# Productivity differences: the importance of intra-state black-white schooling differences across the United States, 1840-2000

Turner, Chad; Tamura, Robert and Mulholland, Sean
Clemson University

31. January 2008

# Productivity Differences: the Importance of Intra-State Black – White Schooling Differences Across the United States, 1840 - 2000

Chad S. Turner Robert Tamura Sean E. Mulholland*

**Abstract**

Using newly created data containing real output per worker, real physical capital per worker, and human capital per worker for US states from 1840 to 2000, Turner et. al (2007) analyze the growth rates of aggregate inputs and total factor productivity (TFP). We continue this line of work by documenting the importance of TFP differences in explaining cross sectional variation in the *levels* of (log) output. We construct plausible upper bounds on the fraction of the variance in output levels that can be explained by TFP and inputs. Similar to the growth rate analysis, we find that TFP can, on average, explain nearly 90% of output variance while inputs can explain up to only 50% of output variance. We then consider the possibility that one major institutional difference across states, the extent to which blacks were denied access to formal education, might explain TFP differences across states. To this end, we generate and present a years of schooling measures, by race, at the state level from 1840 to 2000. While directly exploiting this series has very little impact on the upper bound of the fraction of output variation that can be explained by inputs, we do find that the size of the gap between white and black years of schooling is negatively related to TFP in the period from 1840 to 1950. We also consider the extent to which time-varying rates of return on education alters the upper bound

*Chad S. Turner (Nicholls State University), Robert Tamura (Clemson University and Federal Reserve Bank of Atlanta) and Sean E. Mulholland (Mercer University). This paper was prepared for the 4th Economic Development Conference at Clemson University in honor of Robert E. Lucas, Jr.'s "On the Mechanics of Economic Development." We wish to extend our gratitude to Isaac Ehrlich, Robert E. Lucas, Kevin M. Murphy, and Nancy L. Stokey for their valuable suggestions. Errors and deficiencies that have to this point survived this counsel are most assuredly ours alone. All views expressed herein are the authors and do not represent the views of the Federal Reserve System nor the Federal Reserve Bank of Atlanta.

on the fraction of output variation that can be explained by inputs, finding that time-varying rates have little impact. Finally, we find some evidence for external effects of higher education and physical capital.

Past cross-country studies indicate that variation in total factor productivity (TFP) accounts for a large fraction of the cross-sectional variation in output, whether levels or growth rates are analyzed, e.g. Klenow and Rodriguez-Clare (1997) and Hall and Jones (1999). In searching for explanations of the large TFP variation across countries, researchers often appeal to institutional heterogeneity, e.g. Acemoglu, Johnson, and Robinson (2001).[1] Using a newly created data on real output per worker, real physical capital per worker and human capital per worker at the state level for the United States from 1840 to 2000, Turner, Tamura, and Mulholland (2007), henceforth TTM, analyzes the growth rates of aggregate inputs and TFP. While noting institutions are not entirely homogeneous across states, TTM suggests institutional differences across states are likely to be smaller than across countries. Therefore, state input variation may explain a larger fraction of output variation across states when compared to the cross-country studies. Surprisingly, TTM finds that the vast majority of the cross-sectional variation in output growth rates are explained by TFP differences, even across states.

In this work, we extend this analysis in four main dimensions. First, we examine the variation in the *levels* of (log) output per worker across states. Just as with the analysis of growth rates, the allocation of the correlation between aggregate input and TFP is a central issue. We follow the methodology of Baier, Dwyer, and Tamura (2006) and TTM to construct plausible upper and lower bounds on the fraction of variation in output per worker that can be explained by variation in TFP and variation in inputs. While the variance decomposition results are somewhat sensitive to the assumed rate of return to schooling, we find that TFP can explain as much as 90% of output variation, while inputs can explain only as much as 50% of output variation.

Second, we consider the effect of time-varying rates of return on schooling. TTM chose parameters for the rate of return on schooling to match those used in the cross-country growth literature. Particularly in the last sixty years examined, the diminishing rates of return used in TTM may understate the return to higher education, and thus potentially understate the ability of variation in inputs to explain variation in output. In this work, we allow for time-varying rates of return on schooling.[2] While there is evidence of time-varying rates of return, they have only a small impact

[1] For a greater discussion on the difficulties associated with addressing institutional heterogeneity in cross-country analyses see Temple (1999).

[2] We thank Isaac Ehrlich for suggesting this avenue of research.

on the upper and lower bounds of income variation that can be explained by TFP or inputs.

Third, while institutions are likely to be relatively homogeneous across states, one major institutional difference that existed in the United States was the discrimnatory provision of public schooling for blacks, particularly in slave states. We follow the methodology of Turner, Tamura, Mulholland and Baier (2007), henceforth TTMB, and create years of schooling at the state level, by race, for 1840 - 2000. We subsequently incorporate these measures into development accounting exercises. We examine if inputs can account for a larger fraction of income variation. We begin by requiring the rate of return to schooling be the same for blacks and whites, but subsequently allow returns to vary across races. We find that incorporating these measures directly as inputs has some impact on the upper and lower bounds of income variation that can be explained by TFP or inputs. However, the assumptions required to directly incorporate these measures into accounting exercises, motivated by data availability, are less than ideal. As a result, we further examine the importance of black-white schooling differences by constructing a measure of the gap between years of schooling of white and blacks. We find the size of this schooling gap is negatively related to TFP levels from 1840 to 1950.

Finally, following Lucas (1988), Romer (1986, 1990) and Tamura (2002, 2006), we search for external effects of physical capital and higher education. We find support for external effects of physical capital throughout the time period examined. We find some evidence for external effects arising from higher education exposed workers for both whites and blacks, but it depends on the time period. For the first period, 1840-1950, there is evidence that black workers exposed to higher education contribute positively to a states TFP, whereas in the latter period, 1950-2000 the white workers exposed to higher education contribute positively to state TFP.

Section 2 briefly describes the physical capital per worker, human capital per worker, and output per worker measures created in TTMB and TTM. In Section 3, we conduct development accounting exercises. Section 4 examines time-varying rates of return on education. In Section 5, we introduce the years of schooling measures by race, and examine how incorporating schooling measures that vary by race impacts the results of the development accounting. Section 6 examines the effect of the schooling gap between whites and blacks and provides evidence of external effects of physical and human capital. The final section concludes and outlines plans for future work.

## DATA DESCRIPTION

Before we present the development accounting results, we summarize the measures of output per worker, years of schooling, and human capital created in TTMB and the measures of capital per worker created in TTM, paying special attention to how the cross-sectional variation in inputs and output evolves across time. One thing that will be apparent from examining the output series is that the greatest inequality arises in 1850, and that inequality declines for the remaining century and a half. A similar pattern emerges for human capital and physical capital.

### Output per Worker

Table 1 reports real output per worker by census region created in TTMB. The region with the highest real output per worker in each year is in bold. While the details of the construction of the series are available in TTMB, the data is a combination of work done by Richard Easterlin, BEA state personal income, and original output estimates from TTMB.[3,4] Output measures are expressed in 2000 dollars.

Placing our work in the context of the cross-country income comparisons, we report cross-state and cross-region income comparisons. Figure 1 graphically displays the maximum and minimum values of real output per worker, the coefficient of variation, and the mean value across states for each year. Figures 2 - 4 display the average values of output per worker in each census region. The tremendous convergence across the regions is evident from the graphs. The coefficient of variation peaks in 1850 and falls essentially uninterrupted throughout.

While the ratio between the output level of the state with the highest value of output and the state with the lowest is not close to the 50 to 1 ratio observed in cross country data, it is still meaningful. In 1840, with 28 states, the average worker in the most productive state is 3.4 times more productive than a worker in the least productive state. As new states enter the dataset, a few have extremely high mining output, but most entering with low output levels. Thus the ratio ranges from just below 6 to 1 to more than 19 to 1 over the latter half of the 19th century, before declining throughout the 20th century and settling at just under 2 to 1 in 2000.

[3] For a detailed explanation of the calculation of state real output per worker see TTMB. In particular, Appendix B contains an explanation how output from various sectors and sources were aggregated.

[4] The data from 1840-1920 are at the decadal frequency and are computed from sectoral output. They are output per worker measures. The data from 1929-2000 are annual frequency and come from the BEA as income per worker.

The gaps between the income levels across census regions are smaller and more stable. In 1840 with 28 states and 7 regions, the typical worker in the most productive region was about 2.5 times more productive than a worker in the least productive region. While there is a dramatic increase in the ratio to 13 to 1 in 1850 fueled by the large output level observed in California, the ratio falls to just over 3 to 1 in 1900. By 1950 the ratio has stabilized to around 1.5 to 1, and by 1980 had fallen to about 1.25 to 1, where it has remained. As noted above, the high level of output per worker in the Mountain and Pacific regions early on is quite clear in Figures 3 and 4.

### Physical Capital per Worker

Table 2 presents the reports the physical capital per worker measures created in TTM.[5] It is useful to mention that physical capital per worker are not constructed directly through a perpetual inventory method. Rather TTM uses data on national capital stocks for each sector and state level output in each sector.[6] For years after 1900, state level capital stocks are created by allocating BEA (national) industry level capital stock measures using state level output data and the assumption of a constant capital-output ratio across states within each industry.[7] Prior to 1900, TTM utilizes

[5] The details of the construction of this series are outlined in TTM. These measures are expressed in 2000 dollars. As with our output values, region leaders in each year are in bold.

[6] In the 1840-1920 period TTM constructs output per worker from principally three sectors, agriculture, manufacturing, and all other sectors. For the 1963-2000 period, gross state product (GSP) for each state for nine sectors is avaiable from the BEA. While the BEA does not provide estimates of GSP from 1929-1962, or from 1998-2000 using identical industry classifictions, the BEA does provide measures of wages and salary disbursements in each industry at the state level. There is a very high correlation between wage and salary disbursements and gross state product, and therefore wage and salary disbursements are used to estimate gross state product for 1929-1962 and 1998-2000. The result of combining this data is state level output measures for 3 sectors from 1840 - 1920 and for 9 sectors from 1920 - 2000.

Gallman (1986) reports national measures of capital output ratios for 1840-1900 at the decadal frequency for six sectors as well as sectoral shares. This enables the amount of (national) capital in each sector to be calculated. For 1902 through 2000, data are provided by the Bureau of Economic Analysis in the Fixed Reproducible Tangible Wealth series. This source provides an estimate of the capital stock at the industry level for 1947 through 2000. While this BEA series does not provide data on physical capital stocks for the period 1902 through 1946, it does provide figures on gross investment flows into all industries (except government and residences) which are used to derive estimates of the capital stock. The results of combining this data is national capital stocks measures for 6 industries prior to 1900 and for 9 industries after 1902.

[7] Assuming that a common capital-output ratio holds for each sector in a given year across states is equivalent to assuming that factor returns are equalized across states within sectors. It does not imply that factor returns are equalized across sectors within a state, because we are allowing capital output ratios to vary across sectors, but hold

work from Gallman (1986) that includes capital-output ratios and sectoral shares, using the state level output data from TTMB and again assuming a constant capital-output ratio to allocate capital to each state.

Figure 5 displays the maximum and minimum values of real physical capital per worker, the average across states, and the coefficient of variation in each period. Parallel to what is seen with output per worker, there has been substantial convergence in physical capital per worker. The coefficient of variation peaks in 1850 and falls essentially throughout the remaining 150 years. Figures 6 - 8 display the regional averages. In 1840 the census region with the most capital had roughly 2.5 times more physical capital per worker than the region with the least. In 1850, the ratio peaks at 10 to 1 before falling to just under 2.5 to 1 by 1900. The ratio continues to fall to 1.56 to 1 in 1950, 1.43 to 1 in 1980, and just 1.25 to 1 in 2000. The variation across regions of capital per worker is quite similar to the variation of income per worker.

**Human Capital per Worker**

While details are available in TTMB, the years of schooling measures were calculated using a perpetual inventory method based on enrollment rate data collected from variety of census reports, *Reports of the School Superintendant in the Interior Department*, *Statistical Abstracts of the United States* and *Digests of Education Statistics*.[8] The result is a migration-adjusted measure of years of schooling ($E_{it}$).[9] Table 3 presents years of schooling per worker for each region from TTMB, with region leaders in bold. While the typical worker in the United States had 1 year of formal schooling in 1840, variation across regions is considerable, ranging from 0.25 years to 2.5 years.

Given years of schooling by state and data on the average age of the state population not enrolled

capital shares constant across sectors.

[8] Using a perpetual inventory method, the fraction of the labor force exposed to each category of schooling was computed, as were the years of schooling conditioned on being in each educational category. Multiplying the fraction of the labor force in each category by the years of schooling in that category, then summing across educational categories results in the average years of schooling for each state, $\widetilde{E}_{it}$. TTMB considers this measure a "closed economy" measure as it does not consider the effect of post-education migration. They proceed by combining this measure with census data on place of birth and state of residence to adjust for migration. They assume individuals are educated in the state of their birth and assume that year t residents of state i born in state j have the closed economy average years of schooling of state j in year t. Overlapping data from census years 1940-2000 are used to fit educational attainment of international migrants in each state.

