins and Margo (2006) show a divergence of the racial gap after Reconstruction ended in 1877. As pointed out by historians in American education, politically influential Southern planters had strongly opposed the public supports and expenditure for formal schooling. They economically benefited from the vast supply of unskilled labor and did not want to expand universal public education. Therefore as the black's political clout had waned with the end of the reconstruction era, less investment in education for blacks was made by white-dominated local

school boards. In addition, Pritchett (1989) argues that the tax burden of public education expenditure for blacks increased in some Southern states like North Carolina particularly during the first decade of the twentieth century, as the incidence of school taxes shifted from property and land owners (mostly whites) to renters and workers (disproportionately blacks).

Over time, the impediments to public schooling in the South eroded with philanthropic efforts and later the civil rights movement. Collins (2007) showed that the racial and regional gap strongly converged after the 1940s, as the post-1920 birth cohorts entered the labor market and as inter-regional migration had been more frequent. He suggests that as more educated workers migrated from the North to the South because wages in skilled occupations in the South approached those in the Northeast and Midwest, this was helpful for not only increasing the convergence of educational attainment, but also facilitating technological transfers and capital transfers.

On the other hand, some recent studies highlight the significance of disease environment and its eradication in improving educational attainment. Bleakley (2007) shows that areas with higher level of hookworm infection rates experienced a large increase in school enrollment, attendance, and literacy after the successful eradication of the disease from the American South, which was conducted throughout the 1910s. By matching the outcomes of cohorts found in the 1960 census to the state-level malaria risk at the year of birth, Barreca (2010) finds that more malaria exposure lowered educational attainment and increased the poverty rate in later life.

Finally, a related literature has examined the long-term effects of war and war-related deprivation on child outcomes. This literature typically focuses on early-childhood or in-utero exposure. The seminal studies in this area are of the Dutch winter famine (Stein et al. 1975). (See Akresh and de Walque (2011) for a more recent study following up instead on the Rwandan Genocide.) The evidence below shows a protracted decline in the postbellum South, however, which could not plausibly be explained by early-life exposure to war for a few cohorts.

III. Descriptive Statistics

III.A. Data Sources and Key Variables

The first goal of this study is to quantify the change in schooling before and after the Civil War. We measure it with a variable of school attendance, which is available in historical censuses. In each census, the respondents were asked whether they attended school or not during

a specified period, which was generally defined as the year preceding the enumeration date.[2] The variable is similar to a modern concept of "enrollment" in that only the students who went to school for a day were coded as yes, but we use the terms of "attendance" to match the historical sources.

In Section III.B, we examine the trend of US school-attendance rates among white populations by region in the second half of the nineteenth century. To accomplish this, we use the 1850-1900 IPUMS (Integrated Public Use Microdata series) data,[3] which were drawn from the manuscript census schedules. We select white males and females aged between 5 and 20. For each year, we collapse the dummy for attending school by region or state to estimate average school-attendance rate. Moreover, we utilize county census tabulations (ICPSR #2896), which were digitized by Haines (2005), to construct a map that shows the percentage change in school attendance by county in 1850-1870.

Section IV estimates the significance of the decline in school attendance after the Civil War and seeks its explanations from individual, household, and local characteristics. We conduct individual-level analyses focusing on white males and females (again aged between 5 and 20) in the 1860-1870 IPUMS. In the regression analysis, the dummy of school attendance is used as the dependent variable.

Control variables used to account for the decline in Section IV can be classified into three types. First, individuals' age and sex will reflect demographic characteristics. Second, we obtained their household characteristics from the census records in the IPUMS data, such as dummy variables that show whether they lived with father or stepfather; the value of the household head's real estate wealth and personal property wealth; a dummy variable that shows whether the household head was literate or not; and the number of the household head's own children (of any age or marital status). Third, clustering counties by SEA (state economic area), we examine the role of county-group characteristics in 1860 and 1870: wealth per capita, value of total farm output per capita, population density per acre, ratio of slave populations, number of

[2] The official enumeration date of all the censuses in 1850-1900 is June 1. Thus, a bias in measuring and comparing school attendance rate across the years by seasonality is unlikely in the analysis based on population census. In addition, each census in 1850-1900 has used the same questioners: "attended school within the year". (source: http://usa.IPUMS.org)

[3] The sample for 1890 is not available because the 1890 manuscript schedules were destroyed by a fire.

large-sized farms, and two ecology indexes that measure local malaria and hookworm risk. We obtained the county-level information on wealth, farm output value, the size of total and slave populations, and number of farms by size from ICPSR #2896 (published census tabulations digitized by Haines (2005)). The ecology indexes were estimated in the county level on the basis of environmental risk factors such as weather variables, elevation and soil type, which are estimated by Hong (2007) and Bleakley and Hong (2012), respectively. Those variables are originally measured or estimated in the county level; we converted them into the values at the SEA level using a within-SEA county-area weighted average calculation.

Finally, we use two state-level variables in 1860 and 1870: public school income per capita and total tax revenue per capita. Both were obtained from ICPSR #2896. We primarily consider the role of public school because the type of schools in the period was dominated by public school, mostly known as common school; 91.4% of total pupils in 1860 and 86.4% in 1870 attended public schools, respectively. Public school income is the local fund for maintaining the system of public school, mostly known as common school. It was generally financed from three sources in the mid-nineteenth century: local property taxes, permanent school endowment fund, and tuition fees paid by parents of public-school pupils. Out of them, local property taxes took 54.4% of total public school income in 1860, and this share increased to 91.9% in 1870. To examine the role of taxes, we use total revenue per capita as another key state-level variable, which sums state, county, city, and town taxes. All the variables in monetary unit such as wealth, public school income, and taxes are converted in constant 1870 dollars.[4]

On the other hand, this paper focuses on the school attendance rate between 1860 and 1870. Thus, we drop 15 states that had not achieved statehood before the battle of Fort Sumter in April 12, 1861. In addition, we will compare white school attendance rates between specified state groups throughout this paper, such as Northern, Southern, Confederate, slave, Border, and frontier states. In particular, slave states are the 16 states that allowed slavery by 1860; the Border States are five slave states that were not part of the Confederacy[5]; frontier states are those located west of the Mississippi River. (In Section V, we will use longitudinally-linked samples to test the role of school-demand shocks; we will discuss the datasets in that section.)

[4] The converting deflator is 5.39 for 1860 and 7.20 for 1870 (source: measuringworth.com).

[5] The Border States are Delaware, Missouri, Kentucky, Maryland, and West Virginia.

III.B. Maps and Time-Series

Figure 1 graphically describes the 1850-1900 trend of school-attendance rate among white populations by region and state. We first examine the trend by region in its upper panel, where the region is mainly divided into the Northern and former Confederate states as of 1861. The main feature of the figure is that the Confederate states experienced a substantial decline in school attendance between 1860 and 1870. Although the average attendance rate of Northern states also declined in the period, the decline was slight; the rate had been relatively stable around 60% over the second half of the nineteenth century. However, the rate of the Confederacy declined from 41% in 1860 to 26% in 1870.[6]

[Insert Figure 1 Here]

At its most ambitious, the study sets out to gauge what the counterfactual path of school attendance might have been in the former Confederacy, absent the events during and just after the Civil War. We recognize that attempting this with just the time series is fraught with problems, but we offer some simple analyses here as a starting point. These results should be taken with a grain of salt. The first exercise that we conduct, shown with the thin gray line in the graph, is to simply project forward the antebellum trend in school attendance. We can see that the postwar states of the former Confederacy were left approximately three decades behind their antebellum trend.

A second version attempts to construct a counterfactual using re-weighting of the sample of Northern counties. The resulting sample was more similar to the South in terms of human-capital investment for whites, but was dissimilar in that it was not exposed to the effects of the Civil War and Reconstruction.[7] The results of this exercise are shown in the dashed gray

[6] The decline of the Southern states including Border States was milder than that of the Confederacy. The rate declined from 42% in 1860 to 34% in 1870.

[7] We construct the sample weight, which would make the Northern states' attendance rate the same as that of the confederate states in 1860, at the SEA level. Suppose that we cluster the sample in 1860 into *i* number of SEAs (state economic areas) by the within-sample average of SEA attendance rate. In each group *i*, the average rate between the Northern and confederate states will be approximately same (i.e., $S_{Ni} \approx S_{Ci}$). Then, the average rate of the Northern states (S_N) is calculated by population-weighted average, i.e., $S_N = \sum_i \left(\frac{P_{Ni}}{P_N}\right) S_{Ni}$, where P_N, P_{Ni}, and S_{Ni} denote total sample size in the Northern states, the sample size of SEA group i in the North, and the Northern SEA group i's average attendance rate, respectively. Similarly, the confederate states would have SEAs that belong to each group *i*.

line in the graph (still the upper panel of Figure 1). Note that the decline in 1860-1870 is not observed in the re-weighted sample from the North. Indeed, if anything, the antebellum trend in the South understates the rate of convergence relative to Northern counties with similarly low school-attendance rates.

The lower panel of Figure 1 shows a scatter plot between school attendance rates in 1860 and 1870 at the state level. Most Southern states (labeled in *italics*) are far below the 45-degree line, whereas most Northern states are located around the line. This also suggests that a substantial decline in school attendance in 1860-1870 occurred among those Southern states. It is worthy to note that the plots of five Border States (i.e., Southern slave but non-confederate states such as Delaware, Kentucky, Maryland, Missouri, and West Virginia) are above or relatively near the line. This implies that the postbellum decline in white school attendance rate was more severe in the Confederate states. Therefore, we will compare between Northern states and the Confederate states in the regression analysis in Section IV, excluding those five Border States.

Figure 2 provides a closer look at the change of school attendance at the county level. Because the IPUMS data include an insufficient number of observations at the county level, the county-level school attendance rates were calculated from county census tabulations in ICPSR study #2896. Further, the antebellum county information was adopted from the 1850 census because the 1860 census tabulations do not report the number of students at the county level. The counties that experienced more substantial decline during the period are depicted by brighter colors in the map. There is no particularly strong pattern that emerges within the South. There are a few pockets in which school attendance rises, but generally speaking the decline is widespread and roughly homogenous across the region. There are a few exceptions of note, however. The somewhat inaccessible mountainous areas of West Virginia, eastern Kentucky, and the northwestern half of Arkansas all show increases in school attendance. The similarity among these areas suggests a role for geographic factors, but presumably political factors are important

---

The average rate of the confederate states (S_C) can be obtained by $S_C = \sum_i \left(\frac{P_{Ci}}{P_C}\right) S_{Ci}$, where C denotes the Confederacy. Therefore, if a weight (W_{Ni}) such that $W_{Ni} = \left(\frac{P_N}{P_{Ni}}\right) \cdot \left(\frac{P_{Ci}}{P_C}\right)$ is multiplied in calculating S_N, the re-weighted average rate of the Northern states will be approximately same with S_C because S_{Ni} and S_{Ci} are assumed to be approximately same above. To obtain the weight used in Figure 1, we clustered the sample in 1860 into 35 SEA groups and applied it to the sample in 1870-1900. The number of groups was not critical.

too. Most notably on the political side is that, even if we had not drawn the state boundaries, we could still see the boundary between Virginia and West Virginia (where many of the Virginia-side counties are also mountainous) and the northern border of Kentucky (where the counties share a riverine orientation on either side of the Ohio). In the North, essentially all of the counties show increases in school attendance rates from 1850 to 1870. The largest percentage increases are seen in the extreme northwest, such as in Minnesota.

