Separate and Unequal in the Labor Market: Human Capital and the Jim Crow Wage Gap

Celeste K. Carruthers, *University of Tennessee*

Marianne H. Wanamaker, *University of Tennessee and National Bureau of Economic Research*

Competing explanations for the long-standing gap between black and white earnings attribute different weight to wage discrimination and human capital differences. Using new data on local school quality, we find that human capital played a predominant role in determining 1940 wage and occupational status gaps in the South despite entrenched racial discrimination in civic life and the lack of federal employment protections. The resulting wage gap coincides with the higher end of the range of estimates from the post–Civil Rights era. We estimate that truly "separate but equal" schools would have reduced wage inequality by 29%–48%.

I. Introduction

The American labor market has long exhibited a sizable gap in wages awarded to black and white workers, motivating a large body of research

We thank William Collins, Robert Margo, Don Bruce, John Parman, Michael Kofoed, Price Fishback, and seminar participants at the University of Tennessee, the University of Pittsburgh, the College of William and Mary, Harvard University, York University, the University of California, Davis, the University of Florida, the University of California, Irvine, the 2013 Southern Economic Association annual meetings, and the 2014 NBER Summer Institute for helpful comments and suggestions. Ye Gu, Jim Sheffield, Mary Elizabeth Glenn, Andrew Moore, James England III, and Nicholas Busko provided outstanding research assistance. Funding for

[*Journal of Labor Economics*, 2017, vol. 35, no. 3]

0734-306X/2017/3503-0007$10.00
Submitted June 1, 2015; Accepted April 29, 2016; Electronically published March 23, 2017

devoted to disentangling the role of human capital, or "pre-market," factors from more structural labor market issues and, chiefly, wage discrimination.[1] This decomposition exercise has important policy implications. If pre-market skill gaps are largely responsible for pay differences, policy solutions should prioritize disparities in human capital accumulation over direct labor market interventions. At the same time, if residual gaps unexplained by measurable human capital can be confidently labeled discrimination, direct labor market interventions are apt. Isolating the contribution of pre-market factors also holds historic importance for understanding the long-term consequences of segregation as well as the late twentieth-century convergence of black and white wages. Leading explanations for this convergence include declining wage discrimination—and in particular, the effect of the 1964 Civil Rights Act—and the rise of black public school quality and, in turn, black human capital.

Detailed measures of individual skill are necessary in order to deconstruct the relative importance of human capital versus wage discrimination. In this respect, literature on the post–Civil Rights era has had the advantage of observing rich data on workers' schooling, aptitude, and earnings. We utilize recently developed historic data to extend the time series of wage decomposition to the 1940 US South, a setting renowned for racial discrimination and disenfranchisement. In addition to hosting discriminatory mores in general, this time period lacked employment protections for black workers.[2] We answer two questions fundamental to understanding racial wage inequality early in the twentieth century. How important were human capital differences for racial labor market gaps in 1940? And how large was the corresponding

data collection was provided by the University of Tennessee Office of Research. Additional support includes a grant from the Spencer Foundation (no. 201200064) and a grant from the University of Kentucky Center for Poverty Research via the US Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation (5 U01 PE000002-06). The opinions and conclusions expressed herein are solely those of the authors and should not be construed as representing the opinion or policy of the Spencer Foundation, the US Department of Health and Human Services, or any agency of the federal government. All errors are our own. Contact the corresponding author, Celeste Carruthers, at carruthers@utk.edu. Data used in this article are available as supplementary material online.

[1] Lazear (1991); Oettinger (1996); Darity and Mason (1998); Altonji and Pierret (2001); Lang and Manove (2011). Also see Lang and Lehmann (2012) for a more complete survey of the racial discrimination literature.

[2] Our focus on the 1940 South is also appropriate for practical reasons. Importantly, the 1940 Census is the first to report individual earnings. The vast majority of black males resided in the South in 1940, and regional wage gaps were still large in the early years of the Great Migration. Furthermore, school quality metrics for non-southern states with integrated schools are not reported by race, preventing us from approximating race-specific human capital outside the South.

conditional wage gap, including wage discrimination, in this pre–Civil Rights era?

The US South in this period was characterized both by local control of public schools and separate and de jure unequal schools for black and white students. The result was wide variation in school resources both across and within races.[3] Consequently, an adequate profile of individual human capital in 1940 should include educational attainment, which is readily available for all adults in the 1940 Census returns, as well as local information on the quality of schooling available to black and white students. It is the latter that has previously inhibited estimation of the conditional wage gap in this context. We develop a new panel of county-by-race school quality statistics for each year between 1920 and 1940 for 10 southern states. We use these data to assign young men in the 1940 Census Public Use Microsample (Ruggles et al. 2010) a school quality metric specific to their race, age, and probable county of education. In addition to years of schooling and the quality of available schooling, we utilize a known oddity in the World War II enlistment records to impute Army General Classification Test (AGCT) scores for southern males in 1940 as a third measure of human capital.

We find that pre-market human capital disparities are the predominant determinant of southern racial gaps in 1940 for a number of labor market outcomes. For young, employed males in our sample, the age- and location-adjusted difference between black and white wages in 1940 was a substantial 46.7 log points on an annual basis and 49.0 log points per week worked. These gaps attenuate by 71% and 61%, to 13.7 and 19.1 log points, respectively, when we condition on educational attainment and school quality. Differences in educational attainment and differences in school quality account for similar proportions of the attenuation. It does not appear that school quality measures are simply proxying for local labor market discrimination. School quality per se had little effect on wages but rather served to enhance returns to years of schooling, and we find little evidence of race-dependent

[3] In the 1930 South, according to public reports of state education departments, annual spending per enrolled black pupil was typically \$9, vs. \$61 in spending per white pupil. White school years measured 156 days typically, whereas black term lengths were 20.5% shorter at 123 days (authors' calculations using county-level school resource data described in Sec. III). Although school quality in the post–Reconstruction South was relatively similar across races, and although *Plessy v. Ferguson* (163 U.S. 537 1896) was premised on equal access to schooling, the erosion of black voter protections in the latter part of the nineteenth century resulted in a parallel reduction in public support for black schools (Margo 1990). Regional variation in local control of schools and the tax base with which to fund schools led to wide within-race variance in education resources across the South. Looking again to 1930, we find that per pupil black expenditures had a standard deviation of \$6 (two-thirds of the mean) across counties and per pupil white expenditures had a standard deviation of \$181 (nearly three times the mean). From 1910 through 1940, these differences in interracial and intraracial gaps in school quality converged and diverged at widely varying rates.

returns to school resources that would be indicative of school quality standing in for local wage discrimination.[4] When we impute standardized test scores for our sample, the conditional annual earnings gap falls to 1–11 log points, while the weekly wage gap falls to 12–15 log points. At the same time, and echoing the previous literature (Wright 2013), we find that conditional wage gaps within occupations were smaller (6.7 log points annually and 15.1 log points per week worked). Human capital gaps had less bearing on occupational status gaps than wage and income gaps, implicating occupational sorting by race as one barrier to higher wages.

These results highlight the predominant role of public sector discrimination, as opposed to wage discrimination, in determining wage differences for black and white workers in 1940. We simulate counterfactual racial wage gaps under separate and equal school resources within southern counties. Blacks were disproportionately located in areas with weaker school funding overall; thus, the mandate only partially closes the regional black-white gap in average school quality. Still, the wage gap falls by 23%–43%. When we further allow educational attainment to be endogenous to school quality, the wage gap is reduced by up to 48%.

Our empirical design necessarily has limitations, two of which we highlight here. First, in order to accurately assign school quality, we restrict our analysis to young men and test for returns to human capital at the origin of the age-earnings profile. Results for older men may differ (Smith 1984; Smith and Welch 1989; Carneiro et al. 2005). Second, although we find little evidence of race-dependent returns to human capital (see Sec. V), we cannot completely rule out the possibility that human capital was endogenously determined in anticipation of lower wages for blacks. That is, if black students expected to receive little return to additional schooling or if local school authorities anticipated low returns to black school quality and allocated resources accordingly, we underestimate the level and the nuance of racial wage discrimination. Still, local decisions regarding school quality were highly political and plausibly disassociated from anticipated labor market returns (Goldin 2001; Carruthers and Wanamaker 2013, 2015; Cascio and Washington 2013). In cases where gains in school quality were credibly exogenous (e.g., the Rosenwald school-building campaign or women's enfranchisement), educational attainment rose substantially in response, and it did more so for blacks than for whites (Aaronson and Mazumder 2011; Kose et al. 2015). In other words, educational attainment for blacks, although perhaps limited by anticipated discrimination, was also limited by the supply of local education.

[4] Throughout the paper, we eschew county and state fixed effects as they may be correlated with unobserved discrimination. Still, we show in the appendix (available online) that results are robust to the inclusion of both county and state fixed effects.

Keeping these caveats in mind, the conditional wage gaps we estimate imply that discrimination was only somewhat more crippling for racial wage equality in 1940 than it was much later, when equal employment protections were in place. The discriminatory preferences of white southerners were powerful in limiting black public school quality and reducing the wages of young black men through the human capital channel. But they appear to have been less powerful in affecting outcomes through wage discrimination, at least in the occupations in which we observe southern black males in 1940. These findings do not rule out a role for federal employment policy interventions later in the century, but they do highlight the vast potential for schooling equality to remediate unequal pay.

II. Literature on the Black-White Wage Gap

Ex ante, one might expect the unexplained portion of wage differentials to be greater in 1940 than that observed later in the century. When driven by discrimination, racially separable wage equilibria depend on the number and size of discriminatory firms relative to the group being discriminated against. The size of the wage gap is a function of the disutility of employing workers in this group. Thus the gap is increasing in the prejudicial preferences of the general population because it both increases the number of discriminatory employers and the disutility of employing black workers.[5] In the period in question, extensive discrimination was evident in racially segregated job listings and stark differences in publicly reported salaries for black and white teachers.[6] Our analysis pre-dates the Civil Rights Act of 1964 and the associated employment and pay protections that outlawed labor discrimination against black Americans. Evidence that racially discriminatory views in the United States have declined over time gives rise to the idea that discrimination plays a smaller role in the black-white wage gap than it once did (Fryer 2011), though not so small as to go undetected in modern surveys or randomized audit studies (Bertrand and Mullainathan 2004; Charles and Guryan 2008).[7]

Still, economic historians have many times failed to detect racial wage discrimination in the early part of the twentieth century, observing that black and white workers received equal pay within particular occupations, even in

[5] These implications are true both in Becker's (1957) original framework and in adaptations to a search model as discussed in Lang and Lehmann (2012).

[6] See Goldin (1990). Across 10 southern states in 1930, white (black) teacher salaries averaged \$5.89 (\$2.55) per day in session. See Sec. III for sources and Margo (1984) for further discussion.

[7] See Lang and Lehmann (2012), figure 3, for evidence of a decline in prejudice measures after 1956. The decline continues through the racially charged 1960s. To our knowledge, no analogous data for 1940 exist.

the US South.[8] But this literature stops short of determining whether black and white workers across skill levels and occupations received equal wages conditional on human capital and is therefore not comparable to modern estimates of wage discrimination.[9] Largely due to the absence of data on workers' schooling, the empirical validity of equal pay for equally productive human capital prior to the Civil Rights era remains unknown.