[9] In addition, the average age of those betwen 6 and 65 and not enrolled in school is calculated from census data for later use in calculating the average potential labor market experience level of workers.

in school ($AGE_{it}$), TTM uses the standard labor economics transformation to construct a measure of potential experience ($EX_{it}$) in the labor market:

$$EX_{it} = AGE_{it} - 6 - E_{it} \tag{1}$$

They then create a human capital index for each state from years of schooling and potential labor market experience:

$$h = h_0 \exp\left(\phi_P E_P + \phi_I E_I + \phi_S E_S + \gamma_1 EX - \gamma_2 EX^2\right) \tag{2}$$

where $h_0$ is the level of human capital with no schooling or experience, $\phi_P$, $\phi_I$, and $\phi_S$ are parameters on years of primary, intermediate, and secondary and higher education, and $\gamma_1$and $\gamma_2$ are parameters on experience and experience squared.[10] They follow Hall and Jones (1999) and assign $\phi_P = 0.134$, $\phi_I = 0.101$, and $\phi_S = 0.068$ and use estimates for the return to experience and experience squared from Klenow and Rodriguez-Clare (1997), assigning $\gamma_1 = 0.0495$ and $\gamma_2 = 0.0007$.

The results of these computations are displayed in Figures 9 - 12. Figure 9 contains the average human capital across states, the maximum and minimum value in each period and the coefficient of variation. Similar to the previous series, maximum dispersion occurs in 1850 and diminishes relatively continuously for the remaining 150 years. Figures 10 - 12 display the measures of human capital per worker for each census region.

**Land per Worker**

Finally we present evidence on the reduction of arable land per worker in the United States in Figure 13.[11] For brevity we do not display the land per worker measures by region.

**Growth Rates**

Our focus in this paper is on levels; nonetheless, given that these data spans 160 years, we find it useful to summarize the data by providing information on the growth rates of output and inputs.

[10] Primary schooling is assumed to last 4 years, while intermediate schooling is also assumed to last 4 years. The assumption is made that primary schooling must be completed to attend intermediate schooling, intermediate schooling must be complete to attend secondary, and secondary schooling must be completed to attend higher education.

[11] TTM had not originally intended to include land as an input, but without a measure of land, TFP growth was near zero for 1840 - 1940. After some consideration about the appropriate choice, TTM used a series of land in farms from *Historical Statistics of the United States*, assuming a constant unit price per unit of land. See TTM for additional details.

Table 4 presents the growth of output per worker, $y$, physical capital per worker, $k$, human capital per worker, $h$, land per worker, $\ell$, and TFP, $a$. The assumptions made are that human capital receives 66.7 percent of output, land receives 5 percent of output, and physical capital receives the remaining 28.3 percent of output.[12] We report results for the nation as a whole, census regions, and finally group states into three broader regions (North, South, and West).[13] Average real output per worker growth is 1.45 percent per year for the United States as a whole, while TFP grows at a rate of 0.50 percent per year. The census regions show considerable variation, particularly the western regions, as the Mountain and Pacific display the slowest growth rates of output per worker. This is heavily driven by the high values of real output per worker in the earliest years in California and Nevada where the overwhelming majority of gold and silver was extracted in the US. The final column in the table divides TFP growth by output growth, resulting in the fraction of output growth that is accounted for TFP growth. Even though there is substantial variation in growth rates of real output per worker across regions, on average, input growth accounts for the majority of growth.

Table 5 examines growth rates using data only from 1880 to 2000. This reduces the impact mining in the early years of Western states and also allows us to have complete confidence in the years of schooling measures of each state, as by 1880, the effects of initial conditions on schooling are completely mitigated.[14] Output per worker growth increases from 1.45 percent per year to 1.58 percent per year.[15] Nearly all of the increase in real output per worker growth arises from more rapid TFP growth. As a result, the share of output per worker growth accounted for by TFP growth rises from 34.7 percent to 43.8 percent. The Southern census regions (South Atlantic, East South Central and West South Central) have real output per worker growth in excess of the US average. Despite the fact that physical capital and human capital in the South grow more rapidly than for the country, the share of growth in per worker output accounted for by TFP growth (46%) is similar to the North (50%).

[12] TTM assume that land's share of income is constant across time and equal to 0.05. Later in this work, we relax this assumption.

[13] A list of states contained in each census region and each broader region is available in Appendix A.

[14] It is quite likely we are not capturing the capital value of the ore not mined in the early years of the data. With finer measures of the gold in California and silver in Nevada, one could imagine constructing the capitalized value of the ore in the mines. With this input captured, the rate of growth of TFP would be larger for these regions as the capital value of the ore would be slowly depleted with extraction.

[15] The rising labor force particpation rate of the population accounts for the remaining 0.2% per year difference between the growth rate of output per worker reported here and the familiar 1.8% growth rate of output per *capita*.

## DEVELOPMENT ACCOUNTING

In this section we analyze the determinants of cross-sectional output differences. A key feature in this type of analysis is the allocation of the covariance term between TFP and inputs. One method employed is to assign half the covariance term to each of inputs and TFP, and compare this augmented variance term with the variance of output. An alternative method is to rely on competing economic theories as a means to allocate the observed correlation in the data. The neoclassical growth model of Solow (1956, 1957) and the endogenous growth model of Romer (1990) suggest that technological progress, whether exogenous or endogenous, is the driving engine of growth. In a levels analysis it suggests that high TFP induces high levels of factors, and thus factor accumulation is driven by TFP growth. According to these models, the correlated portion of input growth should be assigned to TFP. In Romer (1986), Lucas (1988) and Tamura (2002, 2006), however, factor accumulation induces TFP growth. Or in the levels, high levels of factors imply high levels of TFP. Hence these models suggest that the correlated portion of TFP should be assigned to inputs. We use a technique first used by Baier, Dwyer and Tamura (2006) and later applied in TTM.[16] We alternately assign the entire correlated variance term to each potential source, and in doing so, create a plausible upper and lower bounds for each of inputs and TFP.

Before beginning the accounting exercises, the variance of log output per worker across the states over time is presented in Figure 14. It is clear that the variance rises dramatically early on with the addition of the western states in the Mountain and Pacific regions. After 1870 there is a trend toward greater equality of output per worker. The Great Depression produced an increase in the variance, reaching a local peak in 1932. From 1940 onward the trend downward continues until around 1980 and then is essentially flat.

We assume that output is produced using a Cobb-Douglas technology combining human capital ($H$), land ($L$), and physical capital ($K$) to produce output ($Y$). Letting lowercase letters represent per-worker variables, we begin with:

$$y_{it} = A_{it} h_{it}^{\alpha} l_{it}^{\gamma} k_{it}^{1-\alpha-\gamma} \tag{3}$$

where $A_{it}$ is the level of TFP in state $i$ and period $t$. We follow TTM and assume $\alpha = 0.667$. We allow $\gamma$ to vary from 0.136 in 1840 to 0.025 in 2000 to capture the decreasing importance of land

[16] Both of these works decompose the variation in the growth rates of output per worker. However, the techniques utilized can be easily applied to conduct a variance decomposition of level of (log) output per worker.

in production across time.[17] We then combine all the factor inputs into a single term, $x$. Thus, output can be represented by:

$$y_{it} = A_{it} x_{it} \tag{4}$$

Taking logs produces:

$$\ln y_{it} = \ln A_{it} + \ln x_{it} \tag{5}$$

We now ask what proportion of the log variance is explained or captured by variation in the two terms on the right hand side of the equation. By the definition of variance, we have:

$$\sigma^2_{\ln y} = \sigma^2_{\ln x} + 2\sigma_{\ln x, \ln A} + \sigma^2_{\ln A} \tag{6}$$

where $\sigma_{\ln x, \ln A}$ is the covariance between $\ln x_{it}$ and $\ln A_{it}$. Dividing by $\sigma^2_{\ln y}$, using the definition of covariance, and rearranging terms results in:

$$1 = \frac{\sigma^2_{\ln x}}{\sigma^2_{\ln y}} + \frac{\sigma^2_{\ln A}}{\sigma^2_{\ln y}} + \frac{2\sigma_{\ln x}\sigma_{\ln A}}{\sigma^2_{\ln y}} \rho_{\ln x, \ln A} \tag{7}$$

where $\rho_{\ln x, \ln a}$ is the correlation between $\ln x_{it}$ and $\ln A_{it}$. If TFP and aggregate inputs are uncorrelated, the first term is the fraction of variance of log income explained by log input variance, while the second term is fraction of the variance of log income explained by variance of log TFP. However, this correlation is not zero empirically, and theoretically it should not be.

The first alternative in dealing with the covariance term, consistent with Romer (1990) and Solow (1956, 1957) assumes the correlation between log inputs and log TFP reflects unmeasured effects of TFP. Assuming a positive correlation, this assumption creates an upper bound on the fraction of the variance of log output that can be explained by variance of log TFP, and therefore creates a lower bound on the fraction of the variance that can be explained by variation in log inputs.

$$\frac{\left(1 - \rho^2_{\ln x, \ln A}\right) \sigma^2_{\ln x}}{\sigma^2_{\ln y}} + \frac{\left(\sigma_{\ln A} + \sigma_{\ln x}\rho_{\ln x, \ln A}\right)^2}{\sigma^2_{\ln y}} = 1 \tag{8}$$

[17] We assume a linear time trend in $\gamma$ from 1840 to 2000. We arrived at our 1840 value by the following reasoning: 80 percent of workers are in agriculture and mining in 1840, and 78 percent of value added comes from agriculture in 1840, this latter figure comes from Table 1 of "Trends in the Structure of the American Economy Since 1840," by Robert E. Gallman and Edward S. Howle in *The Reinterpretation of American Economic History* eds. Robert Fogel and Stanley Engerman.; we assumed that all other output does not utilize arable land, and that land's share in production in agriculture is .17, which comes from Herrendorf and Valentinyi (2007). Hence .136 = .17*.8. Actually the .17 figure comes from their slides, their paper actually reports .18. For the 2000 figure we used the share of land to be .025, which is close to the .03 value that Herrendorf and Valentinyi use for the overall US economy. TTM did not utilize a time varying factor share of income, as doing so would have made the process of decomposing the variation in output growth rates as a function of input growth rates and TFP growth rates much more cumbersome.

The second alternative, consistent with Lucas (1988), Romer (1986), and Tamura (2002, 2006) assumes the correlation between log inputs and log TFP reflects unmeasured effects of log inputs. Assuming a positive correlation, this assumption creates an upper bound on the fraction of the variance of log output that can be explained by variance of log inputs, and thus creates a lower bound on the fraction of the variance of log output that can be explained by variation in log TFP.

$$\frac{\left(\sigma_{\ln x} + \sigma_{\ln A}\rho_{\ln x, \ln A}\right)^2}{\sigma^2_{\ln y}} + \frac{\left(1 - \rho^2_{\ln x, \ln A}\right)\sigma^2_{\ln A}}{\sigma^2_{\ln y}} = 1 \tag{9}$$

Using the data from 1840 to 2000, we can compute these relative upper bounds for both log TFP and log inputs. We display the results in Figures 15 and 16. The effects of the Great Depression and World War II are apparent.[18] From 1950 to 2000 the average plausible upper bound on the importance of variance in log TFP for capturing variance in log real output per worker is 90 percent, while the comparable figure for variation in log inputs is only 46 percent. Even with relatively homogeneous institutions across states, it is striking that TFP explains so large a fraction of output variation. Yet, TTM also reports the high importance of TFP when performing a growth rate analysis. The upper bound of the fraction of output variance that can be explained by inputs decreases steadily after 1940, while the fraction that can be explained by TFP increases Of course this later time frame is a period of small variance in log output per worker.

## TIME VARYING RATES OF RETURN

TTM selects the parameters on years of schooling for ease of comparison with the cross-country literature. However, these international parameters on the returns to schooling may not be appropriate for a *cross-state* comparison. There are two potential criticisms of these parameters. First, it is plausible that diminishing returns to schooling are not as rapid as assumed. In the specification reported above, all years of schooling above eight years return only 6.8 percent per year. Accordingly, a high school graduate would earn only 31 percent more than an eighth grade graduate, and similarly, a four year college graduate would earn only 31 percent more than a high school graduate. A back of the envelope check of median household income levels by head of household education

[18] Clearly input growth like schooling may not be much impacted by the Great Depression. Still one would expect that there would have been substantial deviation of returns on capital both physical and human capital during the period 1929-1940. This would suggest that the Solow residual or TFP would be quite important in picking up the variation of impact of the Great Depression across states. Likewise, with full mobilization during World War II, it is not surprising that factor returns would rise dramatically during this period as there would be no excess capacity.

level suggests that the returns do not decline as precipitously as assumed in TTM.[19] One possibility is the return is roughly constant per year, as posited by Card and Krueger (1992).

The second criticism is whether it is appropriate to assume that these returns are constant over the entire time period. There have been times of compression, as noted by Goldin and Margo (1992) and Freeman (1976) and times of divergence, as shown in Murphy and Welch (1992). We suspect that the returns to higher education may be more variable than the returns to lower levels of education, and thus this criticism to be more pertinent to higher education. Thus, when average schooling is below 4 years, this issue is irrelevant, and when average schooling is below 8 years, it may be relatively harmless. But when average schooling in the US exceeds 8 years, as is the case after 1940, it is likely that the variability across time may play a significant role.

