Separate and Unequal in the Labor Market: Human Capital and the Jim Crow Wage Gap Appendix: For Online Publication

1 Interpreting the Mincer model

1.1 County and Age Variables

Our primary focus is on black-white gaps in earnings and occupation and, for brevity, the main text omits other results from the Mincer model. In this section we discuss the importance of age and county controls before diving deeper into the estimated effect of schooling and school quality. Table 1 lists parameter estimates for race, age, and county economic conditions from our preferred specification of Equation 1 in the main text, where cubic functions of school quality and attainment are fully interacted. County conditions pertain to 1940 counties of residence. Black-white gaps in Table 1 can also be found in Table 4 of the main text.

Earnings, occupational standing, and weeks worked all rise with age. Twenty-five-year old men, for instance, earn a weekly wage that is 48.2 log points greater than the weekly wage for 18-year old men, as well as an annual income that is 85.9 log points greater.[1] The age-earnings profile provides new context for the black-white gap in each labor outcome. Table 1 shows that the gap in weekly wages and annual income is equivalent to no more than two years of potential experience, but that the gap in occupational standing is equivalent to 6 - 7 years of potential experience.

The bottom portion of Table 1 indicates that urbanicity had no significant bearing on labor outcomes, but that men in counties with more crop output (conditional on urbanicity)

[1] Some of this may be driven by selection issues if higher-skilled individuals are more likely to be in school and out of the labor force at younger ages.

realized lower wages, lower annual earnings, and weaker occupational status. But the magnitude of wage differences between crop-heavy and crop-light counties is modest: a \$100 increase in crop value per capita (more than doubling the mean) is tied to a decrease in weekly wages of 15.2 log points. Similarly, the wage premium in places with more manufacturing is small relative to the mean of \$69.6 per capita.

1.2 Interacted Human Capital

Figure 1 visualizes the estimated effect of interactions between school quality and years of schooling. To generate these figures, we evaluated Equation 1 results at 320 combinations of the school quality index (SQI, divided into ventiles, lower right axes) and years of schooling attained (1 - 16, lower left axes). Predicted labor outcomes are represented on the vertical axis. All three outcomes generally rise with both SQI and years of schooling, with the slope with respect to attainment noticeably steeper. An exception to this pattern is found for estimated annual income, which is maximized at 16 years of schooling and the 75^{th} percentile of school quality before sloping somewhat downward for higher values of SQI. Ninety-five percent of men in the sample have no more than 13 years of schooling, and our cubic functional forms may do a poor job estimating the returns to post-secondary attainment. In the more relevant range of schooling, the implicit iso-wage curves in Figure 1 show that better school quality can substitute to some degree for lower attainment. Men with 12 years of schooling from 10^{th}-percentile counties, for instance, have estimated log weekly wages equal to 2.16, as do men with 8 years of schooling from 40^{th}-percentile counties. Similarly, men with 8 years of schooling from median-quality counties have occupational standing on par with men who have just 6 years of schooling from 95^{th}-percentile counties.

1.3 AGCT Results

In the same vein, Figure 2 presents the relationship between World War II enlistees' AGCT, educational attainment (measured as years beyond eighth grade), and deciles of the school quality index in their county of enlistment. Aptitude rises steeply with educational attainment and, to a lesser degree, with school quality. An enlistee with no more than an eighth-grade education in a county with very high school quality (in the 90^{th} - 100^{th} percentile) typically scored just one point less than his peer with ten years of education in the second decile of local school quality.

Recall that imputed AGCT in the census sample drew on known AGCT at race-specific combinations of educational attainment and school quality in the World War II enlistee data. This represents a more general functional form than pooled Mincer equations used throughout our analysis of earnings, as it allows the AGCT return to human capital to vary by race (Oaxaca decompositions do not support this functional form for wage regressions). Importantly, the enhanced functional form allows us to separately identify the role of AGCT in determining wage gaps which would be impossible if we imputed AGCT scores using pooled Mincer equations from the main analysis. Figure 3 plots local polynomial smoothing estimators for enlistees' AGCT against educational attainment (Panel I) and school quality (Panel II), separately by race. Blacks and whites of equivalent AGCT scores in this sample exhibit both a schooling and school quality gap favoring blacks.

2 The Importance of County-Level Data

A key contribution of this paper is a county-level dataset to study how differences in school quality across locations and races in the South ultimately affected labor market outcomes. Prior to the construction of these data, investigation of this question would have been limited to the use of state-level aggregates of school resources.

These state-level averages are suboptimal for two reasons. First, as discussed in Section 2 of the main text, the level of aggregation for school quality data matters in the modern literature (Betts, 2010), and the use of aggregate data flattens important variation in school quality within states. Second, because multiple measures of school quality are reported (term length, teacher salaries, class size, and so forth), and because states did not unanimously report any particular statistic, relying on state-level data would limit the scope of our study to uniformly available metrics or a smaller set of states that reported most of the same metrics. With county-level annual data on several domains of school quality we have the breadth and depth of information necessary to produce a standardized index of school quality for all ten states, as described in Section 3 of the main text.

Table 2 quantifies the advances from utilizing county-level versus state-level data. First, we highlight the aggregation impact coming from state-level data. In Column 1, we regress the outcomes of interest (here limited to weekly and annual wages) on cubic functions of county-level measures of two statistics with close to universal coverage in the data: average teacher salary and teachers per pupil. We continue to include a cubic function of educational attainment, age fixed effects, and a vector of local characteristics, but do not interact these values given the number of covariates now entering the estimating equation. Although estimation of Equation 1 in the main text utilizes a summary Z-score of these (and other) measures, the conditional black-white wage gap under this disaggregated specification matches closely what is observed in Table 4 of the main text (18.0 versus 19.1 log points of weekly wage and 13.3 versus 13.7 of annual wage). We then run the same regression with state-level aggregates of salaries and class sizes and present the results in Column 2 of Table 2. The estimated conditional gap increases by between 12 and 22 log points for the two outcomes of interest, a sizable change. Perhaps state-level data contribute less information to the wage model than county-level data, widening the residual wage gap. An important point to emphasize is that, by the same logic, more granular data at the school level may yield an even smaller conditional wage gap. Diagnostic statistics listed below each black-white gap estimate favor disaggregated data. Though the explained variance is nearly equivalent across Columns 1 and 2, Wald tests of the joint significance of school quality data strongly favor county-level quality.

In Columns 3-8 we show that restricting the measure of school quality to a particular metric (average teacher salary, term length, or teachers per pupil) and then estimating the earnings function illuminates differences between county-level and state-level data but also demonstrates how our results change when we depart from an index of multiple measures and characterize school quality with a single metric. Again, measurement error is a likely factor, as a single domain of school quality (at any level of aggregation) would be expected to do an inferior job of profiling individual human capital. As before, we include age fixed effects in these regressions, so identification comes from within-year variation in the school quality metric of interest across counties (odd columns) or states (even columns). Note, also, that conditional wage gap estimates with term length controls are restricted to a smaller set of states.

