# Education, Migration, and Regional Wage Convergence in American History

William J. Collins

[May 2, 2006]

Prepared for the Conference on the New Comparative Economic History,
in honor of Jeffrey G. Williamson

Collins is an Associate Professor of Economics at Vanderbilt University and a Research Associate of the National Bureau of Economic Research. Martha J. Bailey, Leah Platt Boustan, Robert Driskill, Stanley Engerman, Timothy Hatton, Robert A. Margo, Jacob Vigdor, and Gavin Wright provided helpful suggestions. Emily Boleman supplied excellent research assistance. NSF support is gratefully acknowledged (SES 0095943).

## I. Introduction

The comparative perspective in the "New Comparative Economic History" often entails looking across countries to discern how differences in factor endowments and changes in factor and commodity market integration have influenced long-run economic phenomena (e.g., Williamson 1996). But the logic and force of the approach are every bit as compelling when comparing regions within a country, especially when the regions are as diverse as those in the United States (cf., Williamson 1965, 1980). In that spirit, this chapter describes the twentieth-century economic ascendance of the American South with a focus on convergence between the South and the rest of the country in terms of workers' formal education and earnings.[1]

Before the Civil War, southerners made comparatively small investments in formal education, even among white children. For white children, the economic shock and social dislocation associated with the War appear to have further reduced educational attainment compared to any reasonable no-war counterfactual. For slaves (who had been forbidden to learn to read) and their offspring, the Civil War obviously had different implications – from an extremely low base, African Americans rapidly increased their literacy rate. But even 35 years after the War, the North-South gap in educational attainment and institutions was still very large. The median southern child (age 7 to 17) attended school for five months less than the median non-southern child in 1900.

C. Vann Woodward, in his landmark history *Origins of the New South*, writes, "The public schools of the South at the opening of the new century were for the most part miserably supported, poorly attended, wretchedly taught, and wholly inadequate for the education of the people. Far behind the rest of the country in nearly all respects, Southern education suffered

from a greater lag than any other public institution in the region" (1951, p. 398). Nevertheless, by the end of the twentieth century, the regional gap in educational attainment had nearly disappeared.

After documenting regional trends in educational attainment and children's school attendance with micro-level census data, this paper pursues two main questions. First, how can one account for southern workers' twentieth-century educational catch-up? The rate of educational improvement did not slow down outside the South; rather, the southern labor force accumulated schooling at a consistently faster rate than the rest of the country from at least 1940 onward. I find that part of the regional convergence was driven by migration flows. On net, migrants to the South had relatively high levels of human capital, whereas migrants to the rest of the country (including the foreign born) had relatively low levels of human capital. However, the bulk of the regional convergence in workers' education is accounted for by local investment in regional natives. The educational gains of each new cohort of southern-born African American workers are significant in this regard.

Second, how much did convergence in average educational attainment contribute to convergence in average earnings? To gauge the direct contribution, I estimate counterfactual inter-regional income gaps for each census year from 1950 to 2000, supposing that there had been no educational convergence in the preceding decade. From this perspective, which yields conservative estimates, it appears that education's direct contributions were economically significant, accounting for (on average) roughly 40 percent of the income convergence that occurred in each decade from 1950 to 2000.

This paper does not directly address the Civil Rights Movement, but many connections

between southern educational history and the Civil Rights Movement are manifest here. To wit, civil rights leaders and many of the rank and file stepped forward from a growing, well-educated segment of the southern black population. In this sense, blacks' pre-1950 educational gains set the stage for the Civil Rights Movement, even if such gains were not a sufficient condition for the Movement's success. Civil rights organizations, in turn, pursued litigation that motivated improvements in black schools long before the Supreme Court's *Brown v. Board of Education* ruling in 1954 (Tushnet 1987, Margo 1990). And in the 1960s, changes in federal policy directly targeted discrimination in the South, improved labor market opportunities for black workers (Donohue and Heckman 1991), and may have boosted incentives for young blacks to acquire more and better schooling. The success of the Civil Rights Movement lowered the likelihood of out migration by well-educated blacks and made the region more attractive for well-educated native-born and foreign-born migrants. Thus, although the South's educational convergence on the North began some time before the Civil Rights Movement's most famous events, the histories of schooling, race, and income in the South are at all times intertwined.

## II. Regional Divergence and Convergence in Education, 1850-2000

Regional convergence in educational attainment is clearly evident in the second-half of the twentieth century, but the nineteenth and early twentieth-century evidence is much more mixed. Analysis of individual-level data from the Integrated Public Use Microdata Series (Ruggles et al. 2004, henceforth IPUMS) reveals that in the decade before the Civil War, the South/Non-South gaps in school attendance among white children (ages 7 to 17) and in literacy

rates among young white adults (ages 20 to 25) narrowed slightly.[2] The attendance gap fell from 22 to 20 percentage points between 1850 and 1860, and the literacy rate gap fell from 10 to 7 percentage points. There are indications, however, that southern whites' school attendance and literacy deteriorated around the time of the Civil War and Reconstruction in both absolute terms and relative to the rest of the country. Southern white school attendance fell from 52 to 38 percent from 1860 to 1870, and literacy among young southern white adults fell from 89 to 81 percent from 1860 to 1880. The regional gap in school attendance (for whites) jumped to 33 percentage points in 1870, and the literacy rate gap jumped to 14 percentage points in 1880 (20 to 25 year olds in 1880 would have been of school age during the Civil War and Reconstruction). This represents a remarkable educational setback for a generation of southern white children.

Despite this setback, and although Woodward (1951) dates the "great educational awakening" in the South to the turn of the century, southern children of both race categories sharply increased their rate of school attendance after 1870. Between 1870 and 1900, the proportion of southern children (ages 7 to 17) who attended school for at least part of the year jumped from 28 to 58 percent (from 9 to 44 percent for black children, and from 38 to 65 percent for white children). Comparable figures for all non-southern children are 70 percent in 1870 and 72 percent in 1900.

The census's school attendance question set a very low bar for an affirmative answer. Typically, the question asked whether the child attended school at all in the previous school year or calendar year. To infer the real educational implications of rising attendance rates, it helps to examine measures of outcomes such as literacy and highest-grade-of-attainment. Prior to 1940, unfortunately, literacy is the only gauge of educational attainment available in nationally

representative samples. Figure 1A plots state-level literacy rates from 1870 to 1930 calculated using samples of 18 to 64 year olds from all race categories and the foreign born. Each open circle marks a state average, and markers for southern states are additionally labeled with an "S". The dashed lines trend through the southern and non-southern average literacy rates. The overall southern literacy rate was just 55 percent in 1870 compared to the non-southern rate of 90 percent.[3]

The South's enormous deficit in literacy in the late nineteenth century is largely, but not solely, attributable to slavery and its preclusion of African Americans' educational opportunities. In the antebellum South, it was illegal to teach slaves to read because it was thought that educated slaves were more likely to run away or lead insurrections (e.g., literacy makes it easier to forge passes, write letters, and follow complex directions). Although some slaves did learn to read and write (Anderson 1988, Williams 2004), as far as one can tell from the census data, there was little or no upward trend in literacy among slaves prior to the Civil War (Collins and Margo 2006). The overwhelming majority of black workers had little or no formal schooling in 1870. Only 12 percent of black workers were literate, and they comprised 42 percent of the southern labor force. Southern white workers continued to lag behind non-southerners, too, but the inter-regional gap among whites was much smaller (about 10 percentage points in 1870).

By 1930, when the census last inquired about literacy, figure 1A shows there had been considerable interstate convergence in literacy rates. But literacy, as defined by the census, was an extremely low educational threshold, typically achieved after approximately two years of schooling (Margo 1990). Therefore, figure 1A might mask divergence in more finely tuned measures of educational attainment. The "high school movement", which accelerated outside the

South nearly a full generation before it did in the South (Goldin 1998), is a central concern in this regard.

As noted above, the 1940 census was the first to inquire about highest grade of attainment. Measurement issues associated with ungraded schools and regional differences in school quality will tend to overstate the level of southern educational attainment in 1940 relative to non-southern attainment, and the degree of bias should lessen over time.[4] Therefore, the educational convergence documented here, which does not attempt to adjust for quality differences, is a conservative description of the true magnitude of change.[5] In the 1940 IPUMS sample, individuals who were born in the South between 1880 and 1920 averaged about two grades less education than non-southern-born natives. Thus, in contrast to the literacy-based perspective from figure 1A, there is little or no evidence of convergence in average highest grade of attainment among natives who entered the labor force in the early twentieth century.