To investigate the effect of assuming diminishing returns to schooling, we reconstruct human capital assuming a constant 13.4 percent return per year of schooling.[20] We then repeat the development accounting exercises. The results are shown in Figures 17 and 18, along with the results from the baseline, diminishing returns case. As can be seen in Figure 17, there is an increase in the share of the variance of log output per worker explained by variance in log inputs from 1929 onward. The increase is most dramatic before 1980. Over the entire 1840 to 2000 period, the assumption of a constant 13.4 percent rate of return increases the average upper bound of the share of log output per worker variance that can explained by variance of log inputs by just more than 6 percentage points to 52 percent, while the average upper bound of the share explained by variance of log TFP falls slightly to 87 percent. A comparison of the average of the upper bounds that can be explained by each source, and for various subperiods, is available in Table 9 for this and all other specifications outlined below.

Perhaps a portion of the TFP importance arises from time varying returns to schooling and not

[19] We use data from Table 675 of the *Statistical Abstract of the United States: 2006* which reports median income of households by education level of head of household. In 2003 median income of a household with a head with less than 9th grade education was $18,787, the median income of a high school graduate household was $36,835, and the median income of a college graduate household is $68,728. These figures imply an annual return of 16.8 percent per year of high school and an annual return of 15.6 percent per year of higher education. Using data from Table 681 of the same Statistical Abstract, but for families instead of households, produces annualized returns of 14.2 percent and 14.9 percent, respectively.

[20] We chose a return of 13.4% for schooling by assuming that the rate of return is identical to value that Hall and Jones (1997) used for primary schooling. It is also consistent with TTMB estimates of returns to schooling from 1840-2000.

simply a higher average rate of return. We estimate time-varying returns, then use these returns to recompute human capital inputs for each state, finally determining the effect on the development accounting results.[21] We estimate both a linear time trend as well as a cubic time trend. We estimate time-varying returns by regressing log real output per worker against years of schooling in the state and interactions of years of schooling with time.[22] In addition we included a dummy variable for 1870, in order to capture the effect of the Civil War on output, and individual year dummies for years 1930 to 1945 inclusive. Thus our specification was:

$$\ln y_{it} = \beta E_{it} + \delta_1(year - 1840)E_{it} + \delta_2(year - 1840)^2 E_{it} + \delta_3(year - 1840)^3 E_{it} + \gamma EX_{it} + \theta \mathbf{Z} \quad (10)$$

where $E_{it}$ is years of schooling, $EX_{it}$ is average experience and $\mathbf{Z}$ is a vector of year dummies.[23] The results of the regression with a linear time trend to returns are contained in the first column of Table 6 below. Figure 19 presents the rate of return on schooling implied by the regression results. We note that the returns to schooling in 2000 still sizably exceed the 9.7 percent average rate of return implied by the original TTM parameters.[24] As the cubic specification is sensitive to the initial return on years of schooling in the first period, we imposed a constraint on the initial 1840 return to schooling of 0.2, but allowed the linear, quadratic and cubic time interactions to freely vary.[25] We kept the return to experience constant over time as the ability to identify higher order time interactions on experience proved difficult. The second column of Table 6 and Figure 19 contain the results of this estimation and the implied time path of the rate of return to schooling.

[21] We generously thank Isaac Ehrlich for suggesting time-varying rates of return to schooling and experience.

[22] Our initial consideration was to simply use the parameters that would result from cross-sectional regressions of output on years of schooling and experience in each year. We rejected this procedure because it results in a stock of human capital that was unreasonably volatile. A small change in the estimated rate of return on schooling between two adjacent years could (and did) result in very large changes in the stock of human capital between those two years. To generate reasonable evolutions of the time path of the stock of human capital, while still exploiting the time variation in the rate of retun, we must in some fashion smooth the rate of return estimates across years. The methodology we outline above is a procedure that does just that.

[23] TTMB found that it was only possible to identify the linear term on experience, as the quadratic term on experience was not identified. Here, we chose to weight by gross state product in order to place more weight on larger and more modern economies than smaller and more agrarian economies.

[24] In 2000, the average number of years of schooling for the United States is 13.5. In TTM, the first 4 years of schooling were given a 13.4% rate of return, the next 4 years of schooling were given a 10.1% rate of return, and all remaing years (in this case 5.5) were given a 6.8% rate of return. This implies the average rate of return observed in 2000 would be 9.7%.

[25] The 0.20 value was obtained by running a regression of log output on years of schooling and experience in the 1840 cross-section.

Again, we note that the returns to schooling varying across time and never fall below 11 percent, suggesting that TTM parameters are understating the return to schooling.

We recompute the share of log output per worker variance that is plausibly explained by variation in log TFP and variation in log inputs and present the results for both the linear time trend rate of return to schooling as well as the cubic time interaction with schooling. We present the results in graphical format in Figures 20 - 21, as well as those of the baseline case. There is no difference in the results between the assumption of linear trend in schooling returns and the assumption of a cubic time interaction with schooling. Interestingly, allowing for a time-varying rate of returns (either linear or cubic) changes the results in nearly the same fashion as introducing a constant rate of return. Further, as is reported below in Table 9, the results from the constant rate of return, linear time-varying rate of return, and cubic time-varying rate of return are nearly identical across all subperiods. The evidence implies that the original diminishing returns to schooling is not consistent with these state data. Both the linear and the cubic time specifications increase the upper bound share of the variance of log output per worker plausibly explained by variation in factors.

## DO BLACK WHITE SCHOOLING DIFFERENCES EXPLAIN THE VARIATION?

Although time varying returns to schooling increases the upper bound on the importance of variation in inputs, the importance of variation in TFP is only slightly reduced. In international comparisons the importance of variance of TFP or the variance of TFP growth for explaining variance in output or variance in output per worker growth is quite well established.[26] After establishing the importance of TFP, most economists comment on the importance of institutional differences. Differences in property right enforcement, the ability of governments to expropriate capital or income, variation in inflation, variation in productive geography or market geography are all examined as possible explanations for differences in cross country productivity, e.g. Acemoglu, Johnson, and Robinson (2001,2002), Baier, Dwyer and Tamura (2007), Sachs (2001) and Temple (1999). The unique feature of US state data is that most of the plausible institutional differences are limited by a common Constitution, language, and the like. All states trade in the common currency (once private circulating money disappeared), share a common international trade policy, interstate trade policy, and federal fiscal policy.[27] One clear institutional difference across states was the existence of slavery in the southern states of the US prior to 1865. Furthermore, post-Reconstruction availability

[26] For levels, see Hall and Jones (1999) and for growth rates, see Baier, Dwyer and Tamura (2006).

[27] We examined the relative size of state governments and found very little practical variation.

of public education for the typical southern black youth was greatly curtailed relative to his or her white counterpart.[28] This difference in the distribution of education among black and whites within a state may play a role in state TFP differences. Thus while the schooling in each state may be consistent with the average schooling of whites and blacks reported in TTMB, the possible importance of within state variation in years of schooling are allocated to TFP in TTMB.

Even if there are schooling differences by race, these differences are unlikely to have an impact on state income levels unless a significant portion of that state's population is black. We present the fraction of state population that is black in ten of the former Confederate states in Table B1 in Appendix B. We note that from 1840 to 1920, blacks were the majority population in several states. Table B2 in Appendix B reports the fraction of black population in the nation as a whole and the fraction of black population that is living in the South.

Using census data on enrollments by race, we provide the first estimates of black and white schooling differences by state from 1840 to 2000 to determine what role the variation in years of schooling within a state has on state income levels.[29],[30] We follow the methodology of TTMB, with only slight modifications.[31,32] Data on the racial composition of enrollment rates, labor force, educational attainment, and population are acquired from decennial census records available through IPUMS.

We report the results of the years of schooling calculations in Table 7. Not surprisingly, there are meaningful differences between blacks and whites. These differences are not only large in the south but in the northern US states as well. Amongst the black population, the New England census region has the highest average years of schooling for nearly the entire period, yet New England blacks still trail New England whites by 1.25 years in 2000. Figures 22 - 25 display black and white years of schooling graphically for each of the census regions. We find there is quite a bit of similarity

[28] For more on this see Canaday (2003), Canaday and Tamura (2007) and Margo (1990).

[29] Prior to the end of slavery it was illegal in slave states to educate a slave. After the end of Reconstruction, around 1876, a system of disenfranchisement of blacks occurred throughout the South, see Canaday (2003).

[30] Where possible, we divide the population into those black and those that are non-black. We refer to the non-black popluation as white.

[31] Those familiar with the methodology in TTMB may recall that in adjusting for migration, we must assign a value of the years of schooling for the foreign born. We are unable to determine the racial makeup of the foreign born. We therefore assume that the foreign born population is comprised only of whites. This assumption requires that the share of the population that is foreign born is set to zero when using the algorithm to correct for black migration.

[32] As the racial composition of the enrollment rates and other demographic variables is based on a sample, we occasionally get noisy values for black years of schooling in states with very small black populations. In these cases, we set the years of schooling for the black population equal to the value for the white population.

within the groupings of census regions as displayed. For example, the years of schooling of whites in each census region in the South are similar, as are the years of schooling of blacks in each census region in the South. The largest gap between black and white schooling for the nation as a whole is 3.5 years and occurs in 1900 and 1910.

To determine how incorporating black and white years of schooling into the production function again requires us to assume whether human capital accumulates at a diminishing rate or a constant rate. We will construct both. First we assume that black and white workers accumulate human capital in the fashion first assumed in TTM, that is declining returns to years of schooling. Further, we assume that workers produce together in a Cobb-Douglas function in which labor is perfectly substitutable measured in efficiency units. We then construct human capital for each race.[33] We assume that the production technology is given by:

$$y_{it} = A_{it}(s_{it}h_{bit} + (1 - s_{it})h_{wit})^{\alpha}\ell_{it}^{\gamma}k_{it}^{1-\alpha-\gamma} \tag{11}$$

where $s_{it}$ is the share of the population that is black and $h_{bit}$ and $h_{wit}$ are the human capital levels of black and white workers, respectively, in state $i$ and period $t$.[34] Because of the convexity in human capital in years of schooling, it is possible that a society that specializes investment in one group versus another group and hence has unequal years of schooling attainment across these groups, may actually have higher levels of human capital than a society without specialized schooling investments. We reexamine the results of the log levels variance decomposition, recalculating the upper and lower bounds of log output per worker variance that can be explained by inputs and TFP. Directly incorporating black-white schooling differences has no effect at all on explaining the variation in log output per worker.[35] While this attempt to directly incorporate black white schooling measures into accounting results and hence output measures is less than satisfactory, it may be the case that this gap is meaningfully in other ways in which we can not measure with overall state income alone, such as within state income variation or state sectoral composition.

[33] We have computed data on the average age of the population by race. Thus, human capital differences between black and whites will come about not only from differences in years of schooling, but also from differences in average experience.

[34] The assumptions made above are forced by the lack of data availability on income by race.

[35] We do not include the graphical display of the upper and lower bounds for brevity. They are indistinguishable from the baseline case.

**Time Varying Returns to Schooling, by Race**

Relaxing our assumption of identical diminishing returns to schooling for whites and blacks, we now allow for distinct time varying returns to schooling for blacks and white, respectively. In order to estimate these returns separately we assume that state human capital is the harmonic mean of black and white human capital in the state. For ease of notation we define $time = year - 1840$, thus we assume:

$$\begin{aligned} h_{it} &= h_{bit}^{s_{it}} h_{wit}^{1-s_{it}} & (12) \\ \ln h_{wit} &= \beta_w E_{wit} + \gamma_{1w} time E_{wit} + \gamma_{2w} time^2 E_{wit} \\ &\quad + \gamma_{3w} time^3 E_{wit} + \delta EX_{wit} & (13) \\ \ln h_{bit} &= \beta_b E_{bit} + \gamma_{1b} time E_{bit} + \gamma_{2b} time^2 E_{bit} \\ &\quad + \gamma_{3b} time^3 E_{bit} + \delta EX_{bit} & (14) \end{aligned}$$

Thus our estimating equation is given by:

$$\begin{aligned} \ln y_{it} &= (1 - s_{it})[\beta_w E_{wit} + \gamma_{1w} time E_{wit} + \gamma_{2w} time^2 E_{wit} + \gamma_{3w} time^3 E_{wit} + \delta EX_{wit}] \\ &\quad + s_{it}[\beta_b E_{bit} + \gamma_{1b} time E_{bit} + \gamma_{2b} time^2 E_{bit} + \gamma_{3b} time^3 E_{bit} + \delta EX_{bit}] + \theta \mathbf{Z} \qquad (15) \end{aligned}$$

where **Z** is again a vector of dummy variables controlling for the Civil War and the years 1930 - 1945, inclusive. We imposed an initial constraint on 1840 returns to schooling of .20 for white schooling and .04 for black schooling.[36] Column 1 of Table 8 presents our estimates from the linear specification, and column two presents the results from the cubic specification. Figure 26 displays the implied rates of return on schooling for both the linear and cubic specifications and for whites and blacks.