[Insert Figure 2 Here]

III.C. Summary Statistics

Table 1 continues the presentation of summary statistics for the years 1860 and 1870, labeled "prewar" and "postwar", respectively. For the variables described in Section III.A, we present averages for states that joined the Confederacy during the Civil War and compare them with averages for states that remained in the Union but were not Border States.

[Insert Table 1 Here]

Panel A shows that average school attendance rate among white young populations aged between 5 and 20 decreased from 41.2% in 1860 to 26.0% in 1870 in the Confederacy (15.2%-point reduction), and from 62.8% to 60.3% in the Union (2.5%-point reduction). In other words, the absolute rate of reduction was larger by 12.7% points in the Confederacy. In proportional terms, the white school attendance rate substantially declined in the Confederacy by 33%-40% across age and gender groups, whereas the reduction rate is only about 2%-6% in the Union. In both regions, the postbellum decline is observed higher among those aged between 5 and 12, and boys. The age and gender composition is much similar between two regions and across the years, although within-sample average age increased by 3.2% in the Confederacy.

Panel B compares various characteristics of family, which each observation belonged to, between two regions in 1860-1870. In the Confederate states, the fraction of youth without a father or with a stepfather slightly increased after the Civil War by 19%-28%. Average value of household heads' real- and personal-property wealth considerably decreased in the Confederacy after the Civil War (by 51% and 84%, respectively), but it increased in the Union (by 40% and 47%, respectively). In particular, the reduction in personal property wealth in the Confederacy can be largely explained by emancipation because of slaves counted as personal property in the

1860 census.[8] Household heads' average literacy rate and average number of own children also slightly declined in both regions during the decade, but the reduction rate was a bit higher in the Confederacy.

Statistics by county group (grouped into State Economic Areas or SEAs) are reported in Panel C. We use the SEAs as county groups to manage changes in county boundaries; while county boundaries often changed in the mid-nineteenth century as counties were merged or split, those of SEAs were often defined by state borders or prominent natural features, and thus were more stable than county borders. In Panel C, it is worth noting that the Confederacy was richer than the Union before the War at least in terms of wealth per capita held by white households with school-aged children. However, while the Union SEAs experienced small increases, the SEAs in the Confederacy experienced a large reduction in wealth (-68%) and farm output value per capita (-28%). In 1860, the average ratio of slave populations in the Confederacy is about 36.8%. As discussed above, the emancipation is a key factor of local average wealth in the Confederacy. Similarly, the Confederacy experienced a 50% reduction in the number of large-sized farms (i.e., planters), which heavily relied on the slave labor force.[9]

Population density increased in both regions (21% in the Confederacy and 29% in the Union). Ecology indexes of malaria and hookworm can be considered as annual likelihood of being infected with the diseases. (Because we estimated the average indexes over the second half of the nineteenth century, we do not report their difference between 1860 and 1870.) The figures in the table suggest that people in the Confederacy region were exposed to 134% and 339% higher risk of malaria and hookworm infections, respectively, than those in the Union region on average.

Finally, Panel D shows that average financial status of Confederate states, which might be closely related to the supply of public schools, was poor compared with Union states. Moreover, the gap widened after the Civil War. The average of public school income per capita

[8] The direct effect on wealth of emancipation might be best thought of as a transfer of wealth from slave owners to freedmen, which need not change aggregate wealth. But it would of course change the wealth of the white households, the focus of the present study. At the county or state level, such a transfer might influence the tax base in that it was more common to raise tax revenue through assessments on property rather than on labor income.

[9] The postbellum decline in relative incomes was persistent. Southern per capita incomes did not converge with the rest of the country throughout the remainder of the 19th century (Wright 1986).

decreased by 4% in the Confederacy, but it increased by 81% in the Union. If we normalize public school income by the average school-attendance rate shown in Panel A, both regions saw increases (53% in the Confederacy and 88% in the Union). This suggests that school expenditure, conditional on school-attendance, rose in the Confederate states. Although the growth rate of tax revenue per capita is observed in both regions, the growth of the Confederacy (70%) is far behind that of the Union (169%).

IV. Accounting for the Decline: Regression Evidence

IV.A. Baseline Estimates: Magnitude of the Decline

To estimate the significance and magnitude of postbellum school attendance rate among white young populations in the Confederate states, we first use a difference-in-difference estimation model as described in the following equation (1).

$$Sch_i = \alpha + \beta Post_i + \gamma Post_i \times Confed_i + \delta_{Age\times Sex} + \delta_{SEA} + \varepsilon_i \quad (1)$$

In the equation, Sch_i is a dummy variable that indicates whether individual *i* attended school within one previous year or not. Thus, this is a linear probability model based on a pooled OLS regression. $Post$ is a dummy variable that indicates whether he or she is in the postwar (i.e., 1870) IPUMS sample. Then, the variable is interacted with another dummy that shows whether the individual was living in the Confederate region. Accordingly, the coefficient β measures how much average school attendance changed after the Civil War, and the coefficient γ will measure its difference between the Confederate and Northern states. As standard controls, we include age-by-sex dummies ($\delta_{Age\times Sex}$) and SEA fixed effects (δ_{SEA}). (The variable of $Confed$ itself is omitted due to multicollinearity with SEA fixed effects.)

Panel A of Table 2 reports the estimation results based on equation (1), where we report only the coefficients of key control variables and standard errors clustered on the SEA level. First, models (1)-(2) do not use SEA fixed effects so that the control of $Confed$ is feasible. Model (1) says that white school attendance declined by 5.0% points after the Civil War, and the Confederacy experienced a 28.3%-point decline on average. Model (2), which includes the dummy of postwar interacted with the dummy of the Confederacy, shows that the postwar decline of school attendance was more substantial in the Confederacy than in Northern states. Models (3) and (4) include SEA fixed effects. The result is similar. In terms of the baseline

estimates in model (4), we can say that white school attendance declined by 13.0% points more in the Confederacy than that in the Union, which approximately equals a third of the 1860 rate of the Confederacy (41.2%). This figure is close to what we show in Table 1.

[Insert Table 2 Here]

Models (5)-(9) employ subsamples clustered by the location of states, age, and gender. The same implication is consistently found across these subsamples, but the magnitude is estimated a bit higher among the states located in the east of Mississippi River, younger populations aged between 5 and 12, and boys.

On the other hand, Panel B compares the effect between the 1861 slave and non-slave states so that we include the observations in the Border States. Although slave states experienced larger decline in school attendance after the Civil War than did non-slave states (8.7%-point difference in terms of the baseline estimate), the estimated magnitude is smaller than that of Panel A. The pattern across subsample is almost identical with that of Panel A. For the remainder of the paper, we focus on the difference between Confederate and Union States, excluding the Border States.

IV.B. The Role of Various Observable Factors in Explaining the Decline

Now we try to seek the potential explanations of the postwar decline in white school attendance particularly in the Confederacy, using various observable factors at the family, county-group, and state level. First, we add the characteristic variables of the family (Fam_i), to which each observation belongs, in equation (1) as follows:

$$Sch_i = \alpha + \beta Post_i + \gamma Post_i \times Confed_i + Fam_i \boxplus + \delta_{Age \times Sex} + \delta_{SEA} + \varepsilon_i \qquad (2)$$

In Table 3, we report the estimation results only for the key coefficient γ and those of familial variables. How γ changes as new controls are added is of interest. First, model (2) shows that if the father is not found in the census manuscript schedule or observations lived with stepfathers, they are less likely to have attended school than otherwise. This partly reflects the role of parental investment in children's human capital accumulation. Compared with model (1), which is the baseline estimate in the previous subsection, model (2) estimates a smaller magnitude of the coefficient for '*Post* × *Confed*'. This suggests that the decline in the Confederacy can be explained very slightly by change in presence of father. This can be

connected to the statistics in Table 1 that the ratio of samples without fathers increased in the Confederacy.

[Insert Table 3 Here]

In model (3), we consider the role of household heads' wealth controlling for two wealth variables in a quadratic form. We use the demeaned wealth variable for the squared term to figure out the threshold point of wealth. The result says that as both types of household heads' wealth increases, the likelihood of school attendance significantly increases, but it gradually declines beyond the sample average of wealth. This suggests that parental financial condition was a key factor for children's schooling, but it might be more significant among relatively lower-level SES households.[10] On the other hand, the magnitude of '*Post* × *Confed*' is lowered from model (1). This implies that a large decline in household wealth in the Confederate sample, which is reported in Table 1, probably led to a decline in children's school attendance. But the power for explaining the overall decline in school attendance appears to be small.

In models (4) and (5), we examine the role of household head's literacy and fertility. Those whose household heads were literate attended school by about 10% points more than otherwise, which supports the significance of parental education levels in children's education. In households with more own children (or probably observations had more siblings), more school attendance is seen. This suggests that younger children presumably benefited from older children's support for family finances. However, the addition of these two variables does not change the magnitude of '*Post* × *Confed*'.

Finally, model (6) includes all the familial variables. The coefficient of '*Post* × *Confed*' changed from -0.1296 in model (1) to -0.1045. A simple calculation suggests that those familial variables approximately account for about 20% of the decline in white school attendance in the Confederacy after the Civil War. In particular, the devastation of individual wealth and the disorder of family structure in the Confederacy after the Civil War are considered as key family-level factors.

[10] Note that rich households may have educated their children in private ways rather than through the public school system.

In Table 4, we examine whether various local conditions or their changes in 1860-1870 explain the postwar decline of white school attendance in the Confederacy. In particular, county-group (SEA) variables are added to equation (2) as follows:

$$Sch_{ij} = \alpha + \beta Post_{ij} + \gamma Post_{ij} \times Conf\ \square\square_{ij} + Fam_{ij}\text{H} + SEA_{ij} \times Post_{ij}\Gamma + \delta_{Age\times Sex} + \delta_{SEA} + \varepsilon_{ij} \quad (3)$$

In equation (3), SEA_{ij} denotes the variables of SEA j where individual i resided. It was interacted with the dummy variable that indicates the postwar sample; SEA itself is dropped because of the use of SEA fixed effects. From this specification, we can figure out how the marginal effect of local conditions on school attendance changed after the Civil War.

[Insert Table 4 Here]

Model (3) in Table 4 uses the change of the logarithms of average wealth per capita as county-group characteristics. The key finding is that the addition of the variable considerably changes the coefficient that measures the decline of school attendance in the Confederacy. Compared with that of model (2), which controls for only family characteristics, the magnitude decreases by more than a half. In addition, the coefficient of change-of-wealth variable suggests that the positive effect of local wealth increase on school attendance became more influential after the Civil War. Considering the postwar devastation of per-capita wealth in the Confederacy (Table 1), this result implies that the devastation of wealth strongly resulted in the substantial decline of school attendance in the area. A similar implication is also found from model (4), where we measure the change of local economic conditions with farm output value per capita. But the variable resulted in a smaller change in the magnitude of school-attendance decline in the Confederacy.

On the other hand, as seen from the summary statistics in Table 1, the devastation of wealth in the Confederacy largely occurred in personal-property wealth, which includes the value of slaves. Therefore, the postwar decline in school attendance would be more substantial among counties with more slaves. Higher slave populations might also predict lower postbellum school attendance among whites if educating blacks crowded out some resources for white schools. However, this hypothesis is not supported as estimated in model (5), where we control for the ratio of slave populations in 1860. In addition, the emancipation of slaves heavily affected Southern planters who held many slaves and cultivated large-sized farms. Model (6) tests this

possibility using the change in the number of farms with 500 acres and more from 1860-1870. The estimated coefficient suggests that the postwar school attendance rate rose among the SEAs where the number of large-sized farms increased. The lowered magnitude of '*Post* × *Confed*' is in accordance with the fact that the number of farms decreased in the Confederacy (Table 1).