For the modern era (after 1960 in this context), table 1 presents a limited review of papers measuring the contributions of schooling, school quality, experience, ability, and family background to the overall wage gap. The second column of the table lists the data source and cohorts used in each analysis. The third column indicates which human capital variables are included in the study, and the fourth column indicates what percentage of the overall gap they explain. The fifth column reports the log conditional black-white gap. The overwhelming indication from this literature is that pre-market factors matter for determining wage differences, and, in many cases, the wage gap potentially attributable to labor market discrimination is minimal after controlling for these factors. Estimates from later decades of the twentieth century draw on measures of educational attainment and school quality, as well as standardized test scores.[10]

In evaluating the importance of measured human capital for labor market wages in 1940, this paper speaks to a broader literature on the causes of racial income convergence over the twentieth century. Welch (1974), Smith (1984), Margo (1986), and Smith and Welch (1986, 1989) represent the earliest wave of research highlighting the rising quality and quantity of black education as important drivers of change in the black-white earnings ratio, but to date, direct tests of the impact of improving black school quality have relied on state-level school quality measures or have focused on a time period after employment protections were in place (Link and Ratledge 1975; Link, Ratledge, and Lewis 1976; Nechyba 1990; Card and Krueger 1992a, 1992b; Ashenfelter, Collins, and Yoon 2006). Notably, Card and Krueger (1992b) find a differential return to schooling across blacks and whites in the Census, which, in turn, can be attributed directly to differences in state-level school quality metrics. They conclude that 20% of the narrowing of

[8] The evidence is particularly consistent prior to the 1920s (Smith 1984; Fishback 1989; Smith and Welch 1989). Whatley and Wright (1994) and Wright (2013) cite evidence of a more substantial wage differential for entry-level workers by 1937, a difference they attribute at least in part to a yawning racial gap in human capital related to schooling. Another exception is a large within-occupation racial gap for black teachers relative to whites (Margo 1990).

[9] Alternatively, "(un)equal rewards to otherwise identical workers" (Whatley and Wright 1994).

[10] There are examples of standardized achievement microdata for particular states in the Jim Crow era, but we do not know of multi-state pre-market test data that could be incorporated into 1940 Census respondents' human capital profile.

Table 1
Decomposition of the Black-White Pay Gap: Role of Human Capital

Study	Data	Proxy for Human Capital	Percentage of Total Wage Gap Explained by Human Capital (%)	Log Gap Remaining after Controlling for Human Capital
Altonji and Blank (1999)	1980 CPS, all workers	Years of schooling	25	−.11
Altonji and Blank (1999)	1996 CPS, all workers	Years of schooling	43	−.12
Altonji and Blank (1999)	NLSY1979, age 29–37 in 1994	Years of schooling, AFQT	61	−.06
Altonji and Pierret (2001)	NLSY1979, age 27–34 in 1992	Years of schooling, AFQT, father's education, labor force experience	NA	−.133
Card and Krueger (1992b)	1960–1980 Censuses, age 21–60	Years of schooling, state-level school quality	20% of gap narrowing	NA
Carneiro, Heckman, and Masterov (2005)	NLSY1979, age 26–28 in 1990 and age 36–38 in 2000	Years of schooling, eighth-grade equivalent AFQT	53–65	−.144 – −.241
Fryer (2011)	NLSY1979, age 42–44	AFQT	Women: all; men: 72	Women: .127; men: −.109
Fryer (2011)	NLSY1997, age 21–27	AFQT	Women: 71; men: 39	Women: −.044; men: −.109
Fryer (2011)	College and Beyond 1976, approximate age 38	SAT	Women: 53; men: 44	Women: .286; men: −.152
Lang and Manove (2011)	NLSY1979, age 32–38	AFQT, years of schooling, school inputs	70	−.11
Neal and Johnson (1996)	NLSY1979, age 26–29	AFQT	100; men: 70	Women: insignificant; men −.072
Oaxaca and Ransom (1994)	1988 CPS, age 25 and up	Years of schooling	43	−.125
O'Neill, Sweetman, and Van de gaer (2006)	NLSY1979, age 29–31	AFQT, father's education, noncognitive skills	46–114, rising with wage quantile	Falls with quantile

NOTE.—Authors' calculations are from data presented in each cited work.

the black-white earnings gap between cohorts born in the 1920s and those born in the 1940s (measured between 1960 and 1980, in the midst and wake of the civil rights movement) is attributable to rising school quality.[11] This story of steady, continuous change is countered by Heckman and Payner (1989) and Donohue and Heckman (1991), who identify the Civil Rights era as a watershed in reducing entrenched labor market discrimination that hampered black economic progress. Wage gains for black Americans sharply deviated from underlying trends in the 1960s, implying that the federal anti-discrimination policies of that time were more important than rising human capital.

Finally, we note that a consistent feature of the literature is that estimated returns to school quality differ depending on whether data are at the state, district, or school level. The question of which level of aggregation is best in a standard wage model does not have an obvious answer. A primary limitation of the state level would seem to be aggregation (attenuation) bias, which is discussed explicitly in Morgenstern (1973). Local school quality should do a better job of characterizing school resources that were experienced by respondents. If so, and if school quality is indeed a component of marketable human capital, local data will explain more of the variation in individual wages. In other contexts, however, more granular school quality indices do not necessarily increase the estimated returns to schooling (Hanushek, Rivkin, and Taylor 1996; Betts 2010).[12] In the appendix (available online), we compare conditional wage gaps derived from state-level and county-level school quality measures in our sample, concluding on empirical grounds that county-level data are superior in this application. And in practical terms, county-level data allow us to fully populate a distribution of normalized school quality and capture nonlinear returns to Jim Crow era human capital.

[11] Several others have quantified the impact of school quality on earnings, per se, without a particular focus on the black-white gap. For our period, the most relevant of these are Morgan and Sirageldin (1968), Johnson and Stafford (1973), and Morgenstern (1973). Each use state-level data on school resources. Wachtel (1975, 1976) documents positive returns to school quality for a selected sample of individuals likely restricted only to whites. See Betts (2010) for a summary of the literature on the effect of school quality on earnings. See also Rizzuto and Wachtel (1980) for an estimate of the social rate of return to investments in school quality for whites and blacks separately in the 1960 and 1970 Censuses. Finally, see Orazem (1987) for evidence that differences in district-level school characteristics were the primary driver of differences in both attendance and student achievement test scores in Maryland between 1924 and 1938.

[12] Quoting Betts (2010, 313): "Most of the studies that find no link or a weak link between school inputs and student outcomes measure school inputs at the level of the actual school attended; studies that do find a strong effect typically measure school resources at the level of the state." More generally, it is not obvious that school-level metrics are the "correct" level of observation for this exercise as there may be endogenous within-county household sorting.

III. Data

The Public Use Sample of the 1940 US Census (Ruggles et al. 2010) is one of the earliest available micro-level data sets on wages for a cross section of the US population.[13] Prior to 1940, labor market measures in Census returns include occupation and industry of employment but no individual earnings data.

Census enumerators recorded labor market wages but not nonwage income. Consequently, almost all of the self-employed (including a substantial number of farmers and farm tenants) do not report income in this sample. As such, we exclude individuals without recorded earnings from our main results.[14] Occupational score results are robust to including some of these individuals, as we show in supplemental analyses discussed in the online appendix.[15]

In addition to earnings from the Census, we generate an occupational score for each individual in the sample based on their 1940 reported occupation. Our constructed occupational score variable is the average 1950 income, including nonwage earnings, among white males for each reported occupation in the 1940 manuscripts, mapped to a three-digit code.[16] The occupational score is best thought of as a potential wage or an index of occupational standing that is comparable across races, ages, genders, and geog-

[13] There are precious few sources for labor market earnings other than the Census prior to the advent of the National Longitudinal Surveys in 1966. Notable exceptions are the 1915 Iowa Census (Goldin and Katz 2000), where perhaps 1% of respondents were black, and the NBER Thorndike-Hagen sample exploited by Wachtel (1975, 1976), where the sample is believed to be limited to white males. A 100% sample of the 1940 Census has recently been made available to researchers, but several key variables have not been standardized in accordance with IPUMS practice, limiting the data set's usefulness for this application. We utilize the 1% Public Use Sample in this paper.

[14] Farm laborers remain in the sample. We acknowledge that we are not modeling selection into wage-earning occupations, although the appendix illustrates that human capital has a substantial effect on the likelihood of farming employment. Elsewhere, the literature has identified larger wage gaps for skilled occupations (Wright 2013). Thus, excluding lower-skilled farmers from the analysis may lead us to overstate the underlying black-white earnings gap. In the appendix, we perform a crude imputation of farmer wages as a sensitivity test.

[15] Another drawback of the 1940 Census is that we cannot adequately measure payments-in-kind as part of wages. This limitation is problematic given that there was substantial agricultural employment in this time period, even after omitting farm owners and tenants without reported earnings. Payments-in-kind may have been more prevalent for agricultural workers, and in particular, for black agricultural workers. Robustness checks described in the appendix show that findings are insensitive to the exclusion of respondents with more than $50 in nonwage earnings, which includes payments-in-kind as well as interest income and self-employment income.

[16] Again, we rely on Ruggles et al. (2010). We do not use the IPUMS-constructed "occscore" variable contained in their data because it is calculated using all workers, regardless of race, and thus encompasses wage discrimination.

raphies, abstracting from race-dependent occupational or regional sorting. Keeping this in mind, we are less interested in the score's cardinal properties, which are affected by the Great Compression leading up to 1950 (Goldin and Margo 1992), than its relative attenuation once we condition on human capital.

The 1940 Census contains measures of the highest grade completed by each individual. Census enumerators were instructed to record the "highest full grade that the person has successfully finished." Despite this instruction, there is some concern that the question was interpreted differently across races, especially for (predominantly older) black individuals who were educated in ungraded schools (Margo 1986). In that case, the Census instructions directed enumerators to record the number of years the person was in school.[17] Our focus on younger men in the data avoids much of this form of mismeasurement.[18]

We identify working men aged 18–25 from the Census Public Use Sample who reside in one of 10 southern states for which we have school resource data (described below): Alabama, Arkansas, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, and Texas. The age controls are designed to ensure that we can more accurately assign individuals to their county of schooling and reasonably ignore differential on-the-job training or experience.[19]

To measure the quality of schools available to each individual in our sample, we utilize transcribed county-level measures of race-specific school resources in the years leading up to 1940 for these 10 southern states. Over much of the twentieth century, each US state's department of education or equivalent office published a series of reports containing statistics on revenues and expenditures, disaggregated by county and by race. With the exception of a small number of biennial editions, education reports allow us to measure at least one race-specific school quality statistic for each year for each county. The data and data collection process are described more fully in the appendix.

The school quality data can be matched to each individual in the Census data after making some assumptions about where individuals lived when

[17] "If this cannot readily be determined, [enter] the number of years the persons attended school."

[18] To the extent that there remains overstatement of "highest full grade" for men with an ungraded education (perhaps because grades were typically completed in more than 1 year), it will serve to bias downward the contribution of differences in human capital to the black-white earnings gap and overstate the role of labor market discrimination. See Margo (1986) for evidence that using respondents' highest grade completed as a proxy for educational attainment understates the contribution of human capital to the 1930–70 decline in the wage gap.

[19] We cannot calculate labor market experience as the difference between age and years of schooling (plus 6 or 7) because age-in-grade differed significantly between black and white students (Collins and Margo 2006).

they were young. In 1940, Census takers inquired as to the location of respondents 5 years prior, in 1935. This detailed geographic information is newly released in the microsample files. We assume this 1935 location is the county of residence during an individual's potential schooling years. Because our analysis focuses on individuals aged 18–25 in 1940, this amounts to assuming that individuals aged 13–20 in 1935 reside in their county of schooling.[20] The ability to identify the probable county of education for individuals allows us to assign a race-specific school quality measure more proximate to the actual education experience of individuals than previously possible.

Depending on the year and state, school resource data are composed of one or more of the following eight metrics: expenditures per enrolled pupil, expenditures per pupil in average daily attendance (ADA), teachers per enrolled pupil, teachers per pupil in ADA, certified teachers per enrolled pupil, certified teachers per pupil in ADA, term length, and average teacher salary. Reported measures vary by state and year, but definitions are consistent across races within a state-year (and county-year). States had leeway in the metrics they chose to document, and many changed the format and set of resources reported over time. But, with rare exceptions, if a white-specific metric is reported in a given year, a corresponding black-specific metric is as well. An inventory of school quality statistics available for each state and year can be found in the appendix.