From the 1920 birth cohort onwards, however, educational convergence was significant, and as the new cohorts entered the labor force (and as the old cohorts departed), the inter-regional gap in workers' human capital narrowed. This is depicted in figure 1B, which plots state averages calculated using each IPUMS sample from 1940 to 1980 (all workers, aged 18 to 64 are in the samples). Again, the dashed lines represent southern and non-southern averages. The regional gap in workers' average years of education was large in 1940. Among southern workers, the average was only 7.9 years, compared to the non-southern workers' average of 9.5. Again, the African-American educational deficit is a significant factor in the inter-regional gap. In 1940, southern black workers averaged only 5.4 years of education, compared to 8.8 years for southern white workers. After 1940, the regional gap in average educational attainment fell from

1.6 years to less than 0.50 years by 1980 and to less than 0.25 years by 2000.

Figure 2 summarizes the post-1940 pattern of convergence in educational attainment for each state by graphing changes in average educational attainment among workers from 1940 to 1980 against initial (1940) values. Southern states are clearly grouped in the upper left corner of the graph – they had low initial values and high growth rates. Workers in the deep South (Alabama, Mississippi, Louisiana, Georgia, and South Carolina) had extremely rapid gains between 1940 and 1980, adding about five grades to a 1940 base of only 7$^{th}$ grade attainment.

## III. The Roots of the Southern Schooling Deficit

### *History, Political Economy, and Southern Public Schooling*

The historical literature on the relatively low level of southern investment in public goods and public education in particular describes an evolving balance of power in the region's political economy. At the risk of oversimplification, both before and after the Civil War, the economic interests of wealthy white landowners (planters) were a roadblock to the expansion of the southern public education system, especially in rural areas and for blacks (Woodward 1951, Nicholls 1960, Tindall 1967, Wright 1986, Anderson 1988, and Margo 1990). Anderson writes,

> Before the war poor [white] children were unable to afford private schooling and only rarely had the opportunity to attend charity institutions. In the immediate postwar years the region's poor whites, in general, were still too closely tied to the

> planters' interests and ideology to pursue a different conception of education and society. . . . The South's white middle classes, unorganized and subservient to planter interests throughout the nineteenth century, did not begin their campaign for universal education until the dawn of the twentieth century (1988, p. 26).

The problem with public education, from the planters' perspective, was fourfold: 1) it required tax revenue; 2) educated workers were more likely to leave agriculture (and the South);[6] 3) children of school-age provided a non-trivial amount of farm labor, especially at periods of peak demand; and 4) in the context of post-Reconstruction disenfranchisement, education had the potential to shift political power toward poor whites and blacks (Anderson 1988, p. 98). With the exception of the Reconstruction period, planters held disproportionate political influence in the South, and a relatively low level of support for public schooling ensued.[7]

During Reconstruction, however, southern states adopted new constitutions that included substantial provisions for tax-supported public education for all children, including blacks. This set a foundation for public school systems throughout the region, but after 1877 (the end of Reconstruction) the southern commitment to public schooling faltered. Southern planters re-established their political dominance, and critics linked public schools with "carpetbaggers" of the Reconstruction era. Woodward explains, "Redemption governments, often describing themselves as the 'rule of the taxpayer,' frankly constituted themselves champions of the property owner against the propertyless and allegedly untaxed masses" (1951, p. 59). In addition, for decades thereafter, Alston and Ferrie claim that appeals to race issues "... enabled politicians to cater to the economic interests of the white upper class while maintaining the

support of whites in general" (1999, p. 39). Because wage laborers and sharecroppers valued schools for their children and would move to take advantage of educational opportunities, an incentive to support schools remained even where blacks and poor whites were politically disenfranchised (Margo 1990).[8] But the average level of quality was low, as emphasized by Woodward and as evident in records of expenditures per pupil, teachers per pupil, length of school year.

The idea that public schooling was socially desirable, especially for white children, slowly gained political momentum in the South. By the end of the nineteenth century, many southern whites, including small farmers, professional educators, and advocates of southern industrialization, were fighting vigorously to expand educational opportunities for southern children (Nicholls 1960). In addition, black parents and northern philanthropists stepped in to improve the meager educational facilities available to black children (Donohue, Heckman, and Todd 2002). As noted above, Woodward dates the "great educational awakening" in the South to the turn of the century (1951, p. 400), but at all times, it was an uphill run to catch up to the rest of the country where educational standards and attainment were constantly increasing. In Goldin's (1998) description of the takeoff of America's high school movement in the early 1900s, the South was left, literally but temporarily, in the dust.[9]

*A Closer Look at School Attendance at the Turn of the Twentieth Century*

The 1900 census is especially useful for characterizing the size and nature of the regional schooling gap because it recorded the number of months that children attended school in the

previous year, whereas other censuses simply asked whether or not the child had attended school at all in the past year (or in past several months).[10] In 1900, nearly half of southern children (age 7 to 17) attended school for less than one month, compared to about 30 percent of non-southern children. Less than one-tenth of southern children attended school for nine months or more, compared to nearly 40 percent of non-southern children.[11] By 1960, racial and regional disparities in children's school attendance had essentially disappeared for those from age 7 to 15, though regional gaps remained at both ends of the school-age distribution.[12]

Do regional gaps in economic characteristics, such as income and urbanization, easily account for the large regional gaps in investment in children's education circa 1900? Or was the South distinctly below the regression line in its investment in children's schooling? To explore these questions, I estimated regressions of the following basic form for samples of children (ages 7 to 17) using the 1900 IPUMS data,

$$\text{Months}_{ijr} = \alpha + \gamma_{age} + \beta_1 \text{South}_r + \beta_2 \text{Urban}_{ijr} + \beta_4 \text{State Income}_j + \beta_4 \text{Parent Income}_{ijr} + e_{ijr}.$$

*Months* is the number of months the child attended school, $\gamma$ is a vector of age dummy variables, *South* is a region of residence dummy variable, log *State Income* per capita is from Mitchener and McLean (1999), and log *Parent Income* is an estimate based on the race-gender-region-occupation category of the household head.[13]

The inclusion of age dummies implies that the coefficient on *South* is identified from within age-cohort differences in months of schooling across regions, conditional on the other variables. Such a regression cannot truly explain the South's attendance deficit, but it can quantify the residual significance of "southernness" that extends beyond regional differences in income and urban residence. The unadjusted regional gap in average months of schooling in

1900 was approximately 2.8 months; the median gap was 5 months. An OLS regression that conditions on the variables described above yields a coefficient on *South* of -0.91, whereas a quantile regression (at medians) yields a coefficient of -1.5.[14] In this sense, about two-thirds of the total inter-regional gap in months of school attendance in 1900 can be accounted for by fundamental economic variables, but the residual regional gap is still economically and statistically significant.[15]

The *South* coefficient is nearly identical if a dummy for race (black = 1) is included in the OLS regression (the *Race* coefficient is -0.203, standard error = 0.096). Alternatively, if only white children are included in the sample, the *South* coefficient is only slightly smaller in magnitude than originally ($\beta_1$ = -0.819). Finally, replacing the "months of school" dependent variable with a binary variable for school attendance (=1 if attended at all) yields a southern coefficient of -0.092 (standard error = 0.037), implying that southern students were about 9 percentage points less likely to be enrolled in school than non-southern students, conditional on the other independent variables. Again, the estimate is only slightly affected by adding a control variable for race to the specification.

Thus, at the beginning of the twentieth century, the South lagged the North in educating its children by a wide margin, and only part of that disparity is attributable to regional differences in income, urban residence, and racial composition. Conditional on a basic set of economic characteristics, southern children attended school for fewer months than non-southern children. How, then, did the South catch up?

## IV. Convergence in Educational Attainment: The Roles of Migration and Local Investment

It has often been noted that higher levels of education facilitated black migration out of the South, particularly before World War II (e.g., Margo 1990, Vigdor 2002). Scholars have paid much less attention to the selective nature of migration *into* the South (Fein 1965, Suval and Hamilton 1965, Wright 1986). In 1940, non-southern-born (but US native) workers residing in the South had an average of 10.6 years of education compared to the southern-born residents' average of 7.8 years and the northern-born "stayer" average of 10.0.[16]

Tables 1A and 1B account for regional differences and changes in all workers' average educational attainment. At a point in time, the average education level in a region can be expressed as a weighted average of the attainment of those who were born in the region and continue to reside there ("stayers") and those who were born elsewhere but moved into the region from either elsewhere in the U.S. or from another country ("migrants"). Columns 1 to 4 of table 1A decompose the average years of schooling for each year from 1940 to 2000 for the South and Non-South.

Three key points emerge. First, migrants to the South have always been better educated (on average) than regional native workers (stayers), but migrants to the Non-South (many of whom were foreign born) have always been less educated than Non-Southern regional native workers. Second, while the Non-South maintained a roughly constant proportion of migrant workers (column 4; again, a combination of inter-regional and foreign-born migrants) at around 20 percent, the South steadily increased its share of migrants, from one-tenth of its labor force in 1940 to one-third in 2000. This is one way of viewing the breakdown of regional isolation emphasized by Wright (1986). Third, over time, both the regional natives and the migrants in

the southern labor force have increasingly resembled workers elsewhere in the country (see the "differences" in columns 2 and 3 of panel C), but nontrivial differences remained – "locally produced" non-southern workers are still better educated than "locally produced" southern workers on average, and the migrants to the South are still better educated than the migrants to the Non-South.

