Explaining Productivity Differences: the Importance of Black – White Schooling Differences

Chad Turner Robert Tamura Sean Mulholland Scott Baier*

Abstract

Using a newly created data set 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 (2007b) analyzed the growth rates of aggregate inputs and total factor producvity (TFP). Given that instiutional differences across states are likely to be smaller than those observed across countries, the result that variation in TFP still accounts for the lion's share of variation in output growth is somewhat surprising. We continue this line of work by documenting the importance of TFP in explaining cross sectional variation in the levels of log income. We also consider the possibility that one major institutional difference across states, slavery, might explain TFP differences across states. To this end, we create and present a new state level measure of years of schooling by race from 1840 - 2000. Again, surprisingly, exploiting this series has very little impact on the upper bound of the fraction of income variation that can be explained by inputs. Thus as in Lucas (1988) and Prescott (1998), the mystery of TFP remains. We do, however, find support for external effects of physical and human capital, as suggested by Romer (1986,1990), Lucas (1988), and Tamura (2002,2006).

Past cross-country studies indicate that variation in TFP accounts for a large fraction of the cross-sectional variation in output, whether levels or growth rates are analyzed. In searching for explanations of the TFP variation across countries, researchers often appeal to institutional heterogeneity as an explanation. Using a newly created data set containing real output per worker,

*Chad Turner (Nicholls State University), Robert Tamura (Clemson University and Federal Reserve Bank of Atlanta), Sean Mulholland (Mercer University) and Scott Baier (Clemson University and Federal Reserve Bank of Atlanta). This paper prepared for the 4th Economic Development Conference at Clemson University in honor of Robert E. Lucas, Jr.'s "On the Mechanics of Economic Development." 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.

real physical capital per worker and human capital per worker at the state level for the United States from 1840 to 2000, Turner et. al (2007b), henceforth TTMB2, analyzed the growth rates of aggregate inputs and total factor productivity (TFP). While noting institutions are not entirely homogeneous across state, TTMB2 suggest institutional differences across states are likely to be much less variable than the institutional differences across countries, and therefore expect input variation to account for a larger fraction of output variation than in the cross-country studies. Surprisingly, TTMB2 find that the vast majority of the cross-sectional variation in output growth rates can still accounted for by TFP differences, even across states.

In this work, we intend to extend this analysis on two main dimensions. First, we continue analyzing this dataset by examining the variation in the *levels* of income across states. Just as with the growth rates, the disposition of the correlation between aggregate input and TFP is a central issue in such a variance decomposition. We follow the methodology of Baier, Dwyer, and Tamura (2006) and TTMB2 to report upper and lower bounds that can be explained by TFP and inputs.

Second, while institutions are likely to be homogeneous across states, one major institutional difference that did exist in the United States was the institution of slavery. We follow the methodology of Turner et. al (2007a), henceforth TTMB1, and create years of schooling at the state level, by race, for 1840 - 2000. We subsequently incorporate black-white schooling differences into development accounting exercises, expecting that inputs will be able to account for a larger fraction of income variation. We find that incorporation of these measure makes little change in the upper and lower bounds of income variation that can be explained by TFP or inputs.

We organize the paper as follow. Section 2 describes the physical capital per worker, human capital per worker, output per worker measures created by TTMB1 and TTMB2. We find the accumulation of factors, physical capital per worker and human capital per worker explains about 60 percent of mean growth in output per worker. The Solow residual, or growth in TFP explains the remaining 40 percent of real output growth. In Section 3, we conduct development accounting exercises. We note, than in an attempt to make their results comparable to the cross-country literature, TTMB2 utilize parameters concerning the rates of return on the different types of schooling that may not be appropriate for United States. We examine the sensitivity of development accounting to these parameters. While the variance decomposition results are somewhat sensitive to parameters on years of schooling, we find that TFP can explain up to 80% of output variation, while inputs can explain up to roughly 55% of output variation. In Section 4, we examine how incorporating schooling measures that vary by race impacts the results of the development accounting. The overwhelming result is that incorporating this information makes very little difference in the variance decomposition In Section 5, we search for evidence of external effects on capital as suggested by Lucas (1988), Romer (1986, 1990), and Tamura (2002, 2006). We find evidence that TFP levels are correlated with the share of the population that has been exposed to higher education, lending empirical support for external effects. The last section concludes and outlines plans for future work.

DATA DESCRIPTION

Before we continue to present the growth accounting results, we summarize the measures of output per worker, year of schooling, and human capital from created in BMTT1 and the measures of capital per worker created in BMTT2, paying special attention to how the cross-sectional variation in inputs and output evolves across times. One thing that will be apparent is that the greatest income inequality arises in 1850, and like the other series, inequality typically declines for the remaining century and a half.

Output per Worker

Table 1 reports real output per worker by region created in TTMB1. While the details of the construction of the series are available in TTMB1, the data is a combination of work done by Richard Easterlin, BEA state personal income, and original output estimates from TTMB1 derived from estimates of manufacturing, agricultural, and non-agricultural non-manufacturing output at the state level.

As we are interested in placing our work in the context of the cross-country growth comparisons, we find it informative to provide information about size of the gap between the states and region with the highest value of output per worker and those with the lowest. To that end, Figure 1 graphically displays the state with the largest value, the state with the smallest value, the coefficient of variation, and the mean across states for each year.[1] Although the gap across regions is not as large as the gap between top and bottom countries in the world today, it is still quite substantial. In 1840 with 28 states, the typical worker in the top region was about 2.5 times more productive than a typical worker in the bottom region. In 1850 when the dataset expands to 34 states, this most productive region is 13 times more productive than the least. While this is not close to the

[1] When construct the average real output per worker for the US, we weight states by the size of their labor force. Whe we calculate the standard deviation of real output per worker across the states, we do without using weights.

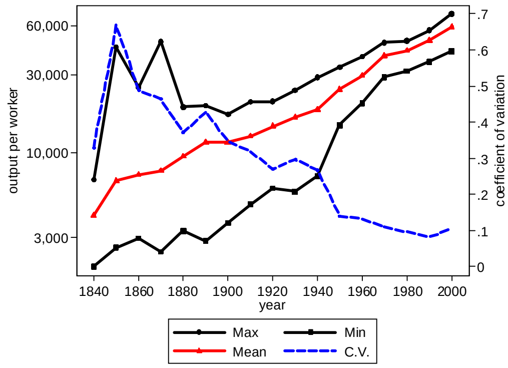
50 to 1 ratio observed in cross country data, it is still very large. The tremendous convergence across the regions is evident from the graphs. The coefficient of variation peaks in 1850 and falls essentially uninterrupted throughout. By 1950 the relative gap has stabilized to around 1.5 to 1, and by 1980 the gap had fallen to about 1.25 to 1, where it has remained. Figures 2 - 4 display the averages value of output per worker in each region. The incredibly high level of output per worker in the Mountain and Pacific regions early on is quite clear.[2]

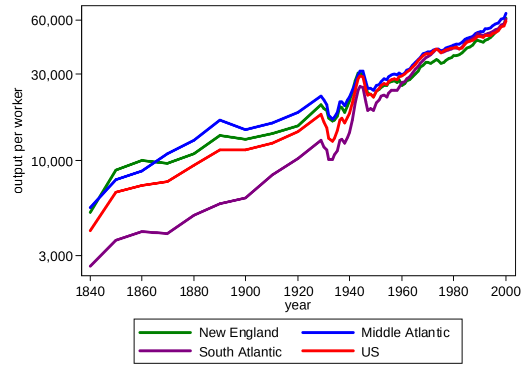
Table 1: Real Income per Worker, Labor Force Weighted

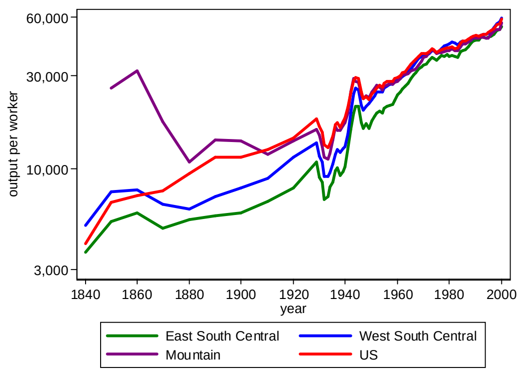
Year	NE	MA	SA	ESC	WSC	MTN	PAC	WNC	ENC	US
1840	5,267	**5,528**	2,342	3,683	5,042	.	.	3,503	4,540	4,114
1850	9,077	7,901	3,302	5,344	7,346	10,250	**43,207**	4,641	7,343	6,691
1860	9,999	8,840	3,647	5,928	7,503	12,606	**24,257**	5,760	7,484	7,297
1870	9,717	10,910	3,728	4,869	6,312	15,299	**16,500**	7,056	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

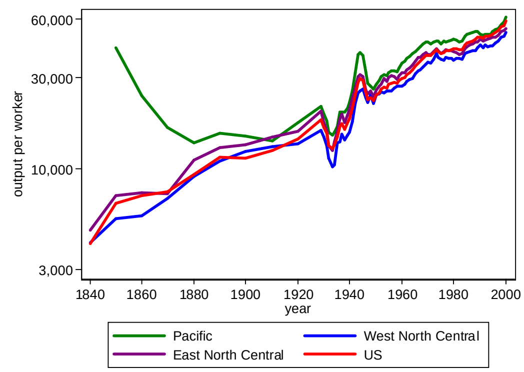
[2] For a detailed explanation of the calculations of state real output per worker see TTMB1. In particular, Appendix B containss an explanation how output from various sectors and sources were aggregated.