Columns 3 and 4 list wage gap estimates conditional on county and state-wide average teacher salaries. The county-level salary data generate an estimated conditional wage gap similar to that in the baseline: 17.8 log points of weekly wage and 13.4 log points of annual wages. But the state-level data generate higher conditional gaps: 30.9 and 35.4 log points, respectively. For term length, the state-level metrics in Column 6 produce *lower* estimates of the conditional wage gap than the county-level data results in Column 5. Wald tests for the school quality function favor county-level term length. In Columns 7 and 8, the ratio of teachers to pupils appears to be a poor standalone proxy of school quality, explaining a subjectively small portion of the wage gap at either aggregation. Nevertheless, Wald tests are considerably stronger for county-level teacher-pupil ratios.

3 Alternate Weighting Schemes for School Quality Measures

Our primary measure of school quality is the mean of up to five race-specific normalized metrics available in state department of education reports: expenditures per pupil, inverse class size, certified teachers per pupil, term length, and average teacher salary. By taking the mean of all available measures, we are giving each measure equal weight in describing school quality. This is arbitrary and may yield an inferior proxy; alternative weighting schemes could potentially represent school quality better and perhaps result in a different conditional wage gap.

A different way of expressing the school quality index is like so:

$$Q_{ctr} = \frac{\sum_{j=1}^{J} \theta_j Z_{jctr}}{\sum_{j=1}^{J} \theta_j}$$

where, as in the main text, j indexes the five quality measures and Q_{ctr} is the quality index for county c, school year t, and race r. Here, θ_j notation represents a weight placed on measure j. The evenly weighted index utilized in the main text has $\theta_j = 1$ for all measures, and the denominator divides the sum of normalized quality metrics by the count of non-missing metrics in that county-year-race cell.

To assess the sensitivity of our conclusions to the use of alternative weights of school quality measures, we conduct a permutation exercise using different allocations of θ_j and build a distribution of conditional wage gap estimates over different weighting schemes. In each of 1,000 iterations, we randomly generate $\theta_j \in [0, 1]$ for $J = 5$ school quality measures and then compute Q_{ctr} according to the equation above. This allows the relative importance of a given metric to increase by up to 500 percent, or decrease to nothing.[2] We then average Q_{ctr} over the years that each cohort was at risk of schooling and re-estimate δ, the conditional black-white gap in weekly wages, occupational score, annual income, or weeks of work.

Figure 4 arrays a cumulative distribution function for each series of conditional gap estimates. Gap estimates from even weighting (also reported in Table 4 of the main text) are indicated by solid blue lines in distribution figures and listed above each panel along with 95% confidence intervals corresponding to the permutation exercise. Gap estimates from specifications that minimize the Akaike Information Criterion (AIC) across all 1,000 iterations are reported above each figure and indicated by red dashed lines within each figure.

Conditional gap estimates under specifications that give equal weight to each school quality metric are not exceptional within the distribution of results from alternative weighting schemes, and the spread of estimates overall is narrow. The question of which weighting scheme is *best* at describing school quality is unsettled; different components of school quality may be more or less important for different labor outcomes. More importantly for our present concerns, AICs for each outcome do not give us a clear ruling on whether gap estimates should be larger or smaller in absolute value, and they point to gap estimates

[2]Weights for missing metrics are set to zero and thus excluded from the denominator.

that are quite similar to our main results, particularly for wage gaps. Weekly and annual wage gap estimates differ by 0.3 and 0.4 log points, respectively, across baseline and AIC-minimized specifications. We conclude that our use of an evenly weighted school quality index does not unduly influence estimated conditional labor market gaps.

4 Difference in the Gap Across Locations

Sundstrom (2007) examined the geography of conditional wage gaps in this era, finding a correlation between wage gaps and historic plantation density, black population shares, urbanization, and voting preferences of the local electorate. We examine whether conditional wage gaps in the 1940 census sample are higher in locations where overall discrimination is also presumably higher, conditioning on heretofore omitted school quality controls. We use each county's black population share as of 1860 to proxy for overall discrimination and race relations. The 1860 black population share is a widely-used measure of the strength of the slave economy and the overall disenfranchisement of blacks in the early 20th century. In more recent literature, this ratio has been tied to lower Democratic vote shares, greater opposition to affirmative action programs, and higher animosity toward black individuals (Acharya et al., 2015).

We bifurcate the sample at the median of 1860 black population shares for respondents' current county of residence and report results for each group in Table 3. Conditional wage gaps in both annual and weekly wages are higher in counties with 1860 population shares above the median value, by 3.2 log points for weekly wages and 12.3 log points for annual income. Occupational standing gaps are very similar in counties with historically higher black population shares, suggesting that occupational sorting was not remarkably more potent in these areas.

5 Robustness Checks

This section outlines the results of several sensitivity checks. Results are reported in Tables 4, 5, and 6. In addition to these checks, unreported specification checks using quartic or linear functions of human capital indicate that conclusions are impervious to the polynomial form of human capital controls.

Table 4 presents results from a number of alternative versions of Equation 1 in the main text. Baseline findings are repeated in Column 1 for weekly wages, Column 8 for occupation scores, Column 15 for annual earnings and Column 22 for weeks worked. In the following three columns, we change the underlying specification to include state fixed effects (Columns 2, 9, 16, and 23), county-of-residence fixed effects (Columns 3, 10, 17, and 24) and county-of-schooling fixed effects (Columns 4, 11, 18, and 25), in turn. Specifications with state fixed effects identify wage gaps from within-cohort, within-state differences in human capital. Conditioning on county-of-residence fixed effects leads us to identify wage gaps primarily from cross-cohort differences in human capital within a particular location, i.e., differences across race and age (note that age fixed effects are omitted for Columns 3-4, 10-11, 17-18, and 24-25). With county of residence fixed effects, additional variation emerges from migrants who were educated elsewhere. When we condition on county of schooling fixed effects, identification comes from school quality variation emerging solely as a result of age and race. If we control for both county-of-residence and age fixed effects (Columns 5, 12, 19, and 26), variation in school quality comes from differences in black and white school quality for individuals of a given age as well as from a relatively small number of migrants (see Table 6 of the main text) who were educated elsewhere. Because migrants are unlikely to be selected at random, we prefer broader identification.[3]

Table 4 also reports on results under alternative constructs of the school quality index. First, in recognition that public teacher salaries could have been affected by the same discrimination we seek to identify, we compute Z-scores without teacher salaries. These results are located in Columns 6, 13, 20, and 27. Finally, we re-generate the school quality Z-score as across-cohort rather than within-cohort to take advantage of tremendous intercohort gains in school quality (Columns 7, 14, 21, and 28).

Under these modifications, conditional black-white differences in labor outcomes are

[3]Recall that our preferred model omits controls for unobserved geographic heterogeneity in labor market outcomes. Discrimination itself is unobservable but to the extent it is concentrated in certain geographic areas, introducing state and local fixed effects would partially obfuscate the effect.

usually within one standard error of baseline estimates. One exception is for models where the school quality index is calculated without average teacher salaries, which widens the conditional gap in weekly wages to 24.0 log points and the gap in annual wages to 18.4 log points. A somewhat wider gap is to be expected in this model, since omitting a key component of school quality renders the index less informative by design.[4] We also see a higher conditional wage gap in annual income but a very similar gap in weekly wages when we include either state fixed effects or the combination of county-of-residence and age fixed effects. These results are driven by a smaller or negative conditional gap in weeks worked arising in these specification (see Columns 22 versus 23 with state fixed effects and 26 with age and county-of-residence fixed effects).