We conduct the development accounting,displaying the upper and lower bounds of output variance that can be explained by TFP and inputs in Figures 27 and 28. For the entire 1840 - 2000 period,

[36] We experimented with various intial rates of return to schooling for blacks in 1840. Our cross sectional regressions in 1840 suggested a low rate of return to black schooling in 1840. However, given the extrememly low values of years of schooling in 1840 of blacks, these results were quite noisy. Further, the regression results indicate the rate of return in 1870 was below that of 1840. Choosing an initial value of less than 0.04 resulted in the rate of return for blacks falling below 0 in 1870. We arrived at 0.04 by increasing the initial rate of return in increments of 0.01 until the rate of return in 1870 was non-negative. We chose 0.20 for whites by returning to our initial assumption across the entire population for 1840. Because of difficulty with identifying returns to experience, we impose the constrant the rate of return to experience be identical for whites and blacks.

this specification reduces the average upper bound of output variance that can be explained by TFP from approximately 90 percent in the baseline case, to just over 80 percent. Over the 1980 - 2000 subperiod, the reduction is most dramatic, reducing the upper bound from over 90 percent in the baseline case to just under 73 percent in either the linear or cubic specification. There is also an increase in the upper bound of output variance that can be explained by inputs of similar magnitude. Over the entire period, there is an increase from 46 percent in the baseline case to 57 percent in the linear specification and 53 percent in the cubic specification. Again, the increase is most dramatic in the 1980 - 2000 subperiod, with an increase from under 36 percent to over 50 percent in the linear specification and 47 percent in the cubic specification.

It seems that allowing non diminishing rates of returns, or time varying rates of return on schooling by race, as specified above, increases the ability of inputs to explain output variance significantly, particularly so from 1980 - 2000. We also note that separate black and white returns to schooling can explain more of the variation in output than the parameterization that allowed for a constant rate of return of 13.4% for both black and whites. Directly exploiting information on schooling by race does result in an increased ability to explain output variation. The overwhelming conclusion of all of the exercises above, however, just as was found by TTM in the growth decompositions, is that the TFP is still capable of explaining the lion's share of output variation, even after experimenting with different rates of returns across time and across races

## EXTERNAL EFFECTS AND SCHOOLING DIFFERENCES BY RACE, REVISITED

Our attempts to exploit variation in schooling by race in the previous section are severely limited by data availability. Even if the unrealistic assumption is made that production were entirely segregated, to fully examine the issue at hand would require data on production and capital by race, which is not available. Having found that the variance of log TFP is still quite important in explaining the variation in log output per worker, we now turn to one final line of inquiry. First, following Lucas (1988), Romer (1986, 1990) and Tamura (2002, 2006) we consider the possibility that there are external effects of human and physical capital. If present, we should observe a positive relationship between log TFP and the levels of human capital and log physical capital. As we expect the external effects to be more likely generated by higher education, we focus our attention on the contribution to human capital made by those exposed to higher education. Second, we consider whether black-white schooling gaps are correlated directly with TFP.

Let $gap_{it}$ represent the white-black schooling gap in state $i$ and year $t$:

$$gap_{it} = \left(E_{wit} - E_{bit}\right)^2 / E_{it} \tag{16}$$

where $E_{wit}$, $E_{bit}$, and $E_{it}$, are years of schooling for whites, years of schooling for blacks, and (overall) years of schooling, respectively. We then define $h_{bit}^{\text{college}}$ and $h_{wit}^{\text{college}}$ as the contributions to state i's human capital in period t made by higher education exposed persons of each race. This contribution is defined (for blacks) using the share of the black labor force exposed to higher education, $f_{bit}^{\text{college}}$, the expected years of schooling for those blacks exposed to higher education, $yrs_{bit}^{\text{college}}$, the time varying rate of return to schooling for blacks, $r_{bit}$, and the share of the work force that is black, $s_{it}$. The contribution to human capital made by white higher education exposed workers is defined analogously.

$$h_{bit}^{\text{college}} = s_{it} f_{bit}^{\text{college}} yrs_{bit}^{\text{college}} r_{bit} \tag{17}$$

$$h_{wit}^{\text{college}} = \left(1 - s_{it}\right) f_{wit}^{\text{college}} yrs_{wit}^{\text{college}} r_{wit} \tag{18}$$

Table 10 presents the results of regressions of log TFP on log of physical capital, the white-black schooling gap, and the contribution of higher education exposed black workers and higher education exposed white workers to human capital. We present the regression results based on two different TFP variables, the first arising from the original TTM parameters, which we label $A_{TTM}$. The upper and lower bounds of output variance that can be explained by $A_{TTM}$ are displayed in Figure 15 and reported in row 1 of Table 9. The second utilizes TFP calculated from the specification that allowed the rate of return to vary across race and time with cubic time interactions.[37] We label this second version of TFP $A_{BW\_3}$, and note the upper and lower bounds of output variance that

[37] It was not obvious which version of TFP to utilize in searching for external effects. As noted above, we ultimately decided to report results for two version of TFP. We chose these specification because they were the specifications that had the highest average upper bound of output variation that could be explained by TFP ($A_{TTM}$) and the specification that had the highest average upper bound of output variation that could be explained by inputs ($A_{BW\_3}$).

We expect we will be more likely to find evidence of external effects on higher education using $A_{TTM}$, as the rate of return is understated. On the other hand, with $A_{BW\_3}$, we expect that time-varying rate of return has more ability to pick up the importance of higher education, particularly, in the latter portion of the sample, and thus we expect this to reduce the likelihood of observing external effects of higher education with this version of TFP. We had no priors on which for which version of TFP external effects would be more likely. Our priors were that a significant negative relationship on the black-white schooling gap would be less likely using $A_{BW\_3}$.

can be explained by $A_{BW\_3}$ are displayed in Figure 27 and reported in row 7 of Table 9. Because there are quite a few states that have very few blacks, particularly early on, we weight by the share of the labor force that is black.[38] In each regression we have dummy variables for each region and each year to pick up regional differences and business cycles, we exclude the District of Columbia, and cluster errors by census regions.[39]

Over the entire time period, we find that the black - white schooling gap is significant and negatively related to TFP produced by the original TTM parameters. When we further examine subperiods, we find the relationship is again negative and significant for the 1840-1950 subperiod, but positive and insignificant in the 1950-2000 period. The results are similar using TFP produced by the cubic time trend. The coefficient on the black - white schooling gap is again negative and significant in the 1840-1950 subperiod. We note that the coefficient on the schooling gap is marginally significant and positive in the 1950-2000 subperiod. We do not control for the potential differential quality of schooling as proxied by the assumption that individuals are schooled in state of birth. Selective migration may be related to the puzzle of a positive relationship between the black-white schooling gap and log TFP in the later years. Together, we feel comfortable interpreting these results as supporting the fact that schooling gaps negatively affected TFP from 1840 - 1950.[40]

As a sense of the magnitude of this effect, we consider two different states in two different years, Mississippi in 1870 and Alabama in 1910. Mississippi whites had 2.55 years of schooling, while blacks had .04 years of schooling. Blacks were about 60 percent of the labor force and hence the

[38] Our procedure to compute black years of schooling is based on a sample. A previous footnote indicated that in states where the black population is small and the sample noisy, we assigned blacks the same years of schooling as their white counterparts in the state. In those cases, the black-white schooling gap has a value of zero. However, because the regressions reported in this section are weighted by black labor force, these observations will receive very little weight in the regression.

[39] The District of Columbia is a clear outlier. Concerns with measuring output in general and measuring education levels (residents of Maryland and Virginia attending DC schools) have made it difficult to to be confident that we are measure TFP correctly in DC.

[40] As mentioned in a previous section, the largest absolute average gap between white and black schooling occurs in 1900 and 1910. In 1910, the absolute gap between whites and blacks is 2.95 years compared to an average years of schooling of 4.32 (both of these figures and subsequent figures in this footnote are weighted by the share of popuation black). In 1950, the schooling gap is 2.69 compared to 6.81 years of schooling. By 1980, the schooling gap is 1.52 years compared to 11.76 years of schooling. Relative to the years of schooling, there is considerably less variation in $gap_{it}$ in the 1950 - 2000 period that there was in the 1840 - 1950 period.

While it is beyond the scope of this paper, we intend to conduct further explorations about the approriate functional form to measure the schooling gap in future work.

overall years of schooling the state was 1.1.. The *gap* for Mississippi was 5.74. Had all residents in Mississippi been educated equally, and possessed 1.1 years of schooling, TFP in Mississippi would have been 10 percent higher, using the $A_{BW\_3}$ specification. The average value of $\ln A_{BW\_3}$ was 5.76, whereas Mississippi had a value for $\ln A_{BW\_3}$ of 5.56. Thus half of the TFP gap is explained by the *gap* in schooling between whites and blacks.[41]

In 1910, whites in Alabama possessed an average of 5.4 years of schooling, while blacks held only 1.9. Blacks comprised almost exactly 50% of the population, resulting in an average of 3.65 years of schooling. This unequal treatment produces a *gap* for Alabama of 3.40. Again using the $A_{BW\_3}$ specification, implies that Alabama's TFP was 5.9 percent lower than if the population of equally educated whites and blacks held 3.65 years of schooling. The average value of $\ln A_{BW\_3}$ was 5.79, but Alabama's $\ln A_{BW\_3}$ was 5.33. Hence *gap* explains 13 percent of the TFP gap.[42]

Table 10 also provides evidence that there is correlation between factor inputs, physical capital and human capital and TFP. The evidence is consistent with both induced factor accumulation arising from exogenous technical progress as well as induced technological progress from factor accumulation. This is most strong for physical capital, where all superiods show a positive and significant relationship between physical capital and TFP. It appears from Table 10 that the magnitudes decline as subperiods become more recent. In finer subperiods examined but not reported, we do indeed find the coefficient on physical capital declines uniformly as the time period becomes more recent, with the little connection between log TFP and log physical capital per worker in the most recent periods.

The evidence for external effects of human capital, particularly higher education is mixed across subperiods. We are not surprised that the results for the 1840 - 1950 subperiods are variable. For instance, even in 1940, while the average years of schooling in the US weighted by the black

[41] However in terms of output, this calculation does not follow for the cubic separate black and white returns to schooling. For under the assumption of different returns, in 1870 the typical black rate of return per year of schooling is only 1.2 percent, whereas for whites the rate of return per year of schooling is 15.2 percent. Hence shifting years of schooling away from whites towards blacks would yield a reduction in the human capital component in Mississippi from 1.684 to 1.551. Even assuming that human capital's share of production is .667, this implies a reduction in inputs of 5.3 percent which almost halves the gain in TFP, for an overall increase in Mississippi output per worker of 4.7 percent.

[42] Again, since the return to black years of schooling is only 5.6 percent in 1910 and 12.9 percent to white years of schooling, this equalization reduces the human capital input from 2.174 to 2.042. With human capital receiving .667 weight, this reduces the inputs by 4.1 percent, and hence combined with higher TFP increases Alabama output by 9 percent.

population share is 7.4 years, the contribution made by higher education is only 0.5 years, and the contemporaneous higher education enrollment rate is less than 7 percent. It is not until after WWII that higher education enrollment rates increase dramatically. We therefore put little emphasis in our discussion in these pre -1950 results. We do find evidence of external effects to higher education to whites in the 1950 - 2000 period using both version of TFP. With $A_{TTM}$, we see a positive and significant at the 5 percent level coefficient on white higher educational contribution, and a positive, though not statistically significant coefficient on black higher educational contribution. In the $A_{BW\_3}$ regression, the coefficient on white higher educational contributions is positive and significant at the 10% significance level, while the black contribution has a negative sign and is insignificant. As alluded to earlier, we are not surprised to see less evidence of external effects using the cubic time trend, as the time trend can capture more variation in output that is correlated to higher education.

## CONCLUSION

Motivated by Lucas (1988), we use a new data set with information on real output per worker, real physical capital per worker and human capital per worker for the states of the United States from 1840 to 2000 to examine cross sectional variation in output per worker. This paper examines the importance of variation in log per worker inputs and variation of log TFP for explaining cross-sectional variation log per worker output. Despite the greater commonalities of states of the US compared with cross country differences, we find that the plausible upper bound on the share of log output per worker variation explained by log TFP variation exceeds the plausible upper bound on the share of log input per worker. As was found in the growth rates analysis by TTM, inputs can explain much less of cross-sectional output variation than can TFP.

We explore the most obvious institutional difference across states, the extent to which blacks were denied access to formal education in southern states. To do so, we produce new estimates of years of schooling for both white and black workers. We find that the addition of this more detailed information provides only minor improvements in the ability of log input variability to explaining log output per worker variability. Although the schooling measures has little impact directly as an input, we show that schooling differences clearly show up in the accumulation of factors as well as directly in TFP measures from 1840 - 1950. While there is evidence of time-varying rates of return, and evidence that rates of return to schooling may be understated in the United States if the human capital parameters of Klenow & Rodriguez-Clare (1997) are used, the main result on the importance of TFP is robust to assumptions on the rate of return of schooling, both across time and across races.