Figures 1 - 4









Physical Capital per Worker

Table 2 presents the reports the physical capital per worker measures created in TTMB2.[3]. It is useful to mention that physical capital per worker are not constructed directly through a perpetual inventory method. Rather TTMB2 uses data on national capital stocks for each sector and state level output in each sector.[4] For years after 1900, state level capital stocks are created by allocating BEA

[3] The detail of the construction of this series are outline in TTMB2.

[4] In the 1840-1920 period TTMB2 constructed output per worker from principally three sectors, agriculture, manufacturing, and all other sectors. For the 1962-2000 periods, gross state product (GSP) for each state for nine sectors are avaiable from the BEA. While the BEA does not provide estimates of GSP from 1929-1962, or from 1998-2000 using identical industry classifiction, 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 combing this data is state level output measures for 3 sectors from 1840 - 1920 and for 9 sectors from 1920 - 2000.

(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.[5] Prior to 1900, TTMB2 utilizes work from Gallman (1986) that includes capital-output ratios and sectoral shares, using the state level output data from TTMB1 and again assuming a constant capital-ouput ratio to allocate capital to each state.

Figure 5 displays the state with the largest and smallest values of physical capital per worker, the national average, and the coefficient of variation. Parallel to what is seen with income 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. In 1840 the top region had roughly 2.5 times more physical capital per worker than the bottom region. In 1850, the ratio of physical capital between top and bottom regions had peaked at 10 to 1. However by 1950 the ratio had fallen to 1.56, in 1980 it was 1.43, and in 2000 the ratio was 1.35. While the variation across regions of capital per worker is larger than variation of income per worker in 2000, it is of the same general magnitude. When raised to a power like .28, the ratio falls to 1.09 much less than 1.25, the income ratio. Figures 6 - 8 display the regional averages.

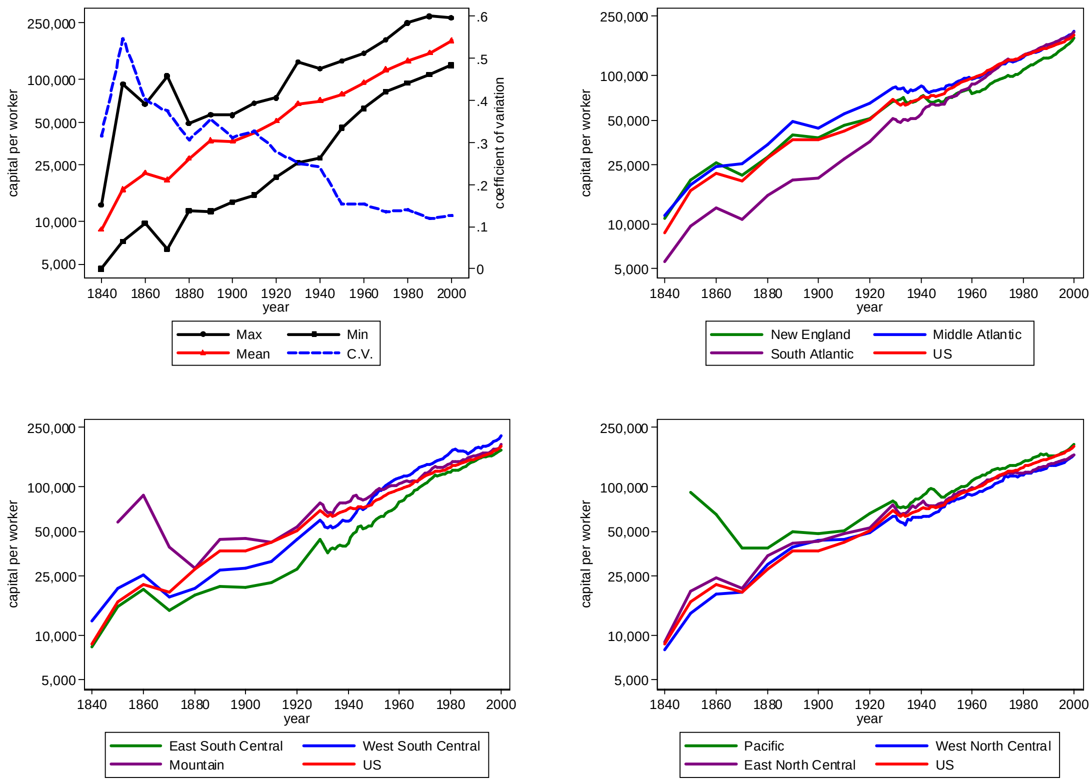
Gallman (1986) reports national measures of capital output ratios for 1840-1900 at the decadal frequency for six sectors as well as sectoral shares. This enables the amount of national capital in each sector to be calculated. For 1902 through 2000, data are provided by the Bureau of Economic Analysis 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.

[5] 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 factors are equalized across sectors within a state, because we are allowing capital output ratios to vary across sectors.

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

Figures 5 - 8: Physical Capital per Worker



Human Capital per Worker

While details are available in TTMB1, enrollment rate data was collected from variety of census reports, Reports of the School Superintendant in the Interior Department, Statistical Abstracts of the United States and Digest of Education Statistics. Using a perpetual inventory method, the fraction of the labor force exposed to each category of schooling was computed, as was the years of schooling conditioned on being in each educational category. Multiplying the fraction of the labor force in each category times the years of schooling in that category, then summing across educational categories results in the average years of schooling for each state, $\widetilde{E_{it}}$. TTMB1 consider 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 use to fit educational attainment of international migrants in each state. The result is a migration adjusted measure of years of schooling (E_{it}).[6] Table 3 presents years of schooling per worker for each region from TTMB1. While the typical worker in the United States had 1 year of formal schooling in 1840, the range was from .25 years to 2.5 years.[7]

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

Given years of schooling by state and average age ($avgage_{it}$) of the state population not enrolled in school, BMTT2 uses the standard labor economics transformation to construct a measure of

[6] In addition, the average age of those betwen 6 and 65 and not enrolled in school is calculated from census data.

[7] The 1840 value depends on the initial condition used. We assumed that the initial stock for each category was half the 1840 enrollment rate. If instead, the initial stock was equal to the enrollment rate for each category in 1840, initial years of schooling is near 2 years.

experience (ex_{it}) in the labor market:

$$ex_{it} = avgage_{it} - 6 - E_{it} \tag{1}$$

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

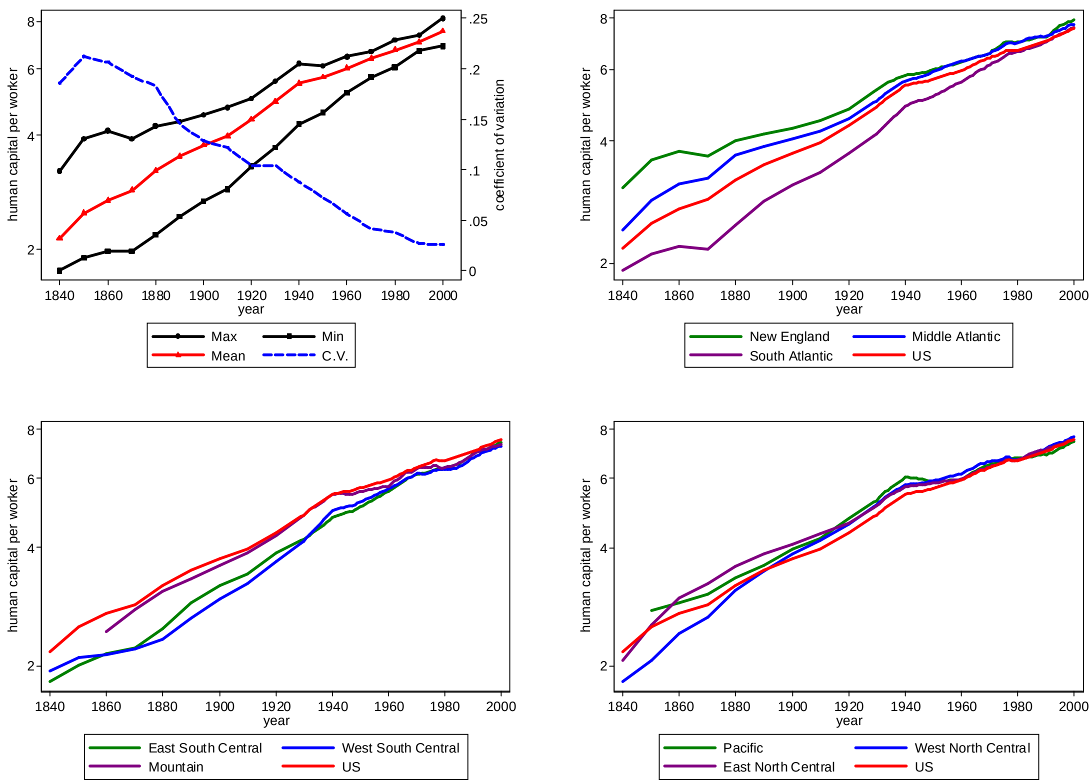
$$h = h_0 \exp\left(\phi_P E_P + \phi_I E_I + \phi_S E_S + \gamma_1 ex - \gamma_2 ex^2\right)$$

where h_0 is the level of human capital with no schooling or experience, ϕ_P, ϕ_I, and ϕ_S are parameters on years of primary, intermediate, and secondary and higher education, $\gamma_1$and γ_2 are parameters on experience, ex, and experience squared.[8] 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 10 - 13. Figure 10 contains the average human capital in the United States, the maximum and minimums over the period as well as the coefficient of variation of human capital. Similar to the previous two series, maximum dispersion occurs in 1850 and diminishes relatively continuously for the remaining 150 years. Figures 11 - 13 display the measures of human capital per worker for each census region.