In Table 5, we check the sensitivity of estimates to various sample limitations. The baseline analysis – results of which are listed in Columns 1, 8, 15 and 22 – limits the sample of 1940 census respondents to young men who reported non-missing earnings, and who may or may not have had substantial non-wage income. To test the sensitivity of results to the omission of in-kind, interest, and self employment earnings, we drop all individuals earning more than $50 in non-wage income (Columns 2, 9, 16 and 23). Resulting black-white wage gaps change very little. In Columns 3, 10, 17, and 24, we limit the sample to exclude agricultural workers and focus only on non-farm sectors. This restriction increases estimates of the black-white gap in weekly and annual earnings (to 25.0 and 20.4 log points, respectively) and increases the occupational score gap to 24.3 log points. Higher discrimination in the non-farm sector is consistent with other evidence on discrimination among skilled versus unskilled men (Wright, 2013) and may be consistent with models of discrimination based on customer preferences and sales penalties imposed on the employers of black workers.

[4]Discrimination in the salary schedule would not have ruled out the possibility that higher pay would attract and retain better teachers. Margo (1984) finds that early 20^{th} century Southern teacher salaries are predictive of teachers' certification, normal-school training, and college attainment. In an unreported extension, we find that 1940 census respondents who list "teacher" as their occupation realized 4-6 percent gains from each year of education. The dominant causal direction between local teacher salaries and local teachers' qualifications is not clear, but regardless, the point stands that teacher salaries likely proxy for teachers' own human capital.

Columns 4, 11, 18 and 25 show results when we restrict the samples of black and white males to a common support defined as school quality and educational attainment contained in the range from the mean to the 95th percentile of observed black values. Much like the non-farm sector, the region of common support is expected to include higher skilled black men, who have been shown to be more affected by labor market discrimination. Indeed, measures of wage discrimination rise by 6 - 13 log points. When we condition on AGCT score for this group (not shown), the estimated gap for annual wages falls to 18.2 log points and for weekly wages to 20.1 log points, attenuation attributable to aptitude that is roughly similar to that for the full sample.

Columns 5, 12, 19, and 26 contain results when we restrict the estimating sample to individuals whose state of birth is equivalent to state of residence in 1935, potentially reducing error in the assignment of county of schooling. We see a reduction in measured wage and occupational standing gaps, although differences from baseline point estimates are slight. We next limit the sample to exclude men who report having had New Deal relief work in the 1940 reference week. These results, located in Columns 6, 13, 20 and 27, indicate higher conditional wage gaps for employment that is more market-based; but again, results are within a standard error of baseline estimates.

Finally, we expand the estimating sample to include a large block of individuals who did not report wages in 1940, namely farm operators. Because we do not observe wages for these men, they must be imputed from other available data. We start with $697, the average farm operator's net income in 1939, a measure which is available only at the national level (U.S. Department of Agriculture, 1957). We then scale this value by the region-to-national ratio of average farm operator net income in 1949, which is $1,554/$2,389 for the South Atlantic census region and $1,799/$2,389 for the South Central region.[5] The resulting figure is a proxy for the typical farmer's income in 1939, but it does not reflect differences by race. For black (white) farmers, we scale this number again by the black-to-all (white-to-all) ratio of average reported income for Southern farmers aged 28-35 in the 1950 IPUMS

[5]Regional farm income is not available for 1940. We cannot use state-by-race average wages from the 1950 Census microsamples as a reference point due to small sample sizes.

sample. The 28-35 age range represents the same cohort of men observed in 1940 when they are 10 years older. The black-to-all farmer income ratio is 0.56, and the white-to-all ratio is 1.14. Results with imputed farmer income, located in Columns 7, 14, 21, and 28,[6] indicate substantially higher conditional wage gaps, as high as 29.9 log points for weekly wages and 26.8 log points for annual income. We can think of these findings as the upper bound of the 1940 black-white wage gap under the assumption that local school quality had no bearing on farmers' effective income (as dictated by the regional income imputation described here). Results for farm employment *per se* presented in the following section, however, indicate that human capital affects the extensive margin of farming employment, and thus we expect the intensive margin of farmers' standard of living to have also been affected by better schools.[7]

In Table 6, we address a small number of potential confounders to our main results and one additional alternative specification. First, we control for the number of missing school quality metrics (Columns 2, 6, 10, and 14). Recall that quality indices are averages for up to five normalized statistics. Specifically, we supplement our primary specification (Equation 1 from the main text) with a quadratic function of the number of missing school quality statistics for each county and cohort. This is meant to address classic measurement error as well as the possibility that the quality of data reporting is correlated with the quality of schooling and/or unobserved labor market mechanisms dictating the black-white wage gap. Another specification includes indicator variables for the availability of each school quality metric (Columns 3, 7, 11, and 15). These two approaches increase weekly wage and annual income gaps by about one log point. Next, Columns 4, 8, 12 and 16 report the main results in our analysis for a group of slightly older men, aged 18-35 rather than 18-25. We hesitate to lend much credence to these results given the increasingly poor assumptions we must make about where each individual attended school (as before, based on 1935 location which, for these men, represents their residence at age 13 to 30). Our estimated wage gap is substantially higher for this expanded group, either because conditional gaps, including the

[6] Note that the occupational score gap in Column 14 is identical by design.

[7] The notion that schooling could improve farmers' productivity was one argument in support of including vocational space in philanthropically-funded black schoolhouses (Hoffschwelle 2006, pp.254-255).

effects of wage discrimination, grew as men aged[8] or because human capital is measured with additional error. We note that unconditional gaps are also far larger for this group and the proportion of the gap explained by human capital is not dissimilar.

6 Other Labor Market Outcomes

In addition to wage earnings and occupational score outcomes, human capital measures may have been influential in determining differences in other labor market outcomes: unemployment, farm employment, and work relief employment. To explore these outcomes, we expand the sample to include individuals who did not report income.[9] Results are found in Appendix Table 7.

Blacks were about as likely to be unemployed as whites in the unconditional view (Column 1), but controlling for human capital reveals a 3.4 percentage-point unemployment gap (Column 3). Columns 4-6 indicate that blacks were more likely to be employed in farming but, conditional on human capital measures, were far less likely to be so-employed.

Our main analysis includes a small share of men who were employed via New Deal work relief at the time of the census (about 6 percent of the analytical sample), which was not market-based employment. Work relief employment, which we cannot observe over the months preceding the census, could bias the implied market-based wage gap toward zero if federal, state, or local work relief countered and offset private sector discrimination (although robustness checks discussed earlier in this Appendix indicate that the gap grows by no more than a standard error when we exclude men with recent work relief employment). It would seem, though, that work relief employment was itself discriminatory. We find that blacks were 2.8 percentage points less likely to be on work relief than whites, and that human capital measures narrowed that gap to 1.6 percentage points.