In model (7), we consider the role of urbanization by controlling for county-group population density in 1860. The result shows that denser (i.e., urban) areas experienced an increase in school attendance after the Civil War. The variable also slightly accounts for the postwar decline in the Confederacy. This reflects the trend that the growth of population density in the Confederacy was slower than in the Union, as shown in Table 1. Finally, frequent infection might have interrupted children's school attendance. Moreover, certain parasitic infections were more prevalent in the South. In model (8), we test whether the postwar decline in school attendance was higher in the regions at high risk of malaria and hookworm. But those ecological environments do not explain the postwar decline.[11] The results for the latter variable are inconsistent with the general argument of Brinkley (1997), who argues that hookworm infection rose after the war and was responsible for a substantial decline in output.

We include all the county-group variables in model (9). Although those variables are highly correlated with each other, the coefficients of the changes in wealth and number of large-sized farms interacted with postwar dummy are still estimated statistically significant. The coefficient for farm output value and hookworm index becomes statistically significant. Nevertheless, size of the coefficient for '*Post* × *Confed*' is similar to that in model (3), which controls for wealth variable only. The coefficient suggests that those county-group characteristics, but particularly the devastation of wealth in the Confederacy, account for approximately 50% of the postwar decline of school attendance rate in the Confederacy. Moreover, the coefficient for slave is still statistically insignificant in model (9). Again, this suggests that the crowding-out effect of increasing black schooling might not be a critical factor that caused the decline in white school attendance after the War.

[11] This does not mean that the diseases themselves do not have impacts on school attendance. In the regression excluding SEA fixed effects, the coefficients [and standard errors] of malaria and hookworm indexes are estimated significantly negative (-0.6526 [0.0683] and -0.2558 [0.0501], respectively).

In Table 5, we consider two state-level variables by interacting them with the postwar dummy as the following equation. The characteristics of state s are denoted by ST_{ijs}; the variables themselves are omitted due to the use of SEA fixed effects.

$$Sch_{ijs} = \alpha + \beta Post_{ijs} + \gamma Post_{ijs} \times Confed_{ijs} + Fam_{ijs}\text{H} + SEA_{ijs} \times Post_{ijs}\Gamma$$
$$+ ST_{ijs} \times Post_{ijs}\Pi + \delta_{Age\times Sex} + \delta_{SEA} + \varepsilon_{ijs} \quad (4)$$

We first add the change of the logarithm of public school income per capita in 1860-1870 in model (4), without controlling for the county-group characteristics used above. Considering the previous results presented in models (1) and (2), the change in public school income substantially explains a large portion of the postwar school-attendance decline in the Confederacy. When the county-group variables are added together in model (5), the magnitude of the decline almost completely disappears. Model (6) also includes the change in taxes per capita, but this does not change the result. The effect of public school income interacted with the prewar dummy is similarly estimated in models (4)-(6). Therefore, we conclude that school attendance rates rose among the states where more public school income was provided after the Civil War as shown in Table 1. Moreover, this is another key statistical explanation of the postwar decline in the Confederacy.

[Insert Table 5 Here]

IV.C. Public School Income and Various Public School Indexes

In the previous subsection, the county-level variable that emerges as key in explaining the decline in school attendance is the change in average wealth per capita, knocking out of the specification variables such as the racial mix, disease environment, and urbanization, among others. This does not mean that those variables other than wealth did not affect educational attainment in mid-nineteenth-century America; note that we focus on the finding that the coefficient for '*Post* × *Confed*' is reduced considerable with the addition of wealth variables. However, this variable is likely subject contamination by reverse causality: school spending is at least in part a function of the number of students enrolled. Moreover, although the role of public school income heavily based on local wealth and property taxes seems to suggest that the decline resulted from supply-side shock by a channel in which lowered public school income constrained the quantity of schools such as the number of schools and teachers, the previous regressions do

not show this channel clearly. Rather, they show only the relationship between public school income/local wealth and school attendance rates. From that aspect, we examine whether the postwar reduction in public school income in the Confederacy caused a supply shock in terms of some public-school indexes.

Because the variable of public school income is available only at the state level, we conduct a panel analysis using the 1860 and 1870 published census tabulations (ICPSR #2856) for 27 states excluding the Border States.

$$Y_{it} = \alpha + \beta Post_{it} + \gamma Post_{it} \times Confed_{it} + \theta Post_{it} \times \Delta ST_i + \delta_{State} + \varepsilon_{it} \quad (5)$$

In equation (5), Y_{it} denotes various state *i*'s public-school indexes at year *t* (i.e., 1860 and 1870), including average school attendance rate, the number of public schools per capita, and the number of teachers per public school. $Post_{it}$ is a dummy variable that indicates whether the dependent variable comes from the 1870 census, and it is interacted with another the Confederate dummy ('*Confed*'). ΔST_i denotes the difference in the logarithm of public school income per capita and tax revenue per capita between 1860 and 1870. It was also interacted with '*Confed*'. We add state fixed effects to capture state characteristics. Thus, the variable of ΔST_i itself is omitted due to multicollinearity.

Panel A of Table 6 shows the results of baseline models that exclude the variable of $Post_{it} \times \Delta ST_i$ from equation (5). So the coefficient γ, which is reported in Panel A, measures how the dependent variable changed after the Civil War. First of all, model (1) uses school-attendance rates averaged from the 1860 and 1870 IPUMS samples across state. The estimated coefficient says that the postwar decline in school attendance was more substantial in the Confederate states. Its magnitude is almost the same as what was estimated in Table 2. Model (2) uses alternative school-attendance rates, which were adopted from the census tabulations. A substantial decline in school attendance is still estimated in Panel A. When we include the change of public school income and taxes in Panel B, it explains a large portion of the decline. Again, this result is similar to that from IPUMS samples.

[Insert Table 6 Here]

Now, we employ other public-school indexes as dependent variables in models (3)-(4). Model (3) of Panel A shows that the Confederate states seem to have had a smaller number of

public schools after the Civil War, but this is statistically insignificant. The change of public school information has a meaningful coefficient itself, but it does not have any meaningful influence on the change of number of schools in Panel B. Similarly model (4) shows that the number of teachers per school was not significantly different between two regions after the Civil War. Moreover, public school income is not associated with the supply of teacher per school.

The key implication found from the above exercise is that the substantial reduction of public school income in the postwar Confederacy does not itself account for the change of school quality or quality among public schools. This may suggest that Southern children did not attend schools due to other factors, such as a demand-side shock in the region after the War.

V. Testing School versus Labor-Market Factors in the Schooling Decision

V.A. A Model of Schooling: Supply vs. Demand Factors

In the remainder of this study, we adopt a more theoretically informed approach to decomposing the decline in schooling into supply vs. demand factors. In particular, we use a standard model of the time-in-school decision and discuss the implications of various shocks that might provoke students to leave school at an earlier age. In a simple way, we can model the time-in-school decision by decomposing marginal benefit into price versus quantity, i.e., $MB = P \times Q$. If a labor-market factor like a decline in demand for skill changes the MB, this will change the price (P), but not affect the quantity of human capital (Q); if the changing factor is the school quantity or quality, it will change the quality of human capital produced at fixed prices. Think of literacy, then, as an indicator of the quantity (again at fixed prices) of human capital produced. (Below we consider the "marginal literacy benefit", i.e., the gain in literacy for an additional unit of time in school). We have the same implication with what we get in the above. For example, the change of demand for skill does not affect literacy.

Alternatively, consider a textbook model in which individuals or parents choose children's time in school so as to maximize lifetime income. Let the benefits of schooling be $B(s)$ and the costs be $C(s)$, where s is time in school. The benefits B are the discounted sum of future earnings, and the costs C include direct cost of education and foregone income while in school (or opportunity cost). The usual assumptions are that the benefits of schooling decline with more time in school because of diminishing returns, and costs rise with the speed of human capital

accumulation. Thus, the marginal benefit curve (*MB* or dB/ds) is downward-sloping, and the marginal cost curve (*MC* or dC/ds) is upward-sloping, as described in Figure 3. A child should stay in school as long as the marginal benefit exceeds the marginal cost; an interior solution *s** will be determined when *MB* = *MC*.

[Insert Figure 3 Here]

It is useful to distinguish between two concepts of marginal benefits: gross versus net. The gross *MB* is simply the *MB*, the additional benefits that accrue (perhaps later in life) from the additional time in school. This is a relatively easy variable to measure, and we will use as a proxy the occupational income score or, as an intermediate proxy of human capital, literacy. We define the net MB as the marginal benefits net of marginal costs. This is a more difficult object to measure because it requires an accounting of a whole host of costs that are difficult to observe. Nevertheless the theory has the strongest prediction about the subject, namely that net *MB* equals zero (or *MB*=*MC*).

Suppose that the Civil War depressed the supply of schools in the former Confederate states for decades after the War. For example, a large decline in local wealth and tax base suddenly shrinks the Southern states' ability to provide schools, which could constrain students from attending as much school as is optimal. As shown in Figure 3 (a), this supply-side constraint would substantially reduce the school attendance in the Southern states. As a result, net and gross marginal benefit would increase.

On the other hand, the postwar sharp decline in wealth and taxes in the Confederate states could depress the instruction received per unit time in school and school quality. To see the subsequent effect, we modify the benefits of schooling and costs augmenting with school quality as follows: $B(s,q)$ and $C(s,q)$, where q is school quality. Additional assumption is that cross-partial derivatives are positive, i.e., $\frac{\partial^2 B}{\partial s \partial q} > 0$ and $\frac{\partial^2 C}{\partial s \partial q} > 0$. In other words, better school quality raises the return to time in school; higher school quality increases past skill attainment for each year of schooling, and so opportunity costs also rise in school quality. Then, as shown in Figure 3 (b), a decline in school quality would shift both the marginal benefit and cost curves downward, and so could cause a decline in average school attendance because the shift of *MB* curve would be more likely larger than that of *MC* curve. Consequently, the net marginal benefit

of schooling would remain unchanged as people adjust their optimal schooling decision, but the marginal benefit would drop.

Now suppose that the Civil War brought about a demand-side shock to the Confederate states rather than the supply-side shock in terms of school quantity or quality constraint. Especially, suppose that the demand for skilled workers decreased in the South after the War due to a decline in labor-market return to schooling. This will definitely shift the *MB* curve downward by lowering the benefits of schooling given time in school. In addition, the cost for an additional year of school will increase because the decline in the skilled-wage premium might increase the opportunity cost for additional schooling. However, such an upward shift of the *MC* curve depends on change of real unskilled wage. If the wage changes little, the *MC* curve would be qualitatively similar to the initial one. In Section V.D, we will provide a trend of real wage by occupation. It shows that the real wage of unskilled works increased by about 5% in the South in 1860-1870. Thus, we propose that the *MC* curve shifts upward, as seen in Figure 3 (C). Consequently, a large decline in average time in school would occur in the Confederate states. The net marginal benefit remains unchanged as people end up at the new equilibrium, s^{**}. Although this depends on the magnitude of the shift of *MC* curve, the gross marginal benefit would not deviate from the initial level much.

Table 7 summarizes the expected change of net and gross marginal benefit by each type of shock. As discussed above, we can test each hypothesis using occupational income score in adult and literacy per unit of time in school or per school attendance, which are proxy variables of the marginal benefit of schooling. Therefore, school-quantity constraint and school-quality decline would increase or decrease literacy per unit of time in school after the Civil War, respectively. If the demand for skills declined, two measures would remain unchanged or change in a small margin.