Columns 1 to 4 of table 1A could be misleading characterizations of the role of net migration in changing the skill mix of the southern labor force. For example, *if* the out-migration of highly educated southerners had been large relative to the in-migration of highly educated workers from elsewhere, then one would *not* want to conclude that inter-regional migration had a net positive influence on southern workers' average educational attainment. Columns 5 to 8 help address this concern. Southern-born emigrants were, in fact, better educated than southern-born stayers in every census sample, although the difference in 1960, 1970, and 1980 is much smaller than in 1940. This is consistent with the idea that southern emigration was more selective on education before World War II than afterwards (Vigdor 2002).[17] But the South's loss of relatively well-educated emigrants (compared to southern stayers) was more than offset by the inflow of well-educated workers in determining the overall average educational attainment of southern workers. The net migration effect, reported in column 8, is simply the difference between the average educational attainment of workers born in the South and that of workers residing in the South. The net effect for the South is positive up to 2000, and is especially strong in the 1960 to 1980 samples.

The net effect of migration for the Non-South is even larger in magnitude, but it is negative (column 8 of panel B). The combined implications for regional convergence are

reported in column 8 of panel C. The regional gap in worker education was about 0.5 years less than it would have been on the basis of place of birth, assuming that each individual's educational attainment is fixed.[18]

Table 1B decomposes changes over time in average educational attainment for workers residing in each region using the following identity,

$$\Delta S = \alpha_1(S_{N1} - S_{N0}) + (1 - \alpha_1)(S_{M1} - S_{M0}) + (\alpha_1 - \alpha_0)(S_{N0} - S_{M0}),$$

where S is the average schooling level in a region's labor force, $\alpha$ is the proportion of workers who were born in the region ("natives"), $S_N$ and $S_M$ refer to the regional native average schooling level and migrant schooling level respectively, and 0 or 1 denotes the beginning or end year of a decade. So, the overall change in a region's stock of educational attainment (among workers) can be partitioned into a weighted *change* in natives' average attainment, a weighted *change* in migrants' average attainment, and the *change* in native/migrant mix in the labor force weighted by the difference in their average attainment.

Within each regional grouping, the lion's share of rising educational attainment is accounted for by investments in human capital by native-born workers. This is not surprising since regional natives comprise a large portion of the labor force (around 80 percent on average) and because educational attainment among natives (both South and Non-South) was rising quickly. The more interesting result is the influence of the changing native/migrant mix in each region's labor force. The South received relatively well educated migrants, and these migrants were increasing their share of the southern labor force over time. The Non-South, received relatively poorly educated migrants, and the migrant share of workers fell from 1940 to 1970 (a positive influence on average attainment) before rising again to 2000 (a negative influence on

average attainment). The combined implications of the changing native/migrant mix are reported in the far-right column of panel C. In the 1940s, 1950s, and 1960s, the changing mix of migrants made little difference to the overall pace of educational convergence. However, in the 1970s, 1980s, and 1990s, the shifting native/migrant composition of the labor force accounts for about half (or more in the 1990s) of the inter-regional educational convergence, though the overall pace of convergence was slower than in earlier decades.

Why did the South attract relatively high human capital workers (on average), while the Non-South attracted relatively low human capital workers? For the period prior to 1940, it is not possible to compare average wages across education groups by region with great certainty. It appears, however, that wages in skilled occupations in the South (especially the South Central subregion) were close to on par with those in the Northeast and Midwest, whereas wages for southern unskilled labor were comparatively low, implying a relatively high skill premium in the South. This pattern is evident at least as far back as 1890 (Coelho and Shepherd 1979, Sundstrom and Rosenbloom 1993), and Margo (2004) has shown that it existed even in the antebellum period.

Using the 1950 IPUMS sample, I calculated median weekly earnings, by education level, for workers who worked at least 40 weeks in 1949, including adjustments to account for state-level differences in the cost of living (from Williamson and Lindert 1980). Again, southern regions offered nearly the same (pre-tax) earnings as non-southern regions for male workers who had twelve years or more of formal education (e.g., $81 per week in Mid-Atlantic compared to $76 in the South Atlantic, in 1949 dollars), and offered much lower wages than non-southern regions for workers with less than twelve years of education (e.g., $59 in Mid-Atlantic compared

to $39 in South Atlantic). This is not merely a reflection of the racial composition of the South – essentially, the same story can be told when looking at separate samples for white and black men.[19] A similar regional pattern is evident for income in 1979 income, but by that time southern workers with more than twelve years of education earned slightly *more* in the South than elsewhere, and the regional wage gap for workers with less than twelve years was much smaller (in percentage terms) than in 1949.[20] Thus, throughout the second half of the twentieth century, real weekly earnings (pre-tax) in the South were as high as those elsewhere in the country for workers with some college education, but earnings were lower in the South than elsewhere for poorly educated workers.

The long-standing regional differences in wage structure have been interpreted as evidence of regional labor market isolation, especially before 1940 (cf. Wright 1986, Rosenbloom 2002). Whatever the extent of southern labor market isolation in the early twentieth century, several factors contributed to its erosion: labor demand booms in northern industrial centers during World Wars I and II; the interruption of the supply of unskilled European immigrants; new agricultural policies and technologies that weakened southern agriculture's hold on unskilled labor; vast improvements in transportation and communication into and out of the South; economic incentives and federal government decisions to locate more industrial, military, and research capacity in the South; and the success of the Civil Rights Movement in revolutionizing southern race relations.[21] The consequent increase in the regional exchange of workers complemented the South's rising investments in its own children in driving the regional convergence in workers' productive characteristics.

## V. Implications for Income Convergence after 1940

The thorough education of all classes of people is the most efficacious means for promoting the prosperity of the South.

Robert E. Lee, 1867[22]

The historical evidence from the late nineteenth century indicates that local policymakers in the post-bellum South did not embrace Robert E. Lee's advice regarding educational investment. Indeed, the South's low level of investment in human capital has been cited in numerous studies of the region's economic shortcomings (cf. Nicholls 1960, Colberg 1965, and Wright 1986). What, then, were the implications of regional convergence in workers' education for regional convergence in average earnings? Did southern prosperity move in lockstep with educational gains?

Two recent papers highlight the role of human capital accumulation in understanding regional income convergence. Caselli and Coleman (1999) argue that the convergence of agricultural on non-agricultural wages and the movement of workers from agricultural to non-agricultural activities (which they group together as "structural transformation") account for most of inter-regional wage convergence between 1880 and 1990. A key aspect of their explanation is that the cost of acquiring "nonagricultural skills" declined over time, primarily due to declining costs of attending school (though not documented directly), and that this led to a labor supply shift away from the agricultural sector. Connolly (2004) argues that human capital accumulation increased the South's relative income both directly, through workers' productivity,

and indirectly, by facilitating technological diffusion from more advanced places.

In this section, I take a methodological approach that is closer in spirit to the wage analyses that are common in labor economics than to the macroeconomic approaches described above. For each decade, I calculate the change in average educational attainment among workers in each region between the census dates ($\Delta$ED). Then, using estimated returns to education for each region ($\beta_{ED}$), I calculate a counterfactual average log wage level by subtracting $\beta_{ED} \times \Delta$ED from the actual average log wage. I estimate the average returns to education using separate Mincerian regressions for workers in the South and Non-South, controlling for state fixed effects, so that the returns to education are measured based on within-state variation in attainment. For simplicity, the regressions assume a linear relationship between log income and years of education, and the estimates are made conditional on a quartic in age, and dummy variables for gender, race, foreign birth, and weeks of work (the IPUMS "wkswork2" variable). The basic idea is to take, say, the 1949 wage structure in each region as given and then to reset the average education level to its 1940 position to predict a counterfactual level of mean earnings for 1949.

This approach might lead to understatements of education's contribution. For example, the incremental increase in educational attainment might lead to faster technological transfer or region-specific innovation that makes all workers more productive; if so, the "subtracted off" portion of wages ($\beta_{ED} \times \Delta$ED) misses an important component of education's contribution, and the counterfactual wage level is set too high. Even so, the calculations provide a conservative basis for sizing up the extent to which incremental improvements in workers' education directly influenced earnings.

To illustrate the factors that drive the subsequent calculations, figure 3 graphs estimated returns to education at a more disaggregated level than just "South" and "Non-South". From a set of separate regressions for each state in 1949, the figure plots returns to education against average educational attainment in the state.[23] The southern states typically had significantly higher returns to human capital and significantly lower initial levels of education. Additional analysis (not shown) reveals that states with high returns to education in 1949 experienced rapid improvements in the average education level of their workers. We know from the previous section that this was accomplished primarily through investments in regional natives and that interregional migration also boosted the rate of educational convergence.