[8] Carneiro et al. (2005) find that wage gaps tend to widen, albeit non-monotonically, as men aged in the NLSY panel. The gap among 1940 census respondents ages 18-35 is 25.1 - 28.8 log points (Columns 4, 12), which is not very different from the 24.1 log point gap among NLSY men aged 32-34 (Carneiro et al. (2005), Table 5), though comparisons should be made with caution since the earlier study included controls for schooling-corrected AFQT.

[9] We exclude farm owners and tenants from the farm employment analysis since we cannot distinguish farm tenants from owners.

7 Data Appendix

7.1 School Finance and Resource Data

This study makes use of data on Southern public school districts between 1910 and 1940 in 10 states (Alabama, Arkansas, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Texas, Kentucky, and Tennessee), including statistics on schools, teachers, students, and expenditures. These states were selected for their consistency in reporting the educational resources of interest. While several researchers have used portions of these data for specific projects, to our knowledge, the assembled dataset is unprecedented in its size, scope, and depth. We have already used these data to estimate the impact of philanthropically-funded Rosenwald schools on public support for segregated schools (Carruthers & Wanamaker, 2013) and to estimate the effect of women's suffrage on local school spending (Carruthers & Wanamaker, 2015). This section describes the data and data collection process in more detail.

Our primary sources of education input data are annual reports from state superintendent offices, departments of education, or equivalent governmental units. Measures of schooling resources reported separately for white and African-American schools typically include:

1. Enrollment, average daily attendance
2. Number of teachers overall
3. Expenditures
4. Teacher salaries
5. Number of schools
6. Average term lengths
7. School revenues drawn from local taxes

See Table 8 for the distribution of data availability across states and time. We transcribed all available statistics for these ten states and assembled county-by-race panels for the years 1910-1940. Data availability is remarkably consistent across states and years. We conducted an informed 0.5 percent audit of each transcribed variable. Specifically, for each school statistic and each state, we regressed transcribed data against county fixed effects and a quadratic function of time, generating predicted values and residuals. We flagged cells in the top 99.5 percentile of residuals, in absolute value. Then, our research assistant verified the accuracy of each flagged cell by consulting the original scanned reports and fixed any transcription errors. The realized error rate from these flags was 14.9 percent. We believe this to be an encouraging signal of the underlying fidelity of these data, considering that our audit focused on the top 0.5 percent of outliers within counties' time series.

Notably, we did not correct items that we thought to be typing or arithmetic errors in the original documents. These errors are, presumably, white noise that will bias our results towards finding no impact of school resources on wages.

7.2 Census Data

Census data are the 1% micro-sample available from IPUMS-USA. We converted top-coded earnings to 140% of the top-coded value. (0.04% of our baseline sample was top-coded.) All individuals reporting a race other than "black" or "white" were deleted from the sample. All individuals with occupation codes corresponding to "unpaid family workers," "farmers," and "farm managers" were eliminated from the working sample, as most individuals in these categories reported no wage income. Educational attainment was top-coded at 17 years of schooling ("5th year of college or more"), and we assume all individuals in this category have 18 years of schooling, reflecting two years of post-college education.

Our results are

7.3 World War II Enlistment Data

World War II enlistment records are available from the National Archives via their website at archives.gov. We downloaded the raw data files (where the weight field is transcribed) for use in this paper. The files contain a large number of errors, and we deleted all records where fields could not have contained the correct information (for example, weight fields containing non-numeric values or with values greater than 301 or less than 1). We eliminated all females from the sample.

For each individual listed March 1943 through the first week of June 1932 (the window during which the weight field was actually each enlistee's AGCT score), we assigned an average school quality metric based on county of residence at enlistment according to the method described in the main text for the census sample. We eliminated all individuals younger than 18 or older than 25 and all individuals where reported race was something other than "black" or "white." The remaining sample contains 66,684 individuals. We estimate AGCT-proxied aptitude as a function of enlistees' observable characteristics that have counterparts in the matched sample of census responses and school quality, and then map predicted AGCT to the analytical census sample. See the main text, Section 4.3, for additional details.

References

Acharya, A., Blackwell, M., & Sen, M. (2015). The Political Legacy of American Slavery. *Working Paper*.

Betts, J. R. (2010). School Quality and Earnings. In P. Peterson, E. Baker, & B. McGaw (Eds.) *International Encyclopedia of Education*, (pp. 313–320). Oxford: Elsevier, 3 ed.

Carneiro, P., Heckman, J. J., & Masterov, D. V. (2005). Labor Market Discrimination and Racial Differences in Premarket Factors. *Journal of Law and Economics*, *48*(1), 1 – 39.

Carruthers, C. K., & Wanamaker, M. H. (2013). Closing the Gap? The Effect of Private Philanthropy on the Provision of African-American Schooling in the U.S. South. *The Journal of Public Economics*, *101*, 53–67.

Carruthers, C. K., & Wanamaker, M. H. (2015). Municipal Housekeeping: The Impact of Women's Suffrage on Public Education. *The Journal of Human Resources*, *50*(4), 837–872.

Carter, S. B., Gartner, S. S., Haines, M. R., Olmstead, A. L., Sutch, R., & Wright, G. (Eds.) (2006). *Historical Statistics of the United States: Earliest Times to the Present*. Cambridge University Press.

Hoffschwelle, M. S. (2006). *The Rosenwald Schools of the American South*. University Press of Florida.

Margo, R. A. (1984). Teacher Salaries in Black and White: The South in 1910. *Explorations in Economic History*, *21*, 306–326.

Ruggles, S., Alexander, J. T., Genadek, K., Goeken, R., Schroeder, M. B., & Sobek, M. (2010). *Integrated Public Use Microdata Series: Version 5.0 (Machine-Readable Database)*. University of Minnesota.

Sundstrom, W. A. (2007). The Geography of Wage Discrimination in the Pre-Civil Rights South. *Journal of Economic History*, *67*(2), 410 – 444.