[8] 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 must be completed to attend higher education. If average years of schooling is less than 4, P = E, I = 0, S = 0. If average years of schooling is between 4 and 8, P = 4, I = E – 4, S = 0. Finally, if average years of schooling is greater than 8, P = 4, I = 4, S = E – 8.

Figures 10 - 13: Human Capital per Worker

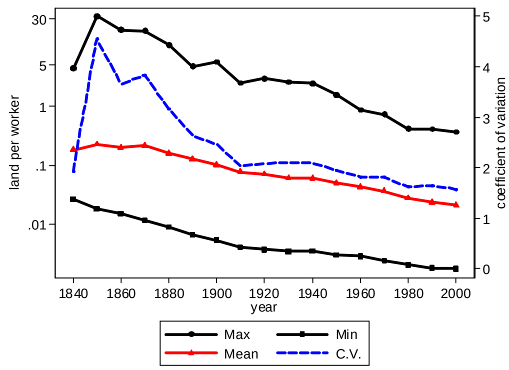


Land per Worker

Finally we present evidence on the reduction of land per worker in the United States.[9] For brevity we do not report the land per worker measures by region.

[9] We debated how to enter land as an input into production. When we ignored land as an input entirely, TFP growth is negative or 0 for 1840-1940. We viewed this as evidence that we had completely captured all productive inputs accurately so we felt that land per worker was a useful input. When we searched for the real price of land by state for each year we were unable to come up with a suitable series. While we did find some estimates, placing those in made the TFP growth problem worse. We ended up with including land per worker assuming a constant unit price per land.

Figure 14: Land per Worker



Growth Rates

Table 4, reproduced from TTMB2, presents the growth of output per worker, y, physical capital per worker, k, human capital per worker, h, land per worker, ℓ, and TFP, a. The assumptions made are that physical capital receives 28 percent of output, land receives 5 percent of output, and human capital receives the remaining 67 percent of output. We report results for the nation as a whole, then group the states into the nine census regions, and final group states into three broader regions.[10] 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 0.56 percent per year. The 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 valuations 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. One contribution of this data set is that, 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.

[10] A list of states contained in each region and each broader region (North , South, and West) is available in Appendix A.

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

	y	k	h	ℓ	a	a / y
All Regions	1.45%	1.66%	0.78%	-2.23%	0.56%	0.385
NE	1.46%	1.70%	0.60%	-1.34%	0.64%	0.441
MATL	1.52%	1.76%	0.75%	-1.79%	0.62%	0.404
SATL	1.92%	2.15%	0.86%	-1.79%	0.82%	0.430
ESC	1.66%	1.89%	0.88%	-1.33%	0.61%	0.367
WSC	1.58%	1.72%	0.84%	-2.02%	0.63%	0.397
MTN	1.26%	1.55%	0.73%	-3.29%	0.41%	0.330
PAC	0.49%	0.69%	0.73%	-3.99%	0.01%	0.021
WNC	1.33%	1.40%	0.83%	-2.26%	0.50%	0.372
ENC	1.44%	1.68%	0.79%	-2.19%	0.54%	0.377
North	1.46%	1.70%	0.70%	-1.74%	0.60%	0.410
South	1.77%	1.98%	0.86%	-1.73%	0.72%	0.408
West	1.16%	1.36%	0.77%	-3.02%	0.38%	0.325

Table 5 examines growth rates using data only from 1880 to 2000. This reduces the impact mining in the early years of western states. It 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. One immediate consequence of the shorter horizon is that output per worker growth increases from 1.45 percent per year to 1.58 percent per year. The rising labor force participation rate of the population accounts for the remaining .2 percent per year difference between real per worker growth and the familiar 1.8 percent per capita real growth.

In comparing using 1880 as an initial condition to using 1840, 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 38.5 percent to 44.8 percent.[11] Notice that two of the broader regions, the North and the West are quite similar in their growth rate of output per worker, TFP growth and share of growth arising from TFP growth. The three regions of the South, South Atlantic, East South Central and West South Central, have real output per worker

[11] It is quite likely we are not capturing the capital value of the ore not mined in the early years of the data. If we had better measures of estimates 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 positive for these regions as the capital value of the ore would be slowly depleted with extraction.

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 is similar to the rest of the nation; 45 percent in the South 46 percent in the North and 44 percent in the West.

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

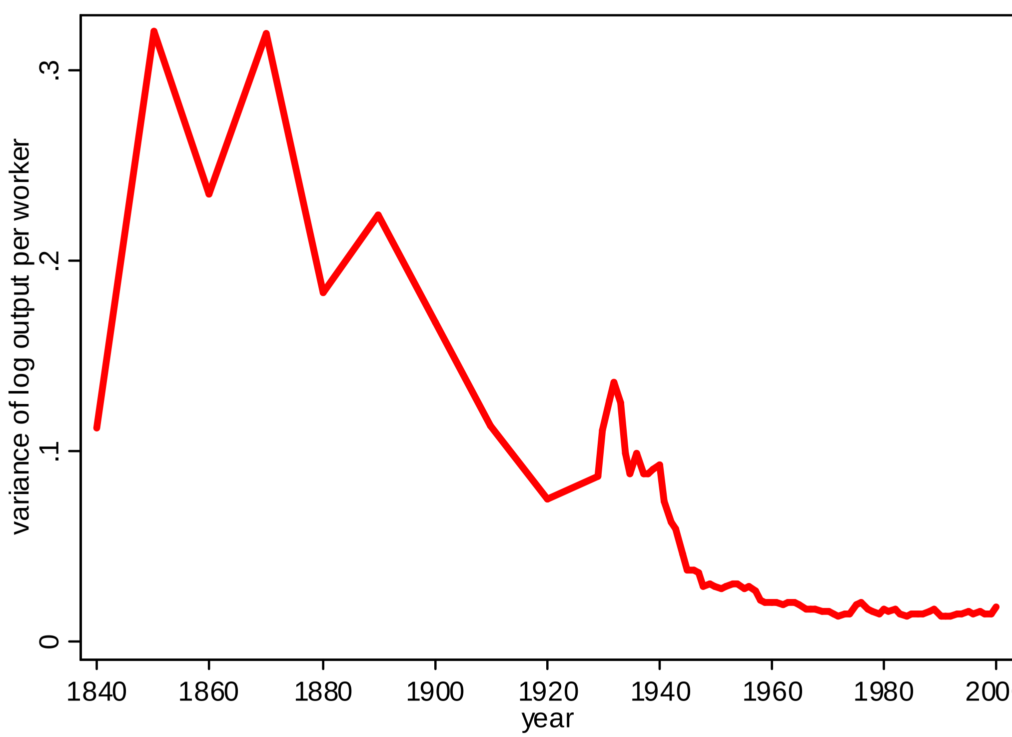
	y	k	h	ℓ	a	a / y
All Regions	1.58%	1.62%	0.75%	-1.80%	0.71%	0.448
NE	1.41%	1.49%	0.57%	-1.20%	0.67%	0.475
MATL	1.34%	1.46%	0.63%	-1.48%	0.58%	0.436
SATL	2.04%	2.07%	0.91%	-1.85%	0.94%	0.460
ESC	1.92%	1.87%	0.91%	-1.21%	0.85%	0.440
WSC	1.87%	1.88%	0.93%	-1.58%	0.80%	0.428
MTN	1.37%	1.57%	0.71%	-2.90%	0.59%	0.432
PAC	1.37%	1.42%	0.68%	-3.26%	0.67%	0.494
WNC	1.41%	1.35%	0.76%	-1.28%	0.58%	0.413
ENC	1.33%	1.31%	0.63%	-1.56%	0.62%	0.467
North	1.37%	1.42%	0.60%	-1.39%	0.64%	0.464
South	1.97%	1.97%	0.91%	-1.62%	0.88%	0.448
West	1.39%	1.46%	0.73%	-2.29%	0.60%	0.435

DEVELOPMENT ACCOUNTING

In this section we examine the determinants of cross sectional income differences. Unlike previous work, we use a technique in Baier, Dwyer and Tamura (2006) and applied in TTMB2. 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 log of output per worker. The key feature in this type of analysis is the disposition of the covariance term between TFP and inputs in the analysis. The typical manner employed is to assign half the covariance term to each of the variances terms and compare this augmented variance term with the variance of the dependent variable in question. 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 growth should be assigned to factor input growth. Our solution here is let the data provide the answer to which of the two factors, inputs and TFP, account for the observed variance in output.

The variance of log output per worker across the states over time is presented in Figure 15. 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.

Figure 15: Variance of Log Output per Worker ($\sigma^2_{\ln y}$)



We assume that output is produced using a Cobb-Douglas technology combining land, physical capital and human capital to produce output. In per worker terms we start with:

$$y_{it} = A_{it}\ell_{it}^{\sigma}k_{it}^{\alpha}h_{it}^{1-\alpha-\sigma} \tag{2}$$

We being by assuming the following parameter values: $\sigma = .05$, $\alpha = .28$. We then combine all the factor inputs into a single term, x. Thus output can be represented by:

$$y_{it} = A_{it}x_{it} \tag{3}$$

Taking logs produces:

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

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. Thus we have

$$\sigma_{\ln y}^2 = \sigma_{\ln x}^2 + 2\sigma_{\ln x,\ln A} + \sigma_{\ln A}^2 \tag{5}$$

Dividing by $\sigma_{\ln y}^2$, using the definition of covariance, and rearranging terms results:

$$1 = \frac{\sigma_{\ln x}^2}{\sigma_{\ln y}^2} + \frac{\sigma_{\ln A}^2}{\sigma_{\ln y}^2} + \frac{2\sigma_{\ln x}\sigma_{\ln A}}{\sigma_{\ln y}^2}\rho_{\ln x,\ln A} \tag{6}$$

where $\rho_{\ln x,\ln a}$ is the correlation between lnx_{it} and $\ln A_{it}$.If TFP and aggregate inputs are uncorrelated, the first term is the fraction of variance of log income caused 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 of course, theoretically it should not be.