Finally we find that there is evidence consistent with the existence of external effects of physical capital, and human capital on productivity. For physical capital it is possible that the higher productive new ideas are embodied in larger capital stocks and this produces the positive correlation observed in these data. Similarly there is evidence that the share of the work force exposed to higher education is correlated with higher productivity. We are left still searching for an answer to Prescott (1998). Why TFP variations across states explain so much of the variation in output per worker across states. In future work we plan to investigate the different output per worker measures by sector by state. Given the manner in which real output per worker and real physical capital per worker are calculated, we have the underlying real output and physical capital by sectors for each state. We seek to determine how much of the aggregate TFP differences arise from differences in TFP at the sectoral level and how much arise from the differences in the composition of output across sectors. Finally we plan on using the more detailed schooling by race to see if the reason for these differences are related to black - white schooling differences.

## REFERENCES


Acemoglu, Daron, Johnson, Simon and Robinson, James A. "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* 91*,* 2001:1369-1401.

Acemoglu, Daron, Johnson, Simon and Robinson, James A. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, 117, 2002: 1231-1294.

Baier, Scott, Dwyer, Gerald, and Tamura, Robert. "How Important Are Capital and Total Factor Productivity for Economic Growth?" *Economic Inquiry* 44, 2006: 23-49.

Baier, Scott, Dwyer, Gerald and Tamura, Robert. "Factor Returns, Institutions, and Geography: A View From Trade," Clemson University working paper, 2007.

Canaday, Neil. "Taxation and the Determinants of Public Spending on the Education of African Americans Under Segregation: Evidence from South Carolina, 1895-1960," unpublished Ph.D. dissertation Clemson University, 2003.

Canaday, Neil, and Tamura, Robert. "White Discrimination in Provision of Black Education: Plantations and Towns," Clemson University working paper, 2007.

Card, David, and Krueger, Alan. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy* 100, 1992: 1-40.

Freeman, Richard. *The Overeducated American* Academic Press, New York:1976.

Gallman, Robert E., and Howle, Edward S. "Trends in the Structure of the American Economy Since 1840," in *The Reinterpretation of American Economic History* eds. Robert W. Fogel and Stanley L. Engerman, Harper & Row, Publishers: New York, 1971.

Gallman, Robert E. "The United States Capital Stock in the Nineteenth Century," in *Long-Term Factors in American Economic Growth*. University of Chicago Press: Chicago, 1986.

Goldin, Claudia, and Margo, Robert. "The Great Compression: The U.S. Wage Structure at Mid-Century," *Quarterly Journal of Economics* 107, 1992: 1-34.

Hall, Robert, and Jones, Charles I. “Why Do Some Countries Produce So Much More Output per Worker than Others?” *Quarterly Journal of Economics* 114, 1999: 83-116.

Herrendorf, Berthold, and Valentinyi, Akos. “Measuring Factor Income Shares at the Sector Level,” Arizona State University working paper, 2007.

Klenow, Peter J., and Rodriguez-Clare, Andres. 1997. “The Neoclassical Revival in Growth Economics: Has it Gone Too Far?”, in *NBER Macroeconomics Annual* 1997, eds. Ben S. Bernanke and Julio J. Rotemberg, pp. 73-103.

Lucas, Robert E., Jr. “On the Mechanics of Economic Development,” *Journal of Monetary Economics* 22, 1988: 3-42.

Margo, Robert A. *Race and Schooling in the South, 1880-1950: An Economic History.* NBER Monograph Series of Long-Term Factors in Economic Development. Chicago: University of Chicago Press, 1990. Pp. vii,164.

Murphy, Kevin M., and Welch, Finis. “The Structure of Wages,” *Quarterly Journal of Economics* 107, 1992: 285-326.

Prescott, Edward C. “Needed: A Theory of Total Factor Productivity,” *International Economic Review* 39, 1998: 525-551.

Romer, Paul. “Increasing Returns and Long-Run Growth,” *Journal of Political Economy* 94, 1986: 1002-1037.

Romer, Paul. “Endogenous Technological Change,” *Journal of Political Economy* 98, 1990: S71-S102.

Sachs, Jeffery D. "Tropical Underdevelopment," National Bureau of Economic Research Working Paper 8119, 2001.

Solow, Robert M. “A Contribution to the Theory of Economic Growth,” *Quarterly Journal of Economics* 70, 1956: 65-94.

Solow, Robert M. “Technical Change and the Aggregate Production Function,” *Review of Economics and Statistics* 39, 1957: 312-320.

Steven Ruggles, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. Integrated Public Use Microdata Series: Version 3.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2004. http://usa.ipums.org/usa/. (viewed: May 2007).

Tamura, Robert. "Human Capital and the Switch From Agriculture to Industry," *Journal of Economic Dynamics and Control* 27, 2002: 207-242.

Tamura, Robert. "Human Capital and Economic Development," *Journal of Development Economics* 79, 2006: 26-72.

Temple, Jonathan. "The New Growth Evidence", *Journal of Economic Literature*, 37, 1999:112-156.

Turner, Chad, Tamura, Robert, Mulholland, Sean and Baier, Scott. "Education and Income of the States of the United States: 1840-2000," *Journal of Economic Growth* 12, 2007a: 101-158.

Turner, Chad, Tamura, Robert, and Mulholland, Sean. "How Important Are Human Capital, Physical Capital and Total Factor Productivity for Determining State Economic Growth in the United States, 1840 -2000," Clemson University working paper, 2007.

Table 1: Real Output per Worker, Labor Force Weighted

| Year | NE | MA | SA | ESC | WSC | MTN | PAC | WNC | ENC | US |
|---|---|---|---|---|---|---|---|---|---|---|
| 1840 | 5,267 | **5,528** | 2,342 | 3,683 | 5,042 | . | . | 3,503 | 4,540 | 4,114 |
| 1850 | 9,077 | 7,901 | 3,302 | 5,344 | 7,346 | 10,250 | **43,207** | 4,641 | 7,343 | 6,691 |
| 1860 | 9,999 | 8,840 | 3,647 | 5,928 | 7,503 | 12,606 | **24,257** | 5,760 | 7,484 | 7,297 |
| 1870 | 9,717 | 10,910 | 3,728 | 4,869 | 6,312 | 15,299 | **16,500** | 7,056 | 7,452 | 7,704 |
| 1880 | 10,998 | 12,954 | 4,752 | 5,447 | 5,971 | 10,951 | **13,786** | 9,248 | 11,147 | 9,449 |
| 1890 | 13,818 | **16,786** | 5,400 | 5,695 | 6,923 | 13,840 | 15,438 | 10,972 | 12,965 | 11,514 |
| 1900 | 13,073 | 14,947 | 5,929 | 5,900 | 7,641 | 13,838 | **14,992** | 12,395 | 13,440 | 11,477 |
| 1910 | 14,230 | **16,234** | 7,909 | 6,774 | 8,633 | 11,789 | 14,188 | 13,167 | 14,682 | 12,554 |
| 1920 | 15,706 | **18,469** | 9,770 | 7,947 | 11,512 | 13,823 | 17,606 | 13,486 | 15,842 | 14,429 |
| 1930 | 19,454 | **21,564** | 11,961 | 9,035 | 11,559 | 14,884 | 19,447 | 14,714 | 17,489 | 16,442 |
| 1940 | 21,518 | **22,639** | 14,278 | 10,240 | 12,993 | 17,247 | 22,302 | 15,515 | 20,512 | 18,328 |
| 1950 | 24,224 | 26,168 | 20,811 | 17,624 | 22,718 | 24,877 | **27,758** | 24,256 | 25,725 | 24,286 |
| 1960 | 26,042 | 29,854 | 26,982 | 24,092 | 28,521 | 28,272 | **35,638** | 26,991 | 31,641 | 29,514 |
| 1970 | 34,919 | 40,110 | 37,781 | 33,949 | 38,449 | 37,353 | **45,806** | 35,770 | 39,605 | 39,139 |
| 1980 | 38,074 | 43,667 | 42,058 | 37,899 | 43,845 | 40,690 | **47,185** | 36,952 | 40,972 | 42,083 |
| 1990 | 45,424 | **51,713** | 49,986 | 46,050 | 48,273 | 46,959 | 50,172 | 44,039 | 47,283 | 48,552 |
| 2000 | 61,426 | **64,758** | 60,216 | 54,134 | 59,833 | 56,277 | 61,374 | 51,527 | 54,162 | 58,791 |

Table 2: Physical Capital per Worker, Labor Force Weighted

| Year | NE | MA | SA | ESC | WSC | MTN | PAC | WNC | ENC | US |
|---|---|---|---|---|---|---|---|---|---|---|
| 1840 | 11,078 | 11,413 | 5,248 | 8,347 | 12,556 | . | . | 6,822 | 8,696 | 8,735 |
| 1850 | 19,934 | 18,268 | 9,047 | 15,521 | 19,703 | 23,641 | 91,162 | 11,971 | 19,871 | 16,737 |
| 1860 | 25,899 | 24,299 | 11,953 | 20,476 | 24,251 | 35,900 | 64,753 | 18,675 | 24,000 | 21,910 |
| 1870 | 21,318 | 25,311 | 10,398 | 14,618 | 17,404 | 34,242 | 38,657 | 19,327 | 20,642 | 19,570 |
| 1880 | 28,288 | 34,269 | 15,248 | 18,718 | 19,822 | 28,630 | 38,774 | 29,721 | 34,099 | 27,681 |
| 1890 | 39,715 | 49,214 | 19,137 | 21,431 | 26,183 | 43,397 | 49,735 | 39,054 | 41,597 | 36,860 |
| 1900 | 38,023 | 44,240 | 19,964 | 20,907 | 27,298 | 45,097 | 48,641 | 43,905 | 42,749 | 36,712 |
| 1910 | 46,154 | 55,078 | 26,618 | 22,650 | 30,012 | 42,405 | 50,455 | 44,483 | 48,407 | 42,319 |
| 1920 | 51,570 | 64,667 | 35,142 | 27,906 | 44,434 | 54,003 | 66,366 | 48,804 | 53,015 | 50,862 |
| 1930 | 67,395 | 83,642 | 50,525 | 41,073 | 56,650 | 76,007 | 78,057 | 62,666 | 71,344 | 67,295 |
| 1940 | 71,662 | 85,122 | 55,391 | 41,914 | 58,369 | 79,066 | 85,279 | 62,042 | 78,388 | 70,770 |
| 1950 | 70,345 | 84,918 | 68,205 | 58,209 | 85,936 | 90,639 | 87,663 | 77,472 | 81,582 | 79,076 |
| 1960 | 74,990 | 93,617 | 86,882 | 78,551 | 113,949 | 104,425 | 110,107 | 87,424 | 98,252 | 95,229 |
| 1970 | 93,481 | 116,576 | 113,415 | 104,430 | 139,152 | 122,870 | 131,981 | 106,661 | 115,310 | 117,295 |
| 1980 | 110,036 | 133,408 | 136,570 | 125,086 | 170,203 | 143,877 | 147,826 | 119,256 | 123,833 | 135,532 |
| 1990 | 131,639 | 156,372 | 160,647 | 148,829 | 180,929 | 160,119 | 159,955 | 137,984 | 141,579 | 154,776 |
| 2000 | 180,011 | 198,041 | 196,498 | 175,629 | 219,668 | 192,082 | 191,901 | 162,149 | 162,722 | 187,992 |

Table 3: Years of Schooling per Worker, Labor Force Weighted

| Year | NE | MA | SA | ESC | WSC | MTN | PAC | WNC | ENC | US |
|---|---|---|---|---|---|---|---|---|---|---|
| 1840 | **2.48** | 1.47 | 0.35 | 0.31 | 0.25 | . | . | 0.46 | 1.04 | 0.97 |
| 1850 | **3.47** | 2.24 | 0.57 | 0.54 | 0.47 | . | 1.92 | 0.86 | 1.79 | 1.50 |
| 1860 | **3.86** | 2.91 | 0.87 | 0.92 | 0.75 | 0.55 | 2.39 | 1.80 | 2.72 | 2.04 |
| 1870 | **4.18** | 3.61 | 1.27 | 1.50 | 1.26 | 2.33 | 2.93 | 2.78 | 3.75 | 2.82 |
| 1880 | **4.69** | 4.54 | 2.02 | 2.24 | 1.94 | 3.23 | 3.63 | 3.77 | 4.69 | 3.64 |
| 1890 | 5.09 | 5.06 | 2.88 | 3.10 | 2.59 | 3.96 | 4.38 | 4.50 | **5.23** | 4.30 |
| 1900 | 5.53 | 5.57 | 3.68 | 4.03 | 3.43 | 4.53 | 5.03 | 5.30 | **5.75** | 4.94 |
| 1910 | 6.15 | 6.07 | 4.21 | 4.57 | 4.02 | 5.25 | 5.76 | 6.03 | **6.24** | 5.48 |
| 1920 | 6.88 | 6.76 | 5.02 | 5.38 | 4.89 | 6.17 | 6.59 | 6.85 | **6.89** | 6.28 |
| 1930 | 7.84 | 7.54 | 6.06 | 6.16 | 6.00 | 7.42 | 7.69 | **7.90** | 7.78 | 7.22 |
| 1940 | 8.79 | 8.23 | 7.43 | 7.25 | 7.59 | 9.17 | **9.68** | 9.16 | 8.92 | 8.41 |
| 1950 | 9.74 | 9.37 | 8.30 | 8.07 | 8.42 | 9.98 | **10.28** | 9.88 | 9.80 | 9.33 |
| 1960 | **10.81** | 10.54 | 9.40 | 9.03 | 9.40 | 10.68 | 10.77 | 10.55 | 10.53 | 10.23 |
| 1970 | 11.22 | 10.89 | 10.37 | 9.92 | 10.21 | 11.35 | **11.46** | 11.24 | 11.07 | 10.87 |
| 1980 | 12.34 | 12.03 | 11.63 | 11.07 | 11.48 | 12.39 | **12.45** | 12.19 | 12.01 | 11.96 |
| 1990 | **13.18** | 12.80 | 12.65 | 12.21 | 12.32 | 13.08 | 12.84 | 12.97 | 12.77 | 12.74 |
| 2000 | **13.93** | 13.59 | 13.45 | 12.96 | 13.18 | 13.60 | 13.59 | 13.61 | 13.44 | 13.48 |