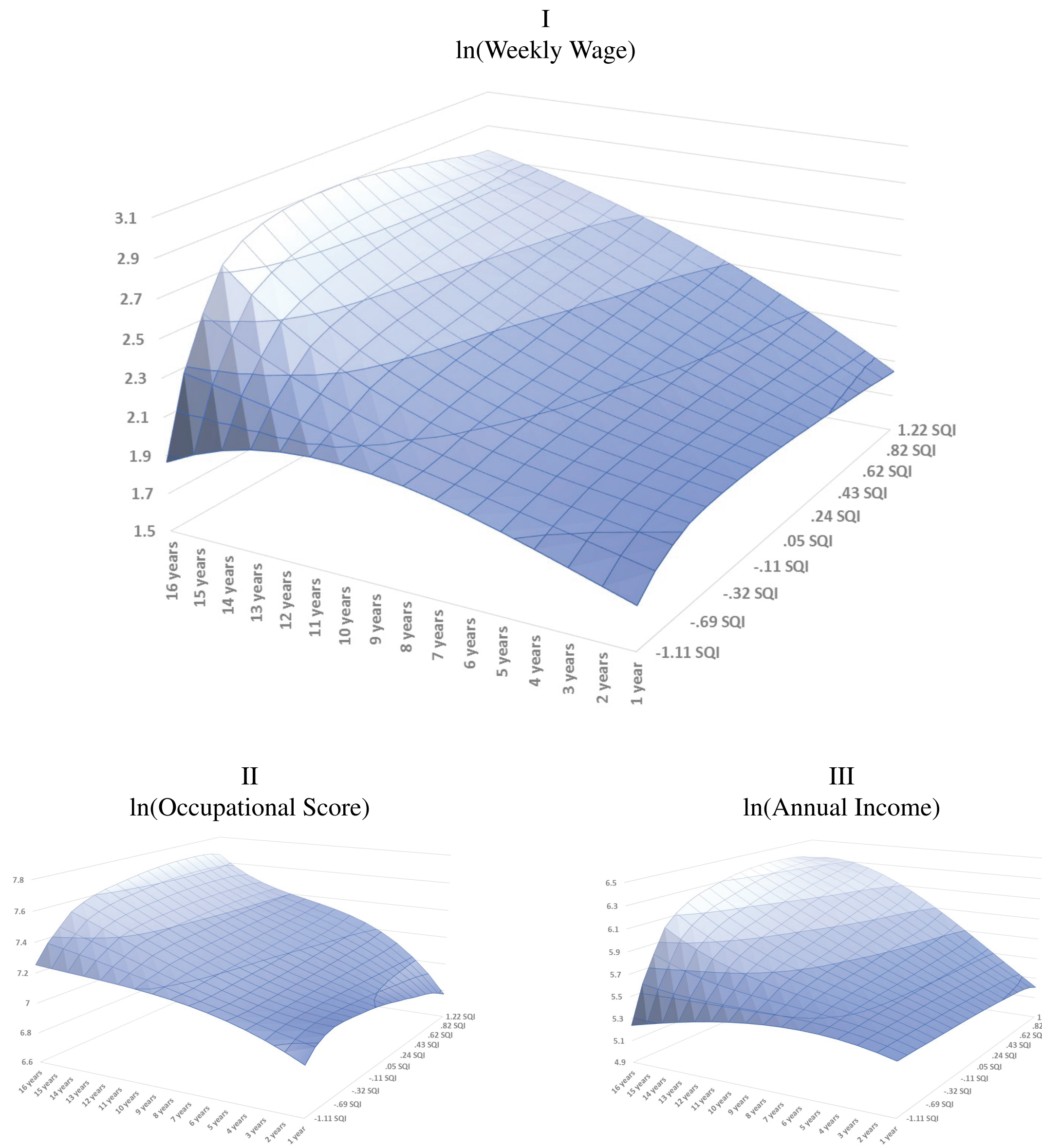
U.S. Department of Agriculture (1957). *Major Statistical Series of the U.S. Department of Agriculture*. Washington, DC: GPO.

Wright, G. (2013). *Sharing the Prize: The Economics of the Civil Rights Revolution in the American South*. Cambridge, MA: Belknap Press.

TABLE 1: Mincer Results (Equation 1) for Race and Other Controls

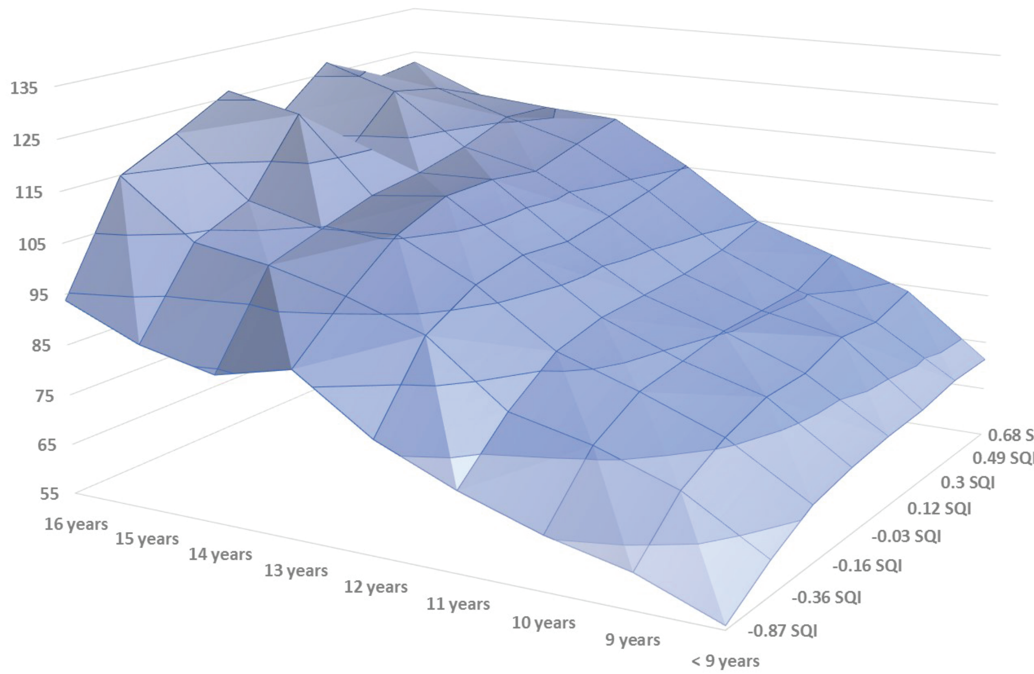
	ln(Weekly Wage)	ln(Occupation Score)	ln(Annual Wages)	ln(Weeks Worked)
BLACK-WHITE GAP	-0.191	-0.168	-0.137	0.054
	(0.032)	(0.022)	(0.038)	(0.024)
Age 19	0.137	0.03	0.239	0.102
	(0.030)	(0.019)	(0.042)	(0.030)
Age 20	0.202	0.064	0.404	0.202
	(0.029)	(0.019)	(0.040)	(0.030)
Age 21	0.251	0.077	0.513	0.262
	(0.029)	(0.019)	(0.037)	(0.028)
Age 22	0.344	0.117	0.631	0.287
	(0.027)	(0.018)	(0.040)	(0.029)
Age 23	0.372	0.123	0.725	0.353
	(0.028)	(0.019)	(0.039)	(0.027)
Age 24	0.450	0.152	0.827	0.377
	(0.027)	(0.019)	(0.037)	(0.027)
Age 25	0.482	0.177	0.859	0.377
	(0.026)	(0.018)	(0.037)	(0.027)
County Economic Conditions				
Percent rural	0.042	-0.050	0.033	-0.008
	(0.074)	(0.049)	(0.081)	(0.053)
Per-capita manufacturing value added (000s)	0.749	0.310	1.068	0.320
	(0.134)	(0.116)	(0.184)	(0.131)
Per-capita retail sales	0.514	-0.229	0.384	-0.131
	(0.156)	(0.117)	(0.200)	(0.131)
Per-capita crop value (000s)	-1.518	-2.205	-1.252	0.266
	(0.304)	(0.221)	(0.344)	(0.232)
Constant	1.537	7.093	4.626	3.089
	(0.092)	(0.064)	(0.110)	(0.079)
R-squared	0.30	0.26	0.28	0.06
Observations	11,394	11,021	11,394	11,394

Notes: Authors' calculations from 1940 IPUMS data (Ruggles et al., 2010), Carter et al. (2006), and annual reports of state education departments. See Section 1 for discussion.