The first alternative in dealing with the covariance term assumes that all differences in log output per worker that could be predicted by log TFP are due to log TFP, or, assumes the correlation 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.

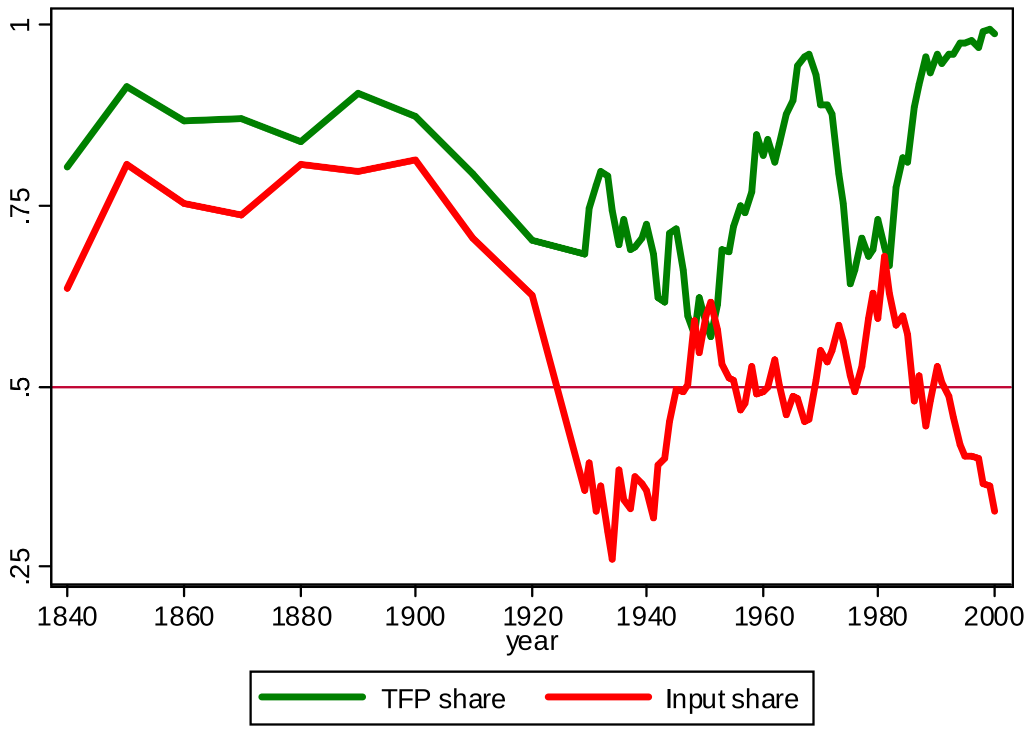
$$\frac{\left(1-\rho_{\ln x,\ln A}^2\right)\sigma_{\ln x}^2}{\sigma_{\ln y}^2} + \frac{\left(\sigma_{\ln A} + \sigma_{\ln x} * \rho_{\ln x,\ln A}\right)^2}{\sigma_{\ln y}^2} = 1 \tag{7}$$

The second alternative assumes that all variation in log output that are predictable by log inputs are due to log inputs, or stated differently, that 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.

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

Using the data from 1840-2000, we can compute these relative upper bounds for both log TFP and log inputs. We display the results in Figure 16. 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 effects of the Great Depression and World War II are apparent. Clearly input growth like schooling may not be much impacted by the Great Depression. On the one hand the opportunity cost of schooling fell dramatically, while on the other hand income dropped dramatically and leisure is a normal good. 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. With full mobilization during World War II, one would expect that factor returns would rise dramatically during this period as there would be no excess capacity. Still, it is quite interesting that 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 80 percent, while the comparable figure for variation in log inputs is only 50 percent.

Figure 16: Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\ln(A_{it})$ and $ln(x_{it})$



BMTT2 chose the parameters on years of schooling primarily to make their methodology comparable to the cross-country literature. An immediate thought, however, is whether or not those parameters for returns to schooling are appropriate for a cross-state comparison. There are two criticisms of these parameters for our purposes. First is it plausible that diminishing returns to schooling are not as rapid as assumed. In the specification above, all years of schooling above eight years return only 6.8 percent per year. According to the parameters, a high school graduate would earn only 31 percent more than an eighth grade graduate, and similarly, a four year college graduate would ear only 31 percent more than a high school graduate. In 2003 median income of a household with a head with less than 9th grade education was $18,787, while median income of a high school graduate household was $36,835.[12] These figures imply an annual return of 16.8 percent per year. The median household income of a college graduate household is $68,728. This implies an annual return of 15.6 percent.[13] This suggests that returns do not decline nearly as precipitously as we

[12] Data comes from Table 675 of the *Statistical Abstract of the United States: 2006.*

[13] Using data from Table 681 of the same statistical abstract but for families instead of households produces the

have supposed. 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. We suspect the returns to higher levels of education are likely to be more variable across time than the returns to elementary schooling. If so, when average schooling is below 4 years, this issue is irrelevant, and while average schooling is below 8 years, it may be relatively harmless. But certainly since 1940, when average schooling in the US exceeds 8 years it is likely, the variability across time is more likely to be a problem. As noted by Goldin and Margo (1992), and Freeman (1976) there have been times of compression, and Murphy and Welch (1992) note times of divergence. It is beyond the scope of this paper to develop time varying rates of return to schooling years, but we plan to examine this issue in future research.[14]

Returning to the concern with the assumption of a diminishing return to schooling, we reconstruct human capital assuming a constant 13.4 percent return per year of schooling. We then repeat the variance decomposition exercises. The results are shown in Figure 17, as are the results from the baseline 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. Of course there are many possibilities, perhaps an even greater rise in return to college after 1980 could pick up what the constant rate of return fails to capture. Table 6 reports the average share of the variance of log output explained by both log TFP and log inputs for various subperiods. The assumption of constant rate of return significantly increases the share of log output per worker variance explained by variance of log inputs to 58 percent, while the share explained by variance of log TFP falls slightly to 77 percent.

---

following statistics, less than 9th grade median income, \$25,313, high school graduate median income, \$44,620 and college graduate only median income, \$81,094. These imply annualized returns of 14.2 percent and 14.9 percent respectively.

[14] We do some exploration of this issue in Section 4 when looking at black-white schooling differences.

Figure 17: Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\ln(A_{it})$ and $ln(x_{it})$

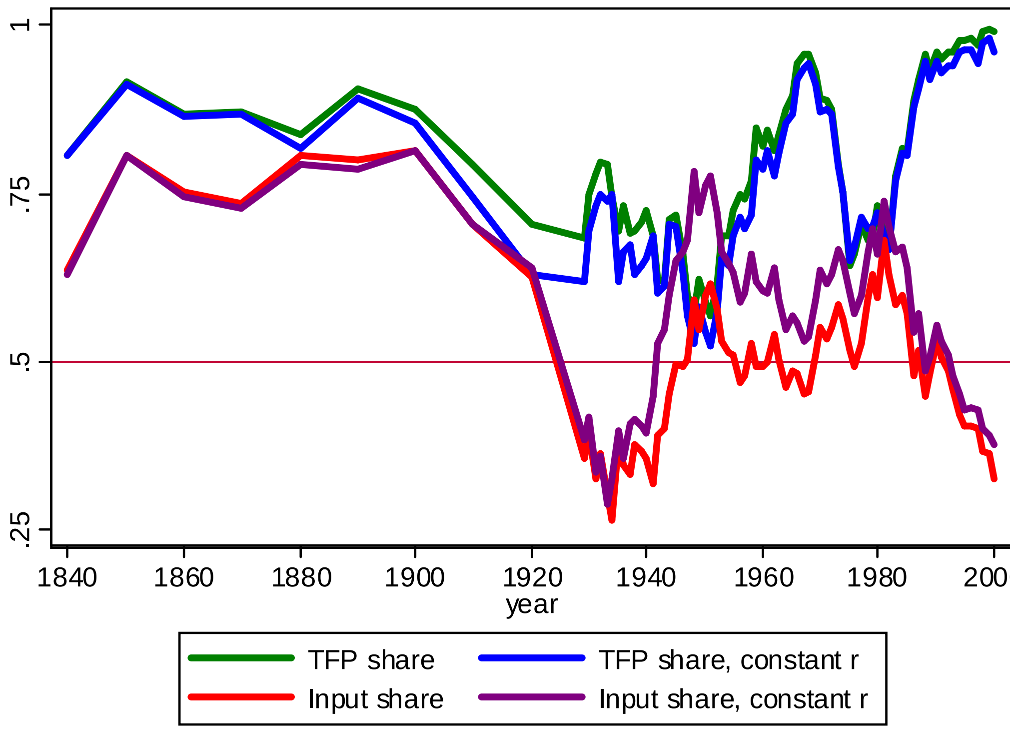


Table 6: Averate Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\ln(A_{it})$ and $ln(x_{it})$

	all years	excl. 1929 - 1949	1929 - 2000	1950 - 2000	1980 - 2000
TFP - diminishing r	.797	.832	.791	.830	901
TFP - constant r	.772	.813	.766	.812	.887
Inputs - diminishing r	.506	.544	.476	.509	487
Inputs - constant r	.576	.609	.556	.587	531

DO BLACK WHITE SCHOOLING DIFFERENCES EXPLAIN THE VARIATION?

In this section we examine the importance in variation in schooling between whites and blacks for capturing variation in real output per worker. 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. For levels, see Hall and Jones (1999) and for growth rates, see Baier, Dwyer and Tamura (2006). After establishing the importance of variance of TFP or variance of growth of TFP, most economists comment on the importance of institutional differences. Differences in property right enforcement, 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. See Baier, Dwyer and Tamura (2007) for a summary.