Selecting a single metric to proxy overall school quality is an untenable strategy. Each metric represents a different view of school resources and, more practically, varying availability of data across states within years and across years within states limits the scope of a given metric. Instead, for each

[20] We note that the 1930s were a decade of extremely low internal migration in the United States; 12% of our sample changed counties between 1935 and 1940, years when they were first entering the labor market. To further gauge the mismeasurement in this assignment, we look to a sample of death records from North Carolina generously shared with us by John Parman. The records include both county of birth and county of death for deaths reported prior to 1976. We examine a subset of males who were born in the relevant years (1914–23) and died between the ages of 7 and 20. For these individuals, we find that 30% lived in a state other than their home state by age 7, as indicated on their death certificate, and 38% did the same by age 20. (The numbers are 26% and 42% using a 3-year moving average.) Thus, by assuming that an individual observed at age 13–20 in 1935 lived in the same location at age 7 when they would have entered school, our methodology will falsely identify the county of education for up to 8% of individuals in the sample (16% using the moving average). The unfortunate assumption in this, and the only one we can reasonably make, is that individuals only move once so that the number who have relocated from their birth county by age 20 less the number who did the same by age 7 captures all migration. Note also that we undertake sensitivity tests on the county-of-schooling assumption in the appendix, finding little change in wage or occupation score gaps when we restrict the sample to respondents whose state of birth matches their 1935 state of residence.

quality measure, we calculate a Z-score relative to all other counties in the data that report the same quality metric for the same academic year. The index computation is:

$$Z_{jctr} = \frac{M_{jctr} - \bar{M}_{jt}}{\sigma_{jt}},$$

where M_{jctr} is the value of metric j (e.g., teachers per enrolled pupil) in county c in year t for race r and $\bar{M}_{jt}$ and σ_{jt} are the mean and standard deviation of measure j across all county-race observations reporting the same metric in year t. We emphasize that the conversion is relative to all county years reporting the same statistic and is across, not within, race. In this computation, Z_{jctr} converts statistic M_{jctr} to a scale with mean zero and unit standard deviation that can be compared across counties and races in year t. In state-years where more than one measure of school quality are reported, we use the average value of Z_{jctr} across all available j's.

$$Q_{ctr} = \sum_{j=1}^{J} Z_{jctr}/J,$$

where J is the total number of available metrics for county c in year t.[21] In cases where data on both enrolled pupils and pupils in ADA are available, we use only the measure per enrolled pupil (i.e., $J \leq 5$).

The index Q_{ctr} allows us to aggregate information about school quality across quality metrics that differ in their distribution and coverage. A Z-score index computed across, rather than within, years would provide a more accurate representation of counties' relative growth in school resources over time. However, a pooled index such as this would suffer more from missing data.[22] We note, however, that if the index is calculated across, rather than within, years, our results are little changed.[23]

[21] Computing Q_{crt} in this way implicitly weights each available school quality metric evenly. In the appendix, we present results from a permutation exercise that tests the sensitivity of results to 1,000 randomly generated alternative weighting schemes. We find that conditional gap estimates to follow are not exceptional within the distribution of possibilities, and moreover, are remarkably close to gap estimates from specifications that minimize the Akaike Information Criterion. Other robustness checks described in the appendix control for the quantity of school data: (i.e., J and J^2) and for the frequency with which each metric was reported during an individual's potential years of schooling.

[22] The current Q_{crt}, since it does not allow for shifts over time in the entire distribution of school quality, will overstate the true quality of school resources earlier in the panel. Age fixed effects ameliorate this issue, and our primary source of identifying variation in school quality is variation within cohorts, across segregated school systems.

[23] See Section 5 of the appendix.

Table 2
School Quality Index: Mean Values and Standard Deviations

	All Schools		Black Schools Only		White Schools Only	
	Mean	SD	Mean	SD	Mean	SD
All states	.029	.853	−.415	.889	.431	.574
Alabama	−.361	.775	−.943	.472	.225	.548
Arkansas	−.169	.649	−.462	.782	.064	.384
Georgia	−.206	.971	−.792	.947	.381	.549
Kentucky	−.018	1.206	−.094	1.652	.056	.448
Louisiana	.095	1.002	−.774	.451	.964	.543
Mississippi	−.223	1.274	−.677	1.089	.273	1.278
North Carolina	.095	.426	−.186	.310	.374	.333
South Carolina	.085	.889	−.759	.241	.929	.312
Tennessee	−.046	.625	−.037	.820	−.055	.362
Texas	.328	.750	−.147	.802	.679	.462

NOTE.—Authors' calculations are from annual reports of state education departments. The table gives the means and standard deviations of Q_{ct}, as defined in the text. The unit of observation is a county-race-year, and statistics are calculated across all years in the panel (1922–40).

Table 2 gives the average within-state variation in Q_{ctr} across all years in the school panel, both overall and separately by race.[24] The sample-wide standard deviation in this measure is 0.85, with higher variability for black schools (0.89) than for white (0.57), and with considerable variation within states as well as across. Even for white schools, where overall variation in the school quality index is smaller, within-state variation approaches or exceeds that for the full sample in the deep South (from Louisiana to Georgia). In order to assign school quality measures to individuals, we must make a final assumption. We observe individuals' ages and years of schooling in 1940, but not their precise years of enrollment, and age-in-grade distributions varied substantially. We assume that all individuals are "at risk" for school enrollment between the ages of 7 and 18, and we measure average school quality across those years.[25] As an example, an individual observed in the 1940 Census who is 25 years of age was a potential enrollee from the 1922–23 school year through 1933–34. For each individual, then, we assign a school quality measure that is the average of the school quality index in the county where he resided in 1935 over the years he could have been in school. Years for which

[24] The index is the average of several variables that are mean zero and standard deviation one but is not, itself, distributed accordingly. The unit of observation is a county-race-year unit.

[25] In practice, across-county variations in school quality measures are far more substantial than within-county differences across cohorts and attendance years, making the county of schooling assignment more important than the years of schooling one.

there are missing data on school quality are excluded from both the numerator and the denominator of Q_{ctr}. The resulting school quality index varies across race, cohorts, and counties and is best thought of as the typical amount of public school resources available to each respondent when they were ages 7–18.

The data linkage generates a base sample of 11,394 men aged 18–25 who report earnings, reside in 10 southern states in 1940, report a discernible residence in 1935 for which school quality metrics are available, and report race of either "white" or "black" to the Census enumerator. A critical issue for the empirical strategy described below is whether there is enough overlap in the school quality index of black and white respondents to justify a pooled regression and a simulated counterfactual. Nonlinearities in returns to education are included in the analysis, but a lack of common support across black and white school quality indices would hinder our ability to simulate black outcomes under a true "separate but equal" schooling system. Figure 1 compares kernel density estimates for black and white school quality indices and illustrates the distribution of black and white educational attainment. In both cases, distributions are distinct but overlapping. Areas of overlap lend support to the empirical strategy described in the next section, although we acknowledge that results will rely heavily on extrapolation. Robustness checks described in the appendix show that restricting our analysis to the common support of these two human capital measures (effectively omitting the bottom half of the black subsample) noticeably increases the conditional wage gap, which is consistent with greater degrees of discrimination among higher-skilled black men.

IV. Empirical Strategy and Results

Table 3 contains summary statistics of labor market outcomes, human capital measures, and other controls.[26] Columns 1 and 2 of the table give average values for all men in the sample, by race.[27] As noted before, a large number of men in the Census have no available income data, and columns 3 and 4 give the average value of these same characteristics for the sample used in the estimation. Due to the loss of nonwage agricultural workers, measures of school quality are slightly higher in the estimating sample, as is manufacturing value-added in counties of residence. In addition, the estimation sample is more urban than the underlying population. Section IV.A contains base-

[26] Labor force participation is measured in the reference week of 1940, while income references all of 1939. Some wage-earners in 1939 became nonparticipants by 1940. Results are robust to their exclusion, as well as the exclusion of a small number of self-employed men who report earning wages in 1939.

[27] The universe is all black and white men from the 1940 IPUMS sample aged 18–25 with a (discernible) county and state of residence in both 1935 and 1940 within our 10-state school quality region.

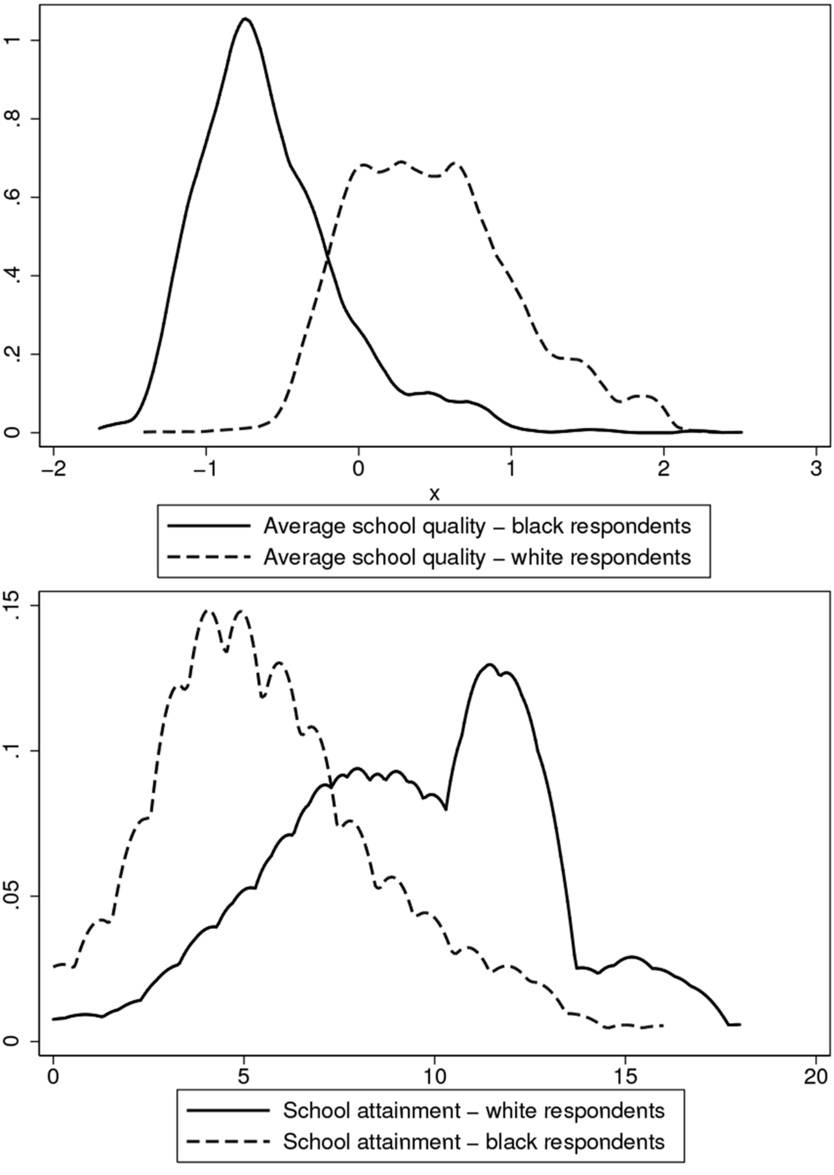

FIG. 1.—School quality and educational attainment kernel densities for black and white 1940 Census respondents.

line results for this working sample. In the appendix, we estimate the impact of pre-market factors on employment per se, agricultural employment, and New Deal work relief employment, all of which can be estimated on a larger sample of individuals.

Table 3
Summary Statistics

	All Black (1)	All White (2)	Baseline Sample Black (3)	Baseline Sample White (4)
Individual:				
Average annual wage income (1939) in natural log	5.42	5.93	5.42	5.93
Average weekly wage (1939) in natural log	1.87	2.40	1.87	2.40
% Reporting	58.8	56.0	100.0	100.0
Occupational score[a] in natural log	6.99	7.35	6.99	7.35
% Reporting	87.0	82.2	97.5	96.4
Average weeks worked (1939)	40.9	40.8	39.0	38.8
Unemployment rate[b]	9.2	9.5	9.9	8.9
If unemployed, duration (continuous weeks)[b]	38.8	45.4	35.4	43.8
Labor force participation rate[a]	88.9	85.0	98.2	97.5
Highest grade completed	5.6	8.9	5.6	8.9
School quality index (standardized (0,1))	−.55	.48	−.50	.55
State of residence in 1940:				
Alabama	12.8	9.0	12.3	8.7
Arkansas	5.6	6.5	4.0	6.0
Georgia	15.5	9.6	18.1	9.7
Kentucky	2.7	12.2	3.0	10.9
Louisiana	11.4	7.5	12.4	7.5
Mississippi	6.6	2.2	3.8	2.2
North Carolina	14.3	12.6	14.3	12.7
South Carolina	13.9	5.6	14.0	6.2
Tennessee	5.8	11.6	6.4	11.0
Texas	11.6	23.3	11.8	25.0
County of residence:				
Percent rural	68.4	67.2	64.2	62.0
Per capita manufacturing value ($1,940)	69.6	73.6	80.1	86.7
Per capita retail sales ($1,940)	.19	.20	.21	.22
Per capita crop value ($1,940)	59.3	53.0	53.4	46.3
Number of observations	5,423	14,849	3,141	8,253

NOTE.—Authors' calculations are from 1940 IPUMS data (Ruggles et al. 2010) and annual reports of state education departments. All variables are measured in 1940 unless otherwise noted. The sample includes all black and white males from the 1940 IPUMS sample aged 18–25 who lived within the 10 southern states covered by our school quality data with reported years of schooling and 1935 county of residence. Columns 3 and 4 contain only those individuals for whom earnings data are available.
[a] As of reference week (March 24–30, 1940).
[b] Up to March 30, 1940.