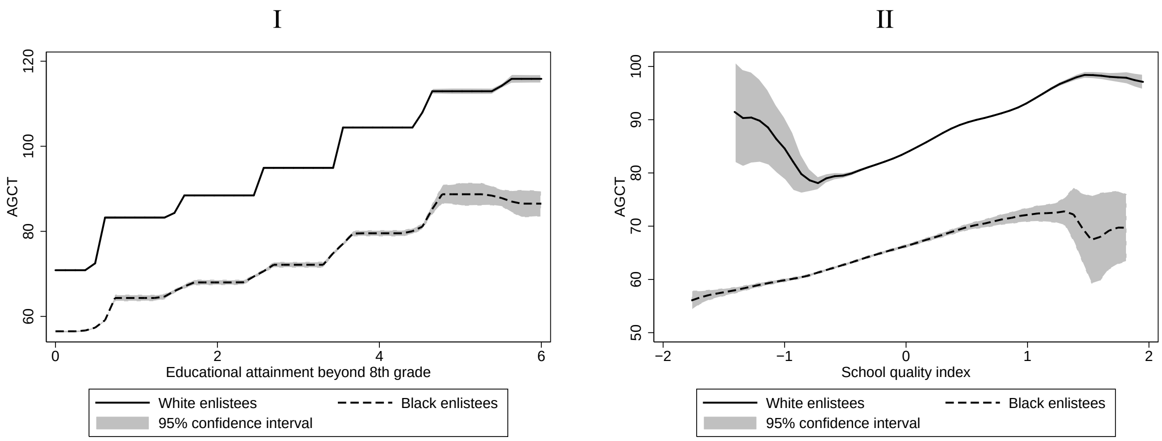
Notes: The figure depicts surface plots for weekly wage (I), occupational score (II), and annual earnings (III). Each plot maps 320 iterations of Equation 1 predictions, where a given iteration measures the mean predicted outcome evaluated at a particular combination of the school quality index (SQI, divided into ventiles) and years of schooling (1 - 16).

FIGURE 1: Conditional effects of school quality and years of schooling on earnings and occupational standing



Notes: The figure depicts a surface plot for AGCT in World War II enlistee records. The plot maps enlistees' average AGCT at 100 combinations of educational attainment and school quality index (SQI, divided into deciles) in the county of enlistment.

FIGURE 2: Observed enlistee AGCT, by school quality and years of schooling



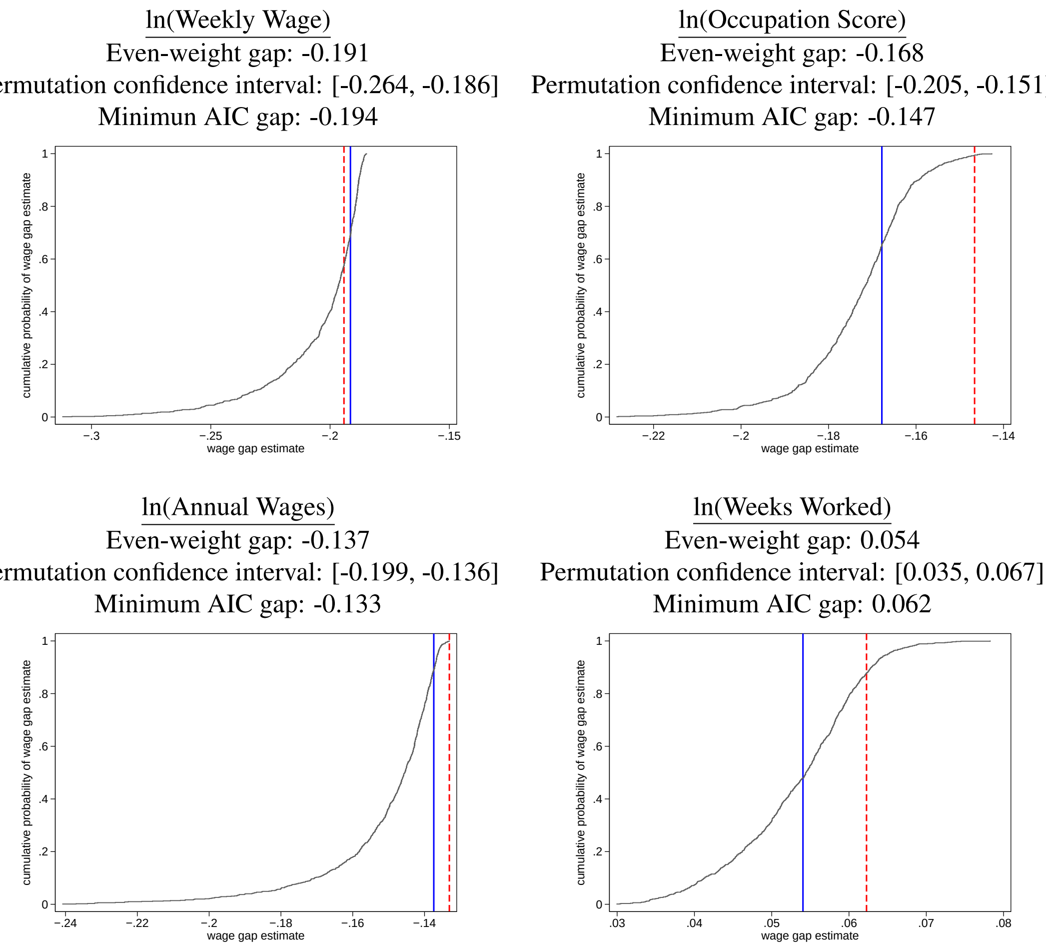
Notes: The figure plots local polynomial estimates for AGCT scores, educational attainment (Panel I), and the school quality index (Panel II).

FIGURE 3: Observed enlistee AGCT scores by race, schooling, and school quality

TABLE 2: Estimates of Black-White Labor Market Outcome Gaps, Comparing County and State Aggregates

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Level of Measurement	County	State	County	State	County	State	County	State
Outcome	Teachers per pupil + Average teacher salary		Average teacher salary		Term length		Teachers per pupil	
ln(Weekly Wage)	-0.180	-0.302	-0.178	-0.309	-0.241	-0.198	-0.312	-0.323
	(0.035)	(0.061)	(0.032)	0.059)	(0.028)	(0.074)	(0.027)	(0.056)
R^2	0.303	0.301	0.302	0.299	0.309	0.307	0.297	0.298
School quality Wald test	19.26***	5.27***	26.88***	8.64***	13.47***	2.41*	17.35***	3.13**
ln(Annual Wage)	-0.133	-0.353	-0.134	-0.354	-0.195	-0.171	-0.243	-0.349
	(0.043)	(0.066)	(0.038)	(0.064)	(0.039)	(0.079)	(0.033)	(0.063)
R^2	0.270	0.270	0.269	0.269	0.269	0.269	0.266	0.268
School quality Wald test	31.83***	6.71***	18.69***	9.76***	8.48***	3.44**	89.50***	3.97***
N	11,130	11,130	11,164	11,164	8,054	8,054	11,182	11,182

Notes: Authors' calculations from 1940 IPUMS data (Ruggles et al., 2010) and annual reports of state education departments. The table lists estimates of the conditional gap, δ, from Equation 1 in the main text using county-level or state-level measures of school quality. Robust standard errors, clustered by 1940 county of residence, are in parentheses. Specific school quality metrics included are given by column headings. See Section 2 for specification details. For Wald tests, *** represents 99% significance, ** 95% significance, and * 90% significance.