The unique feature of the state data being analyzed here is that most of the plausible explanations seem less important or unavailable. All states trade in the common currency once private circulating money disappeared and all workers typically communicated using the same language. There is a common international trade policy, interstate trade policy, and fiscal policy. We examined the relative size of state governments and found very little practical variation. The most obvious difference in institutions across states was the existence of slavery in the United States from its founding until the end of the Civil War in 1865. Prior to the end of slavery it was illegal in slave states to educate a slave. After the end of Reconstruction a system of disenfranchisement of blacks occurred throughout the South. Our earlier measures of schooling assumed that all people within a state are treated equally, irrespective of race from 1840-2000. This is clearly not the case. Using census data on enrollments by race we can construct measures of schooling by race.

Even if there are schooling differences by race, these difference are unlikely to have an impact unless a significant portion of that state's population is black. Table 7 presents the fraction of state population that is black in each of the former Confederate states. We note that from 1840-1920, blacks were the majority population in several states. Table 8 reports the fraction of black population in the nation as a whole and the fraction of black population that is living in a Southern state.

Table 7: 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 8: 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

We then proceed to construct years of schooling by race - black and white.[15] We follow the methodology of BMTT1, with only slight modifications.[16,17] Data on the racial composition of enrollment rates, labor force, educational attainment, and population are acquired from census records available through IPUMS.

We report these result in Table 9. Not surprisingly, there are meaningful differences between blacks and white. Amongst the black population, the New England census region has the highest

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

[16] Those familiar with the methodology in BMTT1 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. In the algorithm to correct for migrations for blacks, the share of the population that is foreign born is 0.

[17] 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.

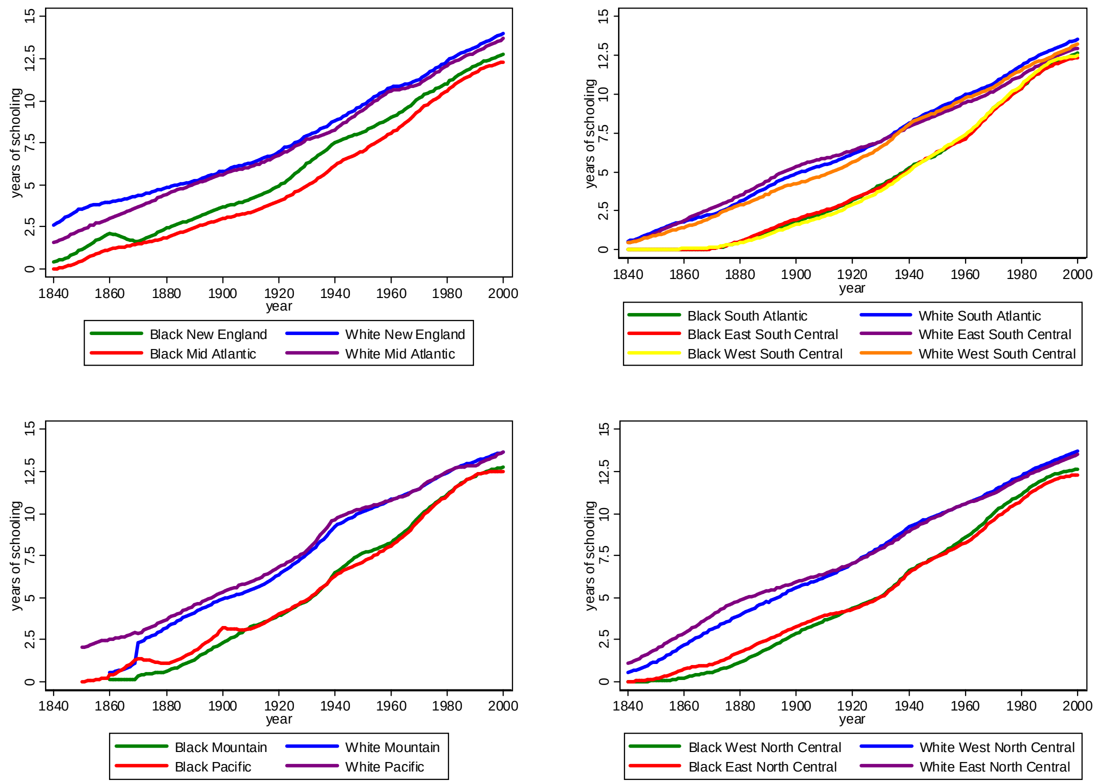
average years of schooling throughout nearly the entire period, but blacks in New England trail whites by 1.25 years even in 2000. Figures 18 - 20 display black and white schooling graphically for each of the census regions. We find there is quite a bit of similarity between the groupings of census regions as displayed. For example, the years of schooling of whites in each census region are similar in the South, as are the black years of schooling of black in each census region in the South. The largest gaps between black and white schooling for the nation as a whole is 3.5 years and occur in 1900 and 1910.

Table 9: Years of Schooling per Worker, 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

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

Figures 18 - 20: Years of Schooling, by race



We now turn our attention to what impact the years of schooling may have on input and TFP calculations. We consider, initially, two possible calculations. The first is to assume that black and white workers accumulate human capital in the fashion first assumed above with diminishing returns to additional years of schooling. We then construct human capital for each race.[18] We assume that workers produce together in a Cobb-Douglas function in which labor is perfectly substitutable measured in efficiency units. Essentially, we assume that the production technology is given by:

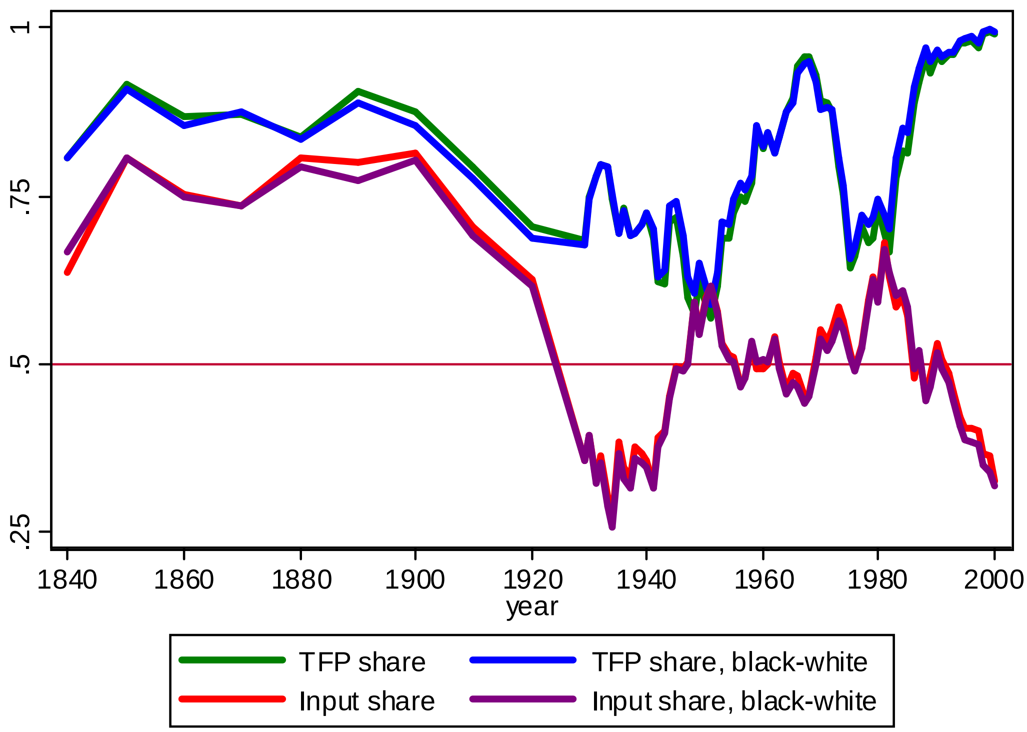
$$y_{it} = A_{it} \ell_{it}^{\sigma} k_{it}^{\alpha} (s_b h_{ibt} + (1 - s_b) h_{iwt})^{1-\alpha-\sigma} \tag{9}$$

We then reexamine the results of the log levels variance decomposition, displaying the upper bounds of log output per worker variance that can be explained by input and TFP in Figure 21. For this,

[18] 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.

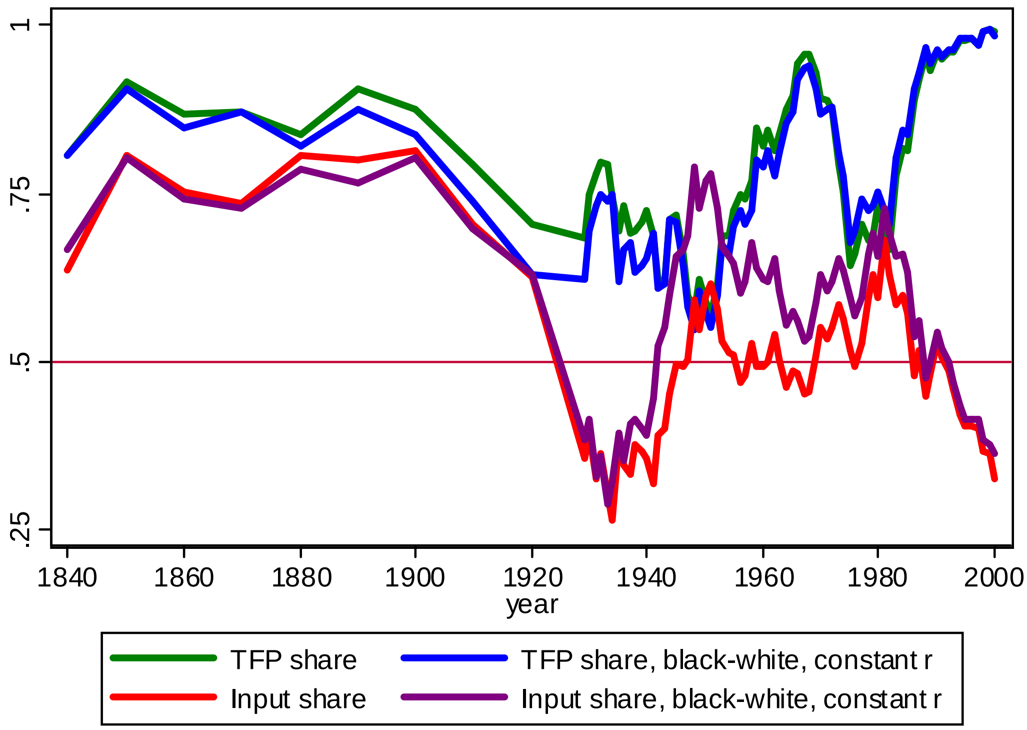
and subsequent versions of the decomposition, we always display the results from the baseline case. Shockingly taking account of black-white schooling differences has no effect at all on explaining the variation in log output per worker!