A. Baseline Results

Among the available labor market measures, it is clear that racial differences in labor force participation and employment rates from table 3 are relatively small. Labor market wages, on the other hand, differ substantially by race on both an annual and weekly basis.

To evaluate the impact of human capital measures on these labor market gaps, we rely on the following estimating equation, with primary focus on the gap in black and white wages, δ:

$$\ln Y_{icra} = \alpha + \delta \text{Black}_i + \beta X_{icra} + \varepsilon_{icra}, \tag{1}$$

where Y_{icra} is the labor market outcome of interest for individual i educated in county c residing currently in county r of age a. In this setting, Y_{icra} measures one of four labor market outcomes: weekly wages, occupation score, annual wages, or weeks worked. Black_i is a binary indicator, and the estimated wage gap is negative when black respondents have lower (conditional) labor market outcomes than whites. In all specifications, we cluster standard errors by the 1940 county of residence.

When X_{icra} is excluded from the estimation of equation (1), the parameter δ measures an unadjusted gap in earnings or occupational scores across races, or the difference in means across blacks and whites. The unconditional mean, reported in table 4 under the heading for each outcome, is 52.9 log points for weekly wages among the 11,394 individuals in the sample. The weekly wage gap reflects the combination of an unadjusted annual wage gap of 51.3 log points and a weeks worked gap of 1.6 (log) weeks favoring blacks, leading to a weekly wage gap that exceeds the annual income gap. The unconditioned occupational score gap measures 35.9 log points.[28]

Unconditional gaps may be attributable to many things: differences in labor market productivity and price levels across locations, differences in observable human capital between black and white workers, unobserved racial differences in human capital, and, finally, labor market discrimination. We first control for differences in labor market productivity and price levels across locations and age fixed effects in order to isolate the role of human capital in the next step. In this case, δ represents the racial gap in labor market outcomes net of differences in the age structure and counties' economic profiles. Our preferred model does not include county fixed effects, since we expect the intensity of wage discrimination to have varied spatially (Sundstrom 2007),[29] although, as we show in the appendix, the conditional wage gap we estimate is robust to the inclusion of county (or state) fixed effects. Instead, we control for characteristics of the individual's 1940 county of residence, which may have affected wage levels in a marginal product sense: urbanicity of the county of residence (to proxy for cost of living differences) as well as per capita manufacturing value added, per

[28] We limit the occupational score analysis to those who also report wages to keep samples consistent, although they differ somewhat due to individuals who do not report an occupation but do report wages. We relax this constraint in robustness tests discussed in the appendix.

[29] A county fixed effect is an unobserved factor that affects average wages in that county, which will include the degree of discrimination for that county. The same can be said for state fixed effects.

Table 4
Estimates of Black-White Labor Market Outcome Gaps

	Outcome							
	ln(Weekly Wage)				ln(Occupation Score)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Unconditional gap (SE)	−.529 (.024)				−.359 (.016)			
Black-white gap	−.490	−.315	−.181	−.191	−.334	−.203	−.160	−.168
	(.022)	(.022)	(.031)	(.032)	(.014)	(.015)	(.021)	(.022)
Contribution of school quality			−.140	−.011			−.044	.136
			(.022)	(.046)			(.016)	(.041)
Contribution of educational attainment			−.168	−.164			−.129	−.129
			(.010)	(.011)			(.007)	(.008)
Contribution of interaction				−.123				−.174
				(.043)				(.038)
Age and county controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Years of schooling?	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School quality?	No	No	Yes	Yes	No	No	Yes	Yes
Interacted HC controls?	No	No	No	Yes	No	No	No	Yes
N	11,394	11,394	11,394	11,394	11,021	11,021	11,021	11,021
Adjusted R^2	.24	.29	.30	.30	.18	.25	.25	.26

	Outcome							
	ln(Annual Wages)				ln(Weeks Worked)			
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Unconditional gap (SE)	−.513 (.027)				.016 (.014)			
Black-white gap	−.467	−.265	−.116	−.137	.023	.050	.065	.054
	(.024)	(.027)	(.037)	(.038)	(.014)	(.016)	(.024)	(.024)
Contribution of school quality			−.155	.080			−.014	.092
			(.026)	(.050)			(.018)	(.043)
Contribution of educational attainment			−.196	−.184			−.027	−.020
			(.013)	(.014)			(.007)	(.010)
Contribution of interaction				−.226				−.102
				(.046)				(.040)
Age and county controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Years of schooling?	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School quality?	No	No	Yes	Yes	No	No	Yes	Yes
Interacted HC controls?	No	No	No	Yes	No	No	No	Yes
N	11,394	11,394	11,394	11,394	11,394	11,394	11,394	11,394
Adjusted R^2	.22	.26	.27	.28	.04	.05	.05	.06

NOTE.—Authors' calculations are from 1940 IPUMS data (Ruggles et al. 2010) and annual reports of state education departments. HC = human capital. Unconditional gaps are found under outcome headings. Columns 1, 5, 9, and 13 represent coefficients from an regression of log weekly wages, log occupational score, log annual wage, and log weeks worked, respectively, on race indicators, as well as age fixed effects and county covariates. Columns 2, 6, 10, and 14 add a cubic function of educational attainment to the regression. Columns 3, 7, 11 and 15 include cubic functions of educational attainment and school quality. Columns 4, 8, 12, and 16 include controls for cubic functions of human capital and their complete interaction. Gelbach (2016) contributions of school quality and years of schooling in attenuation of the black-white gap (i.e., the difference between the age and location-adjusted wage gap in cols. 1, 5, 9, and 13 and the human capital-adjusted model in cols. 3–4, 7–8, 11–12, and 15–16) are listed below coefficient estimates. County covariates include the percent urban population, crop value per capita, retail sales per capita, and manufacturing value added per capita. Standard errors are in parentheses.

capita retail sales, and per capita crop value to represent average productivity in manufacturing, services, and agriculture, respectively. These variables are summarized in table 3. These controls, along with age fixed effects, lower wage and occupational score gaps, albeit very slightly, while raising the gap in weeks worked (cols. 1, 5, 9, and 13 of table 4).

Next, we include a third-degree polynomial function of educational attainment in X_{icra} to capture nonlinearities in the impact of years of schooling on labor market outcomes (cols. 2, 6, 10, and 14). We round out respondents' human capital profiles in columns 3, 7, 11, and 15 with cubic functions of both attainment and school quality. Because our model continues to include age fixed effects, human capital controls account for within-cohort variation in years of schooling and in school quality across counties and races. Table 2 indicates that school quality variation is broad-based and is present both within and between states. The importance of school quality as a predictor of labor outcomes is readily apparent. Comparing columns 1–3 of table 4, for instance, the weekly wage gap narrows by 36% when we condition on years of schooling alone, versus 63% when we condition on attainment as well as school quality.

Although human capital may have narrowed wage gaps, their closure was inhibited by occupational sorting. Results for occupation scores (cols. 5–7) indicate that human capital differences explained only 55% of the column 5 black-white gap in occupational standing.

We note that school quality and individual attainment are highly correlated, and it is not clear from δ estimates alone which measure of human capital is primarily responsible for attenuating the black-white earnings gap. Adding covariates sequentially—as we do in the second and third specifications for each outcome in table 4—is one approach to disentangling the contribution of school quality from that of attainment. But Gelbach (2016) shows that this procedure can lead to misleading results and that such decompositions depend on the order in which controls are added to the model. We use Gelbach's decomposition framework to estimate the relative contribution of years of schooling and school quality (both in cubic functions) to wages and occupational scores.[30] In table 4, the contribution of each term (in log points) is displayed beneath the δ coefficient in columns 3, 7, 11, and 15. Differences in educational attainment account for 16.8 log points of the

[30] Gelbach's (2016) procedure stems from the identity

$$\hat{\beta}_1^{\text{base}} = \hat{\beta}_1^{\text{full}}(X_1'X_1)^{-1}X_1'X_2\hat{\beta}_2,$$

where $\hat{\beta}_1^{\text{base}}$ ($\hat{\beta}_1^{\text{full}}$) is a vector of X_1 coefficients in the limited (fully conditioned) model and $\hat{\beta}_2$ is a vector of X_2 coefficients. In our context, X_1 is a race indicator and X_2 factors are years of schooling and school quality, two contributing factors to race-based differences in labor outcomes.

attenuation of the black-white gap between columns 1 and 3, while school quality accounted for a somewhat smaller 14.0 log points. Similar results obtain for annual wages. In contrast, educational attainment accounts for a clear majority of attenuation in the occupational score gap.

The primary threat to the internal validity of results is that of a classic omitted variable, correlated with both human capital and earnings in a way that falsely attributes labor market gaps to human capital or (implicitly) to discrimination. The foremost concern on this dimension is that school quality measures embodied in X_{icra} in equation (1) are proxies for county race relations in general. If so, it is not at all surprising that including school quality in a wage equation goes a long way toward explaining the racial wage gap. Black southerners could have experienced discrimination in both the labor market and in decisions that affected school quality. If those two factors are highly correlated, local school resources available to blacks may simply be proxies for overall relations and table 4 results understate the role of wage discrimination in explaining the earnings gap and overstate the role of school resource provision. A second, and related, possibility is that public agencies provided substandard black schools because they had reason to be pessimistic about blacks' relative earnings later on.

In recognition of these concerns, our preferred specification includes both domains of human capital and a full interaction between the cubic in school quality and the cubic in educational attainment ("Interacted HC controls" in table 4). This allows us to assess the relative contributions of school quality on its own and school quality as it serves to enhance educational attainment, the former being more indicative of omitted discrimination in the presence of the latter.[31] These results are located in columns 4, 8, 12, and 16. In each case, the conditional gap is little changed and the isolated contribution of school quality toward changing the unconditional wage gap is small, imprecise, and/or positive. The majority of the school quality effect for weekly and annual wages comes through its interaction with educational attainment.[32] This model, which generates larger conditional gaps than the uninteracted model, is our preferred specification for the remainder of the paper.

In table 5, we report the estimated return to educational attainment and school quality by evaluating the marginal effect at the mean of each variable

[31] We further address the underlying identification threat by limiting our analysis to migrant blacks whose county of residence was not their county of schooling (Sec. IV.B), by conditioning on imputed AGCT scores (Sec. IV.C), by conditioning on fixed effects for state of residence, county of residence, or county of schooling (appendix), and by examining an Oaxaca decomposition of the black-white gap (Sec. V).

[32] This is also consistent with work by Card and Krueger (1992a), who find that higher school quality, measured at the state level, is associated with enhanced returns to schooling.

Table 5
Returns to Human Capital in Equation (1)

	ln(Weekly Wage)	ln(Occupation Score)	ln(Annual Wage)	ln(Weeks Worked)
Marginal effect of school quality	.137	.046	.162	.025
	(.048)	(.033)	(.057)	(.043)
Marginal effect of educational	.054	.040	.058	.004
attainment	(.006)	(.004)	(.008)	(.006)

Note.—Authors' calculations are from 1940 IPUMS data (Ruggles et al. 2010) and annual reports of state education departments. The table contains estimated marginal effects for school quality or educational attainment, evaluated at the mean. Bootstrapped standard errors (from 1,000 replications within 10% random subsamples) are in parentheses.