[Insert Table 7 Here]

In addition, the exposure to those potential supply and demand shocks would lead to differences in migration pattern in the long-run, as presented in Table 7. Although the migration depends on various conditions of the local labor market, it is thought that the school-quantity constraint and school-quality decline would not affect migration decisions among Southern students who attended school after the Civil War because this affects the marginal literacy benefit but not price variables. In contrast, if the demand for skill declined in the Confederacy after the War, it is

more likely that Southern skilled workers migrated out of the Confederate states for obtaining better economic opportunity, i.e., brain drain.[12] Although all the cohorts who attended schools before or after the Civil War were presumably exposed to the demand shock, the brain drain among the prewar cohorts in the South would be attenuated because they would have already made their location-specific investment. In the following subsections, we take these implications of the model to the three distinct sets of census data.

V. B. Measuring School Quality by Comparing Literacy and Time in School

In this subsection, we first consider the hypothesis that school quality collapsed in the former Confederacy after the War. School quality here is a productivity concept, measuring output per unit of input. The 19th-century censuses also contain information on literacy, which we use as a measure of output. The input we consider is time in school, which is constructed from the variable on school attendance used above. We combine these variables to ask whether literacy per time in school is changing from before to after the War. Specifically, our strategy is to measure how literacy rises with age (during school ages) and then correlate this with the accumulation of time in school.

One way to conceptualize this analysis is to think of the production function $F(X,T)$ of literacy where T is time in school and X is a vector of other inputs. The question is whether dF/dT is lower after the War, perhaps because of a lower provision of X.[13] We use the snapshot of school-aged children at each census year to form a pseudo-panel under the assumption that the cross-section in a given year is reasonably representative of the behavior of the panel at that point in time. (This is analogous to using a period life table to construct life expectancy when the

[12] There can be some exceptions. If local skill scarcity in the South drove up the return to skill, this could prevent brain drain and might even attract Northern skilled workers into the South. If the skills obtained in school were no longer useful for the urban economy or the frontier, the brain drain might decline.

[13] Note that this is a full derivative rather than a partial, and so should not be thought of as the marginal product of time in school vis-à-vis literacy. With each additional year that the student might be in school, he or she will receive additional inputs in terms of teacher time, use of facilities, etc. For example, if school attendance is expected to be higher, the school authority might hire more teachers or open more school facilities. Thus, there could be two channels by which this full derivative would be different in one regime versus another. This is easily seen by decomposing the full derivatives: $dF/dT = MPT|_X + dX/dT \times MPX$. The first term is the marginal product of time in school, holding fixed all other inputs. The second term incorporates how the other inputs respond to student continuing his studies multiplied by the marginal product of those other inputs. By inspection of this equation, we see that the full derivative could change if any of these three terms were to change.

cohort-based approach might be more appropriate. In this context, we believe that this assumption is less problematic because the time span of exposure covered is shorter.)

The concept of input in this case is the stock of time in school. This presents a measurement difficulty in that the variable on school attendance is a flow rather than a stock. None of the 19th-century censuses contain information on the stock of schooling, but rather the flow of school attendance. This is where the assumption of the pseudo-panel comes into play. We treat flows of school attendance across the observed school ages as if they come from a single cohort, as similarly done in Margo (1986). Then, we can accumulate these flows to estimate the year of schooling (T_i) at age i by summing all the school-attendance rates (*SA*) at ages between 5 and i: $T_{ik} = \sum_{j=5}^{i} SA_{jk}$, where k denotes the region.

A few comments about the sample employed are in order. First, we exclude ages older than 20 in this calculation because any additional school attendance above this cutoff is minimal and any changes in literacy across ages are most likely dominated by a cohort rather than age effects. Second, we do the postwar calculation using 1880 data to avoid using any cohorts whose school years might have been interrupted by having to fight in the War Between the States. (If anything, this exercise makes it appear as if school quality rose for whites after the War rather than fell when conducted using 1870 data.) Third, we take the cumulative sum (by age) of school attendance starting at age 5 because attendance is negligible at earlier ages and probably dominated by measurement error. Fourth and finally, enumerators in 1880 were instructed to record literacy for all those 10 years and above, and therefore our comparison of *F* and *T* starts with age 10, even though the stock of school is accumulated from earlier ages.

We can then estimate d*F*/d*T* using various methods comparing *F* with *T*, but the result is most easily seen graphically in Figure 4. The vertical axis is the literacy rate (fraction that can read and write) for each census/age/region cell. The horizontal axis is the imputed years of schooling cumulated across ages from the flow of school attendance. The "number" points are various ages in the former Confederate states in 1880. The squares are the same concept in Northern states in 1880. (To avoid confusion, the ages are not labeled for the Northern data, but note that the squares cover the same age range for the CSA.)

[Insert Figure 4 Here]

White students in the North are almost all literate after four years in school (seen in the leftmost square near the top of the graph, which corresponds to Northern children at age 10). In contrast, 30% of white students in the former Confederacy are still illiterate at a similar point in their progress through school (between ages 13 and 14). The dashed line is a quadratic fit and extrapolation using the 1880 Confederate data points. (The result of using a linear extrapolation is quite similar. As can be seen, there is not much curvature in the fitted line.) From this extrapolation, for example, it appears that Confederate white children would need approximately seven years of school to get to a literacy rate attained by Northerners after only four years in school.

The regional difference in the productivity of time in school could be due to a whole host of factors. Painting with a broad brush stroke, we could think of this as manifesting regional differences in school quality, which could include anything that reduces the quality of instruction, such as less skilled teachers, less time with the teachers (per student), poorer physical facilities, or lower-ability peers. Some of these differences can be seen in Table 6. In addition, the regional difference might be that Confederate white children had less time in school because either the term or the school day was shorter.

For the present study, the question is: Did time in school become less productive in producing literacy in the Confederate states after the War? An additional data complication is that literacy was only recorded for those aged 20 or older in antebellum censuses. We can nevertheless construct the imputed time in school and compare it with the literacy rate of those 20-24 years of age in 1860. This is the data point labeled "1860, Confederate States" in Figure 4. The 1880 extrapolation passes very close to the 1860 point. If 1880 time in school was less productive, we would expect a shallower slope and the extrapolation would instead pass below the 1860 point. Thus, the pattern of literacy and school attendance is not consistent with a decline of school quality. Instead, it is more consistent with school attendance declining for other reasons, moving along a stable literacy/schooling production function.

V.C. Evidence from Longitudinally Linked Censuses

By examining the relationship between adult outcomes (such as literacy and adult income) and school attendance in earlier ages, we can alternatively evaluate the above argument that the pattern of school attendance in the Confederacy was not consistent with a decline of school

quality in the region after the Civil War. In particular, it is necessary to compare the relationship between cohorts who were of school age before or after the War (denoted hereafter as the prewar and postwar cohorts, respectively). If the school quality in the former Confederacy had collapsed after the War, the postwar cohorts who attended schools would have been more likely illiterate or had a lower level of income in adulthood relative to their prewar counterparts.

On the other hand, the use of longitudinal samples can provide some evidence for the difference in the return to human capital across region and its impact. For example, if the return to human capital, which can be measured by the wage difference between skilled and unskilled workers, were lower in the former Confederate states than in Northern states, well-educated workers would be more likely to migrate out of the Confederacy in adulthood to obtain better opportunities. In addition, if the regional gap in the return to skill became bigger after the Civil War, the above pattern would be stronger among the postwar cohorts. Even though the prewar cohorts would face the same change in the postwar return to skill in the South, they would already have made location-specific investments that would reduce the probability of migration in response to price differentials. We test this hypothesis by comparing the migration pattern between the groups that attended school (potential skilled labors) or not in early ages, and between prewar and postwar cohorts. In Section V.C, we also seek direct evidence that supports this hypothesis by utilizing historical sources.

To conduct the work proposed above, we use IPUMS linked representative samples that link records from the 1880 complete-count database to 1% samples of the 1850 to 1930 censuses (Ruggles et al. 2010). Out of various pairs of years, we utilize two linked samples: 1860-1880 and 1870-1880. To focus on the relationship between school attendance and adult outcomes, we choose a sample of males aged 10 to 18 in the beginning year of the linked samples. Females were excluded because they participated less in the labor force in adulthood. We select 10 years as the youngest age because it was the lowest age for which a majority was not observed in a household with his father.[14] We choose 18 years as the oldest age because school-attendance rates for older ages were negligible. Therefore, the observations in the 1860-1880 linked data

[14] For example, in the 1870-1880 linked data, only 25% of males aged 10 in 1860 lived with fathers in 1870 (equivalently at age 20). The ratio jumped up above 30% for those aged 9 in 1860. The above criteria matters only for the 1870-1880 linked samples; most observations in the 1860-1880 linked samples did not live with fathers in 1880.

were in school before the Civil War (prewar cohort), and those in the 1870-1880 linked data were in school afterward (postwar cohort).

The regression analysis is based on the following equation:

$$
\begin{aligned}
Y_{ij} = {} & \alpha + \beta_1 SA_{ij} && \text{… [Single diff.]} \\
& + \gamma_1 SA_{ij} \times CF_{ij} + \gamma_2 CF_{ij} && \text{… [Double diff.]} \\
& + \theta_1 SA_{ij} \times CF_{ij} \times PW_{ij} + \theta_2 PW_{ij} + \theta_3 SA_{ij} \times PW_{ij} + \theta_4 CF_{ij} \times PW_{ij} && \text{… [Triple diff.]} \\
& + \delta_{AGE} + \delta_{AGE} \times \mathbf{K}_{ij} \times \Gamma + \varepsilon_{ij} && (6)
\end{aligned}
$$

In equation (6), Y_{ij} denotes adult outcome variables of individual *i* belonging to cohort *j*. Four types of variables are used as adult outcomes in 1880: dummy for being literate, logarithm of occupational income score, dummy that indicates whether he lived in the former Confederate states, and dummy for living in urban areas. '*SA*' denotes the dummy that shows whether he attended school in the beginning year of the linked samples (i.e., 1860 or 1870), which is the key explanatory variable in the regression. '*CF*' and '*PW*' are the dummy variables that show whether he lived in the former Confederate states during school ages and whether he is in the prewar cohort sample (i.e., the 1860-1880 linked samples), respectively. We include age dummies, and they are interacted with other control variables in the regression models denoted by a matrix '**K**'.

Our estimation strategy consists of three parts. First, the basic research design compares the adult outcomes depending on whether they were in school 10 to 18 years prior. Its specification is labeled by 'single difference' in equation (6) in which we include only the dummy of school attendance, age dummies, their interaction with the school-attendance dummy, and constant term. We run the regressions separately for four region-by-cohort samples. Second, we also compare this relationship across cohorts that were in school in 1870 versus in the antebellum period. This results in a simple difference-in-difference specification (labeled by 'double difference' in equation (6)) in which the first differences school attendance as a team and the second differences region of residence. Thus, we run regressions for a pooled-sample of prewar and postwar cohorts, but separately by region. Third, we use a triple-difference strategy in which the coefficients in double-difference specification are compared among those who were in the Confederate states in the prior census versus those in the North. All of equation (6) appears

in this specification, and a pooled sample including all the cohorts and regions will be employed in this case.