Table 4: Average Growth Rates: 1840 (or when data becomes available) to 2000

| | $y$ | $k$ | $h$ | $\ell$ | $a$ | $a$ / $y$ |
|---|---|---|---|---|---|---|
| All Regions | 1.45% | 1.67% | 0.79% | -1.15 | 0.50% | 0.347 |
| NE | 1.46 | 1.70 | 0.62 | -2.21 | 0.68 | 0.463 |
| MATL | 1.52 | 1.76 | 0.75 | -2.18 | 0.63 | 0.414 |
| SATL | 1.92 | 2.15 | 0.86 | -1.93 | 0.83 | 0.432 |
| ESC | 1.66 | 1.89 | 0.88 | -1.29 | 0.61 | 0.365 |
| WSC | 1.58 | 1.72 | 0.85 | -1.08 | 0.57 | 0.365 |
| MTN | 1.26 | 1.55 | 0.74 | 0.88 | 0.21 | 0.165 |
| PAC | 0.93 | 1.21 | 0.78 | -2.09 | 0.17 | 0.180 |
| WNC | 1.33 | 1.40 | 0.83 | -0.51 | 0.41 | 0.306 |
| ENC | 1.44 | 1.68 | 0.80 | -1.52 | 0.50 | 0.351 |
| North | 1.46 | 1.70 | 0.71 | -1.96 | 0.60 | 0.413 |
| South | 1.77 | 1.98 | 0.86 | -1.56 | 0.71 | 0.401 |
| West | 1.20 | 1.42 | 0.78 | -0.29 | 0.27 | 0.222 |

Table 5: Average Growth Rates: 1880 (or when data becomes available) to 2000

| | $y$ | $k$ | $h$ | $\ell$ | $a$ | $a$ / $y$ |
|---|---|---|---|---|---|---|
| All Regions | 1.58% | 1.64% | 0.75% | -1.56% | 0.69% | 0.438 |
| NE | 1.41 | 1.49 | 0.58 | -2.67 | 0.73 | 0.520 |
| MATL | 1.34 | 1.46 | 0.62 | -2.43 | 0.64 | 0.474 |
| SATL | 2.04 | 2.07 | 0.90 | -2.39 | 0.97 | 0.478 |
| ESC | 1.92 | 1.87 | 0.90 | -1.67 | 0.88 | 0.457 |
| WSC | 1.87 | 1.88 | 0.92 | -1.27 | 0.79 | 0.424 |
| MTN | 1.37 | 1.57 | 0.72 | 0.29 | 0.42 | 0.311 |
| PAC | 1.45 | 1.65 | 0.74 | -2.22 | 0.60 | 0.413 |
| WNC | 1.41 | 1.35 | 0.75 | -0.63 | 0.56 | 0.395 |
| ENC | 1.33 | 1.31 | 0.62 | -1.78 | 0.64 | 0.480 |
| North | 1.37 | 1.42 | 0.60 | -2.30 | 0.68 | 0.496 |
| South | 1.97 | 1.97 | 0.90 | -1.93 | 0.91 | 0.460 |
| West | 1.41 | 1.51 | 0.74 | -0.71 | 0.52 | 0.370 |

Table 6: Rates of Return to Schooling Estimates (standard errors)

| Variable | Linear | Cubic |
|---|---|---|
| $E$ | 0.1426 | 0.2000 |
| | (0.0061) | |
| $time * E$ | -1.68e-4 | -0.0026 |
| | (3.83e-5) | (1.85e-4) |
| $time^2 * E$ | | 2.59e-5 |
| | | (2.35e-6) |
| $time^3 * E$ | | -7.93e-8 |
| | | (7.72e-9) |
| $EX$ | 0.0223 | 0.0248 |
| | (0.0016) | (0.0016) |
| | | |
| period | all years | all years |
| $\overline{R}^2$ | 0.8816 | - |
| $N$ | 4004 | 4004 |

Table 7: Years of Schooling per Worker by Race, Labor Force Weighted

| Year | $NE_w$ | $NE_b$ | $MA_w$ | $MA_b$ | $SA_w$ | $SA_b$ | $ESC_w$ | $ESC_b$ | $WSC_w$ | $WSC_b$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1840 | **2.57** | **0.44** | 1.53 | 0.01 | 0.56 | 0.00 | 0.45 | 0.00 | 0.44 | 0.01 |
| 1850 | **3.58** | **1.14** | 2.32 | 0.48 | 1.20 | 0.01 | 1.08 | 0.00 | 0.92 | 0.03 |
| 1860 | **3.98** | **2.12** | 2.98 | 1.16 | 1.84 | 0.03 | 1.90 | 0.01 | 1.44 | 0.08 |
| 1870 | **4.36** | **1.61** | 3.70 | 1.47 | 2.28 | 0.11 | 2.68 | 0.09 | 2.09 | 0.16 |
| 1880 | 4.81 | **2.43** | 4.42 | 1.85 | 3.10 | 0.52 | 3.51 | 0.56 | 2.88 | 0.45 |
| 1890 | 5.23 | **3.02** | 5.04 | 2.44 | 4.06 | 1.19 | 4.48 | 1.29 | 3.62 | 0.98 |
| 1900 | 5.79 | **3.69** | 5.62 | 2.99 | 4.91 | 1.85 | 5.38 | 1.99 | 4.25 | 1.62 |
| 1910 | 6.30 | **4.13** | 6.06 | 3.36 | 5.48 | 2.39 | 5.88 | 2.53 | 4.86 | 2.16 |
| 1920 | 6.94 | **4.91** | 6.75 | 4.02 | 6.19 | 3.19 | 6.41 | 3.25 | 5.63 | 2.91 |
| 1930 | 7.92 | **6.26** | 7.68 | 4.93 | 6.99 | 4.16 | 7.01 | 4.06 | 6.58 | 3.81 |
| 1940 | 8.80 | **7.51** | 8.27 | 6.16 | 8.14 | 5.26 | 7.94 | 5.15 | 8.12 | 5.09 |
| 1950 | 9.75 | **8.09** | 9.45 | 7.00 | 9.08 | 6.13 | 8.67 | 6.27 | 8.87 | 6.25 |
| 1960 | **10.82** | **8.99** | 10.63 | 8.07 | 9.99 | 7.24 | 9.49 | 7.16 | 9.79 | 7.41 |
| 1970 | 11.29 | **10.19** | 10.99 | 9.45 | 10.70 | 9.11 | 10.18 | 8.97 | 10.43 | 9.12 |
| 1980 | 12.37 | 11.04 | 12.11 | 10.56 | 11.87 | 10.39 | 11.21 | 10.60 | 11.59 | 10.60 |
| 1990 | **13.22** | 12.12 | 12.90 | 11.71 | 12.83 | 11.81 | 12.25 | 11.82 | 12.38 | 12.06 |
| 2000 | **14.01** | **12.76** | 13.73 | 12.33 | 13.59 | 12.68 | 13.01 | 12.39 | 13.26 | 12.52 |

Table 7: Years of Schooling per Worker by Race, Labor Force Weighted (continued)

| Year | $MTN_w$ | $MTN_b$ | $PAC_w$ | $PAC_b$ | $WNC_w$ | $WNC_b$ | $ENC_w$ | $ENC_b$ | $US_w$ | $US_b$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1840 | . | . | . | . | 0.56 | 0.00 | 1.09 | 0.00 | 1.36 | 0.00 |
| 1850 | . | . | 2.02 | 0.00 | 1.19 | 0.02 | 1.94 | 0.19 | 2.13 | 0.02 |
| 1860 | 0.52 | 0.12 | 2.52 | 0.39 | 2.17 | 0.23 | 2.87 | 0.75 | 2.76 | 0.06 |
| 1870 | 2.29 | 0.36 | 2.86 | 1.37 | 3.11 | 0.56 | 3.97 | 1.02 | 3.45 | 0.19 |
| 1880 | 3.23 | 0.62 | 3.69 | 1.09 | 3.97 | 1.16 | **4.86** | 1.79 | 4.19 | 0.62 |
| 1890 | 4.12 | 1.32 | 4.54 | 1.84 | 4.73 | 1.97 | **5.40** | 2.51 | 4.85 | 1.30 |
| 1900 | 4.91 | 2.29 | 5.32 | 3.20 | 5.61 | 2.87 | **5.94** | 3.30 | 5.50 | 1.99 |
| 1910 | 5.43 | 3.27 | 5.95 | 3.12 | 6.20 | 3.65 | **6.43** | 3.94 | 6.00 | 2.53 |
| 1920 | 6.38 | 3.94 | 6.79 | 3.99 | 7.03 | 4.38 | **7.05** | 4.30 | 6.70 | 3.33 |
| 1930 | 7.56 | 4.80 | 7.78 | 4.85 | **8.04** | 5.04 | 7.92 | 5.03 | 7.60 | 4.27 |
| 1940 | 9.22 | 6.47 | **9.66** | 6.29 | 9.19 | 6.59 | 8.96 | 6.50 | 8.67 | 5.48 |
| 1950 | 10.11 | 7.67 | **10.32** | 7.11 | 9.90 | 7.43 | 9.84 | 7.42 | 9.58 | 6.56 |
| 1960 | 10.85 | 8.28 | 10.82 | 8.08 | 10.61 | 8.57 | 10.62 | 8.25 | 10.46 | 7.66 |
| 1970 | 11.47 | 9.82 | **11.51** | 9.62 | 11.35 | 10.05 | 11.18 | 9.66 | 11.04 | 9.33 |
| 1980 | 12.45 | 11.13 | **12.49** | 11.06 | 12.25 | **11.15** | 12.10 | 10.78 | 12.08 | 10.62 |
| 1990 | 13.11 | 12.26 | 12.87 | **12.32** | 13.05 | 12.23 | 12.84 | 11.92 | 12.83 | 11.91 |
| 2000 | 13.69 | 12.76 | 13.65 | 12.54 | 13.74 | 12.64 | 13.55 | 12.33 | 13.59 | 12.51 |

Table 8: Rates of Return to Schooling Estimates (standard errors)

| Variable | Linear | Cubic |
|---|---|---|
| $(1-s)E_w$ | 0.2000 | 0.2000 |
| | | |
| $time * (1-s)E_w$ | -5.06e-4 | -0.0022 |
| | (2.40e-5) | (2.10e-4) |
| $time^2 * (1-s)E_w$ | | 2.16e-5 |
| | | (2.60e-6) |
| $time^3 * (1-s)E_w$ | | -6.84e-8 |
| | | (8.61e-9) |
| $sE_b$ | 0.0400 | 0.0400 |
| | | |
| $time * sE_b$ | 8.81e-4 | -0.00219 |
| | (3.25e-5) | (2.10e-4) |
| $time^2 * sE_b$ | | 4.46e-5 |
| | | (2.40e-6) |
| $time^3 * sE_b$ | | -1.74e-7 |
| | | (1.09e-8) |
| $(1-s)EX_w$ | 0.0184 | 0.0229 |
| | (0.0016) | (0.0016) |
| $sEX_b$ | 0.0184 | 0.0229 |
| | (0.0016) | (0.0016) |
| | | |
| period | all years | all years |
| $N$ | 4004 | 4004 |