(which enters the preferred specification as an interacted cubic) and bootstrapping standard errors. A one unit increase in the school quality index, evaluated at the mean, brings labor market returns of between 14 and 16 log points with respect to earnings and an insignificant 5 log points with respect to occupational standing. The return to years of education reflects standard measures of the same from elsewhere in the literature: 6 log points of increased earnings and 4 log points of occupational standing. For a more detailed illustration of coefficients on school quality and attainment, we refer readers to surface plots and related discussion in Section 1 of the appendix. Plots indicate that gains in school quality (attainment) conditional on one's attainment (school quality) lead to greater earnings and occupational status, as expected. Coefficients on other equation (1) controls are also given in Section 1 of the appendix.

From our preferred specification, we conclude that differences in human capital measures account for the majority of the black-white wage and occupational score gaps in 1940. The remaining racial gaps, 19.1 log points for weekly wages and 13.7 log points for annual wages, represent 64% and 73% reductions, respectively, from the unconditional wage gaps. To put these residual differences in perspective, note that 11 log points is the modal conditional wage gap estimate listed in table 1 and that 19.1 log points lies at the higher end of the range of these post–Civil Rights era estimates.

B. Collinearity in Public and Labor Market Discrimination

One major impediment to interpreting conditional wage gaps as wage discrimination is the possibility that local wage discrimination is collinear with local school quality. If so, then the measured contribution of school quality to the overall wage gap may be misattributed. As a first step in ensuring that this attribution is appropriate, we showed in Section IV.A that school quality has an impact on wages primarily through its interaction with years of schooling, as would be expected if school quality was not simply proxying for wage discrimination. In this section, we take additional steps to ensure the same. Because decisions about local school quality may have reflected local discriminatory attitudes, we shift our attention to the relative labor market outcomes of black intercounty migrants—those who were likely educated in counties other than their 1940 county of residence. Focusing on these black migrants breaks any collinearity between school quality and labor market discrimination in each individual's 1940 county of residence (although it adds the question of selection into migration).[33] If human capital measures are less successful at explaining the earnings gap of new resident blacks relative to all resident whites (i.e., if the conditional gap is much larger between these two groups than between blacks and whites more broadly), then we would suspect that estimates in table 4 are driven by local race relations more so than local school quality.

About 1 in 10 blacks from the main sample changed counties between 1935 and 1940. Migrating blacks had higher wages than their nonmigrating peers, and this generates lower baseline wage gaps, as evidenced in table 6, columns 1 (weekly wage) and 7 (annual wage).[34]

The remainder of table 6 reports results from equation (1) with nonmigrating blacks excluded. If collinearity between school quality and labor market discrimination was driving our main results, we would expect human capital controls to register little change in the wage gap between whites and migrant blacks. But column 3 lists a nearly equivalent conditional weekly wage gap as the one reported in table 4 (19.2 vs. 19.1 log points). The conditional black-white gap in annual wages falls to 6.2 log points (col. 9 vs. 13.7 log points in the baseline) and is statistically indistinguishable from zero.

We do find some evidence that the conditional occupation score gap grows after eliminating nonmigrant blacks, from 16.8 in table 4 to 27.6 log points in table 6, column 6. The implication is that the baseline occupational score gap closure we observed for this outcome—with variance limited by occupational sorting—may have been driven by unobserved variables more so than human capital.

Still, we take wage results in columns 3 and 9 of table 6 to indicate that school quality measures are not simply serving as a proxy for local race relations in determining pay. The fact that human capital controls result in a null annual pay gap when we limit the black population to intercounty

[33] One concern with this exercise is that migrants are not moving far enough to actually decouple labor market and school quality discrimination. The average distance migrated between 1935 and 1940 for black males in the analytical sample is 83 miles, with a standard deviation of 107. Roughly 15% of intercounty black migrants in this sample moved between states. The correlation coefficient between school quality in migrants' 1935 and 1940 county of residence is 0.54, a subjectively intermediate rate of co-movement. We remind the reader that there is substantial intrastate variation in school quality.

[34] Summary statistics for the migrating sample are available in the appendix, table 9.

Table 6
Estimates of Black-White Labor Market Outcome Gaps, Excluding Nonmigrant Blacks

	Outcome											
	ln(Weekly Wage)			ln(Occupation Score)			ln(Annual Wages)			ln(Weeks Worked)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Black-white gap	−.411	−.428	−.192	−.360	−.354	−.276	−.415	−.439	−.062	−.004	−.011	.130
	(.048)	(.055)	(.072)	(.038)	(.036)	(.050)	(.065)	(.072)	(.097)	(.039)	(.039)	(.057)
Age and county controls?	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Interacted HC controls?	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
N	8,464	8,464	8,464	8,166	8,166	8,166	8,464	8,464	8,464	8,464	8,464	8,464
Adjusted R^2	.01	.17	.26	.01	.12	.23	.00	.20	.27	.00	.06	.07

NOTE.—See table 4 for estimates notes and definitions. Black males who did not migrate between counties over the period 1935–40 are excluded from the estimation. Standard errors are in parentheses. See further discussion in Sec. IV.B.

migrants is itself an interesting conclusion, with the caveat that migrating blacks are very small in number and perhaps positively selected, even after controlling for observable human capital.[35]

C. Adding Proxies for Unobserved Ability

A common concern in the discrimination literature is unobserved human capital heterogeneity that manifests as a wage gap. For example, if unobserved average ability, conditional on observable human capital, differs across blacks and whites, the conditional wage gap does not accurately reflect the depth of labor market discrimination. Using modern data, Lang and Manove (2011) show that black students tend to have more years of schooling for a given AFQT-measured ability. If the same holds true in the pre-war period, omitting ability measures—as we do in Section IV.A—may misstate the conditional pay gap. It may well have been the case that only the highest-ability black students would have achieved higher levels of educational attainment given the pervasive impediments to attendance. On the other hand, it may be that blacks were conditionally less productive due to differences in the intergenerational transmission of human capital or other pre-market investments outside of schooling (e.g., access to health care). If so, we understate the role of human capital and overstate the potential role of labor market discrimination.

A leading proxy for unobserved ability used in the modern literature is performance on a standardized exam.[36] Several studies examining features of the wage gap in the NLSY panels utilize AFQT (Armed Forces Qualification Test) scores as pre-market proxies for ability (see table 1 for examples). Through a historical fluke, standardized test scores are available for a subset of World War II enlistees.[37] Unfortunately, directly linking our

[35] A corollary to our logic here is that school quality from migrants' 1940 county of residence should be unimportant conditional on school quality from assumed counties of schooling. In an extended analysis of migrants' wages and human capital (available on request), we limit the analytical sample to black and white migrants and control for cubic functions of school quality from 1935 and 1940 counties of residence, both fully interacted with a cubic in years of schooling. For all but weeks worked, Wald tests support the importance of migrants' assumed school quality and rule out the importance of school quality in their 1940 locations.

[36] Parental education or family background indicators are other candidate proxies for unobserved ability. In the 1940 Census, however, this information is only available for respondents who were still living with their parents (57% of the analytical sample of 18–25-year-old males).

[37] For a limited time in 1943, World War II enlistment cards contained AGCT scores in place of weight. We know of no evidence that this test was racially unbiased. Enlistment in the armed forces, however, was conditional on a minimum literacy standard, so the test results would not have been racially biased for literacy reasons. Like the modern AFQT, the test appears to measure acquired ability rather than inherent cognition (Zeidner and Drucker 1983). See Troesken, Ferrie, and Rolf (2012) for additional details.

1940 IPUMS (Integrated Public Use Microdata Series) sample of males to the World War II enlistment data for this window of time, matching on name and county of residence, generates too small a sample for meaningful analysis. Instead, we assign human capital measures to each individual in the World War II records where AGCT is recorded, utilizing the fact that World War II records include a measure of highest grade completed, much like the 1940 Census sample, as well as an exact county and state of residence at enlistment. For these World War II enlistees, then, we have a measure of race, educational attainment, school quality, and estimated ability.[38]

We can use this sample to impute AGCT scores for men in the 1940 Census conditioning on educational attainment, school quality, and race. Although these are imputed measures, they do address the fundamental omitted variable question: conditional on observable human capital, does ability add explanatory power to the wage model, and therein, the estimated wage gap? We wish to remain as agnostic as possible about the relationship between race, school quality, attainment, and AGCT scores, recognizing that black-white differences in AGCT conditional on other human capital metrics may differ across the distribution of those metrics. After limiting the age range of World War II enlistees to mirror that of our Census sample, we use three different methods to impute AGCT for the baseline sample of Census respondents.[39]

- *Method 1*: We subdivide school quality into 10 deciles and calculate enlistees' average AGCT scores within each race/educational attainment/school quality decile bin ($R^2 = 0.518$). We then assign an imputed AGCT score to each individual in the Census sample accordingly, with the restriction that bins with fewer than 25 observations in the enlistee data are omitted.
- *Method 2*: We specify that enlistee AGCT is a function of a 5th-order polynomial in school quality within each race/educational attainment bin ($R^2 = 0.547$). We then use the parameter estimates to impute AGCT for the Census sample.

[38] The assumption that allows us to link school quality to individuals is that their county of residence at enlistment is the county where they were educated. This assumption is more likely to produce bad matches than the assumption used in the main analysis due to migration between schooling and enlistment.

[39] Neal and Johnson (1996) are careful to limit the AFQT scores in their analysis to those taken prior to entrance in the labor market, arguing that: "Job experience and post-secondary education surely enhance human capital and will therefore increase test scores. If discrimination limits access to these human capital investments, then postentry discrimination contaminates the test scores (873)." Because the AGCT test was reported at enlistment, the youngest age at which we observe this score is age 18, with a large mass of observations at age 19. Seventy-five percent of the individuals in the World War II enlistment records used for ability imputation are age 20 and younger.

- *Method 3*: We estimate enlistee AGCT by building a regression tree with up to 400 "leaves," twice as many as implied by method 1 (R^2 = 0.534). Branches are determined by a machine learning algorithm that minimizes the mean squared error of predicted versus actual AGCT.

With these three methods for imputing AGCT scores, we revisit equation (1) with imputed AGCT as an additional control variable. Each of the control variables from which AGCT is imputed is present elsewhere in the Mincer model, and it is reasonable to question whether this approach adds any new information to the model above.[40] The estimates presented thus far make it clear that there is a relationship between race and labor market earnings that cannot be explained by school quality or educational attainment. Without additional information, we cannot discern whether this relationship is due to discrimination or due to unobserved human capital. The additional information from AGCT scores allows us to control for the average ability of men in the World War II enlistment records with similar measures of observable human capital, separately by race, narrowing the scope for omitted human capital variables in our wage models.

Results for weekly and annual wages are located in table 7. The black-white wage gap falls substantially when we control for estimated ability. The weekly wage gap falls to 12.5–14.6 log points, depending on the imputation method. For annual wages, residual gaps range from 1.1 to 11.1 log points.

As a falsification exercise, we repeat the second imputation method for World War II enlistees' weight in the same months of 1942, when the weight field should have contained physical weight and not AGCT scores. That is, we estimate enlistees' physical weight as a function of the school quality polynomial by race/attainment bin (method 2) and then map parameter estimates to the 1940 Census sample, "predicting" respondents' weight. Results are found in columns 5 and 10 of table 7. The conditional wage gap rises slightly from the baseline to 20.1 log points and the annual wage gap to 14.2. The contrasts between columns 4 and 5 and between columns 9 and 10 are telling: estimating ability by its known relationship to observables adds important information to the model, whereas adding an anthropometric measure has little bearing on the wage gap.[41]

A final caveat on the imputation of AGCT scores is that the sample of individuals in the 1943 enlistment records may be selected in some way.

[40] A technical note is that the effect of AGCT would not be separately identifiable in eq. (1) if AGCT scores were imputed with the same functional form.

[41] In addition, actual weight is insignificantly predictive of wages in the estimation for 1942 but predicted AGCT is positively, significantly correlated in 1943, further indication that the 1943 "weight" data do measure ability.