Results from the above estimation strategy are found in Table 8. In the table, we report only the coefficients for key variables (β_1, γ_1 and θ_1 in equation (6)) and their robust standard errors. According to models (1) and (2) of Panel A (single difference specification), the positive marginal benefit of schooling in terms of literacy and occupational income score is significant among both prewar and postwar cohorts in the former Confederate states. Panel B shows that the estimated marginal benefit of schooling is not statistically different between two cohorts in the Confederacy. The same implication is also found for Northern cohorts.[15] Panel C shows that the regional difference of the estimated coefficient in Panel B is not statistically significant.

[Insert Table 8 Here]

In summary, Southern students' school attendance predicts literacy and occupational status similarly regardless of whether they were in school prior to or after the War. This result is inconsistent with the hypothesis that school quality in the Confederacy declined after the Civil War, but consistent with what we found in the previous subsection using a pseudo-panel approach. It is again suggested that the postbellum decline in school attendance among Southern white students was not caused by a reduction in the supply of school (in terms of quality).

On the other hand, models (3) and (4) in Table 8 use migration as the adult outcome. Panel A shows that the postwar southern cohorts who attended school in 1870 more likely migrated into the Northern states by 1880 and lived in urban areas in 1880 than those who did not attend school. But this pattern is not found among the prewar cohorts who lived in the Confederate states prior to the War. This difference between prewar and postwar Southern cohorts is estimated significantly in Panel B. But the cross-cohort difference is not found among those who lived in the North in the beginning year. Panel C shows that the triple difference by region is statistically significant for whether they lived in the Confederacy in adulthood, but insignificant for whether they lived in an urban area.

Finally, some may have concerns that the cutoff ages in the beginning year (10-18) are too broad, and that two census-linked samples might not be comparable because each cohort's

[15] The test for Northern states is less informative because literacy rates were so high for those cohorts. We report it nonetheless here for completeness.

adult outcomes are evaluated at different ages. To deal with them, we conduct additional exercise in models (5) and (6). Using the dummy of being in the Confederate states in the ending year, we confine the observations to those aged 13-18 in model (5). Moreover, for model (6), we use newly constructed 1850-1860 census-linked samples for alternative prewar cohort sample.[16] Then, we find the similar implication that Southerners who attended schools in the Confederacy prior to the War more likely migrated out of the area.

The above result shows a brain drain occurring in the former Confederate states after the Civil War. This is indicative of the Northern states having better economic opportunities or a higher result to skill than did the Confederate states, which could induce well-educated Southern students to the North. The above result also suggests that the regional disparity in the return to skill was substantial after the War, but might be relatively small prior to the War. In this case, the lowered skill premium would lead to less schooling in the South in the long run.

V.D. Evidence on the Return to Skill

In the following, we provide direct evidence that a brain drain potentially stemmed from a decline in the return to skill in the Confederate states after the Civil War. This evidence supports that the postwar decline in southern school-attendance rate was caused by demand side rather than by supply side. Furthermore, the evidence for literacy indicates that what changed with the War was the price of skill in the labor market rather than the quality of the education received prior to the market.

For evidence, we compare the real wage by occupation across region and its change over time. In particular, the return to skill is measured by the difference in real wage between engineers (skilled workers) and common laborers (unskilled workers). We use regional wages series for the two types of works constructed by Coehlo and Shepherd (1976) who used data from the Weeks (1886) report.

The upper panel of Figure 5 first graphically presents the percent change of real wages by region and occupation in 1860-1870. In terms of US average, the wage of skilled workers decreased by 1.6% in the period, and that of unskilled workers increased by 1.8%. This small

[16] For the new data, we tried to link all of the aged 13-18 males from selected Georgia counties in the 1850 census manuscript into the 1860 census manuscript. We successfully linked 5,695 individuals.

disparity in wage growth by occupation is also found in New England and Mid-Atlantic regions. But the East-North Central region observed a large gap as unskilled workers' average wage substantially increased by 13.4% and that of skilled workers decreased by 2.7%. The West-North Central region also observed such a large gap because skilled workers' average wage declined by 8.7%. The most astonishing result is found in the East-South Central region. The average wage of skilled workers plummeted by 33.5%, whereas that of unskilled workers increased by 4.6%. In summary, the return to skill measured as real wage by skill-specific occupation substantially declined in the Southern states.[17]

[Insert Figure 5 Here]

In the lower panel of Figure 5, we examine the trend of skilled wage premium over the period of 1857-1880.[18] The graph presents the long-term trends for two regions: East-South Central region including West Virginia, and the region comprising states located in the east of the Mississippi and the north of the Ohio or Potomac rivers.[19] For the graph, we first calculated the wage of engineers relative to common laborers for each year, took a natural logarithm, and then normalized the series to zero in 1860.

The graph shows that skilled-wage premium was similar in the two regions prior to 1855. However, it plummeted with the outbreak of the War, and was cut in half by 1868.[20] Although the skilled-wage premium in the Southern region rebounded throughout the early 1870s, it had been at about 20% below the 1860 level until 1880. Compared with the South, the Northern

[17] Margo (2004) documents that the Civil War led to a dramatic divergence in the regional structure of wages. Using historical data on wage—other than Coelho and Shepherd (1976)—for common laborers, he shows that wages in the South Atlantic and South Central states relative to the North fell sharply after the War and persisted until the early twentieth century, which is consistent with Rosenbloom (1990). He argues that the divergence is consistent with a sharp drop in labor productivity in Southern agriculture rather than the changing racial composition of the Southern wage labor force due to emancipation after the War. On the other hand, he uses carpenters' average wage for skilled labor, and its trend is a bit different with what we show above. But the skill that a carpenter needed would have been less likely acquired in school, relative to the skills used by engineers.

[18] Note that although real wages can change as they are affected by the price change of other resources such as land, relative wage is not affected in this way. For whatever may be the correct deflator, our concern is the relative wage, for which the deflator would be cancelled out.

[19] The second region includes three census divisions: New England, Mid-Atlantic, and East-North Central areas. We calculated unweighted average wage across the three census divisions.

[20] The relative wage decreased from 2.41 in 1860 to 1.34 in 1868.

region shows a relative stable trend of wage premium over time although the region also had a slight drop right after the Civil War. The trend clearly shows that the South provided lower returns to skilled workers after the Civil War, particularly in 1860-1870. This can support the migration pattern estimated in the previous subsection. In addition, the lowered return to skill in the South would lead to less demand for schooling after the Civil War. The regional disparity continued until 1880 also may account for why the Southern school-attendance rates did not return to the antebellum level by 1880, which is shown in Figure 1.[21]

Under the assumption that the change in school attendance is coming only through demand-side factors, we can use these estimates to compute an elasticity of time in school with respect to the labor-market return to skill. Because the other factors that might have shifted attendance after the war probably also operated in the negative direction, this should be considered an overestimate of the true magnitude. The school attendance rate of the former Confederacy declined from 0.4122 in 1860 to 0.2602 in 1870; the school premium (actual relative wage) declined from 2.41 to 1.54 during the same period. Comparing these ΔQ and ΔP yields an elasticity of 1.01. Alternatively, we could perform this calculation using the changes in 1860-1880.[22] Then, the elasticity is between 0.57-0.64. These numbers would inflate to 1.17-1.29 and 1.22-1.31, respectively, if we compared Southern attendance to its antebellum trend instead.[23] It is also possible that people were forecasting some attenuation of the skill premium to its prewar levels, which would in effect reduce the ΔP in the above calculations and result in a larger elasticity.

Finally, although the evidence in this section for skilled workers pertains to manufacturing, there is reason to suspect that the postbellum change in Southern agriculture also would have reduced the return to skill. In antebellum days, plantations could be very large and

[21] Some readers might ask how it could be possible for such dramatic differentials in the price of skill to persist. within the country for so many decades. Various studies (e.g., Rosenbloom (1990)) have noted that the United States was not an integrated labor market during this period. Further, sectionalist tensions were high in this period (e.g., The War Between the States), which must have dampened the willingness to migrate to a different region.

[22] In the case, the decline of school attendance rate is from 0.4122 to 0.3483, and that of relative wage is from 2.41 to 1.83.

[23] That is, we use the counterfactual school-attendance rates in 1870 and 1880, which is estimated from the 1850-1860 antebellum trend, instead of the actual rate in 1860 as the base-year value. The counterfactual values are discussed in Section III.B, and presented in the upper panel of Figure 1: 0.4522 for 1870 and 0.4943 for 1880.

often self-sufficient on many dimensions. This required complicated coordination within the plantation and often various levels of management. In contrast, the system of tenancy (renters and sharecroppers) that emerged after the Civil War would have removed some of these layers of management and reduced the need for coordination. The increasing share of the marginal output that went to smaller farmers (100% in the case of renters and usually 50% for sharecroppers) in effect devolved much of the decision-making from the land owners to those working the land with their own hands. The plantation owner would often provide services to his tenants, but this represented a much more hands-off role than in the antebellum years.

VI. Conclusion

Although the historical gap in schooling between the U.S. North and South has received substantial attention in the literature, few studies have noted that much of this gap emerged with the Civil War, and still fewer studies attempt a quantitative explanation of why Southern white schooling fell behind so drastically in this period. In this study, we show that school attendance among white children substantially declined in the former Confederate states after the Civil War, whereas the Northern states experienced only a slight decline. By examining observable variables, we show that postwar decline of individual and local wealth and tax-based public school income considerably account for the decline. However, our results do not suggest that the Confederacy experienced the postbellum decline in white school attendance because low public support constrained time in school in the region. Moreover, we show that the quality of school (another supply-side factor) did not change after the Civil War by examining the trend of marginal literacy benefit of time in school.

Rather than those supply-side factors, this study proposes the significance of the change in return to schooling, i.e., demand-side factor. We provide some evidence that the return to education, measured by wage premium for skilled workers, substantially declined in the South after the Civil War. In response, well-educated Southerners should have migrated out of the South to seek better opportunities. We support this hypothesis using longitudinally-linked census samples.

References

Akresh, Richard, and Damien de Walque. 2011. "Armed Conflict and Schooling: Evidence from the 1994 Rwandan Genocide." World Bank Policy Research Working Paper 4606. February.

Barreca, Alan. 2010. "The Long-Term Economic Impact of In Utero and Postnatal Exposure to Malaria." *Journal of Human Resources* no. 45 (4):865-892.

Bleakley, C. Hoyt, and Sok Chul Hong. 2012. "Estimation of Historical Hookworm Ecology Index." *Working Paper*.

Bleakley, Hoyt. 2007. "Disease and Development: Evidence from Hookworm Eradication in the American South." *Quarterly Journal of Economics* no. 122 (1):73-117.

Brinkley, Garland. 1997. "The Decline in Southern Agricultural Output, 1860-1880." *Journal of Economic History* no. 57 (1):116-138.

Coelho, Philip, and James Shepherd. 1976. "Regional Differences in Real Wages: The United States, 1851- 1880." *Explorations in Economic History* no. 13 (2):203-230.

Collins, William. 2007. "Education, Migration, and Regional Wage Convergence in U.S. History." In *The New Comparative Economic History*, edited by Timothy J. Hatton, Kevin H. O'Rourke and Alan M. Taylor, 165-191. Cambridge: The MIT Press.

Collins, William, and Robert Margo. 2006. "Historical Perspectives on Racial Differences in Schooling in the United States." In *Handbook of the Economics of Education*, edited by Eric Hanushek and Finis Welch, 107-154. Amsterdam: North Holland.

Cremin, Lawrence A. 1965. *Economic Influences upon Educational Progress in the United States, 1820-1850*. New York: Teachers College Press.

Cubberley, Ellwood P. 1919. *Public Education in the United States: A Study and Interpretation of American Educational History*. Boston: Houghton Mifflin.