Figure 21: Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\ln(A_{it})$ and $ln(x_{it})$



We next recompute human capital under the assumption of a constant rate of return to schooling for whites and blacks. The results of the variance decomposition are those for the baseline version are show in Figure 22. The results are similar to those obtained with the constant rate of return without distinction of race. Excluding the Great Depression years and the years after 1982 appear to make inputs almost as important as TFP.

Figure 22: Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\ln(A_{it})$ and $ln(x_{it})$



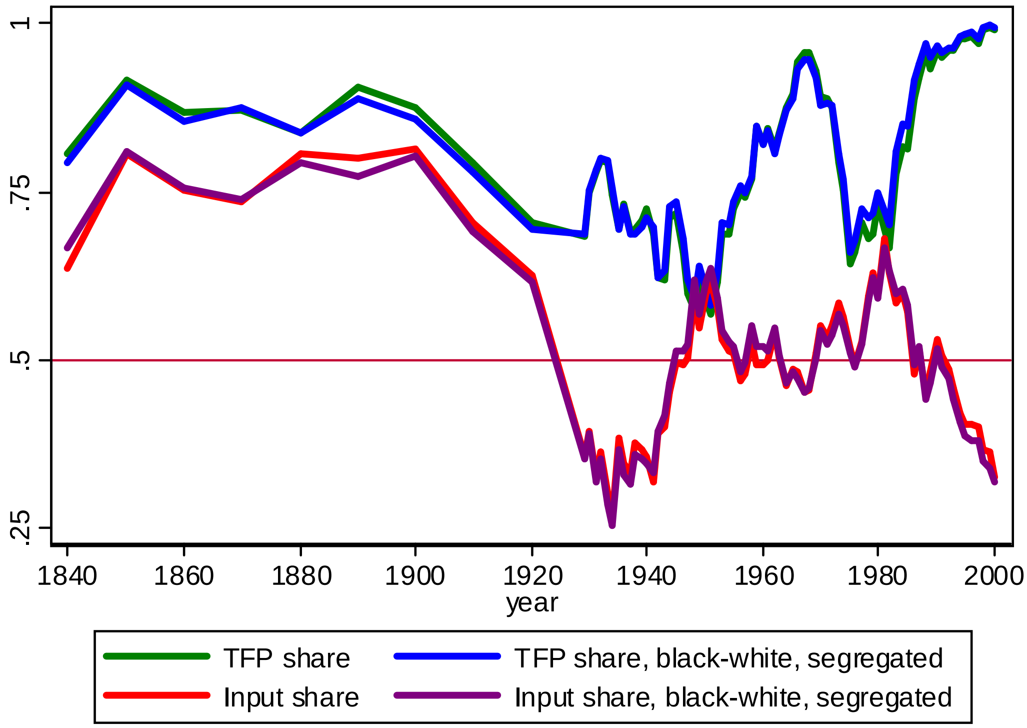
One other possibility is that black and white workers work separately and are segregated by educational attainment as well. We assume that there are four educational types, no education, 0, primary schooling but no more, $0 < E \leq 8$, secondary schooling but no more, $9 < E \leq 12$, and college exposed $E > 13$. Thus workers are setup into 8 different categories. We assign land per worker and physical capital per worker in order to maximize output per worker. Furthermore we assumed that all categories have the same TFP. Ignoring state subscripts, production in each state is given by:

$$y_t = \left\{ \begin{array}{c} s_b \left[s_{0b}y_{0b} + s_{1b}y_{1b} + s_{2b}y_{2b} + (1 - s_{0b} - s_{1b} - s_{2b})y_{3b} \right] \\ +(1 - s_b) \left[s_{0w}y_{0w} + s_{1w}y_{1w} + s_{2w}y_{2w} + (1 - s_{0w} - s_{1w} - s_{2w})y_{3w} \right] \end{array} \right\} \quad (10)$$

where s_{0i} =share of i group with no schooling, s_{1i} =share of i group with at least some primary schooling, but no more, s_{2i} =share of i group with at least some secondary schooling, but no more, and $s_{3i} = 1 - s_{0i} - s_{1i} - s_{2i}$ =share of i group with higher education exposure, and $y_{ji} = A\ell^{\sigma}_{ji}k^{\alpha}_{ji}h^{1-\alpha-\sigma}_{ji}$, for i =black or white.

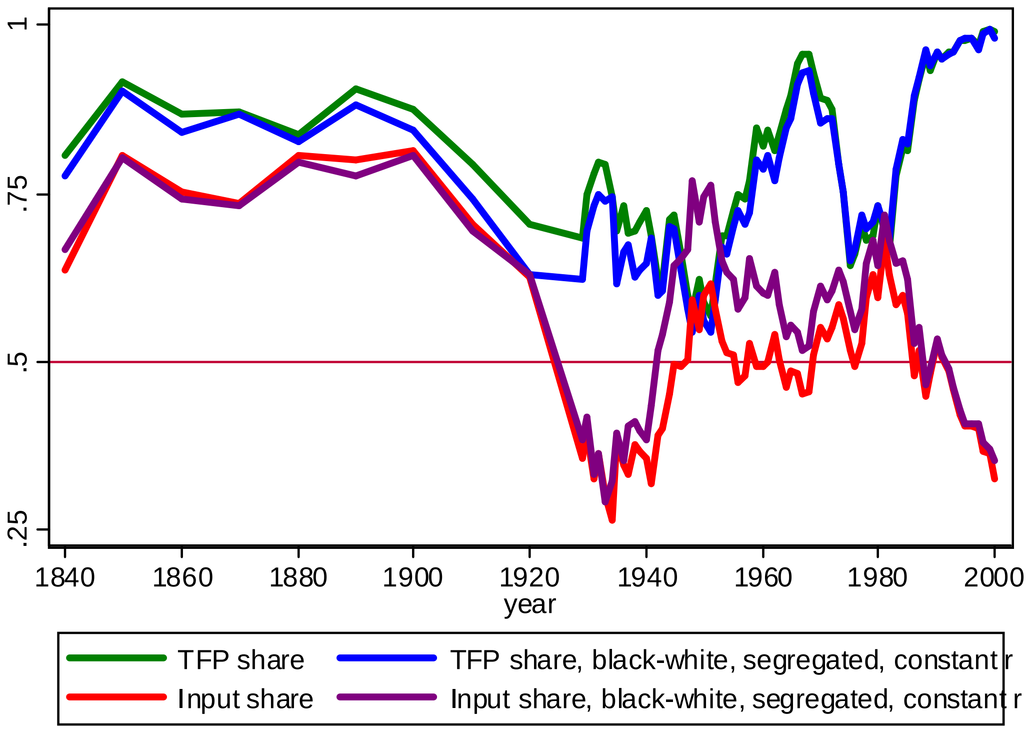
The results are contained in Figure 23. There is essentially no effect of assuming segregated production.

Figure 23: Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\ln(A_{it})$ and $ln(x_{it})$



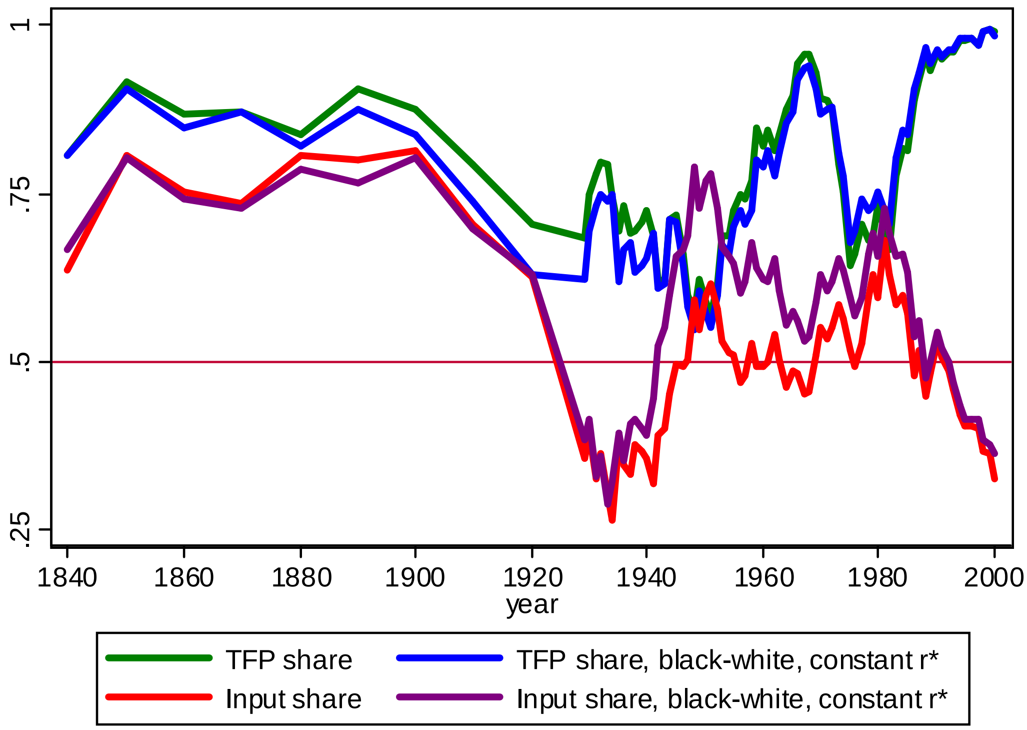
The next possibility is that there are constant returns to schooling but segregated production. This produces the expected result of increasing the amount of variance of log output per worker that can be explained by inputs.