Table 7
Estimates of Black-White Labor Market Outcome Gaps, Including Unobserved Ability Estimates

	Outcome									
	ln(Weekly Wage)					ln(Annual Wages)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Unconditional gap (SE)			−.529 (.024)					−.513 (.027)		
Black-white gap	−.191	−.125	−.138	−.146	−.201	−.137	−.011	−.037	−.111	−.142
	(.032)	(.041)	(.043)	(.037)	(.033)	(.038)	(.049)	(.047)	(.042)	(.037)
Baseline: fully interacted HC	✓					✓				
Method 1: binned AGCT estimates		✓					✓			
Method 2: AGCT predicted by polynomial			✓					✓		
Method 3: AGCT predicted by regression tree				✓					✓	
False AGCT					✓					✓
N	11,394	11,261	11,394	11,394	11,394	11,394	11,261	11,394	11,394	11,394
Adjusted R^2	.30	.30	.30	.30	.30	.28	.28	.28	.27	.28

NOTE.—Authors' calculations are from 1940 IPUMS data (Ruggles et al. 2010), World War II enlistment records, and annual reports of state education departments. HC = human capital; IPUMS = Integrated Public Use Microdata Series; AGCT = Army General Classification Test. Standard errors are in parentheses. See specifications and further discussion in Sec. IV.C.

This would impact the estimate of δ only if selection differed across blacks and whites. We do observe higher educational attainment in the World War II enlistment records than in the Census sample for both blacks and whites, perhaps due to the literacy restrictions on enlistment. The difference in (log) average educational attainment between the enlistment records and the Census records is larger for blacks than whites, indicating more positive selection on observables among blacks. Educational attainment is explicitly accounted for in the imputation, but if the same selection pattern is true for AGCT scores as well, then we overestimate the ability of blacks in the Census sample and our estimates of δ in table 7 are biased away from zero.

V. Racial Differences in the Returns to Skill

We have thus far restricted the role of race in determining labor market outcomes to a constant intercept shift captured by the δ coefficient in the wage regression. In practice, race may interact with other controls to determine wages in ways not accounted for in the preferred specification and, if so, our estimates of discrimination may be understated.[42] In this section, we use Oaxaca decompositions to value racial differences in endowments, racial differences in the returns to human capital, and the interactions of these two in determining the overall wage gap.[43]

We begin by estimating equation (1), separately by race:

$$\ln Y_{icra} = \alpha + \beta X_{icra} + \varepsilon_{icra},$$

where X_{icra}, as before, controls for school quality, years of schooling, the interaction of quality and years of schooling, age fixed effects, and local characteristics.

[42] We also note that there is some evidence that the absolute value of δ rises with worker skill, indicating that controls have less explanatory power over racial wage gaps at higher values of human capital. In appendix table 5, we calculate conditional wage gaps at the common support of black and white human capital, i.e., the highest skilled blacks. The gap in weeks worked disappears at this skill level, so the conditional gaps in weekly and annual wages are indistinguishable from each other. The conditional wage gap rises to 25 log points for this subset of men. A similar pattern obtains if we restrict the analysis to blacks who report an occupation of "operative," a more skilled category than the modal "laborer." (Results are not shown.) The preferred specification allows for differences in the return to skill across the skill distribution, but not for differences in the return to skill by race. The cost of this simplification is summarized in the analysis in this section.

[43] We also evaluate whether the addition of race interactions with each of the covariates improves the predictive capabilities of the model. These results, available upon request, demonstrate that this more general model offers little predictive advantage (as measured by reductions in mean squared error) over our preferred specification.

Table 8
Decompositions of the Black-White Wage Gap

	Outcome	
	ln(Weekly Wage) (1)	ln(Annual Wage) (2)
Black-white gap		
Baseline difference	−.529	−.513
Oaxaca decomposition:		
Difference due to endowments	−.311	−.304
	(.044)	(.053)
Difference due to β's	−.104	.246
	(.099)	(.129)
On cubic in educational attainment	−.128	−.531
	(.294)	(.381)
On cubic in school quality	−.043	.753
	(.392)	(.508)
All other covariates	.067	.025
	(.139)	(.173)
Difference due to interaction	−.114	−.455
	(.108)	(.139)

Note.—Authors' calculations are from 1940 IPUMS data (Ruggles et al. 2010) and annual reports of state education departments. See Sec. V for a discussion. Standard errors are in parentheses.

Black-white gaps in annual and weekly wages are decomposed as follows:

$$\bar{X}_W\beta_W - \bar{X}_B\beta_B = (\bar{X}_W - \bar{X}_B)\beta_B + \bar{X}_B(\beta_W - \beta_B) + (\bar{X}_W - \bar{X}_B)(\beta_W - \beta_B).$$

The first right-hand-side term is the contribution of endowments to the wage gap, the second term is the contribution of coefficients (i.e., race-specific differences in returns to X_{icra} elements including human capital), and the third term is the contribution of the interaction of the two.[44] The value of each is reported in table 8 for both weekly and annual wages, with the contribution of coefficients broken down further by the contribution of school quality, educational attainment, and the remaining covariates.

Table 8 indicates that differences in endowments are the predominant determinant of racial differences in weekly and annual wages, each accounting for 59% of their respective unconditional gap. That leaves 21.8 and 20.9 log points to be explained by coefficients and the interaction of gaps in endowments and coefficients, or 2–7 log points on top of baseline conditional wage gaps reported in table 4. Results for coefficient differences, however, leave us with little guidance as to which (if any) endowments are valued for blacks differently than for whites. Standard errors are large enough to render point estimates statistically insignificant at conventional levels, and

[44] For more discussion of this methodology, see the detail in Biewen (2014).

coefficient differences for some covariates (e.g., school quality with regards to annual wages, local characteristics with regards to both outcomes), if they were precise, would indicate that returns to these variables work against the wage gap. Beyond endowment gaps per se, much of the wage gap is explained by the interaction of endowment and coefficient gaps.

Racial differences in the returns to human capital, or lack thereof, also speak to the outstanding question of whether modest 1940 wage gaps are simply the result of endogenous choices by workers (choosing attainment) or school agencies (choosing school resources for segregated systems). Equivalent returns to skill implied by table 8 are consistent with an optimized simultaneity between the human capital gap accumulated in the 1930s and the wage gap realized in the 1940s, that is, one where black and white human capital investments are made up to the point that marginal discriminatory returns are the same across races. This particular optimization is more plausible for individual workers choosing how long to stay in school than school systems facing the objectives of a political economy. We return to the issue of endogenous attainment in the following section.

Because differences in the returns to human capital appear to be a minor portion of the overall decomposition in table 8, we do not expect allowances for differences in returns to be critical for inference. When we evaluate a counterfactual "separate but equal" scenario in Section VII, we allow returns to school quality and educational attainment to differ by race, motivated by the discussion above. In this exercise, our results are not much changed relative to inference from pooled coefficients, indicating a small role for racial differences in the return to skill across the distribution of school quality and educational attainment modeled below. Whether this is true at even higher levels of skill (i.e., as the quality and quantity of black education rose in the years following 1940) is an important area of future inquiry.

VI. Within-Occupation Wage Gap Results

The fact that human capital explains relatively less of the occupational standing gap than the wage gap raises the question of whether occupational sorting itself was the predominant driver of the black-white wage gap (Higgs 1977).[45]

Figure 2 plots the distribution of black and white workers across nine broad occupation categories, along with average log annual wages for each category. There are substantial differences in earnings across occupation categories, and the wage measures layered on the histogram indicate, unsurprisingly, that wages are higher in the occupations where whites are disproportionately represented.

[45] According to Higgs (1977), racially dependent sorting across high-wage and low-wage firms is another likely source of the overall black-white gap. We do not observe respondents' employers in the 1940 Census.

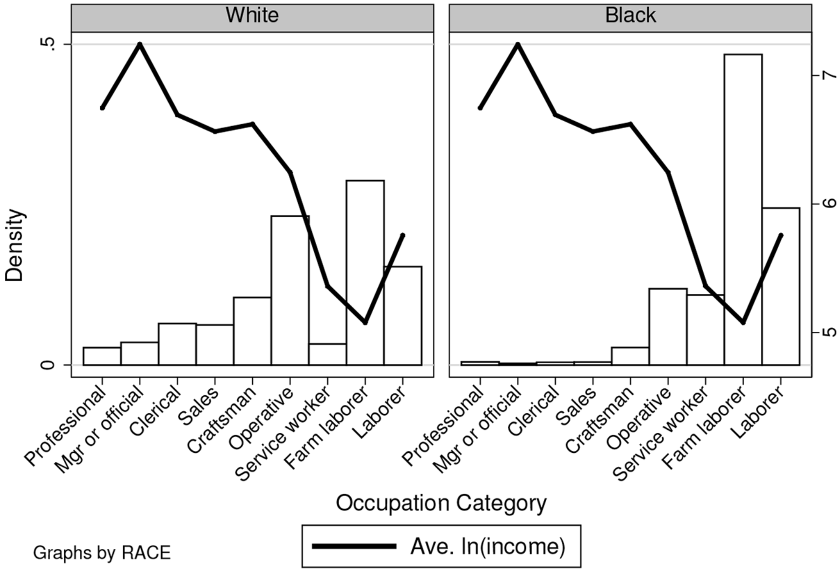

FIG. 2.—Occupation distributions for black and white 1940 Census respondents. Average (log) wages within those occupations, across both races, are reported on the second axis.

Results for occupation score in table 4 reflect the expected size of the racial wage gap if there was no within-occupation wage discrimination and discrimination happened via sorting alone. To calculate the size of the within-occupation gap, we add fixed effects for occupation categories to equation (1) and measure the resulting conditional gap in weekly and annual wages and in weeks worked.[46] We continue to condition on county covariates, cubics in educational attainment and school quality, and the interaction of school quality with attainment measures. Estimates of δ in table 9 show that the within-occupation weekly wage gap is reduced from 28.8 to 15.1 log points following the addition of human capital controls, and that the annual income gap falls from 21.2 log points to an insignificant 6.7 log points. The difference between columns 4 and 8 is reconciled by the conditional gap in weeks worked, which expands with controls for occupation fixed effects. Within occupations, black and white men of equivalent schooling had similar annual earnings, but given that blacks worked 8.4% more weeks than whites with a similar human capital profile, the weekly earnings gap remains significant at 15.1%. These findings support the idea that discrimination manifested in part via occupational sorting, but not so much as to result in equal pay within occupations.

[46] Results are robust to finer occupation definitions.

Table 9
Estimates of Black-White Labor Market Outcome Gaps, with Occupation Fixed Effects

	Outcome											
	ln(Weekly Wage)				ln(Annual Wages)				ln(Weeks Worked)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Black-white gap	−.288	−.294	−.140	−.151	−.212	−.221	−.046	−.067	.077	.073	.094	.084
	(.018)	(.018)	(.028)	(.029)	(.023)	(.022)	(.035)	(.035)	(.015)	(.015)	(.026)	(.026)
Contribution of school quality			−.090	−.027			−.094	.040			−.004	.067
			(.020)	(.042)			(.023)	(.050)			(.017)	(.042)
Contribution of educational attainment			−.061	−.053			−.077	−.068			−.016	−.016
			(.006)	(.008)			(.008)	(.009)			(.005)	(.007)
Contribution of interaction				−.060				−.122				−.061
				(.039)				(.043)				(.038)
Occupation category fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age and county controls?	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
HC (human capital) controls?	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Interacted HC controls?	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
N	11,021	11,021	11,021	11,021	11,021	11,021	11,021	11,021	11,021	11,021	11,021	11,021
Adjusted R^2	.33	.36	.38	.39	.27	.33	.35	.35	.03	.06	.06	.07

Note.—Authors' calculations are from 1940 IPUMS data (Ruggles et al. 2010) and annual reports of state education departments. Standard errors are in parentheses. See discussion in Sec. VI.

Echoing findings for earnings across occupations, Gelbach (2016) decompositions in table 9 indicate that school quality narrowed the within-occupation wage and annual income gaps through its interaction with attainment. This further weakens the case for endogenous school resources but does not speak to the possibility that blacks chose their time in school with the understanding that, at some point, discriminatory labor markets would temper the returns to another grade. If so, we may overstate the portion of the wage gap attributable to the attainment gap.