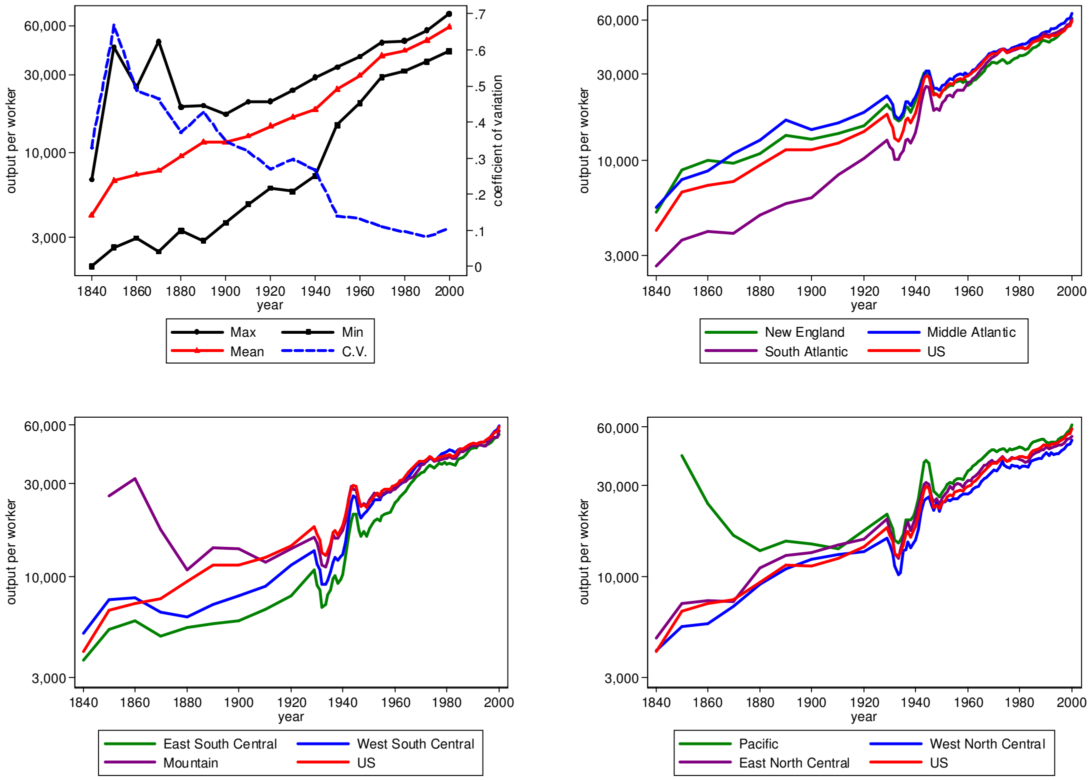
Table 9: Average Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\sigma^2_{\ln A}$ and $\sigma^2_{\ln x}$

| Factor | Fig | Ret. to School | b/w split | 1840-2000 | ex. 1929-49 | 1929-2000 | 1950-2000 | 1980-2000 |
|---|---|---|---|---|---|---|---|---|
| TFP | 15 | TTM | no | .902 | .897 | .907 | .903 | .885 |
| TFP | 17 | .134 | no | .871 | .867 | .875 | .871 | .850 |
| TFP | 20 | linear time trend | no | .875 | .870 | .880 | .875 | .857 |
| TFP | 20 | cubic time trend | no | .875 | .869 | .881 | .876 | .856 |
| TFP | – | TTM | yes | .908 | .903 | .916 | .913 | .903 |
| TFP | 27 | linear time trend | yes | .820 | .802 | .830 | .814 | .727 |
| TFP | 27 | cubic time trend | yes | .825 | .806 | .832 | .812 | .729 |
| | | | | | | | | |
| Input | 16 | TTM | no | .458 | .474 | .447 | .461 | .355 |
| Input | 18 | .134 | no | .520 | .528 | .517 | .524 | .401 |
| Input | 21 | linear time trend | no | .510 | .519 | .506 | .515 | .393 |
| Input | 21 | cubic time trend | no | .499 | .510 | .495 | .506 | .386 |
| Input | – | TTM | yes | .457 | .473 | .446 | .460 | .358 |
| Input | 28 | linear time trend | yes | .567 | .569 | .571 | .576 | .505 |
| Input | 28 | cubic time trend | yes | .535 | .541 | .534 | .541 | .472 |

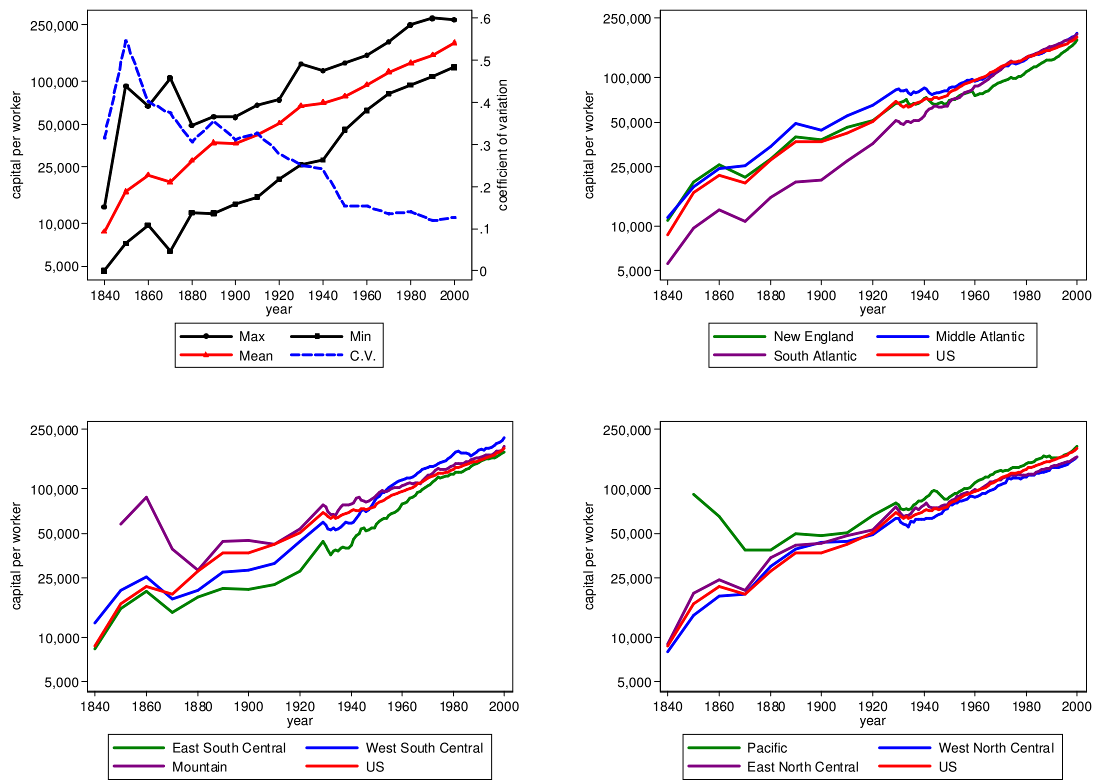
Table 10: Log TFP Regressions *** 1%, ** 5%, *10% significance

| variable | $\ln A_{TTM}$ | $\ln A_{TTM}$ | $\ln A_{TTM}$ | $lnA_{BW\_3}$ | $lnA_{BW\_3}$ | $lnA_{BW\_3}$ |
|---|---|---|---|---|---|---|
| $\ln(k)$ | 0.5240*** | 0.7139*** | 0.3006*** | 0.4840*** | 0.6960*** | 0.2643*** |
| | 0.0390) | (0.0231) | (0.0664) | (0.0334) | (0.0281) | 0.0559 |
| $gap_{it}$ | -0.0278** | -0.0235*** | 0.0041 | -0.0168 | -0.01741*** | 0.0395* |
| | (0.0088) | (0.0052) | (0.0132) | (0.0117) | (0.0038) | (0.0205) |
| $h_{wit}^{\text{college}}$ | 0.1847 | -0.4845** | 0.3717** | 0.0638 | -0.6966*** | 0.2527* |
| | (0.1331) | (0.1742) | (0.1414) | (0.1272) | (0.1862) | (0.1296) |
| $h_{bit}^{\text{college}}$ | 0.2178 | 4.6303*** | 0.2248 | -0.2521 | 5.0695*** | -0.2403 |
| | (0.2081) | (0.9922) | (0.1380) | (0.2258) | (1.0678) | 0.1484 |
| *region dummies* | yes | yes | yes | yes | yes | yes |
| *year dummies* | yes | yes | yes | yes | yes | yes |
| *excludes dc* | yes | yes | yes | yes | yes | yes |
| *weights* | $s_{it}$ | $s_{it}$ | $s_{it}$ | $s_{it}$ | $s_{it}$ | $s_{it}$ |
| *time period* | 1840-2000 | 1840-1950 | 1950-2000 | 1840-2000 | 1840-1950 | 1950-2000 |
| $N$ | 3932 | 1432 | 2550 | 3932 | 1432 | 2550 |
| $R^2$ | 0.9071 | 0.9211 | 0.7898 | 0.8289 | 0.9018 | 0.7445 |

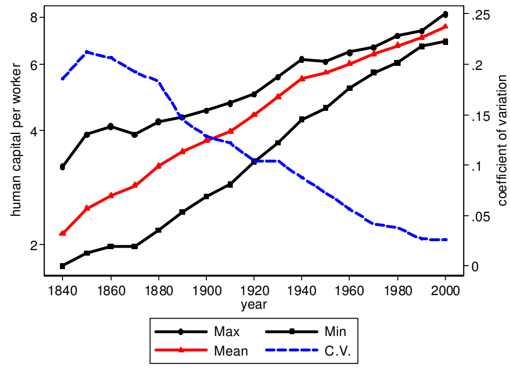
Figures 1 - 4: Real Income per Worker

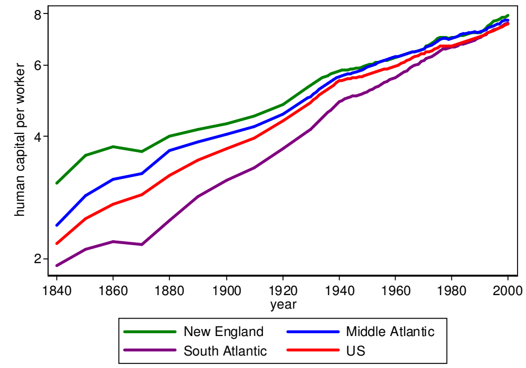


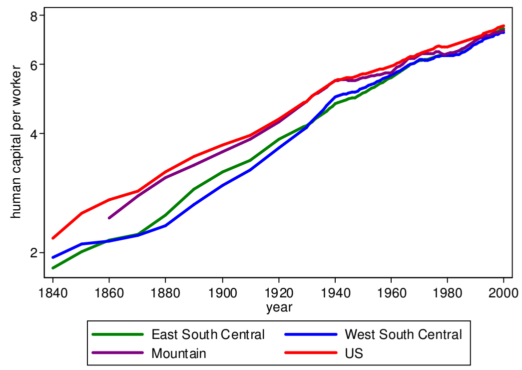
Figures 5 - 8: Physical Capital per Worker



Figures 9 - 12: Human Capital per Worker



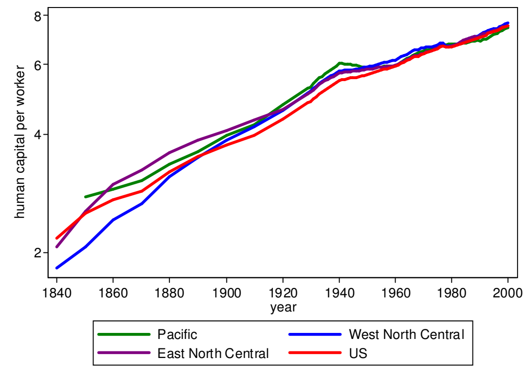



Figure 13: Land per Worker

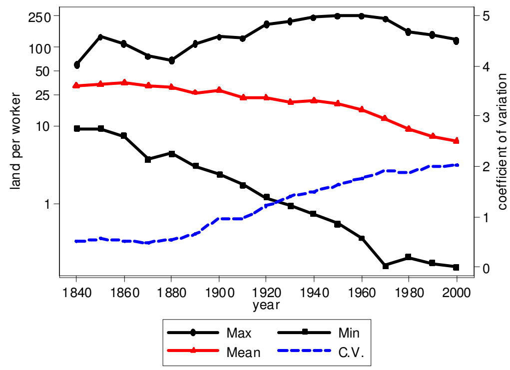
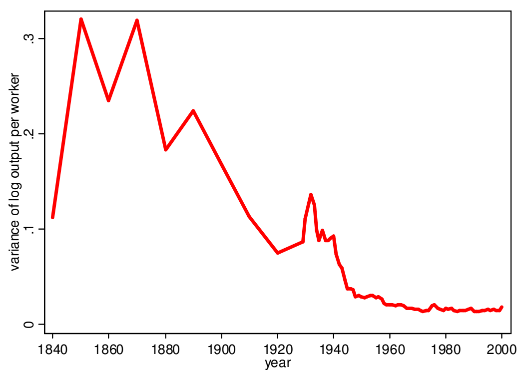
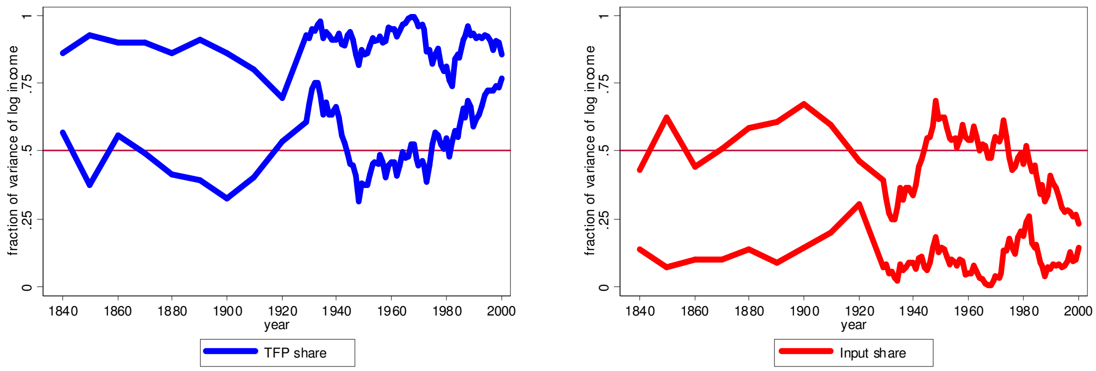