Dabney, Charles W. 1936. *Universal Education in the South: Vol. 1, From the Beginning to 1900*. Chapel Hill: The University of North Carolina Press.

de Carvalho Filho, Irineu, and Renato P. Colistete. 2010. "Education Performance: Was It All Determined 100 Years Ago? Evidence From São Paulo, Brazil."Online at http://mpra.ub.uni-muenchen.de/24494/ MPRA Paper No. 24494, posted 18. August 2010 / 22:06.

Drake, William E. 1955. *The American School in Transition*. Englewood Cliffs: Prentice-Hall.

Go, Sun, and Peter Lindert. 2010. "The Uneven Rise of American Public Schools to 1850." *Journal of Economic History* no. 70 (1):1-26.

Goldin, Claudia, and Lawrence F. Katz. 2010. *The Race between Education and Technology*. Cambridge: Harvard University Press.

Haines, M. R. 2005. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. Interuniversity Consortium for Political and Social Research (ICPSR) no. 2896.

Hong, Sok Chul. 2007. "The Burden of Early Exposure to Malaria in the United States, 1850-1860: Malnutrition and Immune Disorders." *Journal of Economic History* no. 67 (4):1001-1035.

Knight, Edgar W. 1922. *Public Education in the South*. Boston: Ginn and Company.

Knight, Edgar W. 1929. *Education in the United States*. Boston: Ginn and Company.

Lindert, Peter. 2009. *Revealing Failures in the History of School Finance*. NBER Working Paper No.15491.

Margo, Robert. 1986. "Race, Educational Attainment, and the 1940 Census." *Journal of Economic History* no. 46 (1):189-198.

Margo, Robert. 1990. *Race and Schooling in the South, 1880-1950: An Economic History*. Chicago: The University of Chicago Press.

Margo, Robert. 2004. "The North-South Wage Gap, Before and After the Civil War." In *Slavery in the Development of the Americas*, edited by D. Eltis, F. Lewis and K. Sokoloff, 324-352. New York: Cambridge University Press.

Pritchett, Jonathan B. 1989. "The Burden of Negro Schooling: Tax Incidence and Racial Redistribution in Postbellum North Carolina." *The Journal of Economic History* no. 49 (4):966-973.

Rosenbloom, Joshua. 1990. "One Market or Many? Labor Market Integration in the Late Nineteenth-Century United States." *Journal of Economic History* no. 50 (1):85-107.

Ruggles, Steven, Trent Alexander, Katie Genadek, Ronald Goeken, Matthew Schroeder, and Matthew Sobek. 2010. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota.

Sacerdote, Bruce. 2005. "Slavery and the Intergenerational Transmission of Human Capital." *Review of Economics and Statistics* no. 87 (2):217-234.

Stein, Z., M. Susser, G. Saenger, and F Marolla. 1975. *Famine and Human Development: The Dutch Hunger Winter of 1944-1945*. New York: Oxford University Press.

Weeks, Joseph D. 1886. *Report on the Statistics of Wages in Manufacturing Industries with Supplementary Reports*. Washington, D.C.: U.S. Government Printing Office.

Wright, Gavin. 1986. *Old South, New South*. New York: Basic Books.

Table 1. Summary Statistics

Variables	Confederate states		Non-confederate states (Excluding Border States)	
	Prewar (1860)	Postwar (1870)	Prewar (1860)	Postwar (1870)
Panel A: School attendance rate				
age 5-20	0.4122	0.2602	0.6279	0.6030
age 5-12	0.4321	0.2640	0.7395	0.7038
age 13-20	0.3859	0.2557	0.4893	0.4739
Boys	0.4388	0.2654	0.6515	0.6147
Girls	0.3854	0.2550	0.6040	0.5912
Average age	11.72	12.10	11.87	11.86
Ratio of boys	0.5029	0.5016	0.5039	0.5004
Panel B: Within-sample individual and family characteristics				
Family structure				
Ratio of samples without father	0.2096	0.2679	0.1946	0.1857
Ratio of samples with step father	0.0165	0.0198	0.0181	0.0164
Household head's				
Value of real property (10K $)	0.2799	0.1385	0.2382	0.3344
Value of personal property (10K $)	0.4000	0.0639	0.0906	0.1331
Literacy rate	0.8301	0.7597	0.9325	0.9067
Number of children in Household	9.53	8.80	8.50	8.20
Observations	16,568	17,972	63,465	79,200
Panel C: County-group characteristics - SEA(state economic area)				
Wealth per capita ($)	965.0	308.5	676.8	799.4
Farm output value per capita ($)	83.42	59.76	62.31	71.09
Ratio of slave population in 1860	0.3682	NA	0.0000	NA
Number of farms 500 acres +	29.93	14.96	5.76	9.40
Population density (per acre)	11.85	14.37	32.05	41.41
Malaria index	0.3386		0.1449	
Hookworm index	0.3297		0.0751	
Number of SEA	105	105	194	194
Panel D: State characteristics				
Public school income per capita ($)	0.4828	0.4656	1.1419	2.0619
Divided by school attendance rate	1.1712	1.7890	1.8186	3.4196
Tax revenue per capita ($)	2.3532	4.0062	3.2141	8.6395
Number of states	9	9	18	18

Note: The table shows the sample mean of each variable by year and type of state (i.e. Confederate states or not). We excluded the observations in the Border states. Panels A and B report statistics calculated from IPUMS. In Panel C, wealth, farm output population density and ratio of slave population reports are calculated from ICPSR #2896; malaria and hookworm indexes are estimated as Hong (2007) and Bleakley and Hong (2012), respectively. We converted county-level variables into the values at the SEA level using a within-SEA county-area weighted average calculation. Panel D reports state statistics obtained from ICPSR #2896. The value of wealth, public school income and taxes is in constant 1870 dollars. The converting deflator is 5.39 for 1860 and 7.20 for 1870 (*source*: measuringworth.com).

Table 2. Estimates of Reduction in School Attendance in Confederate and Slave States after the Civil War

Dependent variable: Dummy=1 if the sample attended school within the past year.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Without FE		SEA FE	Baseline 1	Excluding frontier states	Age≤12	Age>12	Boys	Girls
Fixed effects	NO	NO	SEA	SEA	SEA	SEA	SEA	SEA	SEA
Controlling for Age and Sex	NO	NO	NO	YES	YES	YES	YES	YES	YES
Panel A: Estimating Reduction of School Attendance in the Confederate States									
Postwar	-0.0499***	-0.0250***	-0.0184***	-0.0312***	-0.0380***	-0.0407***	-0.0205***	-0.0425***	-0.0199***
	(0.0075)	(0.0066)	(0.0061)	(0.0057)	(0.0059)	(0.0053)	(0.0061)	(0.0053)	(0.0053)
Confederate States	-0.2826***	-0.2157***							
	(0.0116)	(0.0148)							
Postwar × Confederate States		-0.1271***	-0.1344***	-0.1296***	-0.1486***	-0.1533***	-0.1102***	-0.1355***	-0.1230***
		(0.0166)	(0.0164)	(0.0161)	(0.0139)	(0.0131)	(0.0146)	(0.0137)	(0.0134)
R-squared or Adjusted R-squared	0.0527	0.0553	0.0754	0.2589	0.2638	0.2532	0.2469	0.2389	0.2779
Sample size	177,205	177,205	177,205	177,205	160,486	98,869	78,336	88,960	88,245
Panel B: Estimating Reduction of School Attendance in the 1861 Slave States									
Postwar	-0.0457***	-0.0250***	-0.0184***	-0.0313***	-0.0381***	-0.0410***	-0.0204***	-0.0426***	-0.0200***
	(0.0072)	(0.0066)	(0.0061)	(0.0057)	(0.0059)	(0.0053)	(0.0061)	(0.0053)	(0.0053)
Slave States	-0.2350***	-0.1949***							
	(0.0146)	(0.0147)							
Postwar × Slave States		-0.0738***	-0.0902***	-0.0874***	-0.1246***	-0.1089***	-0.0677***	-0.0852***	-0.0891***
		(0.0180)	(0.0165)	(0.0160)	(0.0135)	(0.0134)	(0.0127)	(0.0132)	(0.0125)
R-squared or Adjusted R-squared	0.0467	0.0478	0.0725	0.2512	0.2601	0.2515	0.2394	0.2312	0.2701
Sample size	197,906	197,906	197,906	197,906	171,261	110,598	87,308	99,573	98,333

Note: Each regression uses white populations aged between 5 and 20 who are found in the 1860 and 1870 IPUMS datasets. Models (1)-(2) do not include fixed effects; the others use SEA(state economic area) fixed effects and their standard errors are clustered in SEA level. Model (4) where we control for dummies of age and sex and their interactions is the baseline regression, per equation (1) in text. Model (5) excludes frontier states that are located to the west of Mississippi river. Models (6)-(9) run the baseline regressions for each age and gender group. Panel A estimates the reduction of school attendance in the confederate states, in which we exclude five Border states; Panel B does in the slave states in 1861, in which we include the Border states. Models (1)-(2) report R-squared; the others report adjusted R-squared. A single asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%.

Table 3. Estimates of the Effect of Individual, Household-Head and Family Characteristics on School Attendance and its Postwar Reduction in Confederate States

Dependent variable: Dummy=1 if the sample attended school within the past year.

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	Baseline 1	Presence of father	HH's wealth	HH's literacy	Num. of children	Baseline 2
Postwar × Confederate States	-0.1296***	-0.1218***	-0.1138***	-0.1261***	-0.1259***	-0.1045***
	(0.0161)	(0.0160)	(0.0161)	(0.0158)	(0.0159)	(0.0157)
No father		-0.1313***				-0.1099***
		(0.0045)				(0.0046)
Step father		-0.0740***				-0.0715***
		(0.0101)				(0.0101)
Household head						
Real estate wealth			0.0382***			0.0334***
			(0.0051)			(0.0046)
(Real estate wealth - mean)^2			-0.0010***			-0.0009***
			(0.0002)			(0.0002)
Personal property wealth			0.0226***			0.0205***
			(0.0057)			(0.0052)
(Personal property wealth - mean)^2			-0.0006***			-0.0005***
			(0.0002)			(0.0002)
Literate				0.0986***		0.0837***
				(0.0068)		(0.0067)
Number of own children					0.0081***	0.0032***
					(0.0004)	(0.0004)
Adj. R-squared	0.2589	0.2693	0.2625	0.2623	0.2640	0.2756

Note: The number of observations is 177,205. We add individual, household-head and family variables to the baseline model in panel A of Table 2 and in equation (1). Each regression uses SEA fixed effects; standard errors are clustered in SEA level. The variables of no father, step father and literate are dummies that indicate whether each sample had the specified characteristics, respectively. The value of household head's wealth (10K $) is in constant 1870 dollars; the converting deflator is 5.39 for 1860 and 7.20 for 1870 (source: measuringworth.com). A single asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%.