The existence and extent of this form of endogeneity cannot be assessed with available pre-war data (nor can the extent of whites' response to discriminatory returns to schooling), but we can use wage decompositions in conjunction with assumptions about the nature of discrimination to "add back" or decondition the portion of pay disparities that potentially overvalue human capital. Note that attainment, on its own, accounts for 5.3 log points of the 28.8-point weekly wage gap within occupations (table 9, col. 4) and 16.4 log points overall (table 4, col. 4). The difference, 11.1 log points, is the estimated contribution of attainment to occupational upgrading.[47] If, on the margin, discriminatory sorting into jobs would have been inflexible to advances in black educational attainment, then 30.2 log points (19.1 + 11.1) is a truer reflection of the conditional weekly wage gap, rather than 19.1 log points as reported in table 4. The same set of assumptions would serve to increase the annual income gap from 13.7 to 25.3 log points.[48]

VII. Counterfactual Estimates

The implicit counterfactual exercise in our baseline results calculates the 1940 wage gap if blacks had achieved the same level of education in schools of comparable quality to whites. Columns 4 and 12 of table 4 indicate that the remaining gap would have been roughly 19.1 log points of weekly wages and 13.7 log points of annual wages.

But this inherent counterfactual ignores historical realities. First, education budgets differed substantially across counties, as did the geographic distribution of blacks and whites across the South. Assigning all blacks the average education quality of whites presumes perfect mobility of households and education funds when, in reality, education funding was highly decentralized at the county or subcounty level. Second, there were many reasons for differing educational attainment by race, only some of which

[47] This is not very different from the contribution of attainment to the occupational score gap (12.9 log points in table 4, col. 8), although occupational score is derived from the 1950 white analog to a given 1940 occupation.

[48] Comparing these 25–30-point conjectured wage gaps to table 1 estimates from the post–Civil Rights era would require analogous assumptions regarding discrimination in occupational sorting later in the century. This remains a fruitful area for future work.

were within the purview of policy makers. Finally, we again allow for the possibility that the returns to human capital differed by race so that equalizing endowments without equalizing yields would have generated more muted wage impacts.

To address these issues all together, we present a series of counterfactual estimates in table 10. Means and unconditional gaps are located in panel 1. We allow for a balanced budget constraint within counties and equalize black and white school quality at the county average (weighted by black and white school enrollment).[49] We then predict wage outcomes using equation (1) and report results in panel 2.[50] Results in row 2a indicate that counterfactual racial wage gaps would have fallen to between 38 and 41 log points, a reduction of 24%–27% relative to the unadjusted baseline. Note here that although the binding budget constraint effectively lowers the school quality experienced by whites, very modest changes in white earnings between panels 1 and 2 are indicative of diminishing returns to school quality.

In row 2b we allow for race-specific returns to human capital in constructing the separate but equal counterfactual. Echoing decomposition analyses in the previous section, we show here that this more flexible model has limited effects on the weekly wage gap and actually lowers the estimated annual gap by 8.5 log points.[51] Over this range of school quality, then, racial differences in the returns to human capital are biasing our estimates of the racial wage gap upwards.

As a second step to estimate impacts of separate but equal schooling, we allow educational attainment to be endogenous, recognizing that time in school is a function of school quality (Margo 1987), and that equalized school resources would have affected earnings through the years of schooling channel as well as the input quality channel. Of course, these changes,

[49] The imposed restriction is that the weighted average of the Z-scores within a county cannot change where the weights are from enrollment in black and white schools. This weighted average Z-score is then assumed to be the counterfactual level of school quality for both black and white students. This is a practical way to equalize the complete profile of black and white school systems without making ad hoc judgments about how nonmonetary quality domains would be converted to equivalent dollars.

[50] Our preferred specification of eq. (1) conditions on third-degree polynomial functions of attainment, school quality, and full interactions therein. This model describes the wage-setting process well (as we demonstrate in Sec. IV.A), but it does not perform consistently when we extrapolate black earnings under substantially higher school quality. Therefore, we prefer a restricted model—without attainment-quality interactions—for this projection, reminding the reader that the residual wage gap is little changed between the restricted and fully interacted specification (table 4).

[51] To implement race-specific returns to human capital, we re-estimate eq. (1) with interactions between human capital variables and a race indicator. We then equalize school quality, predict counterfactual wages, and report the levels and racial difference.

Table 10
Counterfactual Estimates of Black and White Earnings

	Outcome			
	ln(Weekly Wage)		ln(Annual Wage)	
	Black (1)	White (2)	Black (3)	White (4)
1. As measured in 1940	1.868	2.398	5.415	5.928
Black-white gap	−.529		−.513	
	(.024)		(.027)	
2. Counterfactual "separate but equal" estimates:				
a. With pooled coefficients	1.965	2.370	5.505	5.885
Black-white gap	−.406		−.380	
	(.013)		(.016)	
b. With race-specific coefficients	1.963	2.372	5.589	5.883
Black-white gap	−.409		−.295	
	(.013)		(.016)	
3. Counterfactual "separate but equal" estimates with endogenous attainment:				
a. With pooled coefficients	1.994	2.356	5.540	5.871
Black-white gap	−.362		−.331	
	(.014)		(.017)	
b. With race-specific coefficients	1.979	2.357	5.598	5.866
Black-white gap	−.377		−.268	
	(.013)		(.016)	

NOTE.—Authors' calculations are from 1940 IPUMS data (Ruggles et al. 2010), annual reports of state education departments, and Aaronson and Mazumder (2011) results for the quasi-experimental impact of school quality on years of schooling. The table compares black and white weekly and annual wages under counterfactual levels of school quality and educational attainment. Counterfactuals are estimated by estimating eq. (1), with controls for individual and county covariates, and then altering covariates of interest, and predicting outcomes. The remaining differences are estimated as the coefficient on race in an equation with the predicted values as the dependent variable and no other controls. For each outcome, row 2a lists the counterfactual values under equalized school quality for black and white students at the county average, holding years of schooling constant. Row 2b equalizes school quality, but also allows for race interactions in school quality coefficients in the prediction equation. Row 3a allows black years of schooling to increase with school quality according to the elasticity of time in school with respect to school quality as reported by Aaronson and Mazumder (2011), table 5, col. 1 (1.186 years per Rosenwald exposure). Row 3b again allows for race-specific returns to attainment and school quality. Standard errors are in parentheses.

in turn, likely affected cognitive ability, but we do not seek to model this pathway here.

We rely on quasi-experimental evidence of the effect of school quality from other work. Aaronson and Mazumder (2011) estimate the effect of exposure to Rosenwald schools—in terms of classroom capacity per black school-aged youth—on individual years of schooling, among other outcomes. We convert their Rosenwald exposure measure into a change in our calculated Z-score and then convert the elasticity of educational attainment with respect to Rosenwald schools to an elasticity per unit of school quality in our sample. We then use this elasticity to calculate the counterfactual level of educational attainment after a "separate but equal" mandate. On average, estimated attainment increases by 1.19 years for blacks and falls by 0.43 years for whites.[52] Results are presented in panel 3 of table 10. With endogenous years of schooling and race-neutral coefficients, the wage gap falls to between 33 and 36 log points (row 3a), a reduction relative to the baseline of between 32% and 36%. In row 3b, we again allow for race-specific returns to education quality and attainment and, again, this distinction matters little for our result as estimates differ by 6 log points or less.

Overall, we conclude that a separate but equal mandate would have reduced labor market inequalities substantially, reducing the unconditional weekly wage gap by up to 29% and the annual wage gap by 48%.

VIII. Conclusion

Recent labor market studies have highlighted the importance of human capital in explaining the black-white wage gap. We ask the same question for 1940 workers: How far can human capital inequalities go in explaining the large pre-war racial wage disparity? Incorporating new data on race-specific school quality in 10 southern states, we document a predominant role of school quality and educational attainment in determining wage inequality for young men. Human capital accounts for 73% of the gap in annual wages and 64% of the gap in weekly wage rates. Once we control for estimated AGCT scores imputed from World War II enlistment records, human capital accounts for up to 97% of the gap in annual wages and 80% of the weekly wage gap.

The power of education to drive labor market wages is echoed in a counterfactual exercise whereby school quality is equalized across races in the South. Under this "separate but equal" standard, we estimate a counterfactual gap of between 27% and 38%, far smaller than the 51%–53% gap observed in 1940. Education equality would have been a powerful tool for raising black economic standing in the South; by our estimates, the lost opportunity reduced the earnings capacity of this generation of black southerners by up to 48%. In the end, twentieth-century racial wage convergence

[52] Aaronson and Mazumder (2011) express the effect of school quality on years of schooling with respect to their quality measure "Rosenwald exposure," i.e., the number of classrooms per 45 rural blacks aged 7–17 in a county. Because classrooms are not one of our resource metrics, we assume that each classroom is associated with one new black teacher (as it was intended to be) and then calculate the average change in the (race-specific) Z-score for teachers per pupil when we change Rosenwald exposure from 0 to 1. Equivalently, this raises teachers per pupil over the historic reported figure by an amount implied by full coverage according to Aaronson and Mazumder (2011)'s metric. We lack access to the same population measures, but we substitute the product of the black population aged 10–20 in a county and the percent of the overall county population that is rural. We divide the Aaronson and Mazumder (2011) reported elasticity, which is relative to their exposure measure, by this change in Z-score to get the average change in years of schooling per Z-score unit and use this as our quasi-experimental elasticity of educational attainment with respect to school quality (in standardized Z units).

prior to Civil Rights came through shifts in the national income distribution that favored blacks (Margo 2015), substantial interregional migration (Boustan 2007; Collins and Wanamaker 2014), and gradual increases in the relative quality of black human capital (Card and Krueger 1992b).

Because some portion of human capital and the mechanisms determining local school quality remain largely unobserved, we must be careful in labeling residual wage gaps as the upper limit of discrimination. And our conclusion that wage gaps were relatively small among young men in 1940 does not rule out larger wage gaps and a larger—or perhaps, different—role for discrimination among subgroups of men of the same or later eras. The modern labor market's key departures from the Jim Crow era include a much smaller or null wage gap among higher-skilled men and the emergence of a large, lower-skilled employment and pay gap favoring whites.[53] Alternative sample constructions within the 1940 Census, discussed in the appendix, are consistent with wider wage gaps among higher-skilled men. If the market for higher-skilled men was thinner and less competitive in the pre-war South than in later years, these subgroups could have been more vulnerable to wage discrimination and thus more affected by equal protection laws of later decades. Understanding the factors that contributed to racial wage gaps among unskilled workers in 1940 versus today remains an interesting, and open, area of inquiry. Since part of today's racial wage gap has roots in falling skill prices (Chandra 2000), one notable conclusion to be drawn from this paper and others is that the racial inequities arising from skill-biased economic transitions may have been even greater in the absence of converging black-white school quality and attainment.

By our estimates, conditional wage gaps were perhaps 42% smaller in 2000 than in 1940.[54] We cannot partition the relative contributions of federal intervention, returns to skill, and rising black school quality across the intervening years, nor do we have a comparable metric for the years between World War II and the Civil Rights era. This motivates further research on the dynamics of conditional wage and employment gaps over the transition from the Jim Crow to the Civil Rights era and beyond, particularly by skill level. What we have shown is that the southern labor market of 1940 and the labor market of 2000 suffered a similar disease: vastly different human capital accumulation among black and white workers. In both cases, ameliorating the human capital gap would have substantially reduced racial wage gaps.

[53] See Chandra (2000), Heckman, Lyons, and Todd (2000), Lang (2007), and Johnson and Neal (2011) for details, and see Bound and Freeman (1992), Chay and Lee (2000), Weinberger and Joy (2007), and Heywood and Parent (2012) for evidence on the limits of these generalizations.

[54] This compares the modern literature's modal 11 log point gap to our 19 log point result for weekly wages.