The combination of high returns and rapid educational gains suggests potentially large effects on interstate income convergence. Over time, one would expect the returns to education in the South to fall relative to those elsewhere as the South accumulated more human capital (and they did), but in the meantime educational convergence had an amplified impact on wage convergence because of the interstate differences in returns to education.

Table 2 reports South versus Non-South comparisons for each census year from 1940 to 2000. The first column reports the change in average educational attainment for workers with positive earnings and a full set of reported characteristics for the wage regressions. The second column reports the estimated returns to education in the later census year. Two important points are evident: first, as suggested in figure 3, the returns are higher in the South than elsewhere throughout the second half of the century; second, the estimated returns converge strongly, and by 2000 there was virtually no difference between South and Non-South. The third column is the product of the first two and measures (with caveats noted above) the contribution of

education to average wage growth. These values are always larger in the South than in the Non-South, and therefore help drive income convergence.

In 1960, for example, the actual gap between the South and the rest of the country in average log wages was 0.357, but the counterfactual gap is 0.397 (in the absence of the South's educational catch-up during previous ten years). Overall, there was very little regional convergence in actual wages during the 1950s; the gap was 0.371 in 1950 and 0.357 in 1960. In the absence of the influence of educational convergence (0.040), the calculations suggest that the regional income gap would have actually widened. During the 1960s, there was strong inter-regional convergence in the average log wage (from 0.357 to 0.212), and 0.030 log points of that convergence is attributable directly to educational convergence. Changes of similar magnitudes occurred in the 1970s. After 1980, the pace of inter-regional convergence in both income and education slowed considerably. Summed over the second half of the twentieth century (1950-2000), earnings in the South converged on those elsewhere by 0.286 log points, of which 0.120 can be directly attributed to convergence in workers' educational attainment at prevailing rates of return.

This view of southern convergence emphasizes *changes* in the level of workers' education leading to *changes* in the level of workers' income. In this regard, it is important to note that in standard macroeconomic convergence regressions I find no sign of an "impoverished sophisticate" theme to American convergence after 1940.[24] Conditional on log initial income per worker, states with fast per worker income growth between 1940 and 1980 were *not* those with initially high levels of educational attainment: the 1940 education coefficient is negative, small, and statistically insignificant (beta = -0.028; standard error = 0.022). However, replacing the

level of educational attainment with the change in educational attainment reveals a positive, sizable, and statistically significant relationship (beta = 0.049; standard error = 0.025).[25] Instrumenting for the change in average educational attainment between 1940 and 1980 with the change in the average number of days schools were open between 1910 and 1950 yields a larger coefficient estimate (beta = 0.199, standard error = 0.116).[26] This instrumental variable is not ideal, but it may sidestep the problem of reverse causality running from post-1940 income gains to educational investments. The OLS and IV results are consistent with the argument that educational gains were a key factor in driving southern income gains relative to the rest of the country.

## VI. Conclusion

In the context of educational convergence between blacks and whites, Collins and Margo (2006) emphasize that inter-generational factors have always moderated the pace of convergence. Historically, children whose parents had low levels of education tended to acquire significantly less education than other children, all else equal. In the context of regional education gaps, the same considerations are likely to be important. Even if all else had been equal, the erosion of the large regional gap in educational attainment would have taken decades. But, in fact, all else was far from equal. After the Civil War, the South had a relatively low level of average income, a predominantly agricultural economy, and a high degree of inequality. Income and urbanization alone can account for a significant part, but not all, of southern children's lag in school attendance in 1900. There was long-standing opposition to public school

expenditures by politically influential southerners who benefited from the vast supply of unskilled labor and who did not wish to foot the bill for universal public education. Finally, blatant racial discrimination in southern schooling slowed blacks' educational gains.[27]

Nonetheless, after the first three decades of the twentieth century, measures of school quality, school attendance, and educational attainment did converge between the South and Non-South. Due to data imperfections, the exact timing is difficult to pinpoint, but the convergence appears to be especially strong after 1940, as the post-1920 birth cohorts entered the labor force and as inter-regional migration rates increased. Over time, the opposition to southern public schooling eroded with the emergence of the political and economic forces (both internal and external) that shaped the New South. Early in the twentieth century, philanthropic efforts helped boost educational opportunities for black students, and, later, the Civil Rights Movement ended blatant racial discrimination by state and local governments in the provision of public schooling. Rising urbanization and transportation improvements also facilitated increased schooling in the South.

The paper began by posing two main questions: one about the sources of southern educational convergence, and a second question about the contributions of educational convergence to income convergence. Analysis of micro-level census data reveals that most of the South's relative gains were driven by increasing "local investment" in southern-born children, but there was also a significant net flow of outside talent to the South. While the South was absorbing relatively high human capital in-migrants, the rest of the country was absorbing relatively low human capital in-migrants (some southern-born, many foreign-born).

The data also suggest that the timing and pace of southern workers' educational

improvements strongly influenced the magnitude of inter-regional income convergence. Education was only one mechanism contributing to regional income convergence, but it appears to have been an economically important channel even when gauged in a conservative fashion (by private labor-market returns to education). Additionally, improvements in the educational attainment of the southern work force might have facilitated technology transfers, capital transfers, and relative health improvements.

By 2000, the gaps in average educational attainment, income, and returns to education between southern and non-southern workers were small by historical standards. The massive racial disparities of the late nineteenth century also narrowed markedly by the late twentieth century, though racial convergence has stalled in recent decades.[28] The American story echoes a major theme of the new comparative economic history – that economic integration can lead to factor price and income convergence through factor mobility, accumulation, and trade. But even with no internal political barriers to trade or labor mobility and with relatively strong private incentives for the accumulation of human capital in the South, more than a century passed after the Civil War before regional levels of human capital and earnings approached equality.[29] The history of the long struggle to raise southern educational standards and, especially, to improve educational opportunities for African Americans cautions against interpreting regional convergence in educational attainment as merely an automatic response to high returns. History and politics, both local and national, mattered in this story, and they often worked against the forces that economists stress in theoretical descriptions of the convergence phenomenon. Thus, the American story also underscores how long it can take for regions to converge, how the process can be derailed by politics, and how deeply rooted institutional, legal, and social

arrangements may resist the economics of convergence.

References

Alesina, Alberto, Reza Baqir, and William Easterly. 1999. Public Goods and Ethnic Divisions. *Quarterly Journal of Economics* 114: 1243-1284.

Alston, Lee J. and Joseph P. Ferrie. 1999. *Southern Paternalism and the American Welfare State: Economics, Politics, and Institutions in the South, 1865-1965*. New York: Cambridge University Press.

Anderson, James D. 1988. *The Education of Blacks in the South, 1860-1935*. Chapel Hill, NC: University of North Carolina Press.

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." Forthcoming in *American Law and Economics Review.*

Barro, Robert J. and Xavier Sala-i-Martin. 1992. "Convegence." *Journal of Political Economy* 100: 223-251.

Berry, William D., Richard C. Fording, and Russell L. Hanson. 2000. "An Annual Cost of Living Index for the American States, 1960-1995." *Journal of Politics* 62: 550-567.

Borjas, George J. "Self-Selection and the Earnings of Immigrants." *American Economic Review*

77: 531-553.

Card, David and Alan Krueger. 1992. “School Quality and Black/White Relative Earnings: A Direct Assessment.” *Quarterly Journal of Economics* 107: 151-200.

Cascio, Elizabeth. 2004. “Schooling Attainment and the Introduction of Kindergartens into Public Schools.” Working paper.

Caselli, Francesco and Wilbur John Coleman II. 2001. “The U.S. Structural Transformation and Regional Convergence: A Reinterpretation.” *Journal of Political Economy* 109: 584-616.

Coelho, Philip R.P. and James F. Shepherd. 1979. “The Impact of Regional Differences in Prices and Wages on Economic Growth: The United States in 1890.” *Journal of Economic History* 39: 69-85.

Colberg, Marshall R. 1965. *Human Capital in Southern Development, 1939-1963*. Chapel Hill, NC: University of North Carolina Press.

Collins, William J. and Robert A. Margo. 2006 (forthcoming). “Historical Perspectives on Racial Differences in Schooling in the United States,” in *Handbook of the Economics of Education*, edited by E. Hanushek and F. Welch. New York: North-Holland.

Connolly, Michelle P. 2004. “Human Capital and Growth in the Post-Bellum South: A Separate

but Unequal Story." *Journal of Economic History* 64: 363-399.

Donohue, John J. and James J. Heckman. 1991. "Continuous Versus Episodic Change: The Impact of Affirmative Action and Civil Rights Policy on the Economic Status of Blacks." *Journal of Economic Literature* 29: 1603-44.

Donohue, John J., James J. Heckman, and Petra Todd. 2002. "The Schooling of Southern Blacks: The Roles of Legal Activism and Private Philanthropy, 1910-1960." *Quarterly Journal of Economics* 117: 225-268.