Table 4. Estimates of the Effect of County-Group Characteristics on School Attendance and its Postwar Reduction in the Confederate States

Dependent variable: Dummy=1 if the sample attended school within the past year.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Baseline 1	Baseline 2	Wealth	Farm output value	Slaves	Plantation	Population density	Disease Ecology	Baseline 3
Individual and family variables	NO	YES	YES	YES	YES	YES	YES	YES	YES
Postwar × Confederate States	-0.1296***	-0.1045***	-0.0576**	-0.0964***	-0.0820**	-0.0881***	-0.0997***	-0.1221***	-0.0643
	(0.0161)	(0.0157)	(0.0245)	(0.0188)	(0.0400)	(0.0201)	(0.0156)	(0.0248)	(0.0415)
County group variables									
Postwar × Δln(wealth per capita)			0.0334***						0.0334***
			(0.0104)						(0.0125)
Postwar × Δln(farm output value per capita)				0.0260					0.0456*
				(0.0250)					(0.0266)
Postwar × Slaves per capita in 1860					-0.0666				0.0228
					(0.0888)				(0.1084)
Postwar × Δ # of farms with 500 acres +						0.0010**			0.0010*
						(0.0005)			(0.0006)
Postwar × Population density in 1860							0.0112**		0.0016
							(0.0051)		(0.0049)
Postwar × Malaria index								0.0343	0.0957
								(0.0728)	(0.0811)
Postwar × Hookworm index								0.0634	0.0716*
								(0.0432)	(0.0418)
Adj. R-squared	0.2589	0.2756	0.2758	0.2757	0.2757	0.2757	0.2757	0.2757	0.2761

Note: The number of observations is 177,205. Models (1) and (2) are adopted from Table 2 (model 4 panel A) and Table 3 (model 6), respectively. Wealth per capita, farm output value, population density (per square mile divided by 100), slaves per capita and number of farms with more than 500 acres were obtained from ICSPR #2896 and converted to SEA-level values considering county-boundary changes and using population as weight. Malaria and hookworm indexes were estimated by Hong (2007) and Bleakley and Hong (2012), respectively. Each regression uses SEA fixed effects; standard errors are clustered in SEA level. A single asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%.

Table 5. Estimates of the Effect of State Public School Income and Taxes

Dependent variable: Dummy=1 if the sample attended school within the past year.

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline 1	Baseline 2	Baseline 3	Public school income		Tax
Individual and family variables	NO	YES	YES	YES	YES	YES
County group variables	NO	NO	YES	NO	YES	YES
Postwar × Confederate States	-0.1296***	-0.1045***	-0.0643	-0.0547**	-0.0228	-0.0280
	(0.0161)	(0.0157)	(0.0415)	(0.0277)	(0.0480)	(0.0488)
State Variables						
Postwar × Δln(public school income per capita)				0.0485***	0.0519***	0.0545***
				(0.0179)	(0.0171)	(0.0171)
Postwar × Δln(tax per capita)						-0.0211
						(0.0204)
Adj. R-squared	0.2589	0.2756	0.2761	0.2760	0.2764	0.2764

Note: The number of observations is 177,205. For comparison, models (1), (2) and (3) are adopted, respectively, from models (1), (2) and (9) of Table 4. The state-level value of public school income per capita and tax amount per capita is in constant 1870 dollars; the converting deflator is 5.39 for 1860 and 7.20 for 1870 (*source*: measuringworth.com). Each regression uses SEA fixed effects; standard errors are clustered in SEA level. A single asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%.

Table 6. Public School Income and Postwar Change in Educational Indexes in the Confederate States: State-level Analysis

	Dependent variables			
	State school rate estimated with IPUMS	State school rate reported in the census tabulations	Num. of public schools per 1000 populations	Num. of teacher per school
	(1)	(2)	(3)	(4)
Panel A: Estimating the Reduction in Confederate States				
Postwar × Confederate States	-0.1299***	-0.1532***	-0.5527	0.0443
	(0.0263)	(0.0339)	(0.3959)	(0.1244)
Adj. R-squared	0.9118	0.8978	0.9027	0.5032
Panel B: Estimating the Role of Public School Income				
Postwar × Confederate States	-0.0905**	-0.0935**	0.2269	-0.0660
	(0.0386)	(0.0366)	(0.4286)	(0.1579)
Postwar × Δln(public school income per capita)	0.0459*	0.0877***	1.0419***	-0.0447
	(0.0266)	(0.0257)	(0.3005)	(0.1107)
Postwar × Δln(tax per capita)	-0.0231	-0.0284	-0.1441	-0.2059
	(0.0384)	(0.0318)	(0.3726)	(0.1372)
Adj. R-squared	0.9156	0.9273	0.9306	0.5129

Note: Each regression uses two-year panel data for 27 states. School attendance rate in model (1) is estimated from IPUMS; that in model (2) and dependent variables in models (3)-(7) are calculated from ICPSR #2896. In model (1), the number of within-state samples in IPUMS is used as weights in regressions. The state-level value of public school income and taxes per capital is in constant 1870 dollars; the converting deflator is 5.39 for 1860 and 7.20 for 1870 (source: measuringworth.com). Each regression includes state fixed effects. A single asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%.

Table 7. Expected Change in Marginal Benefits of Schooling and Migration by Type of Shock

Type of shock		Marginal benefits (*MB*) of time in school			Brain drain, by cohort	
		Net *MB*	Gross *MB*	Literacy per unit of time in school	Postwar cohort	Prewar cohort
(a)	School-supply constraint	up	up	up	no change	not exposed
(b)	School-quality decline	no change	down	down	no change	not exposed
(c)	Decline in demand for skill	no change	similar	similar	increases	attenuated

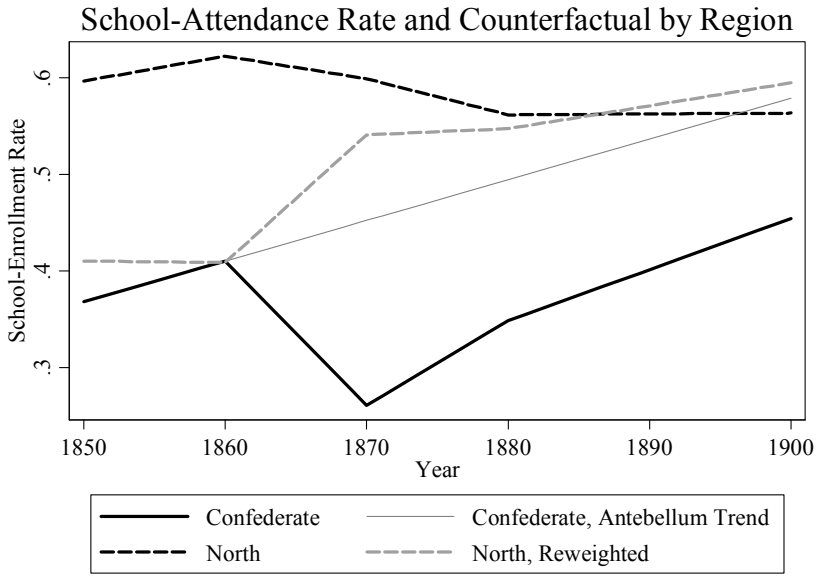
Note: Each shock is discussed in text and Figure 3. See text for detailed explanations.

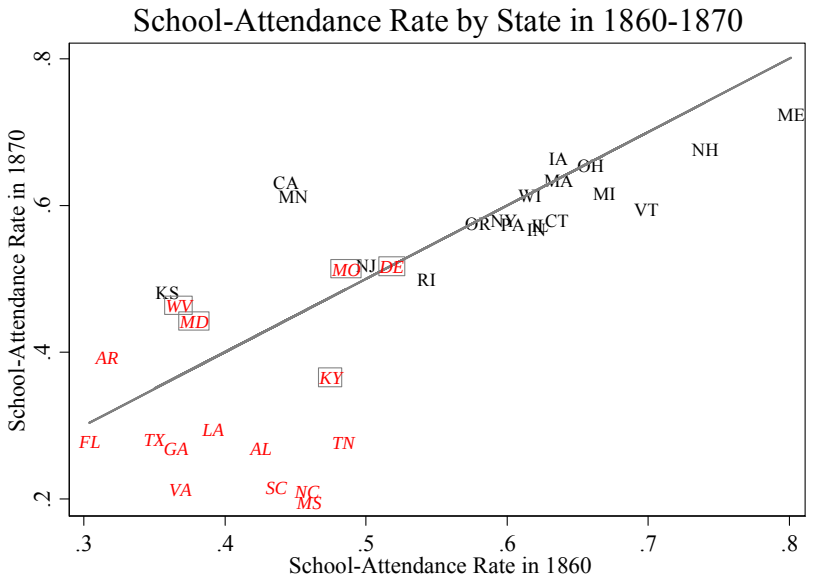
Table 8. Estimating the Relationship between School Attendance and Adult Outcomes Using Linked Census Data

						Living in Confederate states	
		Literacy (1)	ln(Occupation income score) (2)	Living in Confederate states (3)	Living in urban area (4)	Different age group (5)	Alternative prewar cohort (GA 1850-60 linked sample) (6)
Prewar cohort sample		1860-80	1860-80	1860-80	1860-80	1860-80	1850-60
Postwar cohort sample		1870-80	1870-80	1870-80	1870-80	1870-80	1870-80
Ages in the beginning year		10-18	10-18	10-18	10-18	13-18	13-18
State FE		Yes	Yes	Yes	Yes	Yes	No
Panel A: Single Differences by School Attendance at Initial Year							
Prewar	Confederate	0.1493***	0.0789*	0.0256	0.0331	0.0096	-0.0035
		(0.0370)	(0.0438)	(0.0240)	(0.0249)	(0.0279)	(0.0031)
Postwar	Confederate	0.1180***	0.0965**	-0.0303***	0.0903***	-0.0319**	-0.0296**
		(0.0194)	(0.0398)	(0.0094)	(0.0190)	(0.0147)	(0.0135)
Panel B: Double Differences by School Attendance × Postwar Cohort							
Both	Confederate	-0.0357	0.0295	-0.0583**	0.0576*	-0.0430	-0.0251*
		(0.0424)	(0.0587)	(0.0263)	(0.0311)	(0.0309)	(0.0137)
Both	Non-Confed.	0.0056	-0.0457	0.0073	0.0139	0.0105	
		(0.0133)	(0.0375)	(0.0077)	(0.0325)	(0.0100)	
Panel C: Triple Differences by School Attendance × Postwar Cohort× Confederate							
Both	Both	-0.0413	0.0752	-0.0656**	0.0437	-0.0535*	
		(0.0443)	(0.0695)	(0.0273)	(0.0449)	(0.0323)	

Note: We use two IPUMS linked datasets. One is the 1860-1880 linked IUPMS (prewar cohort sample, n=1,368) and another is the 1870-1880 linked IPUMS (postwar cohort sample, n=2,900). Of observations in each sample, we choose males aged 10 to 18 in the beginning year of the linked samples. In addition, we linked all of the aged 13-18 males from selected Georgia counties in the 1850 census manuscript into the 1860 census manuscript, which is named as "GA 1850-1860 linked sample" (alternative prewar cohort sample, n=5,695). Each regression is based on equation (6), depending on single, double, or triple-difference specification. Four types of dependent variables, listed in headings, are used: whether he was literate in 1880, the log value of occupational income score in 1880, whether he lived in confederates states in 1880 or 1860 for GA 1850-1860 linked sample, and whether he lived in urban areas in 1880. Panel A estimates the effects of school attendance in confederate states on adult outcomes by cohort. It controls for age dummies. Panel B is based on a difference-in-difference estimation. Each cell reports the coefficient of school-attendance dummy interacted with an indicator of young cohort, i.e. the 1870-1880 linked samples. We also control for age dummies, respectively, interacted with school attendance dummy and young-cohort dummy. Panel C uses triple differences by pooling two regions. The coefficient of school-attendance dummy interacted with cohort- and region-dummies is reported. The control variables used for Panel B and their interactions with confederate dummy are included. All the regressions use initial-state fixed effects, except model (6). Robust standard errors are reported in parenthesis. Single asterisk denotes statistical significance at the 90% level of confidence, double 95%, triple 99%.