Notes: Main results under even weighting, also found in Table 4 of the main text, are listed above each cumulative distribution function and indicated within each figure by a solid vertical line. Permutation confidence intervals – i.e., gap estimates at 2.5 and 97.5 percentiles – are listed in brackets. Specifications that minimize the AIC across permutations are indicated with dashed vertical lines.

FIGURE 4: Distribution of conditional gap estimates under alternate weighting schemes for school quality measures.

TABLE 3: Conditional Gaps by 1860 Black Population Shares

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	ln(Weekly Wage)		ln(Annual Wage)		ln(Occupational Score)		ln(Weeks Worked)	
	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median
BLACK-WHITE GAP	-0.169	-0.201	-0.108	-0.231	-0.151	-0.162	0.061	-0.030
	(0.043)	(0.051)	(0.052)	(0.055)	(0.031)	(0.034)	(0.031)	(0.034)
N	5,519	5,559	5,519	5,559	5,337	5,385	5,519	5,559
R-squared	0.25	0.33	0.26	0.30	0.21	0.28	0.07	0.05

Notes: Authors' calculations from 1940 IPUMS data (Ruggles et al., 2010) and annual reports of state education departments. The median county black population share is 0.28 among men in the analytical sample. See Section 4 for discussion.

TABLE 4: Alternative Specifications

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Outcome	ln(Weekly Wage)							ln(Occupational Score)						
BLACK-WHITE GAP	-0.191	-0.215	-0.160	-0.137	-0.178	-0.240	-0.196	-0.168	-0.180	-0.164	-0.146	-0.170	-0.162	-0.177
	(0.032)	(0.031)	(0.039)	(0.041)	(0.038)	(0.029)	(0.031)	(0.022)	(0.022)	(0.027)	(0.032)	(0.026)	(0.021)	(0.022)
Baseline	✓							✓						
State fixed effects		✓							✓					
County-of-residence fixed effects			✓							✓				
County-of-schooling fixed effects				✓							✓			
County-of-residence and age FE's					✓							✓		
No teacher salaries						✓							✓	
Pooled school quality							✓							✓
N	11,394	11,394	11,394	11,394	11,394	11,374	11,394	11,021	11,021	11,021	11,021	11,021	11,001	11,021
Adjusted R-Squared	0.30	0.31	0.36	0.36	0.40	0.30	0.30	0.26	0.27	0.35	0.35	0.36	0.26	0.26

Column	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)
Outcome	ln(Annual Wage)							ln(Weeks Worked)						
BLACK-WHITE GAP	-0.137	-0.199	-0.168	-0.060	-0.199	-0.184	-0.153	0.054	0.016	-0.008	0.077	-0.021	0.056	0.043
	(0.038)	(0.035)	(0.047)	(0.053)	(0.043)	(0.036)	(0.037)	(0.024)	(0.024)	(0.026)	(0.031)	(0.025)	(0.024)	(0.023)
Baseline	✓							✓						
State fixed effects		✓							✓					
County-of-residence fixed effects			✓							✓				
County-of-schooling fixed effects				✓							✓			
County-of-residence and age FE's					✓							✓		
No teacher salaries						✓							✓	
Pooled school quality							✓							✓
N	11,394	11,394	11,394	11,394	11,394	11,394	11,394	11,394	11,394	11,394	11,394	11,394	11,374	11,394
Adjusted R-Squared	0.28	0.29	0.29	0.29	0.37	0.27	0.27	0.06	0.07	0.13	0.13	0.17	0.06	0.06

Notes: Authors' calculations from 1940 IPUMS data (Ruggles et al., 2010) and annual reports of state education departments. See discussion in Section 5.

TABLE 5: Alternative Samples

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Outcome	ln(Weekly Wage)							ln(Occupational Score)						
BLACK-WHITE GAP	-0.191	-0.200	-0.250	-0.252	-0.180	-0.208	-0.299	-0.168	-0.166	-0.243	-0.207	-0.161	-0.161	-0.168
	(0.032)	(0.032)	(0.032)	(0.038)	(0.033)	(0.033)	(0.029)	(0.022)	(0.023)	(0.022)	(0.029)	(0.024)	(0.024)	(0.022)
Baseline	✓							✓						
Exclude in-kind earners		✓							✓					
No agricultural workers			✓							✓				
Common support of human capital				✓							✓			
Birthplace check					✓							✓		
Exclude Relief Workers						✓							✓	
Impute Farmer Wages							✓							✓
N	11,394	10,349	9,272	3,671	10,141	10,420	13,447	11,021	10,004	8,899	3,548	9,805	10,080	11,021
Adjusted R-Squared	0.30	0.31	0.25	0.16	0.30	0.32	0.28	0.26	0.26	0.14	0.13	0.26	0.28	0.26

Column	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)
Outcome	ln(Annual Wage)							ln(Weeks Worked)						
BLACK-WHITE GAP	-0.137	-0.152	-0.204	-0.267	-0.111	-0.171	-0.268	0.054	0.048	0.046	-0.014	0.069	0.037	0.031
	(0.038)	(0.039)	(0.038)	(0.048)	(0.041)	(0.037)	(0.031)	(0.024)	(0.026)	(0.027)	(0.033)	(0.027)	(0.026)	(0.022)
Baseline	✓							✓						
Exclude in-kind earners		✓							✓					
No agricultural workers			✓							✓				
Common support of human capital				✓							✓			
Birthplace check					✓							✓		
Exclude Relief Workers						✓							✓	
Impute Farmer Wages							✓							✓
N	11,394	10,349	9,272	3,671	10,141	10,420	13,447	11,394	10,349	9,272	3,671	10,141	10,420	13,447
Adjusted R-Squared	0.28	0.28	0.25	0.19	0.27	0.29	0.25	0.06	0.06	0.07	0.07	0.06	0.06	0.05

Notes: Authors' calculations from 1940 IPUMS data (Ruggles et al., 2010) and annual reports of state education departments. See discussion in Section 5.

TABLE 6: Additional Tests

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	ln(Weekly Wage)				ln(Occupational Score)			
BLACK-WHITE GAP	-0.191	-0.197	-0.198	-0.288	-0.168	-0.167	-0.174	-0.207
	(0.032)	(0.032)	(0.031)	(0.029)	(0.022)	(0.021)	(0.022)	(0.016)
Baseline	✓				✓			
Missing data control		✓				✓		
Resource metric indicators			✓				✓	
Ages 18 to 35				✓				✓
N	11,394	11,394	11,394	25,048	11,021	11,021	11,021	24,472
Adjusted R-Squared	0.30	0.30	0.31	0.39	0.26	0.26	0.26	0.28

Column	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Outcome	ln(Annual Wage)				ln(Weeks Worked)			
BLACK-WHITE GAP	-0.137	-0.151	-0.146	-0.251	0.054	0.047	0.053	0.037
	(0.038)	(0.038)	(0.036)	(0.033)	(0.024)	(0.024)	(0.024)	(0.017)
Baseline	✓				✓			
Missing data control		✓				✓		
Resource metric indicators			✓				✓	
Ages 18 to 35				✓				✓
N	11,394	11,394	11,394	25,048	11,394	11,394	11,394	25,048
Adjusted R-Squared	0.28	0.28	0.28	0.38	0.06	0.06	0.06	0.07

Notes: Authors' calculations from 1940 IPUMS data (Ruggles et al., 2010) and annual reports of state education departments. See discussion in Section 5.