Fein, Rashi. 1965. "Educational Patterns in Southern Migration." *Southern Economic Journal* 32: 106-124.

Goldin, Claudia. 1998. "America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century." *Journal of Economic History* 58: 345-374.

Goldin, Claudia and Lawrence F. Katz. 1997. "Why the United States Led in Education: Lessons from Secondary School Expansion, 1910 to 1940." NBER Working Paper 6144.

Lleras-Muney, Adriana. 2002. "Were Compulsory Attendance and Child Labor Laws Effective? An Analysis from 1915 to 1939." *Journal of Law and Economics* 45: 401-435.

Margo, Robert A. 1986. "Race, Educational Attainment, and the 1940 Census." *Journal of*

*Economic History* 46: 189-198.

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

Margo, Robert A. 2000. *Wages and Labor Markets in the United States, 1820-1860*. Chicago, IL: University of Chicago Press.

Margo, Robert A. 2004. "The North-South Wage Gap, Before and After the Civil War," in D. Eltis, F. Lewis, and K. Sokoloff, eds., *Slavery in the Development of the Americas*, pp. 324-351. New York: Cambridge University Press.

Mitchener, Kris James and Ian W. McLean. 1999. U.S. Regional Growth and Convergence, 1880-1980. *Journal of Economic History* 59: 1016-1042.

Neal, Derek. 2006 (forthcoming). "Why Has Black-White Skill Convergence Stopped?" in *Handbook of the Economics of Education*, edited by E. Hanushek and F. Welch. New York: North-Holland.

Nicholls, William H. 1960. *Southern Tradition and Regional Progress*. Chapel Hill: University of North Carolina Press.

O'Rourke, Kevin H. and Jeffrey G. Williamson. 1997. "Around the European Periphery 1870-1913: Globalization, Schooling, and Growth." *European Review of Economic History* 1: 153-190.

Parker, William N. 1980. "The South in the National Economy, 1865-1970." *Southern Economic Journal* 46: 1019-1048.

Rosenbloom, Joshua L. 2002. *Looking for Work, Searching for Workers: American Labor Markets during Industrialization*. New York, NY: Cambridge University Press.

Roy, A. D. 1951. "Some Thoughts of the Distribution of Earnings." *Oxford Economic Papers* 3: 135-146.

Ruggles, Steven, and Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. 2004. *Integrated Public Use Microdata Series: Version 3.0* [Machine-readable database]. Minneapolis, MN: Minnesota Population Center.

Sandberg, Lars G. 1979. "The Case of the Impoverished Sophisticate: Human Capital and Swedish Economic Growth before World War I." *Journal of Economic History* 39: 225-241.

Sundstrom, William A. and Joshua L. Rosenbloom. 1993. "Occupational Differences in the

Dispersion of Wages and Working Hours: Labor Market Integration in the United States, 1890-1903." *Explorations in Economic History* 30: 379-408.

Suval, Elizabeth M. and C. Horace Hamilton. 1965. "Some New Evidence on Educational Selectivity in Migration to and from the South." *Social Forces* 43: 536-547.

Tindall, George R. 1967. *The Emergence of the New South, 1913-1945*. Baton Rouge, LA: Louisiana State University Press.

Tushnet, Mark V. 1987. *The NAACP's Legal Strategy against Segregated Education, 1925-1950*. Chapel Hill, NC: University of North Carolina Press.

Vigdor, Jacob L. 2002. "The Pursuit of Opportunity: Explaining Selective Black Migration." *Journal of Urban Economics* 51: 391-417.

Williams, Heather Andrea. 2004. *Self-Taught: African American Education in Slavery and Freedom*. Chapel Hill, NC: University of North Carolina Press.

Williamson, Jeffrey G. 1965. "Regional Inequality and the Process of National Development: A Description of the Patterns." *Economic Development and Cultural Change* 13: 3- 45.

Williamson, Jeffrey G. 1980. “Unbalanced Growth, Inequality, and Regional Development: Some Lessons from U.S. History.” In Victor L. Arnold, ed., *Alternatives to Confrontation: A National Policy Toward Regional Change*. Lexington, MA: Lexington Books.

Williamson, Jeffrey G. 1996. “Globalization, Convergence, and History.” *Journal of Economic History* 56: 277-306.

Williamson, Jeffrey G. and Peter H. Lindert. 1980. *American Inequality: A Macroeconomic History*. New York: Academic Press.

Woodward, C. Vann. 1951. *Origins of the New South, 1877-1913*. Louisiana State University Press.

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

Table 1A: Accounting for the Regional Differences in Workers' Education, 1940-2000

| | (1) Average for Residents | (2) Average for Stayers | (3) Average for In-Migrants | (4) In-Migrant Weight | (5) Average for Born-In Region | (6) Average for Out-Migrants | (7) Out-Migrant Weight | (8) Net Migration Effect (1-5) |
|---|---|---|---|---|---|---|---|---|
| Panel A: South | | | | | | | | |
| 1940 | 7.94 | 7.77 | 9.68 | 0.093 | 7.91 | 8.72 | 0.151 | 0.03 |
| 1950 | 8.86 | 8.60 | 10.83 | 0.116 | 8.72 | 9.20 | 0.203 | 0.14 |
| 1960 | 9.92 | 9.63 | 11.75 | 0.137 | 9.68 | 9.86 | 0.231 | 0.24 |
| 1970 | 11.10 | 10.80 | 12.59 | 0.169 | 10.85 | 11.05 | 0.220 | 0.25 |
| 1980 | 12.25 | 11.98 | 13.17 | 0.227 | 12.04 | 12.26 | 0.201 | 0.21 |
| 1990 | 13.05 | 12.87 | 13.51 | 0.287 | 12.94 | 13.31 | 0.175 | 0.11 |
| 2000 | 13.43 | 13.35 | 13.60 | 0.336 | 13.44 | 13.93 | 0.153 | -0.01 |
| | | | | | | | | |
| Panel B: Non-South | | | | | | | | |
| 1940 | 9.51 | 10.01 | 7.76 | 0.223 | 10.03 | 10.63 | 0.035 | -0.52 |
| 1950 | 10.22 | 10.63 | 8.61 | 0.201 | 10.67 | 11.47 | 0.047 | -0.45 |
| 1960 | 10.90 | 11.22 | 9.55 | 0.193 | 11.28 | 12.21 | 0.057 | -0.38 |
| 1970 | 11.81 | 12.02 | 10.83 | 0.177 | 12.08 | 12.90 | 0.070 | -0.27 |
| 1980 | 12.73 | 12.92 | 11.86 | 0.180 | 12.97 | 13.48 | 0.097 | -0.24 |
| 1990 | 13.38 | 13.60 | 12.46 | 0.195 | 13.65 | 14.00 | 0.120 | -0.27 |
| 2000 | 13.67 | 13.90 | 12.86 | 0.219 | 13.95 | 14.28 | 0.132 | -0.28 |
| | | | | | | | | |
| Panel C: Difference, South - Non-South | | | | | | | | |
| 1940 | -1.54 | -2.24 | 1.92 | -0.130 | -2.12 | -1.91 | 0.116 | 0.55 |
| 1950 | -1.36 | -2.03 | 2.22 | -0.085 | -1.95 | -2.27 | 0.156 | 0.59 |
| 1960 | -0.98 | -1.59 | 2.20 | -0.056 | -1.60 | -2.35 | 0.174 | 0.62 |
| 1970 | -0.71 | -1.22 | 1.76 | -0.008 | -1.23 | -1.85 | 0.150 | 0.52 |
| 1980 | -0.47 | -0.93 | 1.31 | 0.047 | -0.93 | -1.22 | 0.104 | 0.45 |
| 1990 | -0.32 | -0.73 | 1.05 | 0.092 | -0.71 | -0.69 | 0.055 | 0.38 |
| 2000 | -0.24 | -0.56 | 0.74 | 0.118 | -0.51 | -0.35 | 0.021 | 0.27 |

Notes: Samples include all observations between 18 and 64 years of age who are in the labor force. The "average for residents" is a weighted average of education for "stayers" and "in-migrants". "Stayers" are those who reside in the region in which they were born. "In-migrants" include both foreign-born workers

and migrants from other US regions. The "average for born-in region" is a weighted average for "stayers" and "out-migrants". "Out-migrants" are those born in one region but residing in the other; workers who left the US would exit the sample all together. Migrant status is determined by place of birth compared with place of residence; it is impossible to identify those who moved as children (and therefore might have been educated in a region other than their birth region). Educational attainment codes changed in 1990. For comparison with earlier years, I assigned midpoint values for workers with low levels of education (e.g., 1st to 4th grade code is assigned a value of 2.5); "some college" and associates degrees holders are assigned two years of college (value of 14); bachelor's degrees, 16; master's degrees, 18; professional degrees, 19; and doctorates, 21.