Figure 24: Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\ln(A_{it})$ and $ln(x_{it})$



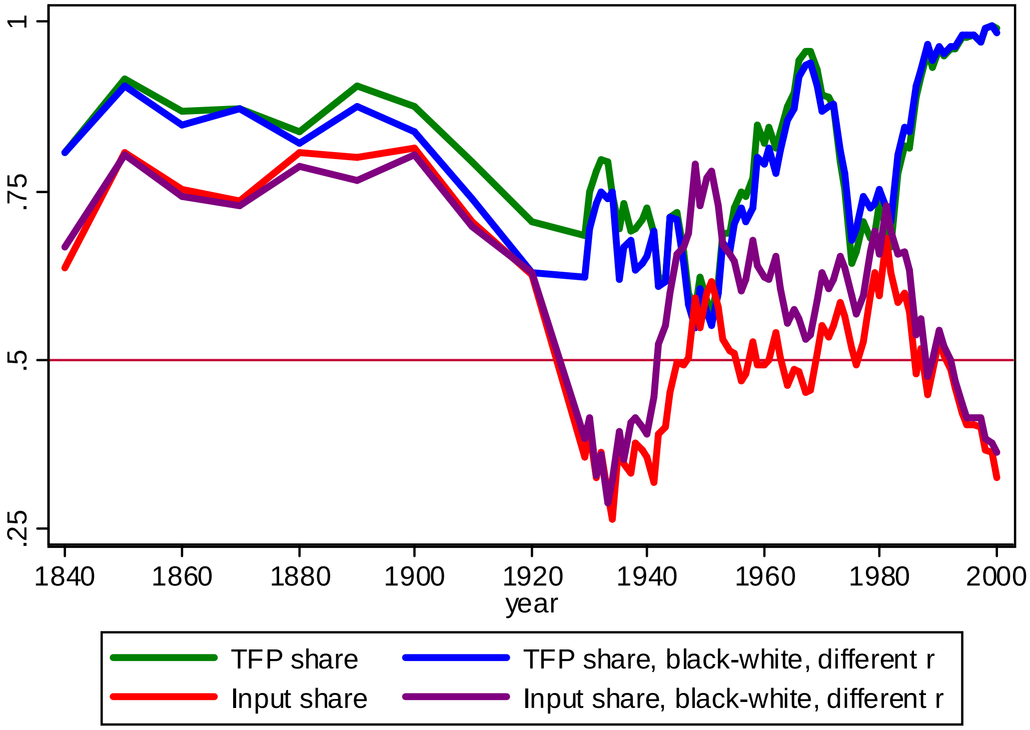
One additional set of tests is to allow the rate of return to higher education increase after 1979. We assume that the return increases to 0.155 for all years after 11 and maintain the assumption that all prior years have rates of return equal to 0.134. Because there is little evidence of improvement from segregated production, we consider the case of black and white schooling as perfect substitutes in production once measured in efficiency units. Figure 25 does not appear to indicate any large change in the explanatory ability of inputs.

Figure 25: Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\ln(A_{it})$ and $ln(x_{it})$



Next, we search for differential returns to schooling by race. We assume that whites receive a constant rate of return to schooling of 0.134, while blacks receive the diminishing returns from the original specification. The motivation would be to account for lower quality of schooling that blacks typically received throughout this period. Figure 26 contains the results of this specification. Again, there is little difference in the results.

Figure 26: Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\ln(A_{it})$ and $ln(x_{it})$



Finally, we again assume that black and white workers receive differential returns to schooling, but this time assume workers segregate from each other. These results are reported in Figure 27.

Figure 27: Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\ln(A_{it})$ and $ln(x_{it})$

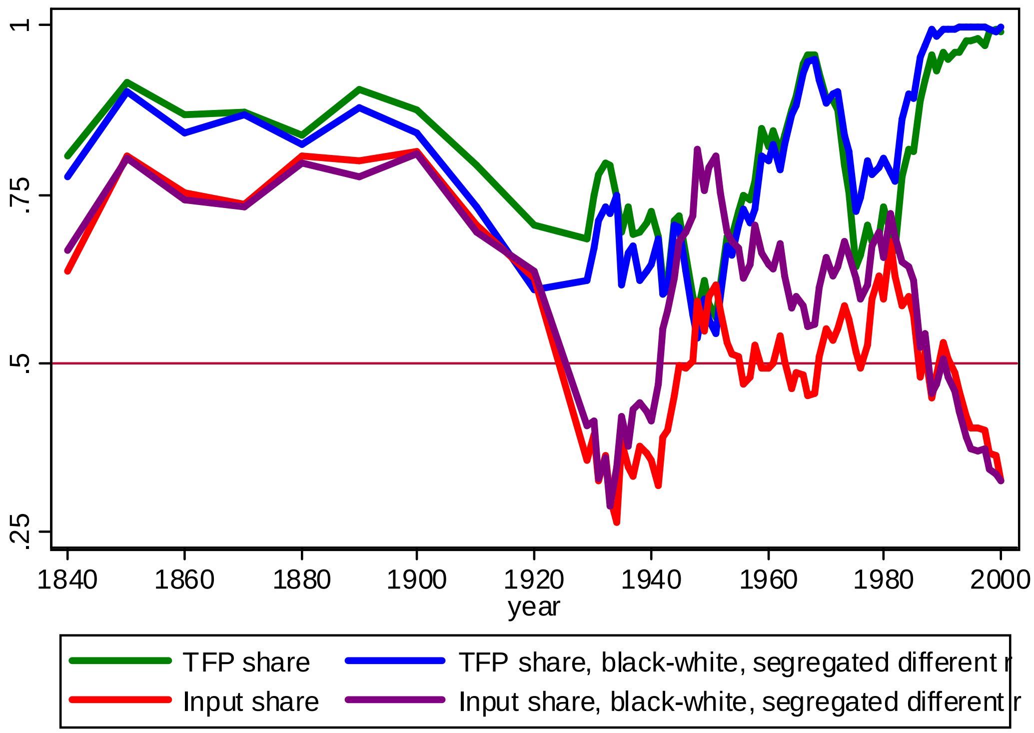


Table 10 summarizes the results of all of the specifications reported above. A constant rate of return to schooling is preferred to the diminishing returns to schooling specification of the original model. It increases the fraction of the log variance that is explained by the variation of log inputs by between .044 to .08. It decreases the share explained by variation in log TFP by .014 to .025. Surprisingly taking account of black white schooling differences does not appear to improve on this specification. However later results will show that there is explanatory power in using this information to capture log TFP variation itself.

Table 10: Upper Bound of Fraction $\sigma^2_{\ln y}$ Explained by $\ln(A_{it})$ and $ln(x_{it})$

Factor	Fig	r	b/w split	segreg.	1840-2000	excl. 1929-49	1929-2000	1950-2000	1980-2000
TFP	16	varies	no	n/a	.797	.832	.791	.830	.901
TFP	17	.134	no	n/a	.772	.813	.766	.812	.887
TFP	21	varies	yes	no	.805	.802	.840	.841	.916
TFP	22	varies	yes	yes	.803	.839	.799	.840	.916
TFP	23	.134	yes	no	.784	.827	.780	.829	.912
TFP	24	.134	yes	yes	.776	.772	.818	.819	.904
TFP	25	*	yes	no	.784	.827	.780	.829	.912
TFP	26	**	yes	no	.784	.827	.780	.829	.912
TFP	27	**	yes	yes	.796	.846	.795	.853	.947
Input	16	varies	no	n/a	.506	.544	.476	.509	.487
Input	17	.134	no	n/a	.576	.609	.556	.587	.531
Input	21	varies	yes	no	.500	.470	.538	.503	.481
Input	22	varies	yes	yes	.505	.543	.476	.508	.479
Input	23	.134	yes	no	.574	.607	.554	.584	.520
Input	24	.134	yes	yes	.564	.542	.595	.570	.511
Input	25	.155*	yes	no	.574	.607	.554	.584	.520
Input	26	**	yes	no	.574	.607	.554	.584	.520
Input	27	**	yes	yes	.582	.610	.562	.587	.492

*indicates that $r = .134$, except $r = .155$ for those with $E \geq 12$ after 1970

**indicates that $r_w = .134$, while r_b varies.

Of course a constant rate of return to schooling changes the growth accounting as well. The results of this are contained in Table 11.