TABLE 7: Additional Outcomes

Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Outcome	Unemployment			Farming Employment			Federal Work Relief		
BLACK-WHITE GAP	-0.004	-2.9E-04	0.034	0.145	0.115	-0.117	-0.028	-0.028	-0.016
	(0.007)	(0.006)	(0.012)	(0.017)	(0.012)	(0.018)	(0.005)	(0.005)	(0.008)
County Covariates?		✓	✓		✓	✓		✓	✓
Age Fixed Effects?		✓	✓		✓	✓		✓	✓
Human Capital Controls?									
Interacted HC Controls?			✓			✓			✓
N	17,439	17,439	17,439	16,914	16,914	16,914	17,439	17,439	17,439
Adjusted R-Squared	0.00	0.02	0.03	0.02	0.22	0.27	0.00	0.01	0.02

Notes: Authors' calculations from 1940 IPUMS data (Ruggles et al., 2010) and annual reports of state education departments. The analytical sample for farming employment (Columns 4 - 6) farm owners and family farm workers. See Section 6 for discussion.

TABLE 8: Availability of Education Quality Variables

year	AL	AR	GA	KY	LA	MS	NC	SC	TN	TX
1920		3 4	1 2 3 4		3 4	3 4		1 2 3 4	3 4	
	5 6 7 8	5 6 7	5 6 7 8		5 6 7 8	8	8	7 8	8	
1921	1 2 3 4	3 4	1 2 3 4		3 4	3 4	3 4	1 2 3 4	3 4	
	5 6	5 6 7 8	5 6 7 8		5 6 7 8	8	7 8	7 8	8	
1922	1 2 3 4	3 4	1 2 3 4		3 4	3 4	3 4	1 2 3 4		
	5 6	5 6 7 8	5 6 7 8		5 6 7 8	8	7 8	7 8		
1923	1 2 3 4	3 4			3 4	3 4	3 4	1 2 3 4		
	5 6	5 6 7 8			5 6 7	8	7 8	5 6 7 8		
1924	1 2 3 4	3 4	1 2 3 4		3 4	3 4	3 4	1 2 3 4		
	5 6	5 6 7 8	5 6 7 8		5 6 7	8	7 8	5 6 7 8		
1925	1 2 3 4				3 4	3 4	3 4	1 2 3 4		3
	5 6				5 6 7 8	8	7 8	5 6 7 8		8
1926	1 2 3 4		1 2 3 4		1 2 3 4		3 4	1 2 3 4		3
	5 6		5 6 7 8				7 8	7 8		8
1927	1 2 3 4	3 4			1 2 3 4		3 4	1 2 3 4	3 4	3
	5 6	8			7		7 8	5 6 7 8	7	8
1928	1 2 3 4	3 4	1 2 3 4		1 2 3 4		3 4	1 2 3 4	3 4	
	5 6	5 6 7 8	5 6 7 8		7 8		7 8	5 6 7 8	7	8
1929	1 2 3 4				1 2 3 4		3 4	1 2 3 4	3 4	3
					7 8		7 8	5 6 7 8	7	8
1930	1 2 3 4	3 4	1 2 3 4		1 2 3 4		3 4	1 2 3 4	3 4	3
		5 6 7 8	5 6 7 8		7 8		7 8	5 6 7 8	7 8	8
1931	1 2 3 4	3 4		3 4	1 2 3 4		3 4	1 2 3 4	3 4	3 4
		7		7 8	7 8		7 8	5 6 7 8	7 8	8
1932	1 2 3 4	3 4	1 2 3 4	3 4	1 2 3 4		3 4	1 2 3 4	3 4	3 4
		5 6 7 8	5 6 7 8	7	7 8		7 8	5 6 7 8	8	8
1933	1 2 3 4			3 4	1 2 3 4		3 4	1 2 3 4	3 4	3 4
				7 8	7 8		7 8	5 6 7 8	7 8	8
1934	1 2 3 4	3 4	1 2 3 4	3 4	1 2 3 4		3 4	1 2 3 4	3 4	3 4
		7 8	5 6 7 8	7 8	7 8		7 8	5 6 7 8	8	8
1935	1 2 3 4			3 4	1 2 3 4		3 4	1 2 3 4	3 4	3 4
	5 6			7 8	7 8		7 8	5 6 7 8	7 8	8
1936	1 2 3 4	3 4	3 4	3 4	1 2 3 4		3 4	1 2 3 4		3 4
	5 6	7 8	5 6 7 8	7 8	7 8		7 8	5 6 7 8		8
1937	1 2 3 4			3 4	1 2 3 4		3 4	1 2 3 4	3 4	3 4
	5 6			7 8	7 8		7 8	5 6 7 8	7 8	8
1938	1 2 3 4	3 4	3 4	3 4	1 2 3 4		3 4	1 2 3 4	3 4	3 4
	5 6	7 8	7 8	7 8	7 8		7 8	5 6 7 8	7 8	8
1939	1 2 3 4			3 4	1 2 3 4		3 4	1 2 3 4	3 4	3 4
	5 6			7	7 8		7 8	5 6 7 8	7 8	8
1940	1 2 3 4		1 2 3 4	3 4	1 2 3 4		3 4	1 2 3 4	3 4	3 4
	5 6	7	7 8	7	7 8		7 8	5 6 7 8	7 8	8

Notes: Data available in annual reports of state education departments, separately for black and white schools (see Section 7.1). Coding 1: Expenditures per enrolled pupil; 2: Expenditures per pupil in ADA; 3: Teachers per enrolled pupil; 4: Teachers per pupil in ADA; 5: Certified teachers per enrolled pupil; 6: Certified teachers per pupil in ADA; 7: Term length; 8: Average teacher salary

TABLE 9: Summary of Variables for Migrant and Non-Migrant Blacks

	Black Non-Migrants	Black Migrants
Average Annual Wage Income *in natural log*	5.41	5.51
Occupational Score *in natural log*	7.0	7.0
Average Weeks Worked	39.0	38.4
Highest Grade Completed	5.5	6.1
School Quality Index	-0.49	-0.58
County of 1940 Residence		
Percent Rural	64.5	60.4
Per Capita Manufacturing Value	79.9	83.2
Per Capita Retail Sales	0.20	0.23
Per Capita Crop Value	53.5	51.0
Number of observations	2,928	211

Notes: Authors' calculations from 1940 IPUMS data (Ruggles et al., 2010). See discussion in Section 4.2 of the main text. Migrants include those who crossed county lines within the South. Non-migrants did not change county of residence between 1935 and 1940.