Sources: Calculated using microdata from the IPUMS (Ruggles et al. 2004).

Table 1B: Accounting for the Regional Convergence in Workers' Education, by Region of Residence, 1940-2000

| | Total Change | Contribution of Δ Native Educ. | Contribution of Δ Migrant Educ. | Contribution of Δ Migrant Share |
|---|---|---|---|---|
| Panel A: South | | | | |
| 1940-50 | 0.915 | 0.736 | 0.133 | 0.046 |
| 1950-60 | 1.061 | 0.890 | 0.126 | 0.045 |
| 1960-70 | 1.182 | 0.971 | 0.142 | 0.069 |
| 1970-80 | 1.151 | 0.915 | 0.132 | 0.104 |
| 1980-90 | 0.799 | 0.630 | 0.099 | 0.071 |
| 1990-2000 | 0.379 | 0.317 | 0.030 | 0.032 |
| *1940-80* | *4.309* | *3.512* | *0.533* | *0.264* |
| *1940-2000* | *5.487* | *4.460* | *0.661* | *0.367* |
| | | | | |
| Panel B: Non-South | | | | |
| 1940-50 | 0.715 | 0.495 | 0.171 | 0.048 |
| 1950-60 | 0.680 | 0.482 | 0.182 | 0.017 |
| 1960-70 | 0.910 | 0.656 | 0.226 | 0.027 |
| 1970-80 | 0.915 | 0.734 | 0.185 | -0.005 |
| 1980-90 | 0.649 | 0.547 | 0.118 | -0.016 |
| 1990-2000 | 0.300 | 0.239 | 0.087 | -0.027 |
| *1940-80* | *3.220* | *2.368* | *0.764* | *0.088* |
| *1940-2000* | *4.169* | *3.154* | *0.969* | *0.046* |
| | | | | |
| Panel C: Difference, S-NS | | | | |
| 1940-50 | 0.200 | 0.241 | -0.037 | -0.003 |
| 1950-60 | 0.381 | 0.408 | -0.056 | 0.028 |
| 1960-70 | 0.272 | 0.315 | -0.084 | 0.042 |
| 1970-80 | 0.236 | 0.181 | -0.053 | 0.108 |
| 1980-90 | 0.150 | 0.083 | -0.019 | 0.087 |
| 1990-2000 | 0.079 | 0.078 | -0.058 | 0.059 |
| *1940-80* | *1.089* | *1.144* | *-0.231* | *0.176* |
| *1940-2000* | *1.318* | *1.305* | *-0.308* | *0.321* |

Notes: The columns report the components of the decomposition described in the text:

$\Delta S = \alpha_1(S_{N1} - S_{N0}) + (1 - \alpha_1)(S_{M1} - S_{M0}) + (\alpha_1 - \alpha_0)(S_{N0} - S_{M0})$. "Migrants" to each region include both foreign-born workers and immigrants from other US regions.

Sources: Calculated using microdata from the IPUMS (Ruggles et al. 2004).

Table 2: Educational Convergence and Regional Income, 1950-2000

| | $\Delta$ Educ. $t_1$-$t_0$ | $\beta$ Educ. $t_1$ | $\Delta$ Ed. $\times$ $\beta$ Ed. | Counterfact. Mean Log Income, $t_1$ | Actual Mean Log Income, $t_1$ | Total Convergence, $(NS\text{-}S)_1$- $(NS\text{-}S)_0$ |
|---|---|---|---|---|---|---|
| 1940 - 1950 | | | | | | |
| South | 0.960 | 0.070 | 0.068 | 7.256 | 7.323 | |
| Non-South | 0.709 | 0.045 | 0.032 | 7.662 | 7.694 | |
| Diff: NS-S | -0.251 | -0.026 | -0.036 | 0.407 | 0.371 | ----- |
| | | | | | | |
| 1950 - 1960 | | | | | | |
| South | 1.045 | 0.075 | 0.079 | 7.729 | 7.808 | |
| Non-South | 0.689 | 0.057 | 0.039 | 8.126 | 8.165 | |
| Diff: NS-S | -0.356 | -0.018 | -0.040 | 0.397 | 0.357 | -0.013 |
| | | | | | | |
| 1960 - 1970 | | | | | | |
| South | 1.180 | 0.075 | 0.088 | 8.291 | 8.380 | |
| Non-South | 0.910 | 0.064 | 0.059 | 8.533 | 8.591 | |
| Diff: NS-S | -0.270 | -0.010 | -0.030 | 0.241 | 0.212 | -0.146 |
| | | | | | | |
| 1970 - 1980 | | | | | | |
| South | 1.183 | 0.067 | 0.079 | 9.022 | 9.101 | |
| Non-South | 0.942 | 0.057 | 0.053 | 9.156 | 9.210 | |
| Diff: NS-S | -0.241 | -0.010 | -0.026 | 0.135 | 0.109 | -0.103 |
| | | | | | | |
| 1980 - 1990 | | | | | | |
| South | 0.807 | 0.085 | 0.069 | 9.559 | 9.627 | |
| Non-South | 0.660 | 0.077 | 0.051 | 9.695 | 9.746 | |
| Diff: NS-S | -0.147 | -0.007 | -0.017 | 0.136 | 0.119 | 0.010 |
| | | | | | | |
| 1990 - 2000 | | | | | | |
| South | 0.383 | 0.091 | 0.035 | 10.001 | 10.036 | |
| Non-South | 0.309 | 0.089 | 0.028 | 10.093 | 10.120 | |
| Diff: NS-S | -0.075 | -0.002 | -0.007 | 0.092 | 0.085 | -0.034 |

Notes: The counterfactual average equals the actual average minus ($\Delta$ ED $\times$ $\beta_{ED}$). In 1940 only wage and salary income is reported. So, convergence in income in the 1940s cannot be calculated on the basis of

total income (as the 1950 to 2000 figures are). Samples include workers from 18 to 64 who are in the labor force, report positive income, and positive weeks worked. Topcoded income figures are multiplied by 1.4. The coefficient on highest grade of attainment is from regressions in which log income is regressed on a quartic in age, state fixed effects, a series of dummies for weeks worked, and sex, race, and foreign birth dummies.

Sources: Calculated using microdata from the IPUMS (Ruggles et al. 2004).