Table 11: Average Growth Rates: 1880 (or when data becomes available) to 2000, black white constant r

	y	k	h	ℓ	a	a / y
All Regions	1.58%	1.62%	1.17%	-1.80%	0.43%	0.272
NE	1.41%	1.49%	0.99%	-1.20%	0.39%	0.277
MATL	1.34%	1.46%	1.02%	-1.48%	0.32%	0.239
SATL	2.04%	2.07%	1.22%	-1.85%	0.73%	0.358
ESC	1.92%	1.87%	1.27%	-1.21%	0.61%	0.318
WSC	1.87%	1.88%	1.29%	-1.58%	0.56%	0.299
MTN	1.37%	1.57%	1.13%	-2.90%	0.31%	0.226
PAC	1.37%	1.42%	1.07%	-3.26%	0.31%	0.226
WNC	1.41%	1.35%	1.13%	-1.28%	0.33%	0.234
ENC	1.33%	1.31%	1.02%	-1.56%	0.36%	0.271
North	1.37%	1.42%	1.05%	-1.39%	0.34%	0.248
South	1.97%	1.97%	1.24%	-1.62%	0.66%	0.335
West	1.39%	1.46%	1.15%	-2.29%	0.32%	0.230

EXTERNAL EFFECTS OF CAPITAL

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 possible explanation. Following Lucas (1988) and Romer (1986, 1990) and Tamura (2002, 2006) we consider the possibility that log TFP is a function of human capital and physical capital. Table 12 presents the results of regressions of log TFP on log of physical capital and the share of workers exposed to higher education (*hishare*). We also consider specifications that include *hishare* interacted with the appropriate race's share of the labor force (*slabor*).

Table 12: Log TFP Regressions - diminishing r

variable	$lnTFP$	$lnTFP$	$lnTFP$	$lnTFP$	$lnTFP$	$lnTFP$
ln(k)	0.3075	0.2853	0.2968	0.3753	0.3489	0.4021
	(0.1065)	(0.0962)	(0.1128)	(0.0698)	(0.0722)	(0.0852)
$hishare$	0.5426			0.6791		
	(0.0670)			(0.1144)		
$hishare_w$		1.0602			0.6612	
		(0.4228)			(0.3759)	
$hishare_b$		-0.7335			-0.2106	
		(0.2944)			(0.2806)	
$hishare_w * slabfor_w$			0.4396			0.6759
			(0.1091)			(0.1023)
$hishare_w * slabfor_b$			1.4438			-0.0861
			(1.0435)			(0.8236)
$year$	0.0001	0.0010	0.0003	-0.0015	-0.0003	-0.0017
	(0.0016)	(0.0015)	(0.0015)	(0.0009)	(0.0010)	(0.0010)
NE				-0.0088	-0.0012	-0.0192
MA				0.0872***	0.0844***	0.0881***
SA				-0.0255**	-0.0297***	0.0088
ESC				-0.0782***	-0.0806***	-0.0508
WSC				-0.1460***	-0.1365***	-0.1272***
MTN				-0.2946***	-0.2590***	-0.2985***
PAC				-0.1305***	-0.0806***	-0.1288***
WNC				-0.1826***	-0.1699***	-0.1947***
N	4004	4004	4004	4004	4004	4004
R^2	.6156	.6403	.6210	.7157	.7104	.7166

Table 13: Log TFP Regressions - constant r

variable	$lnTFP$	$lnTFP$	$lnTFP$	$lnTFP$	$lnTFP$	$lnTFP$
ln(k)	0.3002	0.2761	0.2787	0.3722	0.3483	0.3947
	(0.1064)	(0.0951)	(0.1115)	(0.0678)	(0.0698)	(0.0826)
$hishare$	0.1958			0.3537		
	(0.0649)			(0.1232)		
$hishare_w$		0.9220			0.4924	
		(0.3959)			(0.3442)	
$hishare_b$		-0.8870			-0.3125	
		(0.2729)			(0.2363)	
$hishare_w * slabfor_w$			0.1116			0.3827
			(0.1187)			(0.0973)
$hishareb * slabfor_b$			1.4285			-0.3025
			(1.0718)			(0.7916)
$year$	-0.0008	-0.0002	-0.0006	-0.0026	-0.0017	-0.0029
	(0.0016)	(0.0014)	(0.0014)	(0.0008)	(0.0009)	(0.0008)
NE				-0.0106	0.0035	-0.0182
MA				0.0938***	0.0907***	0.0945***
SA				0.0025	0.0010	0.0288
ESC				-0.0487***	-0.0460**	-0.0260
WSC				-0.1184***	-0.1072***	-0.1042***
MTN				-0.2899***	-0.2535***	-0.2967***
PAC				-0.1320***	-0.0794**	-0.1347***
WNC				-0.1823***	-0.1621***	-0.1912***
N	4004	4004	4004	4004	4004	4004
R^2	.4390	.4918	.4527	.5957	.5964	.5978

Table 14: Log TFP Regressions - diminishing r

variable	$lnTFP$	$lnTFP$	$lnTFP$	$lnTFP$	$lnTFP$	$lnTFP$
$\ln(k)$	0.3047	0.2787	0.2958	0.3717	0.3490	0.4007
	(0.1058)	(0.0983)	(0.1125)	(0.0699)	(0.0760)	(0.0850)
$hishare * \exp hi$	0.0366			0.0430		
	(0.0048)			(0.0078)		
$hishare_w * \exp hi_w$		0.0602			0.0348	
		(0.0271)			(0.0241)	
$hishare_b * \exp hi_b$		-0.0414			-0.0071	
		(0.0187)			(0.0186)	
$hishare_w * slabfor_w$			0.0294			0.0434
$* \exp hi_w$			(0.0071)			(0.0065)
$hishare_w * slabfor_b$			0.0903			-0.0121
$* \exp hi_b$			(0.0686)			(0.0544)
$year$	0.0001	0.0010	0.0003	-0.0014	-0.0001	-0.0017
	(0.0016)	(0.0016)	(0.0015)	(0.0009)	(0.0010)	(0.0010)
NE				-0.0105	-0.0064	-0.0212
MA				0.0846***	0.0829***	0.0855***
SA				-0.0272***	-0.0311***	0.0090
ESC				-0.0793***	-0.0854***	-0.0500
WSC				-0.1457***	-0.1405***	-0.1262***
MTN				-0.2915***	-0.2647***	-0.2970***
PAC				-0.1280***	-0.0911***	-0.1276***
WNC				-0.1806***	-0.1739***	-0.1929***
N	4004	4004	4004	4004	4004	4004
R^2	.6178	.6367	.6221	.7153	.7077	.7165

Table 15: Log TFP Regressions - constant r

variable	$lnTFP$	$lnTFP$	$lnTFP$	$lnTFP$	$lnTFP$	$lnTFP$
ln(k)	0.2983	0.2695	0.2781	0.3704	0.3493	0.3941
	(0.1064)	(0.0971)	(0.1117)	(0.0682)	(0.0732)	(0.0827)
$hishare * \exp hi$	0.0146			0.0223		
	(0.0048)			(0.0084)		
$hishare_w * \exp hi_w$		0.0524			0.0249	
		(0.0252)			(0.0218)	
$hishare_b * \exp hi_b$		-0.0521			-0.0144	
		(0.0171)			(0.0154)	
$hishare_w * slabfor_w$			0.0083			0.0246
$* \exp hi_w$			(0.0078)			(0.0062)
$hishare_w * slabfor_b$			0.0092			-0.0237
$* \exp hi_b$			(0.0703)			(0.0519)
$year$	-0.0009	0.0000	-0.0006	-0.0025	-0.0016	-0.0029
	(0.0015)	(0.0015)	(0.0014)	(0.0008)	(0.0009)	(0.0008)
NE				-0.0115	-0.0005	-0.0193
MA				0.0924***	0.0899***	0.0931***
SA				0.0017	-0.0015	0.0290
ESC				-0.0493***	-0.0499***	-0.0255
WSC				-0.1183***	-0.1107***	-0.1037***
MTN				-0.2883***	-0.2592***	-0.2959***
PAC				-0.1306***	-0.0895***	-0.1340***
WNC				-0.1813***	-0.1663***	-0.1901***
N	4004	4004	4004	4004	4004	4004
R^2	.4404	.4870	.4529	.5955	.5936	.5978

Tables 12-15 show the evidence supporting the connection between installed capital and TFP. If ideas are embodied in capital, it serves to reason that there would be a positive correlation between capital and TFP. Notice that typically the share of the work force that is exposed to higher education is also positively related to TFP. We also interact share exposed to higher education with the expected number of years of schooling of those exposed to higher education. We reason that states with greater values of schooling indicate more professional and graduate education as well as a greater graduation rate from college. The evidence is quite suggestive of a link between human capital and physical capital and measured TFP.

Conclusion

Using 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, this paper examines the importance of variation in log per worker inputs and variation of log TFP for explaining 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. These upper bounds, derived in Baier, Dwyer and Tamura (2006), allow the data to allocate the covariance term in the variance decomposition informed by competing economic theories. Using the greatest institutional difference imaginable, the existence of slavery in southern states from 1840-1865, and the Jim Crow era of racial discrimination in public schooling from roughly 1876-1954, we produce new estimates of years of schooling for both white and black workers. Thus we show that institutional differences clearly show up in the accumulation of factors as well as directly in productivity. Surprisingly we find that the addition of this more detailed information on inputs does not generally improve the ability of log input variability for explaining log output per worker variability. We do find that the typical macroeconomic specification for the construction of human capital using diminishing returns to schooling, see Hall and Jones (1999) and Klenow and Rodriguez-Clare (1997), is not supported by the US experience. There is good evidence that the rate of return to schooling is constant, much more similar to that found by Card and Krueger (1992). 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 the data. Similarly there is evidence that the share of the work force exposed to higher education is correlated with higher productivity.

REFERENCES

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.

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. "The United States Capital Stock in the Nineteenth Century," in *Long-Term Factors in American Economic Growth, 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.

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.

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.

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.

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.

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, Mulholland, Sean and Baier, Scott. "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, 2007b.

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		