Figure 14: Variance of Log Output per Worker ($\sigma^2_{\ln y}$)



Figures 15 - 16: Upper and Lower Bounds of Fraction $\sigma^2_{\ln y}$ Explained by $\sigma^2_{\ln A}$ and $\sigma^2_{\ln x}$




Figures 17 - 18: Upper and Lower Bounds of Fraction $\sigma^2_{\ln y}$ Explained by $\sigma^2_{\ln A}$ and $\sigma^2_{\ln x}$

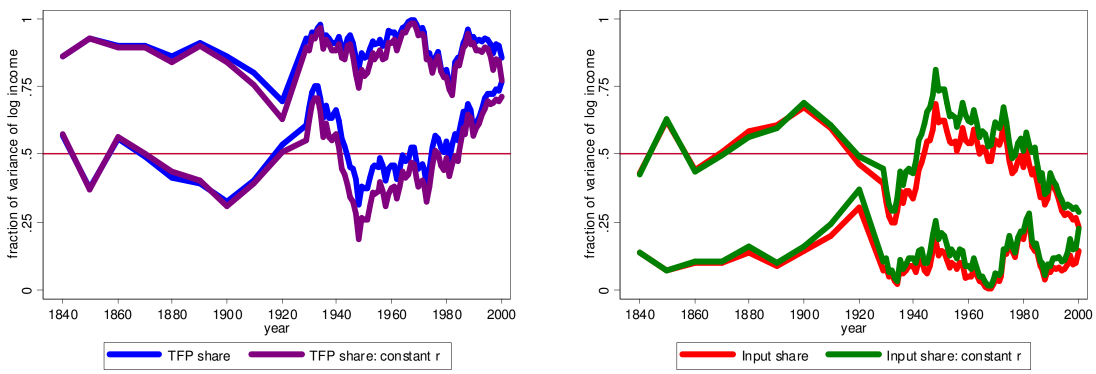


Figure 19: Returns to Schooling, Linear and Cubic Specifications

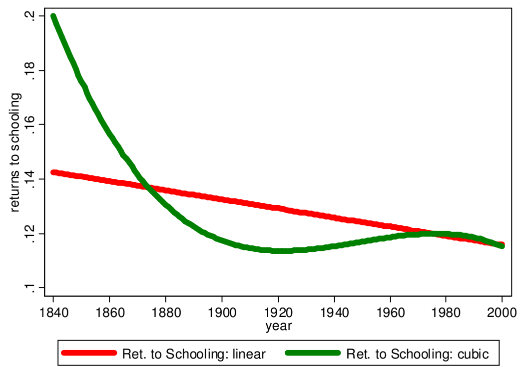
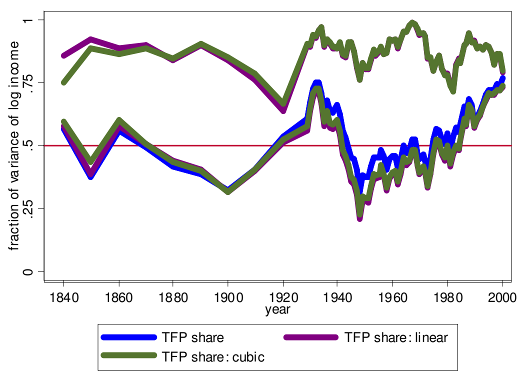
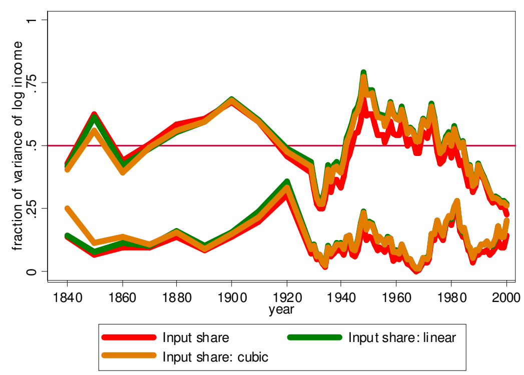



Figure 20 - 21: Upper and Lower Bounds of Fraction $\sigma^2_{\ln y}$ Explained by $\sigma^2_{\ln A}$ and $\sigma^2_{\ln x}$



Figures 22 - 25: Years of Schooling, by Race

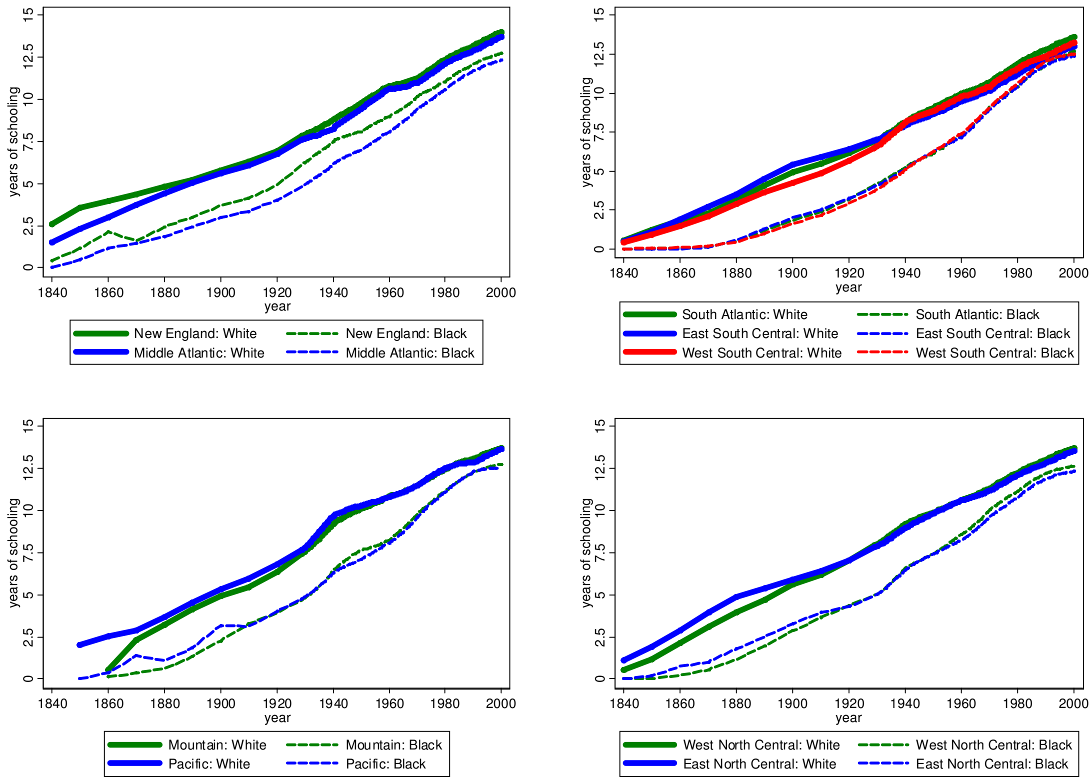
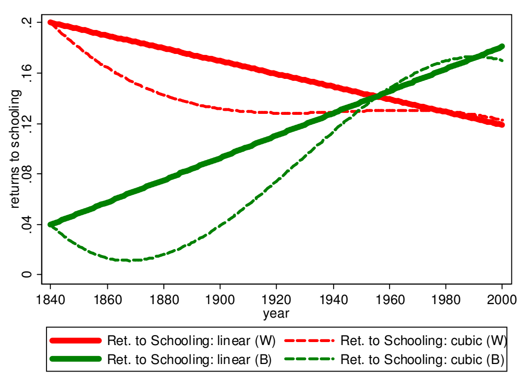
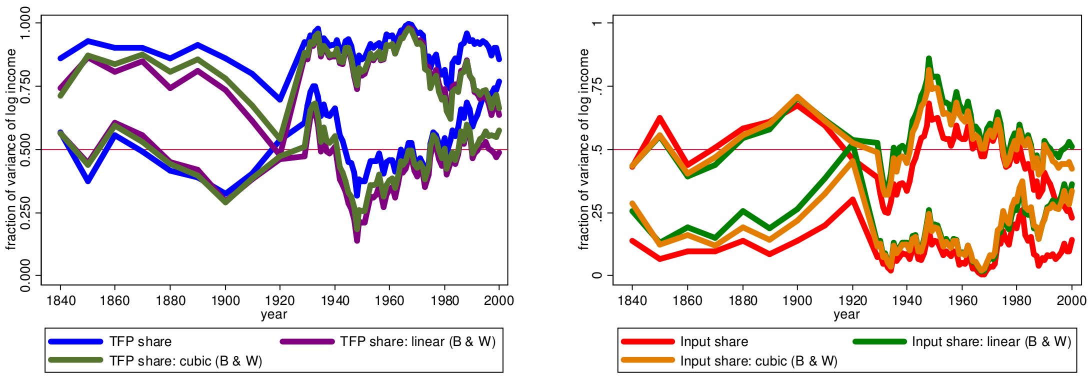



Figure 26: Returns to Schooling, Linear and Cubic Specifications



Figures 27 - 28: Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\sigma^2_{\ln A}$ and $\sigma^2_{\ln x}$




## APPENDIX A

| *New England* | *Middle Atlantic* | *South Atlantic* | *E. South Central* | *W. South Central* |
|---|---|---|---|---|
| Connecticut | New Jersey | Delaware | Alabama | Arkansas |
| Maine | New York | D.C. | Kentucky | Louisiana |
| Massachusetts | Pennsylvania | Florida | Mississippi | Oklahoma |
| New Hampshire | | Georgia | Tennessee | Texas |
| Rhode Island | | Maryland | | |
| Vermont | | North Carolina | | |
| | | South Carolina | | |
| | | Virginia | | |
| | | West Virginia | | |
| *Mountain* | *Pacific* | W. North Central | *E. North Central* | |
| Arizona | Alaska | Iowa | Illinois | |
| Colorado | California | Kansas | Indiana | |
| Idaho | Hawaii | Minnesota | Michigan | |
| Montana | Oregon | Missouri | Ohio | |
| Nevada | Washington | Nebraska | Wisconsin | |
| New Mexico | | North Dakota | | |
| Utah | | South Dakota | | |
| Wyoming | | | | |

| North | South | West |
|---|---|---|
| New England | South Atlantic | Mountain |
| Middle Atlantic | East South Central | Pacific |
| East North Central | West South Central | West North Central |

## APPENDIX B

Table B1: Fraction of State's Population - Black

| year | AL | AR | GA | LA | MD | MS | NC | SC | TN | VA |
|---|---|---|---|---|---|---|---|---|---|---|
| 1840 | .433 | .209 | .411 | .551 | .322 | .524 | .357 | .564 | .227 | 490 |
| 1850 | .447 | .227 | .425 | .506 | .283 | .512 | .364 | .589 | .245 | 471 |
| 1860 | .454 | .256 | .441 | .494 | .249 | .552 | .365 | .586 | .255 | 450 |
| 1870 | .477 | .252 | .460 | .501 | .225 | .536 | .366 | .589 | .256 | 419 |
| 1880 | .475 | .263 | .470 | .515 | .225 | .574 | .379 | .607 | .261 | 418 |
| 1890 | .448 | .274 | .468 | .500 | .207 | .576 | .347 | .599 | .244 | 383 |
| 1900 | .452 | .280 | .467 | .471 | .198 | .585 | .329 | .584 | .238 | 357 |
| 1910 | .425 | .281 | .451 | .431 | .179 | .561 | .316 | .552 | .217 | 325 |
| 1920 | .384 | .270 | .416 | .389 | .169 | .522 | .298 | .514 | .193 | 299 |
| 1930 | .357 | .258 | .368 | .369 | .169 | .502 | .290 | .457 | .183 | 268 |
| 1940 | .347 | .248 | .347 | .359 | .166 | .492 | .275 | .429 | .174 | 247 |
| 1950 | .320 | .223 | .309 | .329 | .165 | .453 | .258 | .389 | .161 | 221 |
| 1960 | .300 | .218 | .285 | .319 | .167 | .421 | .245 | .348 | .165 | 206 |
| 1970 | .262 | .183 | .259 | .299 | .178 | .368 | .222 | .305 | .158 | 185 |
| 1980 | .256 | .163 | .268 | .294 | .227 | .352 | .224 | .304 | .158 | 189 |
| 1990 | .253 | .159 | .270 | .308 | .249 | .356 | .220 | .302 | .160 | 188 |
| 2000 | .260 | .157 | .287 | .325 | .279 | .363 | .216 | .295 | .164 | 196 |

Table B2: Distribution of Black Population

| year | Fraction of Nation's Population - Black | Fraction of Black Population in South |
|---|---|---|
| 1840 | .168 | .791 |
| 1850 | .157 | .806 |
| 1860 | .141 | .823 |
| 1870 | .127 | .807 |
| 1880 | .131 | .826 |
| 1890 | .119 | .817 |
| 1900 | .116 | .814 |
| 1910 | .107 | .807 |
| 1920 | .099 | .770 |
| 1930 | .097 | .707 |
| 1940 | .098 | .690 |
| 1950 | .100 | .602 |
| 1960 | .106 | .523 |
| 1970 | .111 | .451 |
| 1980 | .118 | .454 |
| 1990 | .123 | .453 |
| 2000 | .123 | .473 |