Figure 1A: Literacy Rates for Adults, 1870-1930

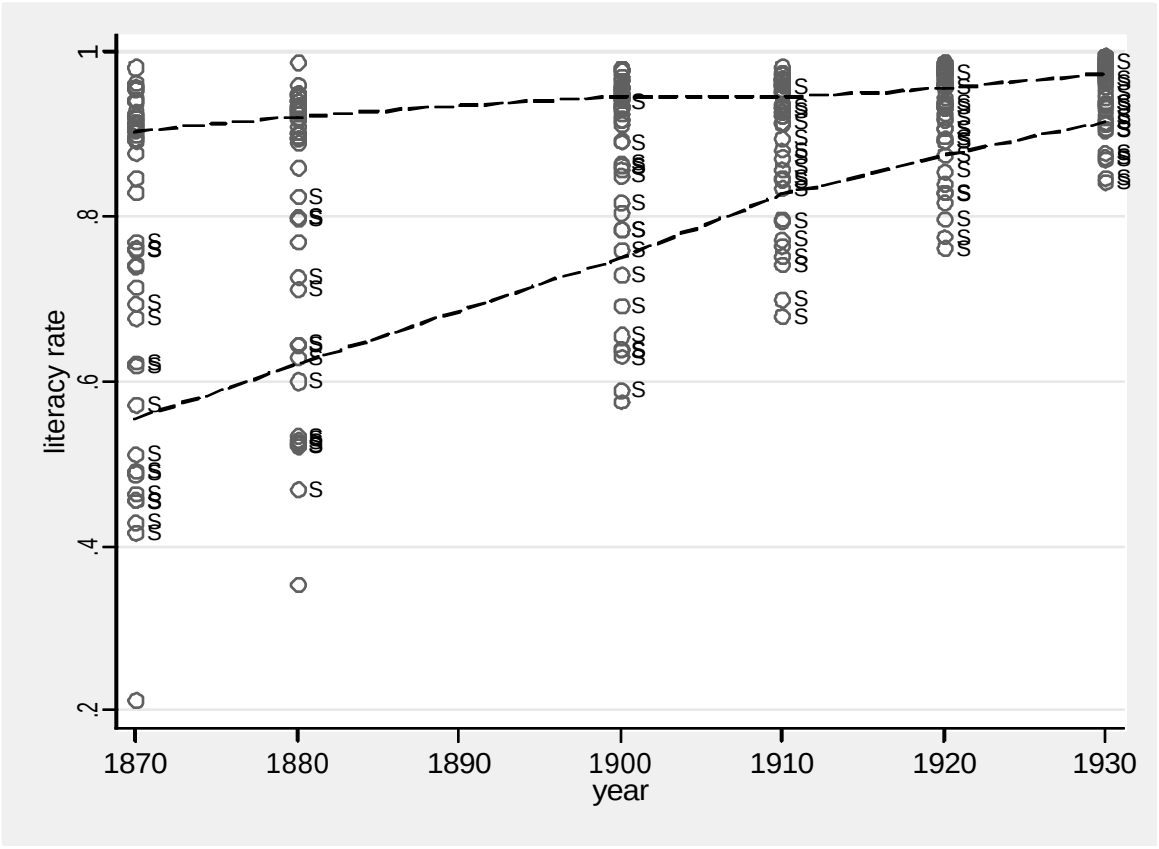
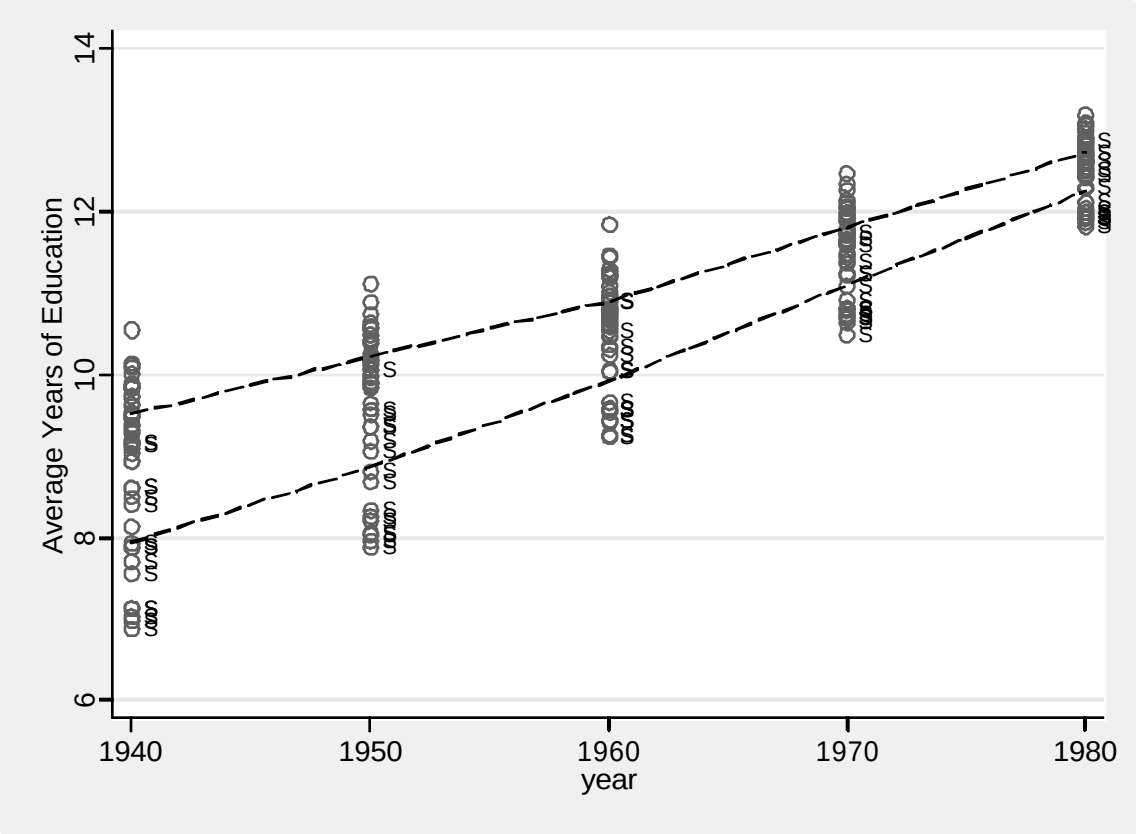


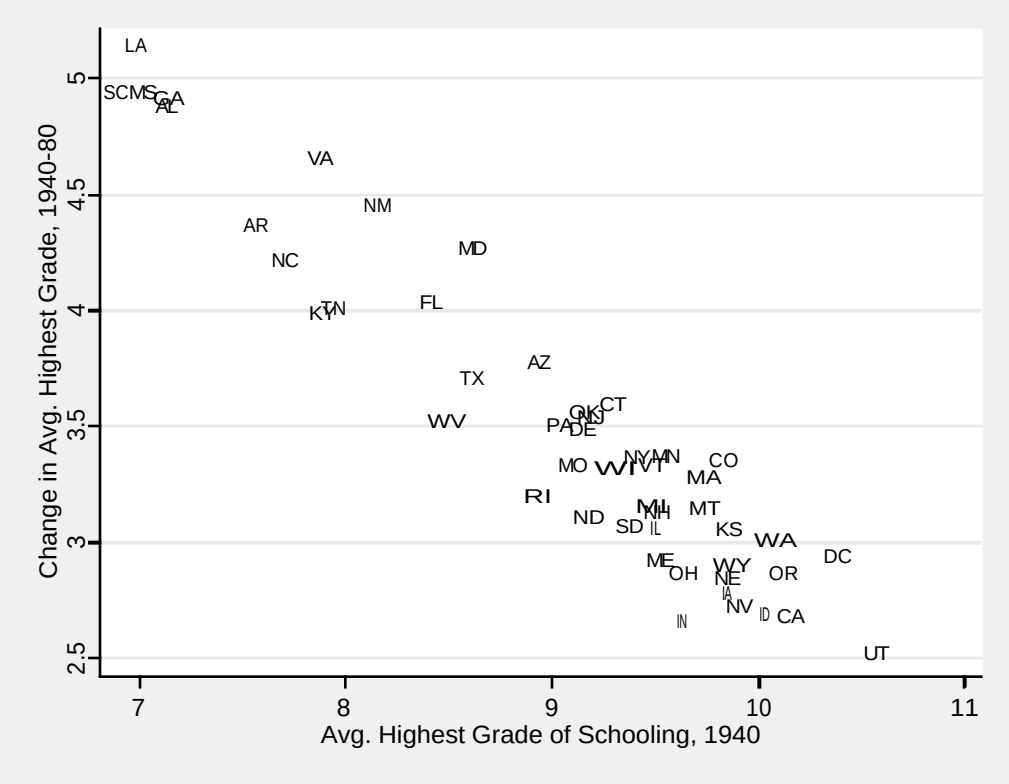
Figure 1B: Average Years of Education, 1940-1980




Notes: Dashed lines connect the southern and non-southern averages from year to year. Each open circle marks a state average in a particular year. "S" is noted next to southern states. Washington, DC is not plotted but is included in the southern average. Samples in figure 1A include everyone from 18 to 64; samples in figure 1B include only those in the labor force, ages 18 to 64. In figure 1A, respondents are counted as "literate" if they can read and write. The low outlier in 1870 and 1880 is New Mexico. In figure 1B, averages are based on respondents' highest grade of completion in 1940 and highest grade of attendance in other years. Person weights are used in calculating state averages.

Sources: Calculated using microdata from the IPUMS (Ruggles et al. 2004).

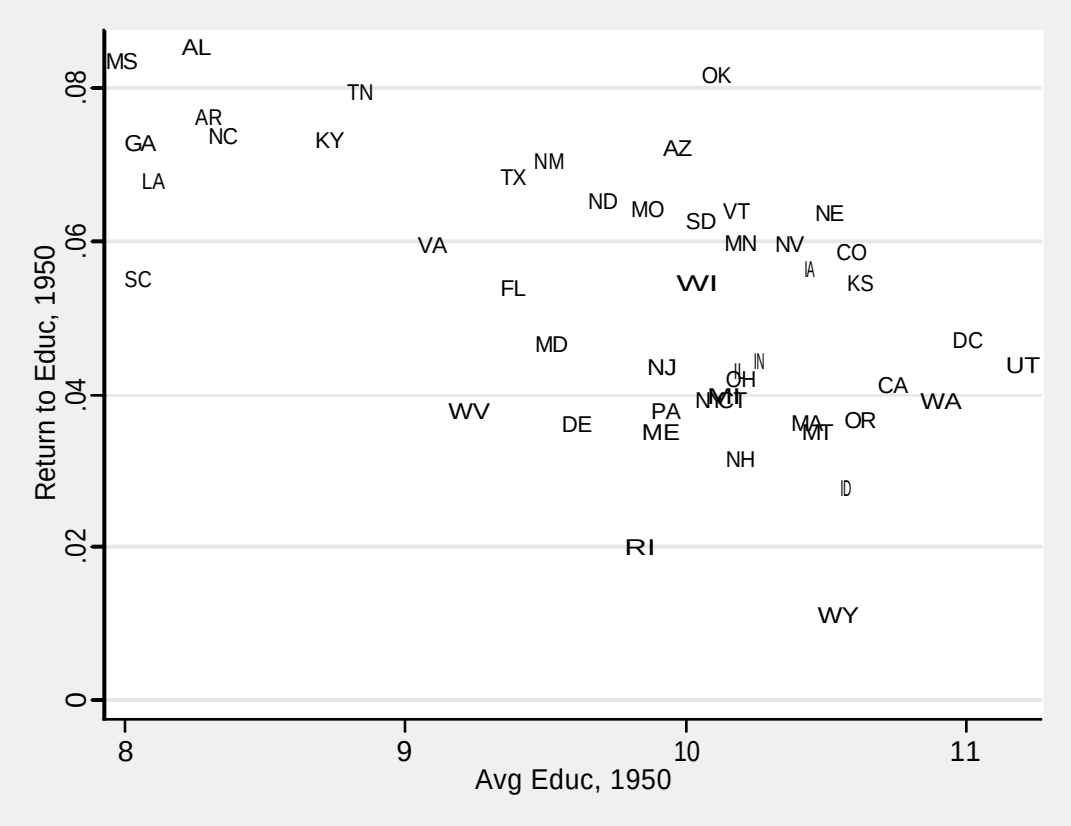
Figure 2: Convergence in Workers' Educational Attainment, 1940-1980




Notes: Averages are based on respondents' highest grade of completion in 1940 and highest grade of attendance in 1980.