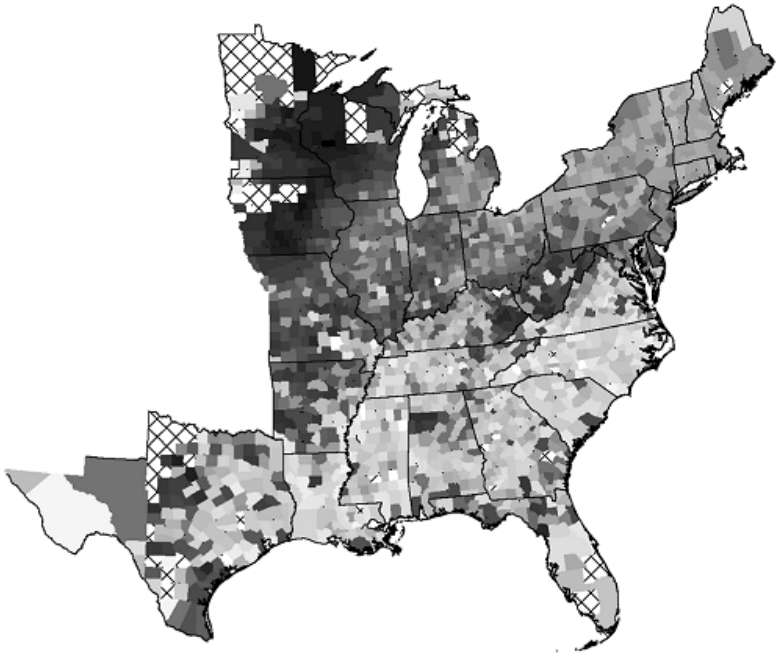
Figure 1. Trends of School-Attendance Rate among White Populations Aged 5 to 20 by Region and State





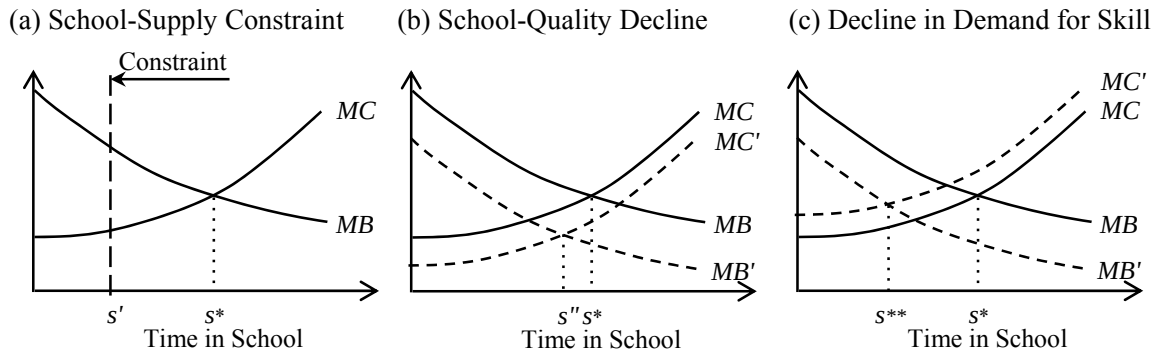
Note: Out of the 1850-1900 IPUMS samples, we use white populations aged between 5 and 20 to calculate school-attendance rate by region, state and census year. Details for counterfactual models are discussed in text. In the lower panel, the line is a 45-degree line; those in italic denote the Southern slave states and those in box are the Border States.

Figure 2. Change in School-Attendance Rates, 1850-1870, by County



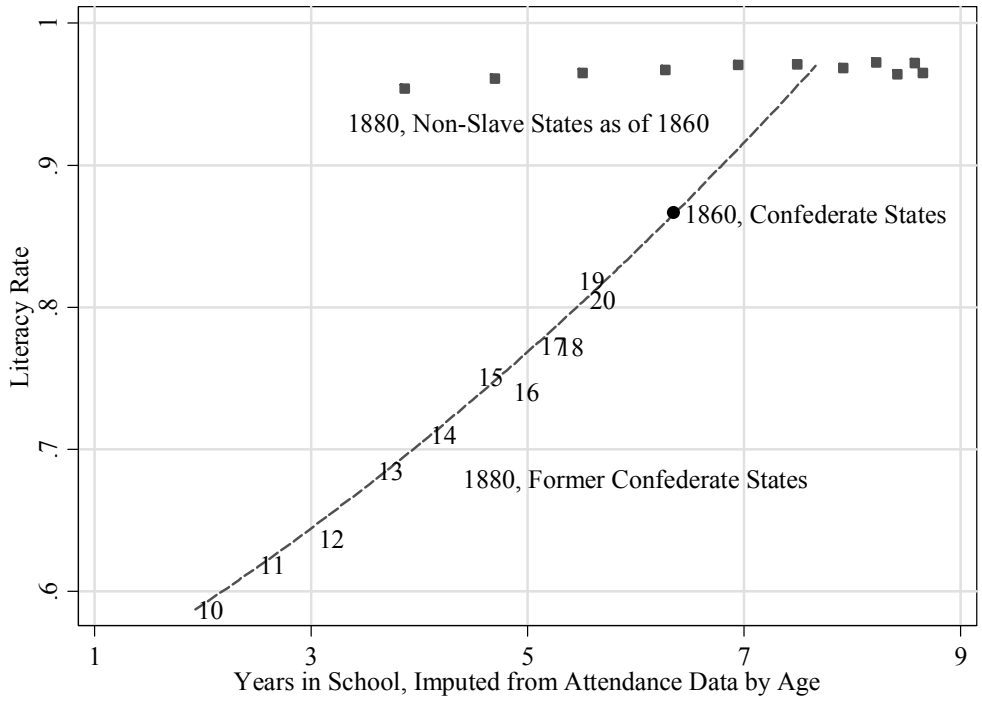
Note: This map presents the changes in white school attendance rates between 1850 and 1870. Note that the antebellum data are from 1850 because of data availability for 1860. We calculated the school attendance rates as the number of total whites in schools divided by white populations aged 5-19 for 1850 and aged 5-18 for 1870. Each variable was obtained from ICPSR #2896 dataset, which provides the published census tabulations. The shades of gray on the 1870-county map measure the increase in school attendance, in natural logarithms, between 1850 and 1870. This is therefore a percentage rather than a level change. Darker colors indicate greater proportional increases. Areas with the hatching have no data in one of the periods. We calculate the difference across time by constructing a raster that encodes the school-attendance rate in the county for 1850. The resolution of the raster is set to 1 km squares, which is much smaller than a typical county. We then take averages of this 1850 raster over the 1870 county boundaries to harmonize the counties across years.

Figure 3. Expected Time-in-School Decision by Supply and Demand Shocks in the Confederate States after the Civil War



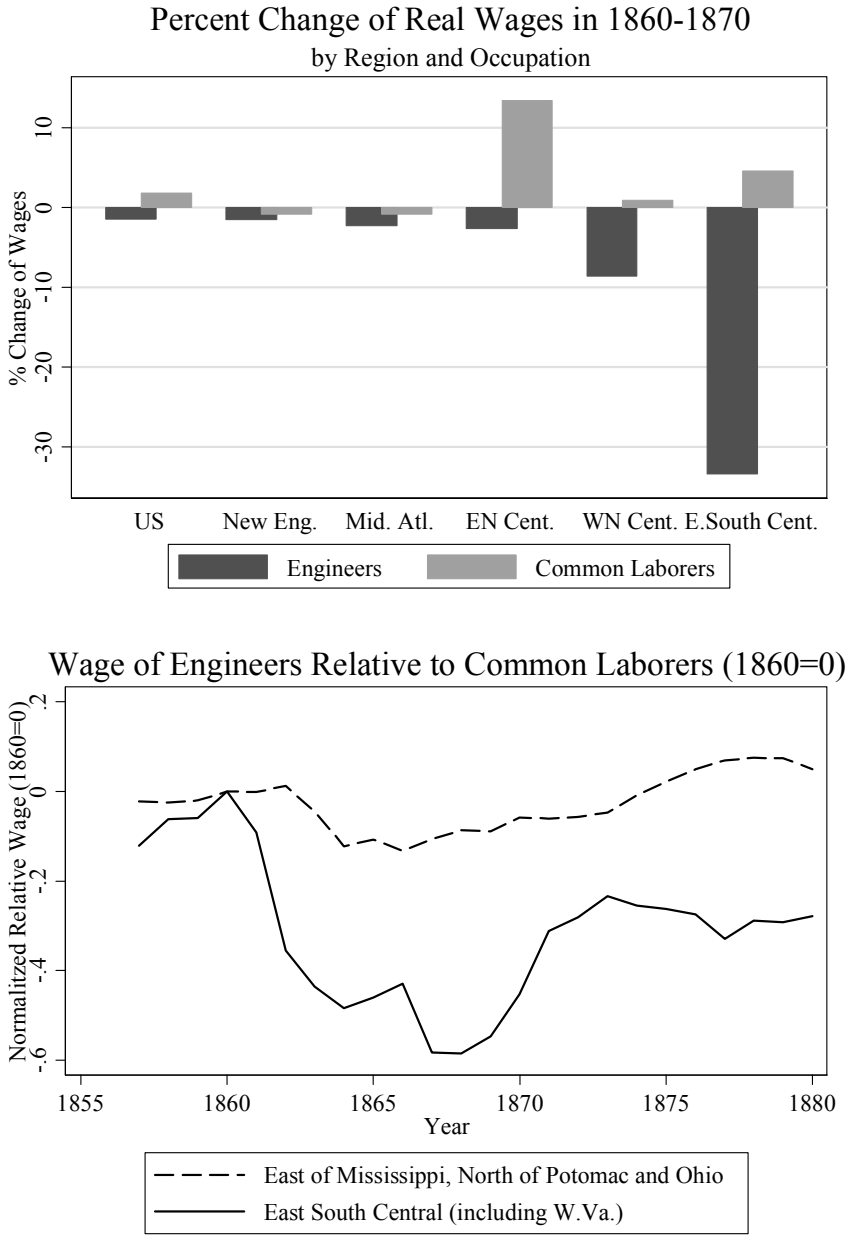
Note: Each model is discussed in text in detail. In each graph, *MB* and *MC* denote the curves of the marginal benefit and cost of schooling, respectively.

Figure 4. Literacy Rate by Cumulative School Attendance in Former Confederate States in 1880



Note: Using the 1880 IPUMS samples, we estimate years in school and literacy rate by age. The year of schooling is the cumulative sum (over ages) of the school-attendance rates. The values for former confederate states are denoted by 'number' points (i.e. ages), and the dashed line is their quadratic extrapolation. Those for former non-slave states are denoted by square points. For a comparison, we also include a point for 20-24 year olds in the 1860 IPUMS samples. Note that the antebellum censuses did not ask about literacy for people under 20 years of age.

Figure 5. Real Wage Change of Skilled and Unskilled Workers in 1860-1870



Note: We obtain the real wages by region and occupation from Tables 6 (for common laborers) and 7 (for engineers) in Coelho and Shepherd (1976). The upper panel shows the percent change of real wages in 1860-1870 by occupation and region. For the lower panel, we calculate the real wage of engineers relative to that of common laborers by two regions. The solid line is the trend of East South Central, including West Virginia. The dashed line indicates an unweighted average of the three census divisions comprising states east of the Mississippi and north of the Ohio or Potomac rivers. We normalize the series to zero in 1860.