Turning back to the time period in question, subjectively narrow wage gaps seem incompatible with what we know about the Jim Crow South. Black southerners were excluded from civil life through a variety of measures that eviscerated their participation in the political process. One result was a denial of black public schooling of the same quality as that available to whites, presenting an enormous roadblock to the accumulation of black human capital. Yet, as we show, blacks participated in economic life, exhibiting labor force participation and employment rates not dissimilar from those of whites and (conditional) earnings ratios not remarkably different from blacks later in the century. Employers of 1940 may have held animus or equality aversion toward black individuals, but the effect of these attitudes on black wages would have been offset to some degree by profit-maximizing objectives. These profit-maximizing values, not necessarily shared by largely white voting constituencies, explain why severe racial discrimination in the provision of public goods could coincide with a more equal (conditional) labor market.

References

Aaronson, Daniel, and Bhashkar Mazumder. 2011. The impact of Rosenwald schools on black achievement. *Journal of Political Economy* 119, no. 5:821–88.

Altonji, Joseph G., and Rebecca M. Blank. 1999. Race and gender in the labor market. In *Handbook of labor economics,* Vol. 3, ed. Orley Ashenfelter and David Card, 3143–3259. Amsterdam: Elsevier.

Altonji, Joseph G., and Charles R. Pierret. 2001. Employer learning and statistical discrimination. *Quarterly Journal of Economics* 116, no. 1:313–50.

Ashenfelter, Orley, William J. Collins, and Albert Yoon. 2006. Evaluating the role of *Brown v. Board of Education* in school equalization, desegregation, and the income of African Americans. *American Law and Economics Review* 8, no. 2:213–48.

Becker, Gary S. 1957. *The economics of discrimination*. Chicago: University of Chicago Press.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review* 94, no. 4:991–1013.

Betts, Julian R. 2010. School quality and earnings. In *International encyclopedia of education,* 3rd ed., ed. Penelope Peterson, Eva Baker, and Barry McGaw, 313–20. Oxford: Elsevier.

Biewen, Martin. 2014. A general decomposition formula with interaction effects. *Applied Economics Letters* 21, no. 9:636–42.

Bound, John, and Richard Freeman. 1992. What went wrong? The erosion of the relative earnings and employment among young black men in the 1980s. *Quarterly Journal of Economics* 107, no. 1:151–200.

Boustan, Leah Platt. 2007. Black migration, white flight: The effect of black migration on northern cities and labor markets. *Journal of Economic History* 67, no. 2:484–88.

Card, David, and Alan B. Krueger. 1992a. Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of Political Economy* 100, no. 1:1–40.

———. 1992b. School quality and black-white relative earnings: A direct assessment. *Quarterly Journal of Economics* 107, no. 1:201–32.

Carneiro, Pedro, James J. Heckman, and Dimitriy V. Masterov. 2005. Labor market discrimination and racial differences in premarket factors. *Journal of Law and Economics* 48, no. 1:1–39.

Carruthers, Celeste K., and Marianne H. Wanamaker. 2013. Closing the gap? The effect of private philanthropy on the provision of African-American schooling in the U.S. South. *Journal of Public Economics* 101:53–67.

———. 2015. Municipal housekeeping: The impact of women's suffrage on public education. *Journal of Human Resources* 50, no. 4:837–72.

Cascio, Elizabeth U., and Ebonya L. Washington. 2013. Valuing the vote: The redistribution of voting rights and state funds following the Voting Rights Act of 1965. *Quarterly Journal of Economics* 129, no. 1:379–433.

Chandra, Amitabh. 2000. Labor-market dropouts and the racial wage gap, 1940–1990. *American Economic Review* 90, no. 2:333–38.

Charles, Kerwin Kofi, and Jonathan Guryan. 2008. Prejudice and wages: An empirical assessment of Becker's *The economics of discrimination*. *Journal of Political Economy* 116, no. 5:773–809.

Chay, Kenneth, and David Lee. 2000. Changes in relative wages in the 1980s: Returns to observed and unobserved skills and black-white wage differentials. *Journal of Econometrics* 78, no. 1:1–38.

Collins, Wiliam J., and Robert A. Margo. 2006. Historical perspectives on racial differences in schooling in the United States. In *Handbook of the economics of education,* ed. Eric A. Hanushek and Finis Welch, 107–54. Amsterdam: Elsevier.

Collins, William J., and Marianne H. Wanamaker. 2014. Selection and economic gains in the great migration of African Americans: New evidence from linked census data. *American Economic Journal: Applied Economics* 6, no. 1:220–52.

Darity, William A., Jr., and Patrick L. Mason. 1998. Evidence on discrimination in employment: Codes of color, codes of gender. *Journal of Economic Perspectives* 12, no. 2:63–90.

Donohue, John J., and James J. Heckman. 1991. Continuous versus episodic change: The impact of civil rights policy on the economic status of blacks. *Journal of Economic Literature* 29, no. 4:1603–43.

Fishback, Price V. 1989. Can competition among employers reduce governmental discrimination? Coal mines and segregated schools in West Virginia in the early 1900s. *Journal of Law and Economics* 32:311–28.

Fryer, Roland. 2011. Racial inequality in the 21st century: The declining significance of discrimination. In *Handbook of labor economics,* Vol. 4B, ed. David Card and Orley Ashenfelter. Amsterdam: Elsevier.

Gelbach, Jonah B.. 2016. When do covariates matter? And which ones, and how much? *Journal of Labor Economics* 34, no. 2:509–43.

Goldin, Claudia. 1990. *Understanding the gender gap: An economic history of American women*, New York: Oxford University Press.

———. 2001. The human capital century and American leadership: Virtues of the past. *Journal of Economic History* 61, no. 2:263–92.

Goldin, Claudia, and Lawrence F. Katz. 2000. Education and income in the early twentieth century: Evidence from the prairies. *Journal of Economic History* 60, no. 3:782–818.

Goldin, Claudia, and Robert A. Margo. 1992. The great compression: The wage structure in the United States at mid-century. *Quarterly Journal of Economics* 107, no. 1:1–34.

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor. 1996. Aggregation and the estimated effects of school resources. *Review of Economics and Statistics* 78, no. 4:611–27.

Heckman, James J., Thomas M. Lyons, and Petra E. Todd. 2000. Understanding black-white wage differentials, 1960–1990. *American Economic Review* 90, no. 2:344–49.

Heckman, James J., and Brook S. Payner. 1989. Determining the impact of federal antidiscrimination policy on the economic status of blacks: A study of South Carolina. *American Economic Review* 79, no. 1:138–77.

Heywood, John S., and Daniel Parent. 2012. Performance pay and the black-white wage gap. *Journal of Labor Economics* 30, no. 2:249–90.

Higgs, Robert. 1977. Firm-specific evidence on racial wage differentials and workforce segregation. *American Economic Review* 67, no. 2:236–45.

Johnson, George E., and Frank P. Stafford. 1973. Social returns to quantity and quality of schooling. *Journal of Human Resources* 8, no. 2:139–55.

Johnson, William R., and Derek Neal. 2011. Basic skills and the black-white earnings gap. In *The black-white test score gap,* ed. Christopher Jencks and Meredith Phillips, 480–98. Washington, DC: Brookings Institution Press.

Kose, Esra, Elira Kuka, and Na'ama Shenhav. 2015. Women's enfranchisement and children's education: The long-run impact of the U.S. suffrage movement. Unpublished manuscript, University of California, Davis.

Lang, Kevin. 2007. *Poverty and discrimination*, Princeton, NJ: Princeton University Press.

Lang, Kevin, and Jee-Yeon K. Lehmann. 2012. Racial discrimination in the labor market: Theory and empirics. *Journal of Economic Literature* 50, no. 4:959–1006.

Lang, Kevin, and Michael Manove. 2011. Education and labor market discrimination. *American Economic Review* 101, no. 4:1467–96.

Lazear, Edward P. 1991. Discrimination in labor markets. In *Essays on the economics of discrimination,* ed. Emily P. Hoffman, 9–24. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Link, Charles R., and Edward C. Ratledge. 1975. Social returns to quantity and quality of education: A further statement. *Journal of Human Resources* 10, no. 1:78–89.

Link, Charles, Edward Ratledge, and Kenneth Lewis. 1976. Black-white differences in returns to schooling: Some new evidence. *American Economic Review* 66, no. 1:221–23.

Margo, Robert A. 1984. Teacher salaries in black and white: The South in 1910. *Explorations in Economic History* 21:306–26.

———. 1986. Race, educational attainment, and the 1940 Census. *Journal of Economic History* 46, no. 1:189–98.

———. 1987. Accounting for racial differences in school attendance in the American South, 1900: The role of separate-but-equal. *Review of Economics and Statistics* 69, no. 4:661–66.

———. 1990. *Race and schooling in the South, 1880–1950: An economic history*, Chicago: University of Chicago Press.

———. 2015. Economic History Association presidential address: Obama, Katrina, and the persistence of racial inequality. *Journal of Economic History* 76, no. 2:301–41.

Morgan, James, and Ismail Sirageldin. 1968. A note on the quality dimension in education. *Journal of Political Economy* 76, no. 5:1069–77.

Morgenstern, Richard D. 1973. Direct and indirect effects on earnings of schooling and socio-economic background. *Review of Economics and Statistics* 55, no. 2:225–33.

Neal, Derek A., and William R. Johnson. 1996. The role of premarket factors in black-white wage differences. *Journal of Political Economy* 104: 869–95.

Nechyba, Thomas J. 1990. The Southern wage gap, human capital, and the quality of education. *Southern Economic Journal* 57, no. 2:308–22.

Oaxaca, Ronald L., and Michael R. Ransom. 1994. On discrimination and the decomposition of wage differentials. *Journal of Econometrics* 61, no. 1:5–21.

Oettinger, Gerald S. 1996. Statistical discrimination and the early career evolution of the black-white wage gap. *Journal of Labor Economics* 14, no. 1:52–78.

O'Neill, Donal, Olive Sweetman, and Dirk Van de gaer. 2006. The impact of cognitive skills on the distribution of the black-white wage gap. *Labour Economics* 13, no. 3:343–56.

Orazem, Peter F. 1987. Black-white differences in schooling investment and human capital production in segregated schools. *American Economic Review* 77, no. 4:714–23.

Rizzuto, Ronald, and Paul Wachtel. 1980. Further evidence on the returns to school quality. *Journal of Human Resources* 15, no. 2:240–54.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. *Integrated Public Use Microdata Series: Version 5.0 (machine-readable database)*, University of Minnesota.

Smith, James P. 1984. Race and human capital. *American Economic Review* 74, no. 4:685–98.

Smith, James P., and Finis Welch. 1986. *Closing the gap: Forty years of economic progress for blacks*, Santa Monica, CA: RAND Corporation.

———. 1989. Black economic progress after Myrdal. *Journal of Economic Literature* 27:519–64.

Sundstrom, William A. 2007. The geography of wage discrimination in the pre–Civil Rights South. *Journal of Economic History* 67, no. 2:410–44.

Troesken, Werner, Joseph P. Ferrie, and Karen Rolf. 2012. Cognitive disparities, lead plumbing, and water chemistry: Prior exposure to waterborne lead and intelligence test scores among World War Two army enlistees. *Economics and Human Biology* 10, no. 1:98–112.

Wachtel, Paul. 1975. The effect of school quality on achievement, attainment levels, and lifetime earnings. *Explorations in Economic Research* 2: 502–36.

———. 1976. The effect on earnings of school and college investment expenditures. *Review of Economics and Statistics* 58:326–31.

Weinberger, Catherine, and Lois Joy. 2007. Relative earnings of black college graduates, 1980–2001. In *Race and economic opportunity in the twenty-first century,* ed. Marlene Kim, 50–72. New York: Routledge.

Welch, Finis. 1974. Education and racial discrimination. In *Discrimination in labor markets,* ed. Orley Ashenfelter and Albert Rees. Princeton, NJ: Princeton University Press.

Whatley, Warren, and Gavin Wright. 1994. Race, human capital, and labour markets in American history. In *Labour markets in history,* ed. George Grantham and Mary MacKinnon, 270–91. Toronto: Routledge.

Wright, Gavin. 2013. *Sharing the prize: The economics of the civil rights revolution in the American South*, Cambridge, MA: Belknap.

Zeidner, Joseph, and Arthur J. Drucker. 1983. *Behavioral science in the army: A corporate history of the Army Research Institute*, Alexandria, VA: US Army Research Institute for the Behavioral and Social Sciences.