Sources: Calculated using microdata from the IPUMS (Ruggles et al. 2004).

Figure 3: Returns to Education and Average Education of Workers, 1950




Notes: Returns to education are estimated from separate regressions for each state in which log total annual income is regressed on highest grade of attendance and several other worker characteristics (age, sex, race, foreign birth, and weeks worked), as described in the text.

Source: Calculated using micro data are from the IPUMS (Ruggles et al. 2004).

## Endnotes

1. Throughout the paper, "South" refers to the Census definition of the region. States include Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia.

2. Accurate measures of literacy and schooling for African Americans prior to 1870 are not available, as the Census did not inquire about slave literacy. The low level of literacy among African Americans in 1870 indicates that very few blacks learned to read and write under slavery. Approximately 19 percent of all 20 to 25 year-old blacks in 1870 could read and write. See Collins and Margo (2006) for more exploration of the "treatment effect" of the Civil War on literacy among blacks.

3. New Mexico is the obvious outlier with a literacy rate of only 22 percent.

4. See Margo (1986), Margo (1990), or Collins and Margo (2006) for more detailed discussions of the accuracy of the years of education data.

5. Rough measures of school quality, such as length of the school term, students per teacher, and expenditures per student, show clear evidence of regional convergence, reinforcing the quantity convergence that is conveyed in the highest-grade-of-attainment measures. See Card and Krueger (1992) and Ashenfelter, Collins, and Yoon (2006) on the labor market implications of racial disparities in school quality within the South.

6.Wright (1986, p. 79) cites an Arkansas planter in 1900: “My experience has been that when one of the younger class gets so he can read and write and cipher, he wants to go to town. It is rare to find one who can read and write and cipher in the field at work.”

7. The low level of support for public education in the South may be interpreted in light of the more general link between ethnic heterogeneity and low public goods provision described by Alesina, Baqir, and Easterly (1999).

8. With respect to blacks, Margo (1990, p. 49) quotes a report by J.W. Joyner in 1910, superintendent of North Carolina’s schools: “There is no surer way to drive the best of them from the state than by keeping up this continual agitation about withdrawing from them the meager educational opportunities that they now have. Their emigration in large numbers would result in a complication of the labor problem. Some of our Southern farms would be compelled to lie untenanted and untilled.”

9. Lleras-Muney (2002) finds that compulsory schooling laws passed in the early 1900s had positive but small effects on educational attainment in both the South and North. The scope for contributing to inter-regional convergence in average attainment appears to be small.

10. The published volumes of the 1890 census also have months of schooling information, but there is no microdata sample for that year.

11. As bleak as that characterization is, the situation was especially poor for black children. Southern states maintained separate school systems for white and black children, and the state

boards of education produced regular reports that describe basic characteristics of the schools (Margo 1990, Card and Krueger 1992, and Donohue, Heckman, and Todd 2002). Around 1910, black school sessions were more than 30 days shorter than white school sessions in the South (on average), and black schools had many more students per teacher than white schools. See Margo 1990.

12. In 1960, non-southern five and six year-olds were about 25 percentage points more likely to attend school than southern children of the same age. See Cascio (2004) on the late expansion of kindergarten programs in the South. Nontrivial inter-regional attendance gaps open up from ages 15 to 17 (on the order of 5 to 8 percentage points) before narrowing for 18 to 21 year olds to about three percentage points. Direct federal involvement in primary and secondary education was minimal until the passage of the Elementary and Secondary School Act of 1965.

13. I thank Mitchener and McLean for sharing the state-level figures. The income estimates for the household heads are based on earnings of workers in the same race-gender-region-occupation category in the 1960 census.

14. OLS results are as follows, with standard errors clustered by state in parentheses: $\beta_1 = -0.907$ (0.249); $\beta_2 = 0.867$ (0.207); $\beta_3 = 0.682$ (0.136); $\beta_4 = 1.250$ (0.062).

15. See Wright (1986, p. 80) and Goldin and Katz (1997, table 3) on this point. Goldin and Katz find that southern states had lower high school graduation rates in 1910 and 1928 than can be

accounted for by differences in basic economic and demographic characteristics; the same is true for the change in the graduation rate between 1910 and 1928.

16. This pattern is consistent with a Roy model of migration in which skilled workers in the area with relatively low returns to education (North) tend to seek employment in places with higher returns (South). See Borjas (1987) for a description of how the Roy model (1951) relates to migration into the U.S. Estimates of regional differences in returns to education are discussed later in the paper.

17. Based on the Roy model, one might expect migrant selection from the South to be negative; i.e., most attractive to the least educated. Indeed, evidence discussed later in the paper indicates that the regional real wage gap was much larger for unskilled than for skilled workers. The higher level of emigration by the educated might reflect greater ease in finding information about distant opportunities, financing for the move, and jobs on arrival.

18. This is not a trivial assumption. For example, Southern-born out-migrants who left as children would have been educated in the Non-South before joining the labor force, and this might have had a positive influence on their educational attainment. This would perhaps work in the opposite direction for children moving into the South. A second consideration is that southern-born children might have invested more in education than they would otherwise to facilitate inter-regional migration.

19. Although the southern wage was low for blacks of all education levels, the inter-regional gap was far larger for workers with low levels of education. But even with a relatively small gap,

more educated blacks could have found it easier to finance migration, to learn about distant employment opportunities, and to get hired for any given job.

20. For the 1980 calculations, I used a cost of living index for each state from Berry et al. (2000).

21.Emphasizing the 1940s as a turning point, Parker writes, "What a Civil War, populist agitation and limited progressivist good will, New South industrialists, New York capital, self-help and cooperatives could not do over nearly a century, was accomplished at last. The back of the Southern institutional peculiarity was broken and its energies and resources released to be merged with the national society" (1980, p. 1046).

22. This is cited in Nicholls (1960, p. 109). Subsequent investigation revealed that the quote is from a letter written by Lee in 1867 to John Brown Gordon.

23.The choice to start with estimated returns to education in the 1950 census is influenced by data constraints in the 1940 census, which reports only wage and salary income rather than total income. The neglect of self-employment income in 1940 is a significant problem because a large portion of southerners were still working in agriculture.

24. The "impoverished sophisticate" reference is to Sandberg (1979) who argues that Sweden's late nineteenth century growth spurt is in large part attributable to its relatively high pre-existing level of education (given its low level of initial income). See O'Rourke and Williamson (1997) for further discussion of Scandinavian growth and education. See Barro and Sala-i-Martin (1992) for a study of income convergence within the United States since 1880.

25. These results are based on simple conditional convergence regressions using state-level data on income per worker (price adjusted) from Mitchener and McLean (1999). The average dependent variable [ln (income1980/income1940)] value is 2.68. The states are not weighted.

26. The rationale for the instrument is that a state's length-of-school-year policy change influences the change in the workers' average educational attainment (F-stat in first stage is 7.1). Because the policy change largely preceded the post-1940 change in income per worker, the approach lowers the likelihood of finding a spurious correlation between contemporary education and income gains due to reverse causality. However, conditional on 1940 income, if states undertook other policy changes that were correlated with the length of school year changes and that had long delayed effects on income growth, then the instrument is not valid.

27. Although not discussed in this paper, the interaction of children's health, cognitive skills, life expectancy, and educational attainment might have influenced the regional gap in human capital.

28. Neal (2005), however, cites evidence of racial divergence among students after 1980.

29. See Margo (2000, p. 45) for wage differences by occupation and region in 1850. See Collins and Margo (2006) for estimated returns to literacy by race and region from 1870 to 1930. See Coelho and Shepherd (1979, p. 76) for wages by occupation and region in 1890 based on the Aldrich Report. Sundstrom and Rosenbloom (1993) also report occupation-based regional wage figures for cities in the late nineteenth century.